Complexity, Concentration and Contagion

Prasanna Gai^{*}, Andrew Haldane[†]and Sujit Kapadia[‡]

 $9 \ {\rm November} \ 2010$

Preliminary – Not for Quotation or Further Circulation

Abstract

This paper develops a model of the interbank network in which unsecured claims and obligations, repo activity and shocks to the haircuts applied to collateral assume centre stage. We show how systemic liquidity crises of the kind associated with the interbank market collapse of 2007-8 can arise within such a framework, with contagion spreading widely through the web of interlinkages. And we illustrate how greater complexity and concentration in the financial network may contribute to fragility. We then suggest how a range of policy measures – including tougher liquidity regulation, macro-prudential policy, and surcharges for systemically important financial institutions – may make the financial system more resilient.

Keywords: Network models; Contagion; Financial crises; Systemic risk; Liquidity risk; Interbank markets; Regulatory policy

JEL classification: D85; G01; G21; G28

Disclaimer: This paper represents the views of the authors and should not be thought to represent those of the Bank of England or Financial Policy Committee members.

^{*}Crawford School of Economics and Government, Australian National University, Canberra, ACT 0200, Australia. *E-mail address*: prasanna.gai@anu.edu.au.

[†]Bank of England, Threadneedle Street, London, EC2R 8AH, United Kingdom. *E-mail address*: andy.haldane@bankofengland.co.uk

[‡]Corresponding author. Bank of England, Threadneedle Street, London, EC2R 8AH, United Kingdom. *E-mail address*: sujit.kapadia@bankofengland.co.uk

1. Introduction

Herbert Simon spent over half a century teaching at Carnegie-Mellon University. In his classic 1962 study of "The Architecture of Complexity", he laid the foundations for evaluating the evolution and resilience of complex systems (Simon (1962)). Typically for Simon, his analysis drew on a wide spectrum of systems thinking – from physical systems and biological systems as well as social systems. Out of this, Simon drew a powerful conclusion.

Even complex systems tended to exhibit a basic simplicity. But this was simplicity in a specific sense. These systems could be arranged in a natural hierarchy, comprising nested substructures. Complex systems were naturally fractal. Simon saw the emergence of such simple, yet complex, structures as an evolutionary one. Non-hierarchical structures would tend to be deselected over time because of their inefficiency or because of their lack of robustness relative to simpler, hierarchical structures. Within complex systems (physical, biological, social), it was a case of survival of the simplest.

Finance has bucked this evolutionary trend. Over the past century, the financial system has evolved into a much more complex, concentrated and interconnected structure (Haldane (2010)). In Simon's words, it has become less modular, less hierarchical. In consequence, the financial system may have become markedly more susceptible to systemic collapse. Recent events rather attest to that. So why have Simon's evolutionary forces not deselected such a system?

One possible answer is policy. Successive crises have seen a progressive broadening and widening of the financial safety net: from last resort lending from the 19th century onwards, to deposit insurance from the 1930s onwards, to the block-buster capital and liquidity support for both banks and non-banks during this century (Alessandri and Haldane (2009)). Until the end of the Bretton Woods era, increasing implicit support went hand-in-hand with tougher financial regulation and greater restrictions on global capital flows. The 25-year postwar period was remarkable for its absence of systemic banking crises.

But widespread deregulation since the mid 1970s has resulted in increasing complexity and concentration, and ever more frequent crises (Bordo *et al*, 2001). Far from deselecting the complex and concentrated, the widening of the safety net in an era of deregulation may have caused them to flourish. There has been no better historical example than the policy interventions during this crisis.

This paper develops a model of the interbank network in which repo markets, amplification from shocks to the haircuts applied to collateral, and liquidity hoarding by banks play a central role in the spread of contagion and the collapse of interbank markets, similar to that experienced during 2007-8.¹ The framework we develop is sufficiently rich to allow explicit consideration of the role of complexity and concentration on the fragility of the financial system. It also facilitates investigation of a range of policy measures aimed at addressing intra-financial system activity to improve system resilience. In particular, our analysis suggests that tougher liquidity regulation, macro-prudential policy and surcharges for systemically important financial institutions may all help to make the system less prone to collapse.

The remainder of the paper is structured as follows. Section 2 presents stylised facts on complexity and concentration and their evolution over time. It also briefly discusses how the interbank market collapse of 2007-8 unfolded. Section 3 relates our paper to the literature. Section 4 outlines the modelling framework and section 5 applies an approximation to the setup to demonstrate analytically how contagion can break out in the model and when it is most likely to occur. We then turn to simulations: after outlining the methodology in section 6, section 7 presents simulations of liquidity crunches. Sections 8 explores the role of concentration in contributing to contagion; while section 9 focuses on complexity. Section 10 then considers how policy might be able to improve outcomes. Finally, section 11 concludes.

2. The Pre-Crisis Evolution of Complexity and Concentration

We motivate our approach with some stylized evidence on financial system complexity and concentration. Network diagrams, such as those depicted in Figures 1 and 2, offer a convenient way to describe the evolution of intra-financial system activity. Figure 1 shows the complexity of the global financial network and its evolution since 1985. The nodes in the graph represent countries, while the links reflect the cross-border banking claims between them and their thickness reflects the value of these claims, normalised by world GDP. It is immediately evident that the complexity of this network has increased considerably over time. And concentration is also evident from the prominent role in the network of the major financial centres, including the UK, US and, at various points in time, Germany and Japan.

Figure 2 illustrates the network of bilateral large exposures between the major UK banks. It also points to significant complexity and concentration at the national banking system level. The data on bank interlinkages only reflects large-scale interbank loans on banks' balance sheets. So it is likely that the "true" financial network may be significantly more complex than the picture presented. On concentration, the distribution of linkages and loan sizes between nodes in the

 $^{^{1}}$ A repo transaction entails borrowing money using securities as collateral. It is structured as the spot sale of a security for cash coupled with an agreement to repurchase the same security at the initial price plus interest at a particular date in the future. When the cash lent on repo trades is lower than the current market value of the security used as collateral, the discount is referred to as the haircut.

inter-bank market appears to exhibit a fat-tailed distribution. As our model demonstrates, this opens the door for rare but catastrophic system-wide breakdowns of the kind seen in 2007-8.²

Complexity in the financial system is difficult to quantify. But it is likely to go hand-in-hand with increased intra-financial system activity. For example, as Shin (2009) notes, the advent of securitization has markedly increased the complexity of financial systems by lengthening the intermediation chain. In many cases, the same security is used repeatedly in repo lending, with the lender using the security received as collateral with other lenders. Increased lending and borrowing activity of this kind between intermediaries is reflected in the dramatic rise in the stock of repos and financial commercial paper as a percentage of broad money in the UK and the US (Figure 3) and, for the US case at least, the subsequent dramatic fall.³ The nature of these transactions can mean that they are subject to amplifying dynamics and cyclical fluctuations linked to the variability of haircuts over the credit cycle.⁴ Figure 4 indicates that the growth in intra-financial activity extended well beyond banks to non-bank financial intermediaries. For example, financial corporate debt (which includes banks and non-banks) accounted for around two-thirds of the total growth in UK debt between 2003 and 2007.

The period preceding the financial crisis was also characterized by increasing concentration within national financial systems, from already high starting points. Figure 5 shows the marked increase in concentration in the UK and US banking systems as measured by the combined assets of the largest three banks by total assets as a percentage of total banking system assets (see also King, 2010, for evidence on this). The size of the banking systems in these countries has also been increasing dramatically, with UK banking system assets accounting for over 500% of GDP just before the crisis (Figure 6). Signs of increasing concentration are also evident in the interbank market. Figure 7 plots the Herfindahl index for the UK interbank network between 2004 and 2008, drawing on the same data as used to construct Figure 2. Again, the results point to the presence of increasingly important key players in the system.

The key aspect of the crisis that this paper tries to capture is the interbank market collapse. This is illustrated clearly in Figure 8 which shows that the cost of unsecured interbank borrowing rose dramatically at the onset of financial difficulties in August 2007, with the premium sought by banks on 3-month interbank loans rising from 10 basis points to around 100 basis points. Precautionary hoarding behaviour among banks combined with counterparty risk concerns to

 $^{^{2}}$ In addition to complexity across countries and across individual banks, there is also a considerable amount of complexity *within* banks. As Herring and Carmassi (2010) document, the largest eight banks in the world all have over 1,000 subsidiaries and operate in anything between 16 and 84 countries. Such organisational complexity clearly brings further risks as it makes it difficult to resolve institutions when they get into distress and can greatly amplify uncertainty on the part of counterparties.

³The US chart is taken from Adrian and Shin (2010b).

⁴For evidence on cyclicality in haircuts, see Geanakoplos (2010), CGFS (2010) and Krishnamurthy (2010).

lead a freeze in interbank borrowing. The ensuing collapse of Lehman Brothers in September 2008 led the interbank premium to increase about 20-fold from pre-crisis levels. And, throughout the crisis, prices did not tell the whole story – the quantity of funding available, especially at maturities longer than overnight, declined dramatically. As a counterpart to this, banks' holdings of reserves with central banks stepped up markedly during this period (Figure 9), as the financial network effectively collapsed to a star network with central banks at its centre.

3. Related Literature

Our paper contributes to a growing theoretical literature on contagion in financial systems. This strand of work has taken two forms: one that emphasizes the role of bank behaviour in driving contagion and a second that applies network techniques from mathematics and physics to clarify the probability and spread of contagion.

The most well-known work on contagion is that of Allen and Gale (2000) who study how the banking system responds to contagion when banks are connected under complete and incomplete network structures. In a stylized four bank model, banks insure themselves against liquidity shocks by exchanging interbank deposits. The connections created by swapping deposits, however, expose the system to contagion. The more complete the network the more resilient it is, since the proportion of the losses in one bank's portfolio is more easily transferred to more banks through interbank agreements.

Freixas, Parigi and Rochet (2000) present a related model in which the connections between banks are realized through interbank credit lines that enable these institutions to hedge regional liquidity shocks. As with Allen and Gale (2000), more interbank connections enhance the resilience of the system to the insolvency of a particular bank. Moreover, the stability of the banking system depends crucially on whether many depositors choose to consume at the location of a bank that functions as a money centre or not.

Dasgupta (2004) also explores how linkages between banks, represented by crossholding of deposits, can be a source of contagious breakdowns. In his model, depositors receive a private signal about banks' fundamentals and may wish to withdraw their deposits if they believe that enough other depositors will do the same. Multiplicity of equilibria is eliminated using global game techniques pioneered by Morris and Shin (1998) and a unique equilibrium, depending on the value of the fundamentals, is identified.

Despite the obvious parallels between financial systems and complex systems in other fields (May et al, 2008; Haldane, 2009), the use of network techniques from the mathematical and physical sciences to the study of financial contagion remains in its infancy. While these techniques place relatively less emphasis on optimising behaviour in the way links are formed, they have the strong advantage that they do not preclude any type of network from the analysis, including those which might arise as the optimal outcome of a prior network formation game or real-world networks such as those depicted in Figures 1 and 2. And they are particularly well suited for the analysis of processes taking place on networks.

Gai and Kapadia (2010a) use these methods to develop a model of default contagion in financial networks, adapting techniques primarily used in the statistical physics literature and the epidemiological literature on spread of disease in networks (eg Newman, 2002; Watts, 2002) to show that financial systems may exhibit phase transitions or 'tipping points'. As with Allen and Gale (2000), they find that greater connectivity can also reduce the likelihood of widespread default. But the 'tipping point' property underpins their finding that when contagion does occur, it has catastrophic consequences for the entire system. These results help to shed light on the contagion dynamics reported by Nier et al (2007) in their earlier numerical analysis. Further analytical insight is provided by May and Arinaminpathy (2010) in an important contribution which draws on the framework used in both of these papers. In particular, through the application of a mean-field approximation, they can identify the exact contribution of different parameters in the model to the likelihood of contagion and its dynamics. Their paper also develops the underlying framework considerably by providing a much deeper analysis of the influence of various types of asset liquidity shocks on the dynamics of contagion.⁵

However, despite their importance during the crisis, this line of literature generally excludes both funding contagion propagating through the asset side of balance sheets due to liquidity hoarding and the impact of haircut shocks on repo markets. And although Gai and Kapadia (2010b) develop a network model of liquidity hoarding, their focus is purely on the unsecured market and their analysis of policy interventions is fairly limited.

One key contribution of this paper is to develop a broader network model of interbank market collapse which integrates a role for repo markets, haircut shocks and spirals, and liquidity hoarding in unsecured interbank markets to all contribute to systemic collapse. While this necessarily involves sacrificing some of the detail provided by Gai and Kapadia (2010b), it allows for a richer story to be told both about the crisis and the role of complexity and concentration in contributing to it. It also facilitates investigation of a much broader range of policy interventions – indeed a

⁵Related work that draws on network techniques to explore contagion includes Hatchett and Kuhn (2009), Giesecke and Weber (2006), and Minguez-Afonso and Shin (2010). The latter use lattice-theoretic techniques to study systemic risk in high-value payment systems. See also Soramaki et al (2006). There is also a related empirical literature which exploits large exposure data to analyse default contagion in interbank markets – see Upper (2007) for a survey.

very strong focus on the analysis of policy is another key contribution of this paper. The final main contribution of the paper is in its novel application of the mean-field approximation used by May and Arinaminpathy (2010) to the context of liquidity hoarding and repo market activity – as we shall see below, this is particularly useful in helping to understand the intuition which underpins the subsequent simulations.

Our paper is also related to a number of recent papers that have sought to explain the "freeze" in interbank markets. Unlike our analysis, these contributions are explicitly behavioural and explore imperfections, other than network effects, that influence banks' decisions to form credit relationships and to generate potential collapses in interbank lending.⁶

Allen et al (2009) present a model of interbank liquidity without counterparty risk or information asymmetry in which banks stop lending and hoard liquidity because incomplete financial markets do not allow them to hedge aggregate liquidity shocks. In Diamond and Rajan (2009), banks can only sell assets to realise cash but face a limited set of potential buyers. Faced with a liquidity shock, banks are deterred from asset sales because the alternative of holding them is more beneficial. The induced illiquidity in the market for these assets acts to depress lending to others in the interbank market, prompting a credit freeze.

Other papers invoke asymmetric information to generate a market breakdown. Heider et al (2009) place counterparty risk at the centre of their analysis of liquidity hoarding. Banks become privately informed about the risk of their illiquid assets after they choose their asset portfolios. They show how the interbank market becomes impaired by adverse selection as suppliers of liquidity attempt to protect themselves from lending to 'lemons'. In Bolton et al (2008), long-term investors cannot tell whether short-term investors sell because of adverse asset quality or genuine liquidity needs. The price discount is exacerbated as potential sellers learn about the asset. When confronted with a decision to sell now to meet a liquidity need or risk future sales at a large discount, late trading equilibria that resemble a freeze become possible.

Exogenous shifts in information structure and Knightian uncertainty have also been emphasised as important factors in recent work. Caballero and Krishnamurthy (2008) suggest that the recent bout of financial innovation in credit markets heightened Knightian uncertainty in financial markets and use this to justify flights to quality and hoarding. In Acharya et al (2009) runs in short-term funding markets arise when the information structure in the market shifts exogenously from being optimistic about asset quality to pessimistic. In their paper, "no news is good news" becomes "no news is bad news", and they show that if good news arrives at a slower

⁶Related papers that study the functioning of the interbank market also include Bhattacharya and Gale (1987), who show how banks underinvest in liquid reserves in the face of moral hazard, and Repullo (2005) who highlights banks' incentives to free-ride on central bank liquidity.

rate than that at which debt is rolled over, interbank lending dries up.

In a related contribution, Cabellero and Simsek (2009) also explore the link between complexity and financial crises. But whereas our paper uses a generalised random graph as a metaphor for the complexity of real world financial networks, Caballero and Simsek appeal to the rising costs of understanding the structure of the network as the basis for complexity and, ultimately, the reason for hoarding. Their model adopts a "ring-like" network structure of identical banks with only one neighbour serving as a funding counterparty. Banks must learn about the health of the trading partners of the trading partners of their trading partners, and so on. If information about the network structure is costless, there is no hoarding. But if, following a liquidity shock, these information costs rise sharply, banks' inability to understand the structure of the network to which they belong leads them to withdraw from their loan commitments.

Finally, our work is related to recent contributions that emphasise the role of collateral in amplifying financial shocks. In their seminal work, Kiyotaki and Moore (1997) show how a shock that lowers asset prices lowers the value of collateral, leading to a decline in net worth, less borrowing, and further declines in the value of collateral. Geanakoplos (2010) discusses how interest rates and collateral are jointly determined. This is in contrast to Kiyotaki and Moore, where the level that can be borrowed against collateral is treated as given. Brunnermeier and Pedersen (2009) also highlight the important role played by credit constraints in amplifying financial shocks, demonstrating how capital and margin requirements of agents depend on the market liquidity of the asset.

Our model takes on the lessons from this literature by recognising that a key risk in report transactions arises from the ease with which the value of collateral can be realized in a sale and that the size of the "haircut" on the securities serving as collateral thus reflects the market risk of the collateral. In particular, following the implications of Brunnermeier and Pedersen (2009), Adrian and Shin (2010a) and Geanakoplos (2010), some of our simulations allow for movements in haircuts to create amplifying dynamics which can generate cycles in repo activity.

4. Modelling Framework

We start by outlining our general modelling framework. While the analytical results on tipping points presented in section 5 are derived by applying a mean-field approximation and other simplifying assumptions to the setup, the simulations of liquidity crunches which follow use all features of the full model.

4.1. The Financial Network

The financial network consists of n financial intermediaries, 'banks' for short, which are linked together randomly by their unsecured claims on each other. Each bank is represented by a *node* on the network, and the interbank exposures of bank i define the *links* with other banks. These links are *directed*, reflecting the fact that interbank connections comprise both assets and liabilities. Figure 10 illustrates a directed financial network with five banks.

More formally, let the links that point into a bank (or node) represent the unsecured interbank assets of that bank (i.e. money owed to the bank by a counterparty) and the outgoing links be the unsecured interbank liabilities (i.e. money borrowed by the bank from a counterparty). Denoting the number of incoming links, or *in-degree*, to bank *i* by j_i , and the number of outgoing links, or *out-degree*, by k_i , we define the *joint degree distribution* of in- and out-degree, p_{jk} , to be the probability that a randomly chosen bank simultaneously has in-degree *j* and out-degree *k*.

The joint distribution of in- and out-degree and, hence, the network structure of interbank exposures plays a key role in determining how shocks spread through the network. Except when applying the mean-field approximation in section 5, we allow for an arbitrary joint degree distribution, though we focus on the Poisson and (more fat-tailed) geometric distributions in our simulations. We also assume that there is no statistical correlation between nodes and mixing between nodes is proportionate (i.e. there is no statistical tendency for highly connected nodes to be particularly connected with other highly connected nodes or with poorly connected nodes). Unlike more stylised network models which assume a particular form, these assumptions imply that no network structures, including those which might arise as the optimal outcome of a prior network formation game or real-world networks such as those depicted in Figures 1 and 2, are precluded.

Since every unsecured interbank asset of a bank is an unsecured interbank liability of another, every outgoing link for one node is an incoming link for another node. This means that the average in-degree in the network, $\frac{1}{n}\sum_{i} j_{i} = \sum_{j,k} jp_{jk}$, must equal the average out-degree, $\frac{1}{n}\sum_{i} k_{i} = \sum_{j,k} kp_{jk}$. We refer to this quantity as the *average degree* and denote it by

$$z = \sum_{j,k} jp_{jk} = \sum_{j,k} kp_{jk}.$$
(1)

4.2. Balance Sheets and Repo Haircuts

Figure 11 presents the composition of individual bank balance sheets in the model. The total liabilities of each bank, L_i^T , are comprised of unsecured interbank liabilities, L_i^{IB} ; repo liabilities (ie borrowing secured with collateral), L_i^R ; retail deposits, D_i ; and capital, K_i . We assume that

the total unsecured interbank liability position of every bank is evenly distributed over each of its outgoing links and is independent of the number of links the bank has (if a bank has no outgoing links, $L_i^{IB} = 0$ for that bank). These assumptions do not affect any of our main results. But they serve to maximise diversification for a given number of links and the extent to which banks can diversify themselves by forming new outgoing links.

Since every interbank liability is another bank's asset, unsecured interbank assets, A_i^{IB} , are endogenously determined by the network links. So, although total unsecured interbank assets equal total unsecured interbank liabilities in aggregate across the entire system, each individual bank can have a surplus or deficit in their individual unsecured interbank position. Apart from unsecured interbank assets, banks hold four further asset classes which make up their total assets, A_i^T . These differ in their suitability for use as collateral in repo transactions: fixed assets (eg individual corporate loans or mortgages), A_i^F , which are completely illiquid and cannot ever be used as collateral; assets which may be used as collateral in repo transactions ('collateral assets'), A_i^C ; reverse repo assets (ie collateralised lending), A_i^{RR} ; and unencumbered fully liquid assets (eg cash, central bank reserves, high-quality government bonds), A_i^L .

We assume that fully liquid assets can always be used as collateral to obtain repo financing if required without any haircut (or alternatively sold without any price discount) ie borrowing can be obtained against the full value of the asset. By contrast, we suppose that the aggregate haircut associated with using collateral assets to obtain repo funding is denoted by $h \in [0, 1]$ – note that this is equivalent to assuming alternatively that A_i^C is made up of many different asset types with an average haircut of h. This haircut is to protect the lender against price moves in the underlying collateral, given that the lender may be left with the collateral if the counterparty defaults. Therefore, it may partly be taken as reflecting the underlying probability of default on the securities used as collateral. It will also be dependent on the market liquidity of the assets as this will affect price discounts upon any sale.

We also allow for the possibility of an additional bank-specific haircut, h_i , so that $(1 - h - h_i) A_i^C$ reflects the maximum amount of repo funding that can be obtained from collateral assets. This could be interpreted as stemming from the fact that a particular borrower offers lower-quality collateral than the typical bank. Or it could be due to the greater default probability of a particular bank – even if a bank is offering identical collateral to another bank, if the lender perceives that there is a higher chance it will fail, then it might demand a higher haircut as extra protection both because it is more likely to end up with the collateral in practice, and because there may be some legal risk in accessing the collateral in a timely fashion. We assume that reverse repo is secured with collateral that has the same aggregate haircut as on A_i^C . Abstracting for simplicity from idiosyncratic haircuts on banks other than bank *i*, this implies that the amount of collateral that bank *i* receives on its reverse repo assets is given by $A_i^{RR}/(1-h)$. We allow for this collateral to be fully rehypothecated to obtain repo funding with the same aggregate haircut, *h*. The maximum amount of repo funding that can be obtained from rehypothecating collateral obtained in reverse repo transactions is then given by $(1-h_i) A_i^{RR}$.⁷ Finally, since fixed assets cannot be used as collateral in repo transactions, the haircut on these assets is one; we also assume that unsecured interbank assets cannot be used as collateral.

4.3. The Liquidity Condition

Throughout this paper, we assume that there are never any systematic retail deposit inflows or outflows (though idiosyncratic inflows and outflows in a given period are one interpretation of the exogenous liquidity shock which is introduced below). We also assume that the central bank never takes collateral at more generous terms than the market in its liquidity operations. For expositional purposes, let us also start by assuming that unsecured interbank deposits are always rolled over but that there is no possibility for banks to recall unsecured interbank loans or sell fixed assets. Under these assumptions, a bank will remain liquid in each period provided that the amount of collateral it has available to obtain repo funding (which includes its unencumbered liquid assets) plus any new unsecured interbank borrowing, L_i^N , is sufficient both to exceed the amount of repo funding it has and to meet any idiosyncratic liquidity shock, ε_i . Given the balance sheet and haircut assumptions above, this implies that the bank is liquid if:

$$A_i^L + (1 - h - h_i)A_i^C + (1 - h_i)A_i^{RR} + L_i^N - L_i^R + \varepsilon_i > 0$$
⁽²⁾

From this expression, we can immediately see that exogenous liquidity shocks, or shocks to aggregate or idiosyncratic haircuts (perhaps due to shifts in the underlying quality of assets, changes in market liquidity or changes in bank-specific risk), have the potential to trigger a liquidity crisis at bank i if it is unable to raise a sufficient amount of new unsecured interbank borrowing. It is also evident that a sufficiently large shock to aggregate haircuts has the potential to trigger widespread liquidity stress, a point which we return to below.

But it is also through this condition that we can see how liquidity hoarding in unsecured interbank markets may emerge. In particular, if a bank does not meet condition (2), then it

⁷Depending on the motivation for the idiosyncratic haircut term on collateral assets, it may not be present here – in particular if the term is motivated by a bank having lower quality collateral on its own balance sheet than the average, that argument cannot extend to the collateral it is receiving on its reverse repo transactions (given that these are the 'typical' transactions). But we retain this term to keep the model as general as possible.

needs to take action to avoid defaulting on required payments. One seemingly obvious option might be for the bank to increase the interest rate that it is prepared to offer on new interbank liabilities until it obtains a sufficient amount of new funding. But as well as being directly costly in terms of future profitability, this may be associated with adverse stigma effects – if a bank is seen to be 'paying up' significantly in interbank markets, it may actually exacerbate the liquidity problem that the bank is trying to resolve as wholesale lenders may view it as a signal of underlying difficulties and run on the bank. In a sense, if banks 'pay up' too much, there may be adverse selection in the spirit of Stiglitz and Weiss (1981) so that the quantity of funding that may be obtained through this route could be constrained.

A second option would be for the bank to try to liquidate its fixed assets, A_i^F , in a fire sale. But selling assets cheaply is unlikely to be an attractive option because if real losses are incurred on the sales, then the capital position of the bank will weaken. Further, as Diamond and Rajan (2009) note, banks may be hesitant to enter a fire sale because the alternative of holding on to assets may be more beneficial. And given that fire sales are highly visible (especially if in high volume), potential stigma effects are again a deterrent.

A third option is for the bank to hoard liquidity by withdrawing its unsecured interbank lending, A_i^{IB} , from others it is connected to in the market and converting these deposits into liquid assets. Hoarding in this context is purely due to the bank's own liquidity needs; concerns over the solvency of its counterparty play no role in its decision. While hoarding may entail reputational costs, it is likely to be less directly costly than 'paying up' for funds or selling assets in a fire sale. And since unsecured interbank transactions are over-the-counter and are sometimes withdrawn even in normal times, adverse stigma effects can be kept to a minimum. It is also worth noting that during the recent crisis, liquidity hoarding was frequently and widely observed, whereas there was less evidence of significant 'paying up' or widespread fire sales.

Finally, the bank may try to cut lending to the real economy, A_i^F . While this was certainly observed during the crisis, we abstract from this possibility as it is only likely to generate liquidity slowly over the long-term. But the potential for this type of credit crunch effect is precisely one of those things which can make systemic liquidity crises so costly for society.

Given this, if a bank does not meet condition (2), we assume that its first course of action will be to raise resources by hoarding liquidity and withdrawing interbank assets held against other banks in the system. Clearly, it is possible that hoarding alone may not allow banks to raise sufficient resources to meet their obligations, at which point fire sales or 'paying up' become options. [And we briefly discuss the possible systemic consequences of fire sales later in the paper - possibly to be added to the simulation section by linking the aggregate haircut to the number of banks who have insufficient resources even after hoarding, on the grounds that fire sales are likely to cause a margin spiral]

But once we recognise that liquidity hoarding may be a response to a failure to meet (2), we can immediately see how network effects come to the fore. In particular, let us suppose that a fraction μ of banks connected to bank *i* in the network 'hoard' liquidity from it that is, they withdraw a portion of their deposits held at bank *i*. In the model presented here, we assume that this represents a genuine drain on the liquidity of the entire banking system which flows outside, for example, ending up as increased reserves at central banks. But, as discussed by Gai and Kapadia (2010b) in a similar model which completely abstracts from repo market activity, hoarding behaviour can also be interpreted as a switch from lending at long maturities (e.g. for three months or one year) to lending at much shorter maturities (e.g. overnight).

Further, let us suppose that, on average, hoarding banks withdraw a fraction λ of the deposits that they hold at bank *i*. In principle, the proportionate amount that each bank hoards will depend on how much liquidity it needs to raise to meet its liquidity condition, ie on its liquidity shortfall. If this shortfall were the only determinant of the amount of liquidity hoarding, then λ would be fully endogenous within the model. At the other extreme, $\lambda = 1$ would correspond to full withdrawal in all circumstances (i.e. lending banks withdraw their entire deposit irrespective of what their liquidity shortfall is).

Clearly, the value of λ is a key determinant of the strength of amplification of shocks in the model – in particular, the higher the value of λ , the larger the shocks that can hit banks further down the chain of contagion. In reality, the value of λ is likely to lie somewhere between the two extremes identified above. In particular, immediate full withdrawal may be unlikely not only because current liquidity needs may not necessitate it but also because contractual obligations may prevent banks from withdrawing their entire deposit straight away. On the other hand, a bank which has suffered a haircut shock may fear larger shocks to come. And if a bank has a liquidity shortfall because it has lost some portion of its deposits from a hoarding counterparty, it may consider it to be only a matter of time before the full amount is lost. In particular, further withdrawals may occur when contractual obligations expire. And time consistency issues make it difficult for the lender to commit not to making further withdrawals in future. Moreover, even if current withdrawal is only partial, a forward looking bank may choose to act immediately as if it had lost its entire deposit, in order to limit the prospect of suffering liquidity problems in the future. As such, assuming a value of λ at or close to 1 offers a simple mechanism for capturing

a rich set of dynamics which may operate through forward-looking expectations.

Under these assumptions, bank *i* loses $\mu \lambda L_i^{IB}$ of its liabilities due to liquidity hoarding by its counterparties in the network. Therefore, its liquidity condition now becomes:

$$A_{i}^{L} + (1 - h - h_{i})A_{i}^{C} + (1 - h_{i})A_{i}^{RR} + L_{i}^{N} - L_{i}^{R} - \mu\lambda L_{i}^{IB} + \varepsilon_{i} > 0$$
(3)

4.4. The Dynamics of Contagion

From (3), it is immediately clear that the decision by any one bank to hoard liquidity makes it harder for the banks which were previously borrowing from it to meet their own liquidity condition without resorting to hoarding themselves. In particular, as each successive bank fails to meet its liquidity condition, its hoarding behaviour has the potential to trigger liquidity stress at any of the banks to which it is connected via an interbank lending relationship, and this process only dies out if either no neighbours to newly distressed banks become distressed themselves or when every bank in the network is in distress.

Thus hoarding can potentially spread across the system, with the structure and connectivity of the unsecured interbank network playing a key role in determining the evolution of contagion. As noted above, the dynamics of contagion also depend crucially on the value of λ . In principle, they also depend on whether, when a bank hoards liquidity, its withdrawals are concentrated on particular counterparties or more evenly distributed. But in all of what follows, we assume that banks raise any resources needed by withdrawing funding equally and proportionately from *all* of their counterparties. As well as being necessary for the analytical solution in section 5 and making the simulations computationally simpler, this assumption also seems reasonably plausible as immediately accessible deposits are only likely to be available in relatively small amounts from each counterparty.

To illustrate the contagion dynamics at work, let us first simplify the structure slightly by assuming that in the initial state

$$(1 - h - h_i)A_i^C + (1 - h_i)A_i^{RR} = L_i^R$$
(4)

for all banks so that all of the collateral obtained via a bank's own reverse repo lending plus all of its collateral assets are pledged as collateral to obtain repo funding. Then changes in the aggregate or idiosyncratic haircuts, Δh or Δh_i , are equivalent to liquidity shocks of $-\Delta h A_i^C$ and $-\Delta h_i \left(A_i^{RR} + A_i^C\right)$ respectively. So we can rewrite the liquidity condition as:

$$A_i^L - \Delta h A_i^C - \Delta h_i \left(A_i^{RR} + A_i^C \right) + L_i^N - \mu \lambda L_i^{IB} + \varepsilon_i > 0$$
(5)

Rewriting gives:

$$\mu < \frac{A_i^L - \Delta h A_i^C - \Delta h_i \left(A_i^{RR} + A_i^C\right) + L_i^N + \varepsilon_i}{\lambda L_i^{IB}} \tag{6}$$

Now suppose that a single bank suffers a haircut or idiosyncratic liquidity shock which is sufficiently large to cause it to start hoarding liquidity. Recalling that interbank liabilities are evenly distributed over different counterparties in the network and that we take bank *i* to have k_i borrowing links, if a single counterparty to bank *i* hoards, $\mu = 1/k_i$. Therefore, bank *i* will continue to meet its liquidity condition provided that:

$$\frac{1}{k_i} < \frac{A_i^L - \Delta h A_i^C - \Delta h_i \left(A_i^{RR} + A_i^C\right) + L_i^N + \varepsilon_i}{\lambda L_i^{IB}}$$
(7)

Therefore, for contagion to spread beyond the first bank, there needs to be at least one neighbouring bank for which:

$$\frac{A_i^L - \Delta h A_i^C - \Delta h_i \left(A_i^{RR} + A_i^C\right) + L_i^N + \varepsilon_i}{\lambda L_i^{IB}} < \frac{1}{k_i}$$

$$\tag{8}$$

If this holds, then contagion starts to spread more widely. And, if we allow for the possibility of banks being exposed to multiple hoarding counterparties, similar equations then determine whether it spreads more widely across the network,

5. Analytical Solution under a Mean-Field Approximation

The structure presented above fully characterises the model that is at the core of our simulations and policy analysis. But to obtain a better intuitive understanding of the dynamics of the system, we first apply a mean-field approximation which, along with some other simplifying assumptions, allows us to obtain stark analytical results.

Specifically, rather than taking the network to be random, we assume that each bank is connected to exactly z other banks as both a lender and borrower (which implies that $j_i = k_i = z$ for all banks). Further, we assume that all banks have identical balance sheets, which allows us to drop all *i* subscripts except for the ones on h_i and ε_i . Given that every interbank asset must be another bank's interbank liability, taken together, these assumptions also imply that $L^{IB} = A^{IB}$ for all banks. We also assume that banks can never raise new unsecured interbank funding (so that $L_i^N = 0$) and that $\lambda = 1$ so that whenever a bank hoards, it withdraws its lending from its counterparties completely (full withdrawal assumption). Under these assumptions, we can rewrite (8) as:

$$\frac{A^L - \Delta h A^C - \Delta h_i \left(A^{RR} + A^C\right) + \varepsilon_i}{A^{IB}} < \frac{1}{z} \tag{9}$$

And if we were to abstract further from idiosyncratic shocks, this would reduce to:

$$z < \frac{A^{IB}}{A^L - \Delta h A^C} \tag{10}$$

This expression is particularly illuminating because under the above assumptions, it is identical for every bank in the network. If it is satisfied, then provided that z is greater than or equal to 1 so that there is sufficient connectivity on the network, any initial case of liquidity hoarding will cause all neighbouring banks to become distressed and start hoarding. But since neighbours of neighbours also have the same liquidity condition, hoarding behaviour will cascade through the entire network. By contrast, if equation (10) is violated, an initial case of liquidity hoarding will have no systemic consequences at all. This starkly illustrates the phase transition or 'tipping point' which is embedded in the model, whereby a very small change in the parameters can lead to a fundamentally different outcome – indeed, as Gai and Kapadia (2010b) demonstrate, although a precise threshold in the mould of equation (10) cannot normally be identified, it is possible to prove the *existence* of such phase transitions in this type of model without making the mean-field approximation and while still maintaining heterogeneous balance sheets (though the $\lambda = 1$ assumption is still necessary in their formal analysis).

Equation (10) also provides important initial insights on the conditions under which systemic liquidity crises may occur. In particular, it points to the fact that low liquid asset holdings, large adverse aggregate haircut shocks, a high volume of assets used as collateral in repo transactions, and high level of unsecured interbank lending are all likely to contribute to the susceptibility of the system to a widespread liquidity crisis. As we shall demonstrate below, all of these results are borne out in the simulations conducted under more general assumptions.

6. Simulation Methodology

The mean-field approximation used in section 5 imposes some stark assumptions. It is therefore natural to consider what happens in more general settings. So we now return to the more general framework developed in section 4 and simulate it under a range of assumptions to demonstrate how systemic liquidity crises may occur and to explore the role of concentration and complexity in contributing to them. In this analysis, we focus on two main degree distributions for the network: the simple Poisson distribution and the more fat-tailed geometric distribution.

To keep the number of simulations manageable and to retain focus on the key areas of interest, we assume throughout that banks can never raise any new deposits in the unsecured interbank market (ie $L_i^N = 0$ for all banks). As Gai and Kapadia (2010b) show, although allowing for replacement of lost interbank deposits can significantly reduce the likelihood of systemic liquidity crises, provided that the probability of replacement is not always very close to 1, it does not otherwise change the fundamental properties of the model and its funding contagion dynamic – in particular, the phase transition is still present and contagion can still to spread to the entire system on some occasions. Apart from in one of the policy experiments, we also assume that $\lambda = 1$ throughout, though, as noted above, this may be a reasonably plausible benchmark assumption. And, although the model readily applies to networks of fully heterogeneous banks, we treat the repo liabilities, interbank liabilities, capital buffers, liquid asset holdings, reverse repo assets, and collateral assets of each bank to be identical for the purpose of illustration.

The interbank network in the simulation comprises 250 banks. We take the liability side of the balance sheet to be comprised of 15% interbank liabilities and 4% capital, with repo liabilities determined as described below and the remainder being retail deposits. Since each bank's interbank liabilities are evenly distributed over their outgoing links, interbank assets are determined endogenously within the network structure which is drawn from a distribution as specified below. And when drawing the network, we allow for the possibility that two banks can be linked to each other in both directions – no netting of exposures is assumed. Liquid assets are set to be 2% of total assets, with reverse repo and collateral assets determined as described below and the asset side being 'topped up' by fixed assets until the total asset position equals the total liability position.

In terms of the repo parts of the balance sheet, we assume that (4) applies in the initial state ie that all collateral assets and assets received as part of reverse repo transactions are used as collateral to obtain repo funding. We also suppose that reverse repo assets are 11% of the balance sheet, collateral assets are 10%, and the initial aggregate haircut, h, is 0.1. This implies that repo liabilities comprise 20% of the balance sheet.

In our simulations, we vary the average degree, z, drawing 1000 realizations of the network for each value and then shock the network in different ways according to the experiment in question. The dynamics of contagion follow the process described in section 4.4., with the liquidity condition at the centre of the propagation dynamics. For each realisation, we follow these dynamics iteratively until no new banks are forced into hoarding liquidity or until every bank is hoarding. We count as "systemic" those episodes in which at least 10% of banks are forced into hoarding liquidity, and our results identify the frequency of systemic liquidity crises and their scale in terms of the average fraction of the system affected in each systemic outbreak (ie how widely contagion spreads, conditional on it spreading to at least 10% of banks in the system).

7. Liquidity Crunches: Basic Experiments

In our first set of experiments, we assume that the joint degree distribution for the network takes the simple Poisson form. Under this distribution, each possible directed link in the network is present with independent probability p. So the network is constructed by looping over all possible directed links and choosing each one to be present with probability p.

Figure 12 shows what happens when there is no aggregate haircut shock and only a single bank receives a very large adverse idiosyncratic haircut shock which causes it to start hoarding liquidity. Contagion occurs for values of z between 0 and 20 and its probability is non-monotonic in connectivity, at first increasing before falling. But when contagion breaks out, it invariably spreads to the entire network. These results accord well with the analytical solution from the mean-field approximation, which is not surprising because this approximation is most reasonable for a Poisson network. Specifically, given that there are no aggregate haircut shocks, equation (10) suggests that when banks have seven and a half times more interbank assets than liquid assets, contagion will occur for z < 7.5 but not for z > 7.5. From Figure 12, it is evident that z = 7.5 is the point around which the probability of contagion starts to fall from one. The reason it remains positive for higher values of z is due to the randomness of the network structure which means that contagion can still break out under certain configurations. And the reason contagion is not always certain for very small values of z is that the initial shock may hit a bank which either has no interbank assets and is therefore unable to trigger any contagion by hoarding liquidity, or is in an isolated subset of the network.

Figure 13 shows the impact of combining the large idiosyncratic haircut shock with an aggregate haircut shock that increases the haircut from 0.1 to 0.2. Given that collateral assets represent 10% of the balance sheet, it is clear from (10) that this shifts the phase transition to around z = 15. And this is borne out by the results. Arguably, this type of experiment reflects the behaviour of interbank markets in the early part of the crisis during August and September 2007. In light of bad news on subprime mortgages and other types of collateral which were being used to back repo and other secured forms of funding, aggregate haircuts increased, possibly amplified by asymmetric information because nobody knew which banks were holding the tainted securities. As a result, some banks found themselves short of liquidity. And this was especially true of banks which were also forced to take back assets from off-balance sheet vehicles for which liquidity had dried up. In response to funding liquidity stress, banks started to hoard in the unsecured interbank market. This led to a collapse in this market as reflected by the sharp increase in spreads depicted in Figure 8. And this may have increased counterparty risk which may have exacerbated problems for some institutions. Although this framework abstracts from the counterparty risk dimension, it makes clear how a seemingly small shock to a limited set of assets which are being used as collateral can lead to a collapse in both secured and unsecured interbank markets.⁸ And provided the fundamental shock to aggregate haircuts is sufficiently large, this is true even in a world without any asymmetric information.

8. The Role of Concentration

As is evident from Figure 2, real-world financial networks do not appear to be Poisson; instead they appear to exhibit fat tails, with a small number of key nodes who are very highly connected both in terms of the number of interbank relationships they have and in terms of the overall value of those relationships. This reflects the underlying concentration in the banking sector.

To explore the impact of concentration on our results, we repeat our first simulation exercise but use a geometric rather than Poisson degree distribution. This distribution exhibits fat tails, with some banks having substantially more connections than the average degree. In the version implemented, we draw in-degree and out-degree separately from the same distribution, implying that there is no correlation between the number of counterparties a bank lends to and borrows from.⁹

Figure 14 presents the results. Contagion is less likely and less severe for low values of z than under the Poisson distribution, which reflects the well-known result that fat-tailed networks tend to be more robust to random shocks (Anderson and May, 1991; Albert et al, 2000). On the other hand, it is also clear that contagion outbreaks can occur for much higher values of z, albeit rarely. So, for a broad range of connectivity, increasing concentration in the network (which would translate in a movement away from a Poisson distribution and towards a geometric one) makes the system slightly more susceptible to a systemic liquidity crisis.

However, this analysis masks what is perhaps the key difference between contagion dynamics under Poisson and geometric distributions. Thus far, we have assumed that when the initial idiosyncratic haircut shock occurs, it hits any bank in the network at random. Now suppose instead that the initial shock hits the bank with the largest number of unsecured interbank lending relationships. Figures 15 and 16 show the results for Poisson and geometric degree distributions. As might be expected, when the shock is targeted at the most connected interbank

⁸In the model, hoarding entails withdrawal of funds while in reality, much of the hoarding behaviour early in the crisis involved banks dramatically reducing the maturity of their lending. But, as noted above, the broad framework can be interpreted as speaking to this behaviour as well.

 $^{^{9}}$ To construct the graph, we also need to ensure that the total number of outgoing links drawn equals the total number of incoming links. We follow the algorithm outlined by Newman et al (2001) to achieve this.

lender, contagion occurs more frequently in both cases. But what is striking is that for the Poisson distribution, a targeted shock only makes a relatively small difference to the results (compare Figures 12 and 15); on the other hand, compared to the random shock case, a targeted shock under a geometric distribution has absolutely catastrophic consequences, making contagion a near certainty for a very wide range of z. Again, this is consistent with the results of Anderson and May (1991) and Albert et al (2000) who both demonstrate how fat-tailed networks are particularly susceptible to shocks targeted at key nodes. The intuition for this is that in the Poisson network, the most connected bank is not that much more connected than the typical bank; on the other hand, under a geometric distribution, the most connected bank is likely to be connected to a very large portion of the other banks in the network, so if it becomes distressed, it has the potential to spread contagion very widely.

If we also consider aggregate haircut shocks in this context, then it is clear that banks who are heavily involved in repo activity are more likely to suffer from this type of shock ie for a given size shock, a bank concentrated in repo activity is more likely to become distressed. Given this, we can see that the most dangerous banks for the stability of the network are likely to be those which are both heavily involved in repo activity and big lenders in the unsecured interbank market – the former because it makes them highly susceptible to the initial shock; the latter because they propagate the shock widely. When we consider that banks that are heavily involved in repo activity are often the same large, complex financial institutions that are also big players in the unsecured interbank market, we can immediately see how the financial system may be particularly prone to collapse. And it also becomes clear why the seemingly small shocks of early / mid 2007 could have had such catastrophic consequences – they affected banks that were both highly susceptible and central to the network. When we further realise that many of these same banks were also heavily involved in lending to the real economy, then we can start to see how credit crunch effects subsequently emerged as well.

9. Complexity, Cycles and Contagion

As discussed above, complexity is difficult to measure but greater complexity in the financial system seems likely to be reflected in greater intra-financial system activity, both in terms of unsecured and secured markets. In Figure 17, we consider a random shock under a geometric distribution but suppose that unsecured interbank liabilities comprise 25% rather than 15% of the balance sheet. As we would expect, contagion occurs more frequently than in Figure 14.

But it is perhaps more interesting to consider complexity in a dynamic setting in which there is

a feedback between falling aggregate haircuts, greater repo market activity and associated balance sheet expansion in an upswing. [Simulation to be added which does the following: Draws an initial network, hits it with an aggregate haircut shock it to determine the frequency of contagion in that period, then goes to the next period in which the initial aggregate haircut falls a bit allowing repo borrowing and balance sheets more broadly to grow, and then hits the network with another aggregate haircut shock. This should demonstrate how the frequency of a collapse increases as the cycle evolves. Possibly also to be augmented with cyclicality in unsecured interbank lending to enrich the story eg a low aggregate haircut might go hand-in-hand with high unsecured interbank lending]

10. Policy Implications

At the heart of the systemic collapses modelled in this paper is an underlying network externality, whereby banks fail to internalise the consequences of their hoarding behaviour on others in the network with potentially devastating system-wide repercussions. The extent of the externality depends on the network and balance sheet structure and some of the assumptions, including over the amount of liquidity hoarding (the value of λ) when a bank is in distress. As the simulations involving targeted shocks make clear, it also varies across banks. Under these circumstances, there is a clear case for public policy intervention, possibly along a range of different dimensions. These include:

(a) Tougher Microprudential Liquidity Regulation One immediate implication of our results is that banks need to have a larger stock of genuinely high-quality liquid assets than they would naturally choose given their own individual incentives. It is clear from equation (10) that an increase in liquid asset holdings makes the system directly less susceptible to systemic liquidity crises. This is supported by the simulation presented in Figure 18, in which 3.5% of assets are assumed to be liquid rather than 2% as in Figure 14. But to the extent that greater liquid asset buffers also reduce the amount of repo activity and overall balance sheet size, then they may also indirectly improve the resilience of the system as banks become less susceptible to haircut shocks – in equations (4) and (10), this would show up via a reduction in the amount of collateral assets and repo liabilities. Implicitly, tougher liquidity requirements may have an ex ante benefit on system resilience by dampening the money multiplier.

In terms of the quality of liquid assets, the model takes liquid assets to be those which can either be sold without any price discount or used as collateral to obtain repo financing without any haircut, clearly distinguishing them from 'collateral assets' which can only be used to obtain repo funding at a positive haircut. This implies that any asset which is likely to have a large haircut in times of stress is much less useful as a buffer against systemic liquidity crises, pointing to the need for microprudential liquidity regulation to maintain a relatively tight definition of what constitutes a liquid asset. In particular, it suggests that holdings of bank debt (eg certificates of deposit; covered bonds etc.) should not be permitted to constitute part of banks' prudential liquid asset requirements because it is precisely when there is systemic liquidity stress that such assets are likely to have large haircuts and thus relatively little value in helping to avert the stress – by contrast, genuine 'outside' liquidity is likely to prove much more useful. From an ex*ante* perspective, the model also suggests that maturing interbank assets should not be allowed to contribute to liquid asset requirements because if they are, then there is again less overall 'outside' liquidity available in stress – indeed, one of the central points of the model is to demonstrate how withdrawals of interbank assets may be a key amplifier in precipitating a systemic liquidity crisis.¹⁰

(b) Macro-prudential Policy and Systemic Surcharges The results on concentration and complexity starkly highlight two different dimensions of risk linked to intra-financial system activity. In particular, it is clear that for a given level of concentration, a more complex network is more likely to be prone to collapse; similarly, for a given level of complexity, a more concentrated network is more vulnerable to shocks to key banks. Further, it is evident both from the underlying dynamics and the stylised facts that fluctuations in the complexity of the network are likely to be at least partially cyclical, while changes in network concentration are likely to be more structural.

This has two important implications: first, by the Tinbergen principle, two distinct policy instruments (one time-varying and one structural) are likely to be needed to help address network risks; second, to be most effective, time-varying macro-prudential policy should not just be targeted at real economy lending but should also react to changes in intra-financial system activity given its inherent cyclicality (and perhaps greater cyclicality than real economy lending – see Figure 4).

Macro-prudential policy aims to lean against the wind of the credit cycle, much like monetary policy does in respect of the business cycle (Bank of England (2009), Hanson et al (2010)). There are a variety of instruments that might be used, including varying headline regulatory capital and

 $^{^{10}}$ This second issue is less clear-cut *ex post* once a liquidity crisis has begun. In this case, again for reasons made clear by the model, it is better from a systemic perspective for banks to use their 'outside' liquidity to meet liquidity needs rather than withdrawing interbank funding. Therefore, care is needed to ensure that banks do not have strong incentives to withdraw interbank funding rather than run down liquid assets ex post, for example because such action helps to maintain their prudential liquid asset buffer at a high level.

liquidity ratios, changing the risk weights which attach to some of the components of lending, and adjusting haircuts on secured financing to financial institutions (CGFS (2010)). With the advent of the Financial System Oversight Committee in the US, the European Systemic Risk Board in the EU and the Financial Policy Committee in the UK, macro-prudential policy is about to become a reality.

It is not difficult to see how macro-prudential policy, appropriately executed, could deliver a less fragile financial network. For example, raising the risk weights which attach to intra-financial system exposures would serve as a disincentive to network complexity. Increasing the amount of liquidity held by banks in the face of higher complexity would reduce the effects of liquidity hoarding in the event of stress and, as noted above, may also help to mitigate the cyclical build-up in complexity by dampening the money multiplier. Similarly, adjusting the haircuts on secured financing in a pro-cyclical fashion would probably neutralise some of the adverse feedback effects of haircut contagion. In particular, increasing haircuts in good times may help to break the feedback loop between falling haircuts and rising leverage described in section 9.: with higher initial haircuts, the likelihood of a big positive shock to haircuts falls; and with less repo activity and smaller balance sheets, any shocks that do occur are less likely to trigger initial liquidity stress at a set of institutions.

It is possible to simulate the effects of such a policy in the model. [Simulation of time-varying haircut to be added]

In terms of the structural risks in the network, one regulatory response to the too concentrated to fail problem has been to explore levying a tax on institutions in line with their contribution to systemic risk. Because systemic risk is an externality, such an approach is equivalent to the imposition of a Pigouvian tax. Various governments, including in the UK, France and Germany, have already announced their intention to impose a direct fiscal tax on financial institutions, levied in line with their size and the riskiness of their balance sheet. This is in the spirit of an externality tax and would tend to nudge in the direction of a less concentrated system.

Within the international regulatory community, the G20 and FSB are drawing up a blueprint for dealing with so-called systemically important financial institutions (SIFIs). Among the options on the table are additional, graduated capital charges for SIFIs ("capital surcharges") and regulatory limitations on the extent of exposures between these SIFIs (a "large exposures" regime). To target concentration effectively and achieve a clean separation from any time-varying requirement linked to intra-financial system activity, such capital or liquidity surcharges could be set according to banks' market *shares* in interbank activity (eg the share of a bank's interbank assets as a fraction of total interbank assets across all banks) and potentially other markets which are deemed to be systemically important – note that since market shares cannot be timevarying in aggregate by definition, this would ensure that such surcharges were not time-varying in aggregate. These surcharges would then help both to enhance the resilience of SIFIs to shocks and should create incentives for the network as a whole to become less concentrated.

To gauge the effects of such policies on network stability, we simulate the model under a liquidity surcharge rule where the extent of the surcharge is related to the amount of interbank assets a bank has. Specifically, we assume that each bank holds a minimum of 2% liquid assets plus an amount equal to 10% of its total interbank assets. Since interbank assets comprise 15% of total assets on average, this implies that the average liquid asset holding is 3.5% (thus making the experiment directly comparable to the one in Figure 18) but banks with higher-than-average interbank asset positions will hold more liquid assets and banks with lower-than-average positions will hold less. Figure 19 presents the results. Comparing with Figure 18, it is clear that the surcharge rule is more effective in reducing the probability and spread of contagion than an equivalent across-the-board increase in liquid asset requirements. The intuition is simple – targeting the requirements at the banks which are most instrumental in spreading contagion is more potent than requiring peripheral players to hold extra liquid assets. And this experiment understates the potential benefits of systemic liquidity surcharges because it only considers the benefits from enhanced resilience of key players and does not account for any reduction in concentration that might be incentivised by them.¹¹

(c) Network Transparency At present, very little is known about the dynamic properties of the financial network. This is in part a result of data deficiencies. Historically, financial data has tended to be collected and analysed on an institution by institution basis (Haldane (2009)). That is the essence of micro-prudential regulation. Such data do not allow an effective mapping of the entire financial web or a simulation of its properties, though small parts of its sub-structures such as the domestic payments and interbank networks have been mapped in some countries.

This may be all about to change. There are efforts internationally to begin collecting systematically much greater amounts of data on evolving financial network structure, potentially in close to real time. This would allow the topology of the financial system to be mapped and,

¹¹There is an interesting analogy to be drawn here with the use of targeted vaccination programmes to control the outbreak of disease. It is a well-known result in the epidemiology literature that if vaccine stocks are limited, programmes should be targeted at the most highly connected individuals [insert reference]. The same logic applies to banks here. But there is also an important difference: while vaccinating particular individuals is unlikely to reduce the size of their social networks, applying surcharges to SIFIs may create real costs for them which may affect the structure of the network in a desirable way. Thus targeted approaches are likely to be even more useful in relation to financial networks than they are in epidemiology.

in principle, simulated. The introduction of the Office of Financial Research (OFR) under the Dodd-Frank Act is sure to nudge the United States in this direction. Indeed, one of the motivations for the OFR was precisely to allow such a mapping, enabling finance to catch up with the technology and techniques which are already applied to the dynamics of transport, weather and utility networks.

This transformation in the authorities' capacity to map and simulate the financial network is likely to bring at least two significant benefits. First, it ought to provide the authorities with a much better early warning on emerging vulnerabilities in the financial network; it will allow financial tipping points and cliff edges to be charted. This provides a much sounder, quantitative basis for taking ex-ante remedial policy action to avoid these cliff edges. Over time, this ought to reduce policymakers' reliance on ex-post interventions to prevent the system crash-landing once already over the cliff-edge.

Second, if more data is made available on network structures, this ought to have positive spin-off effects on the behaviour of financial institutions in the network. Armed with greater information on counterparty risk, banks may feel less need to panic and hoard following a disturbance. Banks' liquidity decisions would become less hair-trigger. This would have a collective benefit, lessening the potential for destabilising liquidity spirals. Network information would serve as a public good, reducing the externality of liquidity contagion.

More specifically in the context of the model, greater transparency might mean that haircuts are less volatile. With less severe initiating shocks, this would reduce the frequency of contagion. Alternatively, greater transparency might reduce the value of λ . To illustrate how this can improve outcomes, Figure 20 presents results from a simulation in which, rather than withdrawing all their interbank lending when under liquidity stress, banks only withdraw an amount equal to their liquidity shortfall (i.e. their outstanding shortfall after their liquid asset buffer has been used up) plus half of any remaining interbank assets. As is evident, this greatly diminishes the frequency of contagion.

(d) Netting and Central Clearing The financial system is a dense cats-cradle of exposures. These gross exposures can easily exceed a bank's capital. And it is gross exposures which matter when gauging the virulence of contagion through a network. So one means of reducing this contagion is by netting-off gross exposures between participants within the financial system as this would lower the value of interbank connections relative to balance sheet size (ie it would have a similar effect to moving from Figure 17 to Figure 14). There have already been some attempts to do so in respect of derivatives contracts. In particular, regulators have sought to eliminate,

or net-down, outstanding derivatives contracts which are effectively redundant because they are two offsetting sides of a transactions. This has reduced notional values outstanding in the OTC derivatives markets considerably.

A more ambitious, and far-reaching, regulatory initiative is the drive to centrally clear a much larger proportion of, in particular, OTC products through central counterparties (CCPs). This is one of the stated aims of the G20 and Financial Stability Board (FSB (2010)). The rationale for such action is well-captured by the model developed here. CCPs have two effects, which speak directly to concentration and complexity. On complexity, a CCP radically simplifies the catscradle of bilateral exposures, condensing it down to a simple hub-and-spoke configuration. Higherorder, unobservable counterparty credit risk, is replaced by first-order, observable counterparty risk with respect to the CCP. This, too, ought to reduce the sensitivity of banks' liquidity hoarding to disturbance, thereby lowering the fragility of the system as a whole.

On concentration, a CCP also alters radically the structural configuration of the network. This now resembles a star formation. From a resilience perspective, this is a double-edged sword. On the one hand, concentration among the key financial institutions is effectively eliminated, reducing the contagion risk from too-big-to-fail. On the other, concentration risk in an important sense is relocated, rather than eliminated, from too-big-to-fail firms to the CCP. In the language of the model, a world of centrally cleared transactions might be more susceptible to targeted attack.

This underscores the importance of ensuring that CCPs are bullet-proof moving forward as they clear larger numbers of transactions. This is the only way of ensuring systemic resilience is bolstered by the enhanced role for CCPs. That means CCPs may need to be operate as quasi or actual public utilities. Risk management and default arrangements will need to be many times stronger than that of even the largest financial firm. It is for this reason that the regulatory community is currently upgrading its risk management standards for CCPs (CPSS-IOSCO (2010)). Provided these standards are sufficiently high, the effect will be to boost systemic resilience.

At the height of the crisis, the financial network degenerated to a hub-and-spoke configuration with the central bank the de facto central counterparty. To avoid that outcome ex post, a set of bullet-proof CCPs could enforce a simplified hub-and spoke configuration ex ante. From an incentives and counterparty uncertainty perspective, this change in ex ante (rather than ex post) structures is much the preferable. (e) System Structure With the exception of a tougher large exposures regime among SIFIs and the ideas on netting and central clearing, all of the above measures seek to alter the structure and transparency of the financial system by acting on price incentives – for example, capital or liquidity requirements or taxes. A more directive approach would involve intervening in the underlying structure of bank's balance sheets – a quantity rather than price approach. In line with Weiztman (1974), this directive approach may be optimal if there is uncertainty about the elasticities associated with price-based instruments.

In the spirit of Simon (1962), the aim of such structural interventions would be to create a financial system that was ultimately decomposable and modular. In other words, disruption to one part of the network posed less risk to the network as a whole. To be effective, this modularity would need to operate both within financial firms, especially those comprising many thousands of legal entities, as well as across firms. This is the essence of Simon's hierarchy for complex systems if they are to exhibit simplicity and resilience. A number of regulatory initiatives and ideas are currently in train which speak to this agenda.

One is "living wills" or, more formally, recovery and resolution plans (RRPs). These require a bank to draw up plans for how they would be wound-down in an orderly fashion in the event of failure. The newly-passed Dodd-Frank Act in the US requires that US firms draw up credible RRPs, or else face a set of possible sanctions including break-up of the group. Pilot RRPs are currently being drawn up for the world's largest financial institutions, with a view to determining whether existing institutional structures would allow continuity of banks' key operational functions, while allowing the remainder of the business to be detached and liquidated. The evidence in section 2 suggests that considerable structural adjustment of global banks' business models may be necessary to deliver credible RRPs.

A second idea, which is being debated within academic and policymaking circles, is the potential for ring-fencing of financial activities, either within firms or across them. The Volcker rule, enacted in US Legislation recently, is one example of such a forced separation, in this case in respect of proprietary trading activity. The Glass-Steagall Act, which separated investment and commercial banking in the US between 1933 and 1999, is a second. The McFadden Act, which limited interstate banking in the US from the 1930s, would be a third. Narrow banking, as proposed by a range of academic economists over the past eighty years (including Irving Fisher (1936), Milton Friedman (1960) and James Tobin (1987)), would be a fourth. And limited purpose banking, with banking assets separately hypothecated and backed fully by equity, would be a fifth (Kotlikoff (2010)).

Enacting any of these structural solutions unilaterally would be difficult for a single country. That presents an operational obstacle. Against that, these structural measures have the advantage that they act directly on the topology of the network. As such, they are less subject to implementation error. Firebreaks and firewalls are familiar fail-safe devices against systemic problems in other networks, from infectious diseases spreading across people to infectious viruses spreading across the web, from transport networks to meteorology maps, from forest-fire control to military control (Haldane (2009)). These disciplines have found ways round the operational obstacles.

11. Conclusion

[To be added]

Acknowledgments

We thank Barry Willis, Clare Rogowski and, particularly, Jason Dowson for outstanding research assistance. We are also grateful to Lavan Mahadeva and seminar participants at the Bank of England for helpful comments and suggestions.

References

Acharya, V, Gale, D and T Yorulmazer (2009), Rollover risk and market freezes, mimeo, NYU.

- Adrian, T and H S Shin (2010a), Liquidity and leverage, Journal of Financial Intermediation, 19 (3), 418-437.
- Adrian, T and H S Shin (2010b), The changing nature of financial intermediation and the financial crisis of 2007-09, Annual Reviews: Economics, 2, 603-618.
- Albert, R, Jeong, H and A-L Barabasi (2000), Error and attack tolerance of complex networks, *Nature*, 406, 378-382.
- Alessandri, P and Haldane, A G (2009), Banking on the state, Paper from material presented at the Federal Reserve Bank of Chicago/World Bank event – International Financial Crisis: Have the Rules of Finance Changed?, 24 September 2009.
- Allen, F, Carletti, E and D Gale (2009), Interbank market liquidity and central bank intervention, Journal of Monetary Economics, 56, 639-652.
- Allen, F and D Gale (2000), Financial contagion, Journal of Political Economy, 108, 1-33.

- Anderson, R and R May (1991), Infectious diseases of humans: dynamics and control, Oxford University Press.
- Bank of England (2009), The role of macroprudential policy. November.
- Bhattacharya, S and D Gale (1987), Preference shocks, liquidity and central bank policy, in W Barnett and K Singleton (eds), New Approaches to Monetary Economics, Cambridge University Press, 69-88.
- Bolton, P, Santos, T and J Scheinkman (2008), Inside and outside liquidity, mimeo, Princeton University.
- Bordo, M, Eichengreen, B, Klingebiel, D and M S Martinez-Peria (2001), Is the crisis problem growing more severe? *Economic Policy*, 16(1), 51-82.
- Brunnermeier, M. K. and Pedersen, L. H. (2009), Market liquidity and funding liquidity, *Review* of Financial Studies, 22(6), 2201-2238.
- Caballero, R and A Krishnamurthy (2008), Collective risk management in a flight to quality episode, *Journal of Finance*, 63, 2195-2230.
- Caballero, R and A Simsek (2009), Complexity and financial panics, mimeo, MIT.
- CGFS (2010), The role of margin requirements and haircuts in procyclicality.
- Dasgupta, A (2004), Financial contagion through capital connections: a model of the origin and spread of bank panics, *Journal of European Economic Association*, 2(6), 1049-1084.
- Diamond, D and R Rajan (2009), Fear of firesales and the credit freeze, NBER Working Paper #14925.
- Fisher, I (1936), 100% Money, Revised edition, New York: Adelphi Company.
- Freixas, X, Parigi, B and J-C Rochet (2000), Systemic risk, interbank relations and liquidity provision by the central bank, *Journal of Money, Credit and Banking*, 32, 611-638.
- Friedman, M (1960), A Program for Monetary Stability. New York: Fordham University Press.
- Gai, P and S Kapadia (2010a), Contagion in financial networks. Proceedings of the Royal Society A, 466, no. 2120, 2401-2423.
- Gai, P and S Kapadia (2010b), Liquidity hoarding, network externalities, and interbank market collapse. Mimeo, Bank of England.

- Geanakoplos, J (2010), Solving the present crisis and managing the leverage cycle, Federal Reserve Bank of New York Economic Policy Review, August.
- Giesecke, K and S Weber (2006), Credit contagion and aggregate losses, Journal of Economic Dynamics and Control, 30, 741-767.
- Haldane, A G (2009), Rethinking the financial network, Speech at the Financial Student Association, Amsterdam, 28 April.
- Haldane, A G (2010), The \$100 Billion Question, Speech at the Institute of Regulation & Risk, North Asia (IRRNA), Hong Kong, 30 March.
- Hanson, S, Kashyap, A K and J C Stein (2010), A Macroprudential Approach to Financial Regulation, *Journal of Economic Perspectives*, forthcoming.
- Hatchett, J and R Kuhn (2009), Credit contagion and credit risk, Quantitative Finance, 9, 373-382.
- Heider, F, Hoerova, M and C Holthausen (2009), Liquidity hoarding and interbank spreads: the role of counterparty risk, *mimeo*, European Central Bank.
- Herring, R and Carmassi, J (2010), The Corporate Structure of International Financial conglomerates: Complexity and Its Implications for Safety and Soundness," in *The Oxford Handbook* of Banking, ed. by Allen Berger, forthcoming.
- King, M (2010), Banking: From Bagehot to Basel, and back again, Speech at the Second Bagehot Lecture Buttonwood Gathering, New York City, 25 October.
- Kiyotaki, N and J Moore (1997), Credit cycles, Journal of Political Economy, 105, 2, 211-248.
- Kotlikoff, L J (2010), Jimmy Stewart is Dead: Ending the World's Ongoing Financial Plague with Limited Purpose Banking, John Wiley & Sons.
- Krishnamurthy, A (2010), How Debt Markets have Malfunctioned in the Crisis, *Journal of Economic Perspectives*, forthcoming.
- May, R, Levin, S and G Sugihara (2008), Ecology for bankers, Nature, 451, 893-895.
- May, R and N Arinaminpathy (2010), Systemic risk: the dynamics of model banking systems, Journal of the Royal Society Interface, 46, 823-838.
- Minguez-Afonso, G and H Shin (2010), Precautionary Demand and Liquidity in Payment Systems, *Journal of Money, Credit and Banking*, forthcoming.

- Morris, S and H Shin (1998), Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks, American Economic Review, 88, 587-597
- Newman, M (2002), The spread of epidemic disease on networks, *Physical Review E*, 66, 016128.
- Newman, M, Strogatz, S and D Watts (2001), Random graphs with arbitrary degree distributions and their applications, *Physical Review E*, 64, 026118.
- Nier, E, Yang, J, Yorulmazer, T and A Alentorn (2007), Network models and financial stability, Journal of Economic Dynamics and Control, 31, 2033-2060.
- Repullo, R (2005), Liquidity, risk taking and the lender of last resort, International Journal of Central Banking, 1, 47-80.
- Simon, H (1962), The architecture of complexity, Proceedings of the American Philosophical Society, 106 (6), 467-482.
- Shin, H (2009), Financial Intermediation and the Post-Crisis Financial System, Paper presented at the 8th BIS Annual Conference, June.
- Soramaki, K, Bech M, Arnold, J, Glass, R and W Beyeler (2006), The topology of interbank payment flows, *Federal Reserve Bank of New York Staff Papers #243*.
- Stiglitz, J and A Weiss (1981), Credit Rationing in Markets with Imperfect Information, American Economic Review, 71(3), 393-410.
- Tobin, J (1987), The Case for Preserving Regulatory Distinctions, in *Restructuring the Financial System*, Federal Reserve Bank of Kansas City, 1987, pp. 167-183.
- Upper, C (2007), Using counterfactual simulations to assess the danger of contagion in interbank markets, Bank for International Settlements Working Paper #234, August.
- Watts, D (2002), A simple model of global cascades on random networks, Proceedings of the National Academy of Sciences, 99, 5766-5771.
- Weitzman, M L (1974), Prices vs. Quantities, Review of Economic Studies, 41, 477-91.





Figure 2 : Network of large exposures $^{(a)}$ between UK banks, 2008 $^{\rm (b)(c)}$



Source: FSA returns.

(a) A large exposure is one that exceeds 10% of a lending bank's eligible capital during a period. Eligible capital is defined as Tier 1 plus Tier 2 capital, minus regulatory deductions.

(b) Each node represents a bank in the United Kingdom. The size of each node is scaled in proportion to the sum of (1) the total value of exposures to a bank, and (2) the total value of exposures of the bank to others in the network. The thickness of a line is proportionate to the value of a single bilateral exposure.(c) Based on 2008 Q1 data.

Figure 3a:Repos and Financial CP as a Figure 3b:proportion of M2 in the USproportion of

Figure 3b: Repos and Financial CP as a proportion of M2 in the UK



Source: Adrian, T and Shin, H S (2010b), "The changing nature of financial intermediation and the financial crisis of 2007-09" *Annual Reviews: Economics*, 2, 603-618.



Source: Bank of England and Bank calculations





Source: Bank of England calculations and ONS Blue Book

Figure 5: Concentration of the UK and US banking systems



Source: FDIC and Bank calculations. Top 3 banks by total assets as a percentage of total banking sector assets. Data is to January 2009





Source: Sheppard, D. K (1971), FDIC and Bank of England calculations. Top 3 banks by assets as a percentage of National GDP

Note: The definition of UK banking sector assets used in the series is broader after 1966, but using a narrower definition throughout gives the same growth profile.

(a) Red line shows US banking assets as a percentage of US GDP (left-hand scale).

Figure 7: Herfindahl Index of Concentration in UK Interbank Market, 2004-2008



Figure 8 Three-month interbank rates relative to expected policy $\mbox{rates}^{(a)}$



Sources: Bloomberg and Bank calculations.

(a) Spread of three-month Libor to three-month overnight index swap (OIS) rates. Five-day moving average.

Figure 9: Reserve Holdings at Central Banks



Source: Federal Reserve Bank and Bank of England calculations Data is to March 2009 before the introduction of Quantitative Easing

Figure 10: A Directed Network with Five Nodes



Figure 11: Stylised Balance Sheet



Figure 12: Systemic Liquidity Hoarding (Poisson Distribution; Single Random Idiosyncratic Haircut Shock)



Figure 13: Systemic Liquidity Hoarding (Poisson Distribution; Aggregate Haircut Shock and Single Random Idiosyncratic Haircut Shock)



Figure 14: Systemic Liquidity Hoarding (Geometric Distribution; Single Random Idiosyncratic Haircut Shock)



Figure 15: Systemic Liquidity Hoarding (Poisson Distribution; Single Haircut Shock Targeted to Bank with Highest In-Degree)



Figure 16: Systemic Liquidity Hoarding (Geometric Distribution; Single Haircut Shock Targeted to Bank with Highest In-Degree)



Figure 17: Systemic Liquidity Hoarding (Geometric Distribution; Single Random Idiosyncratic Haircut Shock; 25% Unsecured Interbank Liabilities)



Figure 18: Systemic Liquidity Hoarding (Geometric Distribution; Single Random Idiosyncratic Haircut Shock; 3.5% Liquid Asset Holdings)



Figure 19: Systemic Liquidity Hoarding (Geometric Distribution; Single Random Idiosyncratic Haircut Shock; 3.5% Average Liquid Asset Holdings but set via Systemic Liquidity Surcharge Rule)



Figure 20: Systemic Liquidity Hoarding (Geometric Distribution; Single Random Idiosyncratic Haircut Shock; Partial Withdrawal)

