Untying the Gordian Knot:
Interconnectedness, stability and the post-crisis reforms

Dermot Turing
Visiting Fellow
Kellogg College,
60-62 Banbury Road,
Oxford OX2 6PN, UK

Non-publication draft [11 Nov 2019]
Comments welcome: please address to dermotturing@btinternet.com

Abstract

‘Interconnectedness’ was considered to be a cause of the 2008 financial crisis, stimulating a number of studies into the topology of financial markets. Yet the analysis of instability within networks has tended to focus on a type of ‘contagion’ which imagines serial insolvencies, with non-performance of due obligations causing solvency issues for connected institutions. A more realistic assessment of the 2008 crisis was that it was due to a drying-up of available cash. A taxonomy of contagion is proposed, and the illiquidity model of contagion is then analyzed with reference to the observed core-periphery structure of financial market networks. Finally, the post-crisis reforms are judged against the view of ‘interconnectedness’ which emerges.

I Introduction

As politicians and regulators struggled to make sense of the European-American financial crisis in the autumn of 2008, one thing seemed reasonably clear. What had begun as doubt about the ability of firms like Bear Stearns and Lehman Brothers to meet their margin calls - to pay their debts on time – had spread across the financial system. The failure of Lehmans led to the bail-out of American International Group and the absorption of Merrill Lynch, Wachovia and Halifax-Bank of Scotland into other, possibly stronger, banks. The Royal Bank of Scotland and Lloyds Banking Group were partly nationalized in order to assure their solvency. These failures and near-misses, and the market-wide bailout of American banks under the ‘troubled assets relief program’ were all ascribed to ‘interconnectedness’: the failure of one had contaminated the others. Dismantling interconnectedness became one of the central themes of the reformed post-crisis financial system.

Interconnectedness thus deserves thorough study. In the first place, one needs to understand the configuration of connections that bind the participants in the marketplace to one another – the topology of the market. And secondly, one needs to understand what those
connections consist of, so as to understand how a disease may spread from one entity to another. A good deal of work has been done in the decade since the crisis, yet the thinking on both network structure and on the varieties of interconnection is incomplete. The picture is further complicated by the introduction, as part of the portfolio of post-crisis reform measures, of structural changes which alter the topology of the market and, possibly, the nature of inter-institutional dealings. It is the purpose of this paper to take stock of the situation.

The structure of this paper is as follows. Section II reviews network theory as applied to financial markets, in an attempt to understand the meaning of ‘interconnectedness’ of financial firms. In section III the problem of contagion is analysed, leading to the conclusion that many studies of stability may be focused on a problem which was not actually observed in the 2008 crisis. A taxonomy of contagion is proposed, with ‘Type 2 contagion’ – propagation of illiquidity in the sense of unwillingness to lend – being put forward as a more likely candidate for the cause of the crisis. Type 2 contagion is then used in section IV to judge the reforms introduced since the crisis, in particular the changes to prudential regulation and the move to mandatory clearing of derivatives transactions. The conclusions, set out in section V, are that the reforms have not addressed the underlying interconnectedness problem, as exposed through Type 2 contagion, and may even have been counterproductive. Finally it is suggested that, alongside further research, a review of Lender-of-last-resort policy may be desirable.

II Topology and Interconnectedness

It is axiomatic, at least as far as the widely-understood analysis of the 2008 crisis is concerned, that ‘contagion’ spreads through a network of financial institutions, from one to another. (Exactly what it is that spreads is unpacked in the following section.) The shape of the network, that is to say its topology, may critically affect the speed and extent of the contagion. So to begin, it may be appropriate to recapitulate some basic concepts from network science.

- Classical network analysts call entities ‘nodes’ and their interlinkages ‘edges’. In this paper, the focus is on financial firms and their counterparties and clients. They are referred to as ‘entities’ or ‘institutions’ rather than ‘nodes’, and the terminology of ‘edges’ is not followed in the assessment of different types of interlinkage.

- The number of linkages which an entity may have is referred to as its ‘degree’. But different entities may have different degree, so a fundamental question for a network modeller is to envision how the variance in degree is distributed between the entities. It may be random or it may be ‘scale-free’, that is to say disproportionately weighted so that a small number of entities have very large degree but a large number of entities have very small degree. (The scale-free phenomenon is observed in all kinds of networks occurring in the real world, such as website popularity or followers on social media.)

- One variety of scale-free networks is the core-periphery network. A core-periphery network should thus be understood as having a small number of entities forming a densely interconnected ‘core’, whose participants have interconnections to other core participants with extremely high degree, and a large number of entities which sit outside the core, have few interconnections with each other, but which are each connected to one or several participants in the core (Borgatti and Everett, 1999; Csermely et al, 2013).

Figure 1 illustrates these different types of network.
Studies of financial market networks have been undertaken. Setting aside a tiny number of papers pre-dating the crisis, the specialism of network study in finance would seem to begin with the identification of ‘interconnectedness’ as a cause of the crisis. However, it must be noted that a substantial portion of these post-crisis studies is of an entirely theoretical nature, for example answering questions about the robustness of a theoretically-modelled network following removal of one or more entities.

Unfortunately, theoretical models of networks can be misleading. For one thing, actual network structures observed in the financial sector (like network structures in other walks of life) may not conform to the ideals of models. Then there is the danger that the shocks imposed on the modelled network are not the type of shock which could actually occur (the question addressed in the following section of this paper): whether it is a valid approach to imagine the random or targeted removal of financial entities, from the network, as the source of a crisis, is open to question. Next, one has to deconstruct the concepts of ‘stability’ and ‘robustness’ which are being looked for in determining an optimum network structure. That also is not easy: should financial stability mean re-establishment of the network in a form similar to that before the shock, or does it mean minimizing the number of infected entities? Different research groups have approached this issue in different ways. Taking all these criticisms on board, it may be difficult to interpret the conclusions, especially when they show apparently contradictory behaviours (e.g. financial networks are ‘robust but fragile’ (Gai and Kapadia, 2010) or ‘connectivity is both a driver of contagion and a hedge against contagion via diversification’ (Chinazzi et al., 2015)).

While it may be useful to take note of the model-based studies, a more fruitful approach may be to look at empirical studies which observe the actual shapes of financial network. These show some interesting similarities. Fiedor et al (2017) used data to analyse the European interest-rate swaps market. They find that the topology is highly asymmetric and has a core-periphery pattern, confirming and amplifying observations of earlier researchers (e.g. Markose (2012), Borovkova and El Mouttalibi (2013)). Developing markets, lending, and international activity between financial institutions show similar characteristics (Cont et al., 2012; Silva et al., 2017; IMF, 2018; León et al., 2018). What seems to be happening in the wholesale financial market is that a relatively small group of significant institutions (nowadays recognized as systemically important banks or SIBs) is both highly connected to other institutions in its peer-group, and to provide services to a very much larger group of less significant client institutions which have few connections between each other and, individually speaking, not many connections to the core institutions. One should thus perceive the network in layers, where the upper layer is thinly populated but densely interconnected, and the lower layer is populous but weakly connected.

From here, one might conclude that theoretical model-based studies which look at randomly configured networks are of less value than those which look at scale-free structures; but even so, a model of a scale-free network may be unreliable when it looks at the interlinkages within the core of a core-periphery structure. Care is needed in interpreting all model-based work. It cannot be assumed that any conclusion relating to the stability, robustness or any other criterion for judging quality of a theoretical structure will translate across to the actual core-periphery networks in existence.
Stability and Contagion

The purpose behind much of the theoretical analysis of network structures has been to understand the weaknesses of financial market topology, so as to be better prepared for another shock. It may therefore be somewhat surprising to discover that there is no agreement on what should be regarded as a ‘shock’ or indeed what constitutes ‘contagion’ in a stressed financial market.

The classic example of contagion is ‘one bank failure causing the failure of another bank, even though the second bank initially seemed solvent’ (ECB, 2009). Instantly one imagines, then, that the contagion observed in the 2008 crisis was that the institution of insolvency proceedings in respect of Lehman Brothers caused the insolvency of AIG, Merrill Lynch, RBS and the others. Presumably, then, the interconnecting factor was that Lehman defaulted on its payment or collateral-delivery obligations to these other entities, causing their insolvency and rippling out as those now-insolvent entities defaulted in turn. Many of the network studies which have been undertaken since the crisis assume this style of cascade of defaulted obligations as the background for their analysis. Unfortunately for those studies, this is not what actually happened in 2008.

Contrary to the initial understanding of the policy-makers trying to diagnose the problem in the heat of the 2008 crisis, the trouble lay not with derivatives transactions, but with an impairment of asset values, a run on liquidity and a lack of trust exacerbated by inadequate information about the viability of peer-group institutions (Gorton, 2009; Summe, 2011; Ricks, 2016; Ahnert and Georg, 2018). The symptoms might better be described as ‘panic’ than ‘interconnectedness’ (cf. Ricks, 2016). Moreover, as shown by Adrian and Shin (2010), the natural behaviour of core participants in the repo markets, a primary source of interinstitutional liquidity, was pro-cyclical, so that they would manage their balance sheets to conserve rather than lend cash, depressing the value of collateral and deepening the crisis.

In other words, the problem was with liquidity, meaning cash liquidity: banks would not lend, and the impairment of asset prices meant that the markets for secured interbank lending instruments such as repo became closed. All this was not to say that banks had no cash, or that they held bad assets, still less that banks were insolvent. The problem was all in the perception: that it would be foolish to lend, even against securities, since the default risk of the borrower was unmeasurable (even when enhanced with collateral of unmeasurable value). This kind of contagion is different: it is liquidity hoarding, and it means that a needy borrower, which has an immediate thirst for cash cannot satisfy its need, even if it is rich in high-quality collateral. When the market dries up, nobody will buy collateral which is being liquidated in response to the default, and what is market illiquidity can turn into a solvency problem for a bank which cannot meet its cash obligations. This is a different species of contagion from the chained-insolvencies of the imagined classic example.

One might thus categorize two types of contagion:

- **Type 1 contagion.** This is the classic perception of contagion: the failure of a bank causes the bank to default on its payment obligations, thus causing an unsolvable solvency or liquidity problem for one or more of its creditors, which collapse in turn, spreading the solvency problem throughout the network in a chain of failures.
• Type 2 contagion. This concept of contagion does not need to presume insolvency of any financial firm, but considers the implications of illiquidity. A shock may come from impairment of asset values as much as from a firm’s insolvency. The inability of a firm to obtain liquidity might lead to inability to pay debts, but it the immediate result will be its own unwillingness to lend. That means that the illiquidity propagates through the virtual network of potential counterparties who would ordinarily seek to borrow from the affected lender. (Cf. Brunnermeier and Pedersen, 2008; Di Maggio and Tahbaz-Salehi, 2015). The cause may be ‘information contagion’ (Ahnert and Georg, 2018) as much as classically-understood asset price impairment.

Some observations may be in order at this point. First, if Type 2 contagion is what was observed in 2008, and what one should focus on, the type of network we need to look at, in order to understand the factors shaping instability, may be different. Rather than study the existing loan relationships and draw inferences about who might be next to default (cf. the analyses of Nier et al. (2008), Cont et al. (2012), Glasserman and Young (2015), Acemoglu et al. (2015), Barroso et al. (2018), etc.) we should instead identify the rather more nebulous population of potential lenders to whom an entity in need of liquidity might expect to turn. Such an approach is possible, as in the study by Lenzu and Tedeschi (2012).

Type 2 contagion appears to be the wholesale denial of new interlinkages into the principal candidate lenders. Translating that to the core-periphery network structure which appears to prevail in wholesale money markets, we discover that within the core it may suddenly become impossible for cash to find its own level – to be allocated to the places where it is needed, and thus to be denied to the periphery as well as to the core. That can happen without the failure of a single firm. Indeed, the behaviour of the US repo market in September 2019 appears to have symptoms of cash-aridity and contagion, along the lines of Type 2 contagion (cf. Economist (2019)). León et al. (2018) have identified that certain institutions act as ‘super-spreaders’ of central bank liquidity in a network of banks: perhaps it is such institutions (which, in the global context, may possibly be the SIBs) on which the effort to avoid Type 2 contagions should be focused.

Insofar as Type 2 contagion spreads by information contagion, or what in common parlance is the ‘grape-vine’, one might enquire whether the network structure observed for existing interinstitutional transactions is a valid proxy, or model, for the network for information transmission. Undoubtedly there are additional players in the information network, notably the official sources of news and price-feed providers. But these channels may not be implicated in the contagious spread of a panic. It seems likely that panic propagates during a crisis because the processes by which interinstitutional transactions are created involve requests for quotes along the linkages which are habitually used by the parties to transactions, and it is along those same linkages that hints and indications of market tightness are conveyed. If that is true, then the existing transactional network is a valid model for the network for information contagion. On the other hand, as with other models of contagion, there may be super-spreaders whose role in transmitting the problem (the panic) is disproportionate (for review, see May (2013)), and those super-spreaders may not be identical with the largest market players in terms of transactional volume. Further research may be desirable in this area.
The solution to problems of ‘interconnectedness’ has included, since the financial crisis, some degree of adoption of a different network shape, namely the introduction of various disincentives to interinstitutional transactions and making clearing of certain derivative products mandatory. The impact of these changes is considered in section III below. For now, it may simply be asked whether there might be a better, or even an optimal, network structure for the limitation of Type 2 contagion. The classic exposition of network stability by Baran (1967) demonstrated that redundancy of pathways created reliability and resilience in the event of entity failure. But Baran was working on a problem which is different from that which occurs in a financial crisis: his mission was to ensure the effective delivery of a message from one side of a network to another – something which led to the adoption of packet-switching for transmission of messages over the internet. In other words, Baran’s idea of stability was the opposite of what might be desirable in the case of information contagion: the spread of data (or pseudo-data) across a population is something to be contained, not ensured. A highly interconnected core of financial institutions who are in constant communication with each other does not easily lend itself to isolation and control.

IV The impact of the post-crisis reforms

Having attempted an analysis of financial sector interconnectedness, one might now be in a position to consider the impact of the post-2008 reforms. Various efforts to prevent a recurrence of 2008, which are relevant to network structure and contagion, might be mentioned: an increase in the amount of top-tier capital required of financial firms; the introduction of new prudential measures, especially a liquidity coverage ratio (LCR); and the requirement that standard derivatives trades be cleared at a central counterparty (CCP).

A. Prudential measures

As regards capital requirements, it will be obvious that capital addresses an institution’s solvency, not its liquidity. Insofar as a crisis might involve Type 1 contagion, extra capital will help. As to Type 2 contagion, additional capital may inspire greater confidence among the firm’s potential lenders, thus making the risk of Type 2 contagion less probable than in 2008. In relation to network structure, it may be observed that the additional capital required of firms branded SIBs (BCBS, 2013) will be situated within the core of the network, so while these systemically important firms may be less likely to fail (and may be more willing to lend among their peers or to clients), it is unclear whether regulatory capital improvements in the core would lead to this outcome (cf. Dautović, 2019), especially if information contagion underlies the loss of confidence.

The introduction of a liquidity coverage ratio is intended to ensure that firms have adequate resources to meet cash outflows in a period of crisis. In principle, therefore, a firm will be better equipped not only to satisfy deposit withdrawals or drawdowns on loans, but may be able to respond more positively to approaches for fresh cash liquidity. However, the analysis of LCR is more difficult than that simple theory might suggest. LCR is designed to help a single firm cope with a run – a run on the bank, rather than a closure of the market to requests for liquidity.
Further analysis is required to see the effect of LCR in an arid market, but some tentative suggestions might be inferred. If there is an asset price shock at the heart of the cash aridity problem, LCR may act in a perverse way. Firms will be required to increase their stock of HQLA, with the possible result that they will hoard HQLA and cash. Furthermore, an asset price shock will likely lead to increases in margin calls from bilateral counterparties and CCPs (Turing and Singh, 2018) reducing the amount of HQLA available for repo operations. A margin hike not only leads to HQLA storage but, in relation to lower-grade HQLA, imposes its own LCR requirement on the collateral provider. So it may be that LCR works counter to the objective of resolving a liquidity crisis, even though it is intended to be a liquidity-enhancing tool.

One further change, introduced as part of the package of revisions to the prudential regulatory framework, might be mentioned. The move to mandatory clearing of derivatives could have had various undesirable prudential side-effects, which were mitigated through various adjustments to the Basel 3 framework so as to remove disincentives to clearing members offering clearing services to clients. One of these adjustments was to allow clearing members to disregard their exposure to CCPs on cleared client transactions, if the clearing member secured the client’s agreement that non-payment by the CCP excused the clearing member from making corresponding payment(s) to the client. The pay-as-paid standard will roll a Type 2 contagion, if it results from action such as variation margin gains haircutting (cf. Heath et al. (2015); Turing and Singh (2018)) out to the periphery.

B. Mandatory clearing

A flagship of the post-crisis reforms has been the clearing mandate: that all standardized derivatives transactions should be cleared at a central counterparty (CCP). In terms of interconnectedness, the complex web of interbank derivatives transactions covered by bilateral master agreements would be replaced by a simpler pattern of links between CCPs and their participants, and (insofar as banks or end-users were not themselves CCP participants) links between non-participants and clearing members of CCPs. Vast swathes of interconnections would be removed.

That thinking is not erroneous. The topology of interbank relations is definitively altered by the introduction of mandatory derivatives clearing. The studies of Fiedor et al (2017) show a markedly different market shape, split among three tiers, where CCPs are the most highly-connected, clearing members have few links inter se but are connected to a few CCPs, and end-users are connected typically to only one, or maybe a small number of clearing members, each of whom serves many end-users. This is a very different pattern from the densely interlinked pattern of before.

Various studies have been carried out on the effect of bringing CCPs into the network. The now-classic work of Duffie and Zhu (2011), which showed that clearing would actually reduce netting efficiency, has been reinforced and validated by further work by Garratt and Zimmerman (2015) and Benos et al. (2019). Markose et al. (2017) have argued that CCPs are possibly ‘too interconnected to fail’ and, following Heath et al. (2015), have studied the potential for CCPs to act as conduits for Type 1 contagion unless properly capitalized. Borovkova and El Mouttalibi (2013) carried out a study of Type 1 contagion, looking at the difference which
interposition of a CCP might make, concluding that in heterogeneous networks (such as those observed in actuality) defaults of smaller firms are more probable with clearing than without. Berndsen et al. (2018) looked at ‘financial market infrastructures’, which in the context of their Colombian dataset meant essentially settlement systems rather than CCPs; insofar as their conclusions transpose to other infrastructure types their observation that modularity (substitutability) is diminished, and instability increased, in multiplex structures, the inference is that the change to a more layered structure observed in practice in cleared markets by Fiedor et al. (2017) has rather reduced, than improved, stability. The Fiedor et al. study was done at a time when the process of moving end-user derivatives transactions into clearing was not complete. The change in market topology observed by them will be amplified unless clearable instruments for some reason fall into disfavour.

The general direction of all this work tends to suggest that making clearing mandatory has not unequivocally created better network structures, even if the objective was to reduce Type 1 contagion.

However, the somewhat negative perspective on the drive towards centralized clearing, which may be suggested by these studies, should be understood in its proper context. Nobody doubts that the failure of a CCP would be a catastrophic event, which is why so much effort is currently being devoted to creating a credible framework for CCP resolution. But if the domino-insolvency model of how financial crises unfold is inaccurate, and Type 2 contagion is more realistic, then the conclusions from modelling Type 1 contagion and CCPs should be treated with reserve.

While a vulnerable CCP is undoubtedly a problem, the type of financial crisis being considered in this paper is not one of entity solvency but that of cash aridity – market-wide reluctance to engage in new lending. In relation to Type 2 contagion, the question is whether introducing a CCP into the picture provides a new or alternative way to address cash liquidity shortages. CCPs are not sources of liquidity. CCPs do not lend. Instead, they suck in collateral in the form of cash and HQLA to satisfy their margin and default fund requirements. (It may be noted that CCPs demand collateral from both counterparties to a trade, so their involvement is not a zero-sum game in terms of liquidity drainage.) This behaviour is enhanced in times of crisis (Turing and Singh, 2018). Nor are they marketplaces where it is possible for lenders and borrowers to meet and forge new counterparty relationships: they are post-trade service providers who only engage once a transaction has been agreed. So CCPs will not help with Type 2 contagion; indeed, the additional collateral they may demand could actually worsen cash aridity.

Banks lend to each other, typically through repo structures, and they participate in lending syndicates, place deposits inter se, use each other as custodians and brokers, and have many other relationships. These linkages will not be affected by a move to centralize counterparty relations in derivatives transactions; so it might be the case that pushing derivatives into clearing has not affected the overall interconnectedness of the financial firms in the network’s core, only perhaps the strength of the linkages. These linkages could be relevant in a financial crisis, insofar as deposits are withdrawn or repos unwound. But, if it is accepted that the problem observed in the crisis was cash aridity then these existing relationships are not greatly relevant. A final comment may be made on the existence of CCPs which clear repo trades. Repo clearing was not made mandatory under the post-crisis reforms, so the network topology of a cleared marketplace will be different from that for instruments which are subject
to a clearing mandate. Entities will have a free choice whether to clear or to retain bilateral links within the core of the network. The existence of a clearing option may be a source of confidence for firms wishing to enter into new repo transactions during a period of market turmoil, assuming that CCPs are regarded as much less likely to fail than a peer counterparty.

V Conclusions

This paper has suggested that a fresh look at the relationship between interconnectedness and financial crises is desirable. While interconnectedness has typically been analysed in terms of Type 1 contagion, it is argued that the phenomena associated with Type 1 contagions are not actually observed during a crisis, whereas in fact the problems stem from the drying-up of usual sources of liquidity – and that illiquidity can propagate throughout the network, manifesting as Type 2 contagion. Further research into this topic is considered desirable.

The evidence is that the wholesale financial market has a core-periphery network topology. A relatively small number of banks occupy the highly interconnected core, providing services to each other and to the poorly-connected banks in the periphery. These core banks, which are likely those categorized by the Financial Stability Board as SIBs, are net debtors, who do not seek liquidity from the periphery. Core banks are also net collateral takers. These factors expose participants in the core to Type 2 contagion risk. Their failure would not necessarily cause failures in the periphery, but their unwillingness to lend (whether caused by absolute failure, or by cash aridity in the core) would affect the peripheral banks’ access to cash and thereby spread an illiquidity crisis throughout the system.

The works on network topology, plus other observed trends and incentives in the firms affected would suggest that the risk of Type 2 contagion has in fact risen, rather than been lowered, by the changes of the last 10 years. First, take clearing. The use of CCPs has been mandated in respect of certain derivatives products; yet CCPs have a thirst for liquidity and collateral in times of market stress. So far from being a solution to ‘interconnectedness’ within the core, CCPs may in fact exacerbate a Type 2 contagion problem.

More broadly, the trend since the crisis has been to increase, rather than dilute, concentration of business within the core. As many G-SIBs (Global Systemically Important Banks) have been removed from the list maintained by the Financial Stability Board as have been added to it since 2011. The study (FSB, 2018) of clearing interdependencies shows a high degree of service concentration within the core: notably, only 17 out of the 66 liquidity providers to CCPs are not also clearing members; the FSB analysis implies that a clearing member default could affect not just the CCP’s access to liquidity but to custody and settlement services, and jeopardize cash investment transactions and more.

The introduction of formal prudential requirements for Liquidity Coverage, while helpful at an individual-firm level, may be inadequate to deal with an asset price shock affecting the whole market; the Liquidity Coverage Ratio is not designed to address a whole-market liquidity crisis, just a run on a single bank.

The question then is what could or should be done to reduce the risk of Type 2 contagion. Relaxation of rules which tend to concentrate business within a core of SIBs could lead to decentralisation and a more stable market topology. It appears (cf. Freixas et al., 2000) that an uncontaminated central source of liquidity could reinject confidence, and cause the cash
to flow again – much as the repo market seizure since September 2019 has been addressed. Lender of Last Resort policy could be adjusted where cash aridity is observed.

References


Basel Committee on Banking Supervision (BCBS), 2013. Global systemically important banks: updated assessment methodology and the higher loss absorbency requirement


