Network structure, indirect losses and financial contagion in inhomogeneous, stochastic interbank networks

by

Nadine Mari Walters

under supervision of
Dr Conrad Beyers and Dr Gusti van Zyl

Submitted in partial fulfilment of the requirements for the degree

PhD Actuarial Science

In the Faculty of Natural and Agricultural Sciences

University of Pretoria

Pretoria

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

The financial assistance of the National Research Foundation (NRF) and the Absa chair in Actuarial Science towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF or Absa.

Abstract

We introduce new tiered bank network structures, allowing for many different bank sizes, and compare risk propagation in these structures with the well-known Erdős-Rényi, assortative and disassortative structures. The simulations indicate that in the presence of market sentiment and liquidity effects, the details of the structures in combination with the distribution of assets, the system's interconnectedness and its size are crucially important in determining the risk of major capital loss in the network. In fact, even networks with similar levels of tiering can behave markedly different depending on these factors. In the absence of market sentiment and liquidity effects, the differences between the network structures is smaller. This highlights the importance of considering the network structure in conjunction with network characteristics, market sentiment and liquidity effects. This implies that policy actions aimed at influencing a network's characteristics must consider all aspects unique to that particular system and cannot follow a 'one-size-fits-all' approach. The framework is illustrated with an application using South African bank balance sheet data. Spikes in simulated assessments of systemic risk agree closely with spikes in documented subjective assessments of this risk. This indicates that network models can be useful for monitoring systemic risk levels.

In a large network setting, the study then considers the fraction of nodes that default in stochastic, inhomogeneous financial networks following an initial shock to the system. Results for deterministic sequences of networks are generalized to stochastic networks to account for interbank lending relationships that change frequently. A general class of inhomogeneous stochastic networks is proposed for use in systemic risk research, and we illustrate how results that hold for Erdős-Rényi networks can be generalized to the proposed network class. The network structure of a system is determined by interbank lending behaviour which may vary according to the relative sizes of the banks. We then use the results to illustrate how network structure influences the systemic risk inherent in large banking systems.

Contents

Chapter 1

Introduction

The purpose of this research is to investigate how network structures can influence the way that shocks propagate through financial networks and how this may affect policy decisions. For smaller networks, we investigate how the interaction between network structures, liquidity risk and market sentiment can influence the results obtained. By focussing on banking systems, we propose innovative structures to use for simulating the spread of contagion between banks. This is used to show that different structures behave differently under changes to network characteristics and therefore may react differently to changes in legislation.

The first part of this research is focussed on smaller networks. A new simulation model is developed and used to show how the risk of systemic collapse is affected by differences in the network structure and other characteristics. While the proposed model is highly simplified, it does incorporate certain real-world mechanics, namely heterogeneity between bank asset sizes, market liquidity losses, investor sentiment, asset maturities, differences in asset compositions and to some extent real-world lending behaviour. It can thus be considered as more comprehensive and realistic than previously proposed systems. By applying this framework to South African bank balance sheet data, it is shown that the model is capable of detecting increases in systemic risk over time.

The second part of this research is focused on larger networks and therefore makes use of asymptotic results. Known asymptotic results are shown to be applicable to a newly defined, versatile class of stochastic networks. This is then used to make a similar comparison between network structures as was made for smaller interbank systems.

The rest of this chapter is organised as follows: Section 1.1 discusses the background to the research. Section 1.2 gives a short overview of literature that is relevant in order to motivate the research questions. This includes a brief overview of the necessary network

theory background in section 1.2.1. Section 1.3 then elaborates on the purpose and significance of the study, while section 1.4 is used to set out the structure for the remainder of this thesis.

1.1 Background

Systemic risk and the spread of financial contagion are important considerations for regulators tasked with overseeing stability of banking systems. Banking systems are at the core of a well-functioning financial system. A breakdown of the system would hinder economic growth, which in turn may cause permanent damage to the economy [34]. Therefore, it is important for regulators to prevent such a breakdown from being triggered. Regulatory intervention at a late stage could prove to be costlier than intervention at an earlier stage. The burden of costly bailouts by the regulator are ultimately borne by the taxpayers, which negatively affects the economy. On the other hand, if banks are allowed to fail without any intervention, the economy can be strained by losses on investors' deposits, rising interest rates, possible bank runs etc. Monitoring the level of systemic risk in a financial system is therefore crucial for ensuring long-term stability and growth of an economy.

Liquidity and market sentiment are two key requirements for a working banking system that are also closely related. During times of economic distress, a lack of trust translates into a reluctance of non-bank financial institutions to renew funding to banks. They then impose more stringent lending requirements, which leads to increased risk premia on loans and debentures thereby increasing banks' wholesale funding costs. The higher interest rates charged on servicing new debt means that additional assets may need to be liquidated to service the debt or a reduction in asset origination, reducing (shrinking) the balance sheet sizes of the affected banks. This puts a strain on those banks' liquidity positions as the maturity mismatch between short-term liabilities and assets increases. Ultimately, when the funding costs become unsustainably high the bank may be forced to call in loans or liquidate assets prematurely. This, together with the increased funding costs can substantially reduce the bank's profitability and hence its retained earnings. This in turn reduces its Tier I capital, which may lead to solvency problems [49]. This creates a spiral of distrust.

The complex nature of banking systems remains difficult to replicate and model precisely. Bottom-up approaches using integrated modelling frameworks are very useful, yet they are difficult to calibrate, expensive and not readily available. This is because in practice, such an approach would involve the regulator providing a specified scenario to all banks, after which the banks quantify their own risk position so that the regulator can then

aggregate the risk positions [24]. It is therefore of interest to find simplified models that consider the entire system from the start and that can detect changes in systemic risk. We contribute to this by showing that network models of systemic risk can satisfy this requirement to a large extent. We illustrate how such a top-down model can be used by first applying it to a hypothetical banking system. It is then applied to real-world balance sheet data, and it is shown that changes in risk can be detected under times of market stress for various network structures.

The chain of events that we aim to model is as follows: One bank in the system experiences solvency problems, which may arise because of a significant increase in impairments from non-performing loans. This could be because of a number of events such as unsustainable lending practices or a disruption in its target market (such as the mine closures experienced in South Africa). It is important to note that the applied model does not require us to specify the event that leads to the initial bank's default, nor do we attempt to model it. The equity of the aforementioned bank then declines, and shareholders need to absorb the losses (followed by other subordinated creditors). Now there are three key potential effects on the banking system. Firstly, the bank may start to default on its interbank credit obligations. Secondly, other banks' balance sheets may be affected through a revaluation of assets and impairment provisions and they may need to raise additional impairment provisions (e.g. if the initial bank's troubles were due to increased impairments on a specific type of loan book, other banks may need to raise their impairment provisions for similar books to account for an anticipated rise in impairments). Another possibility is that the bank may ultimately need to resort to forced sales to generate liquidity. The increased supply of those assets in the market may depress their market value, leading to mark-to-market losses for other banks holding similar assets. For this study, the distinction between these possibilities (and hence the effect of the initial default on the banking or trading books of other banks) is not explicitly made. Here, we assume a net reduction in the balance sheets of other banks takes place which could be due to mark-to-market losses, an increase in required reserves, or to a combination of both. This approach is adopted to keep the model simple and consistent with existing models in the literature (see for example [78, 52, 75, 13]). From here on, whenever the approach adopted for this study makes reference to market liquidity risk, it is done with the understanding that losses due to raised provisions is also considered. The second effect of the initially troubled bank affects the liability side of other banks' balance sheets and is more likely to lead to contagion. Funders' trust in the ability of banks to service their debt may decline as they become incapable of distinguishing between financially sound and troubled banks. This leads to liquidity issues, as the cost of rolling forward short-term debt increases for the affected banks as

Table 1.1: Illustration of a simplified bank balance sheet.

Assets	**Equity & Liabilities**
Central bank deposits	Total Loss Absorbing Capacity (TLAC)
Interbank assets	Unsecured creditors
Retail loans	Interbank liabilities
Corporate Investment Banking (CIB) loans	Guaranteed depositors
Other assets	Secured creditors

non-bank financial institutions are reluctant to renew their loans. Banks need to roll forward their short-term debt as they usually invest in long-term assets and take short-term deposits from funders. This gives them the needed liquidity at a low funding cost under normal circumstances. Banks may then be forced to sell assets below their market value to generate liquid funds and avoid maturity mismatches on their balance sheets.

A bank's balance sheet can be illustrated by table 1.1, where the order of the liabilities roughly reflects seniority levels. In the event of a shock on the non-interbank assets of a bank, the TLAC portion of the liabilities will be affected first, followed by the unsecured creditors. If problems in the system spread even further, banks could (depending on the jurisdiction) choose to default on depositors before defaulting on interbank payments to try and salvage confidence in the system. Interbank creditors that must bear the losses of a shocked bank will then become shareholders of that bank.

While no bank has failed before due to interbank exposures, this may be because government intervention has prevented such failures [93, 66]. The fear of direct contagion may cause losses through other channels of contagion and it is therefore important to understand the direct mechanism for losses between banks before exploring indirect losses [66]. We address this in section 2.3 by investigating systemic risk with and without the inclusion of indirect mechanisms for contagion. An analysis of interbank contagion should not focus solely on interbank losses [93]. Contagion within the system would necessarily be underestimated without including all channels of contagion [38], and may lead to low probabilities of default within the system [57]. This is addressed by incorporating market liquidity and market sentiment effects in the model discussed in chapter 2.

When a bank experiences solvency or liquidity problems, it may be forced to sell assets in order to generate liquid funds. This is an important contagion channel to include, as market liquidity risk can have a significant effect on systemic risk in a network setting [35]. Section 2.3 further shows that indirect channels of contagion can have a significant effect on the systemic risk inherent in different network structures, and hence is a key

aspect to consider.

Regulatory requirements such as Basel III aim to influence the system to make it more robust. Considering this, it is important for regulators to be aware of the properties of the system that make it prone to contagion. These properties need to be monitored and controlled by regulators to protect the whole system. For example, regulators might encourage the system to form connections in such a way as to promote a certain overall structure. However, other characteristics of the system may influence which type of structure will be the safest. For example, differences in the size distribution of banks could lead to different structures being considered safe. The extent to which the system is affected by liquidity risks could also affect the relative stability of different structures.

Not all interbank systems are alike in terms of the characteristics of the market players within the system. Banks in different systems will have different distributions of assets, balance sheet compositions, overall levels of interconnectedness, susceptibility to liquidity risk etc. Strengths and available opportunities will differ between banks in a system and between banks of different systems. This will in turn affect the business models adapted by banks, thereby affecting their balance sheet compositions, funding structures and risk appetites [82, 7]. If these differences in interbank networks result in different optimal network structures, then the structure of banking systems need to be studied before investigating what actions regulators might take to improve a system's stability. For example, the large Euro interbank market will likely behave differently than the South African banking system where the number of banks participating in overnight interbank lending activities is generally fewer than 20 [26].

Therefore, it is vital for regulators to understand the influence of interbank structure and characteristics on the stability of the system. Admittedly the structure of banking systems and its role in contagion is but one of many aspects that regulators need to consider [33]. It is nevertheless an important consideration that this research is devoted to.

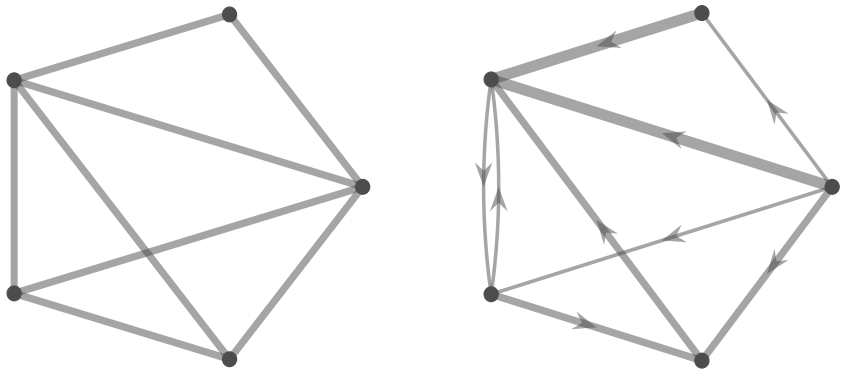
The complexity of financial systems results makes it difficult to precisely model how idiosyncratic shocks can lead to system-wide problems. For this reason we consider simplified banking systems, which is discussed further in section 2.1. Network models of systemic risk are by nature based on assumed parameters since the data for calibrating the parameters do not necessarily exist (e.g. data for calibrating parameters linked to liquidity risk is scarce at best). Therefore, we aim to gauge the relative effect of network characteristics on system stability instead of attempting to determine the precise probability or severity of systemic crises.

1.2 Literature

1.2.1 Background to network models

A network is a system of N interacting agents (called nodes or vertices), where the interactions between them form links between the nodes (called edges) [31]. It can be represented as graph, which is a pair $G = (V, E)$, where $V = 1, 2, \ldots, N$ denotes the collection of nodes and $E = \{\{i, j\}\}$ is the collection of edges. An edge is represented by a set $\{i, j\}$ of two nodes $i, j \in V$ that are connected via a link in the network. For directed graphs, the set $\{i, j\}$ is ordered and each edge starts at the first node i and ends at the second node j.

The banking system is represented as a network where the banks form the vertices. These are the entities being connected to one another. The lending activities between them form the connecting links (edges). Figure 1.1a provides an example of an unweighted, undirected network where the edges all carry equal weight and do not contain any directional information. This is not representative of lending activities, as the directions of lending and the loan amounts are important. For this reason, banking systems are represented as weighted, directed graphs. Figure 1.1b provides a graphical representation of such a system where the arrows indicate the direction of lending. The thickness of the arrows illustrates the varying weights of the exposures between banks.



(a) Unweighted, undirected graph (b) Weighted, directed graph

Figure 1.1: Graphical representation of an unweighted, undirected network and a weighted, directed network.

A commonly used method of modelling the edges between nodes is based on one of the earliest probability models of a graph, studied in [44]. Here, each edge in the network is present with a fixed probability p, referred to as the Erdős-Rényi probability. The resulting

graph is called an Erdős-Rényi network. It can be used for both directed and undirected graphs. We compare this model with extensions of it where the probability that a directed edge exists from node i to j, $p(i, j)$, is dependent on the nodes i and j. Note that for this application, $p(i, i) = 0$ for all $i \in V$, and it is said that the graph does not contain any loops (i.e. an edge cannot start and end at the same node).

The banks are represented by nodes and the mechanism through which a financial loss of one bank spills over to others is represented by the edges. The edges need to be directed to take account of the direction in which the losses propagate through the system. The interpretation of the edges as mechanisms through which uncertainty is channelled is discussed further in section 2.1.

The average probability that any node i is connected to another node $j \neq i$ is given by $\bar{p} = \frac{1}{n(n-1)} \sum_{i \neq j} p(i, j)$. It can be used as a measure of the interconnectedness of a network. This is because higher/lower values of $\bar{p}$ will lead to more/less edges in the network on average, which in turn will make the nodes in the network more/less closely connected to one another on average. It is therefore a simple way in which to assess how closely interconnected the nodes are on average. We also use it to standardise different extensions of the Erdős-Rényi network as described in section 2.2.1. The transmission of losses in the network as described in section 2.1 makes use of nodes' shortest paths. The shortest path from a node i to another node j, say d_{ij}, is the least number of edges that can be used to travel from node i to node j. It is used as a measure of distance between two nodes in a network and takes account of the edge directions.

The way that banks are connected to one another via edges in a network is referred to as the structure of the network. Real-world systems are often more complex than networks where edges are distributed uniformly between nodes such as the standard Erdős-Rényi network [69]. They often exhibit characteristics such as low average shortest paths and nodes that are clustered together (called the small-world property). The Watts-Strogatz model [96] can be used to generate such a network. Other commonly used network models include those where nodes have a power law distribution of the number of edges connected to them (referred to as a power law degree distribution). The Barabási-Albert model [4] can be used to generate networks with this property.

1.2.2 Relevant literature and motivation for the study

With the growing complexity present in financial systems, the task of safeguarding banking systems against collapse has itself become increasingly intricate. Against the backdrop of the 2008 financial crises, a significant increase in systemic risk research has emerged in

the hopes of understanding and protecting against it[1]. It is the goal of this research to contribute to the growing body of literature by expanding on existing simulation models and theoretical results. This is done by introducing a refined simulation model of interbank lending, formally defining a class of networks that can be of great value to systemic risk research, extending existing theoretical results for large networks and investigating the effect of banks' lending behaviour on systemic risk in both a small and a large network setting.

A wide range of methods have been proposed to measure systemic risk and many of these can be categorised according to whether they are based on market information (e.g. [63, 62]) or balance sheet information (e.g. [8, 75]). Studies based on balance sheet data have the advantage that the required data is publicly available, and hence it is not necessary for researchers to gain access to sensitive information to apply their models. Care must be taken in interpreting results obtained based on market information as these rely on certain efficiency assumptions of the market [22, 5]. For example, the use of market information may indirectly assume that market participants act rationally, that no information asymmetries exist, or that all participants have the same risk vs. return preferences.

Many studies that make use of balance sheet data are network driven. Each bank in the network has a simplified balance sheet structure, which may be assumed to change over time (see e.g. [19]) and which determines the bank's initial financial position before any defaults have occurred. An illustration of the balance sheet structure used by the majority of these studies is given in table 1.2.

Table 1.2: Illustration of a simplified bank balance sheet.

Assets	**Equity & Liabilities**
Interbank assets	Capital/Net worth
Other assets	Interbank liabilities
	Other liabilities

Network theory has been applied in a wide variety of disciplines including sociology, computer science, epidemiology, biology, economics and finance [86, 15, 80, 3, 87, 68]. Network models of systemic risk typically represent the banking system as a network of interconnected agents, where interactions between banks are modelled explicitly. An advantage of the network approach is that it emphasizes the relationships between the banks and provides a clear distinction between individual entities and the financial network that

[1] See [85] for a discussion around topics covered by existing systemic risk research, and how the literature has expanded over the last decade. Other notable surveys include [42, 89, 66, 36, 58].

they form [23]. It provides a natural means to model the contractual relationships between financial entities [6, 76].

Network models are flexible and can be adapted to incorporate different aspects of market behaviour e.g. liquidity risk, liquidity hoarding and asset correlation (see e.g. [64, 52, 70, 51]). For example, the network of exposures generated by credit default swaps of the United States of America is investigated in [73], the loans generated by the US Federal Reserve Bank emergency loans program is studied in [20], and in [53] international banking exposures are considered. Network models can be embedded in equilibrium models that take a game-theoretic approach to banks' behaviour (see e.g. [1, 12]). Network models can in general be embedded into larger macroeconomic models such that the effect of important financial variables are taken into account when determining the level of systemic risk. This has been done in [56] and [2] for example, where the former also includes central bank activity as part of the model.

The models proposed here specifically exclude central bank activity for three main reasons. Firstly, the main purpose of the study is to compare risk levels across different networks from a regulator's point of view. In this case it makes sense to compare the network structures independent from central bank attempts to stabilise the system. By excluding it, it is possible to look at the levels of risk that the central bank must manage. Furthermore from a practical perspective, regulators assessing policy responses to banking crises should compare the cost of intervention to the cost of not intervening, since both options can be costly [49]. A useful area of future research would be to consider a set of policy responses and investigate how the network structure affects the risks borne by various stakeholders such as depositors, taxpayers and other banks. One can then consider the risk borne by each stakeholder when the risk is ignored and when the regulator intervenes.

Secondly, it would be an arbitrary and futile exercise to include a central bank's response, since crisis decisions by a central bank are usually made on a case-by-case basis [5]. It would make sense to include a central bank where the purpose is to investigate the effect of possible policy decisions by the central bank [56, 40]. For our purposes however, it is important to separate the pure network effects from the consequences of unpredictable regulatory decisions.

Finally, it is not the aim of this study to produce an accurate estimate of the level of risk present in a financial system. The focus is rather on comparing different network structures and characteristics in order to investigate relative levels of risk. The inclusion of a central bank would make the results difficult to interpret since different conclusions may be reached depending on how the central bank is chosen to interact with the system.

It is for a similar reason that the macroeconomic environment is left out at this point.

For our purpose, we would not mix the effects of modelling the macroeconomic environment with the effects of varying network characteristics as this may influence the conclusions unduly. We do however acknowledge that embedding our network model into a model of the macroeconomic environment and including a central bank remains an important next step for future research.

It is to be expected that different network structures (not only different levels of interconnectedness, but different *types* of structure) should have different effects on systemic risk. This is because of the fact that interconnections between banks serve both as channels of contagion and mechanisms for spreading losses more evenly over a number of counterparties. Heterogeneity between banks' individual degree distributions should therefore affect how contagion is amplified or resisted by the network. Network studies that investigate different types of structure (e.g. [51, 54, 83]) typically investigate systems with homogeneous banks, or systems where the asset sizes are determined uniformly via simulation. We argue that it is important to investigate different types of structure together with a heterogeneous banking system since the size of a failing bank can significantly influence the spread of contagion [70, 29].

Previous studies on network topology [70, 71, 83] have focused on comparisons between scale-free networks (which have power law distributions of degrees) and random networks, as the former type has been found to be representative of real-life systems [25, 84, 50, 39]). However, not all networks follow a power law distribution of degree [32] and for smaller networks such as the South African system, a power law distribution is theoretically unsuitable and difficult to test for. This is because a random variable X is said to follow a power law distribution if it satisfies $P(X > x) \approx cx^{-\alpha}$ as $x \to \infty$ (here, $0 < c < /infty$ is a constant and $\alpha > 0$ is known as the tail index) [72]. With too few data points it is not appropriate to test for the asymptotic behaviour of $P(X > x)$.

This thesis addresses this shortcoming by investigating network structures that are applicable to smaller networks. We restrict ourselves to models where the probabilities $p(i, j)$ are functions of properties of the nodes i and j. This is done firstly to enable us to test a wide variety of structures that are comparable to one another, since they are simulated in a similar way. Secondly, since the $p(i, j)$ probabilities are dependent on the nodes i and j, it enables us to explicitly take account of differences between banks.

Even large systems may more closely represent a tiered network as opposed to a scale-free network. For example, from [41] it seen that the German interbank system represents fits tiered network model better compared to a scale-free network. Other empirical studies such as [67, 48] have also found evidence of core-peripheral structures in interbank networks (see [66, 58]). Core-peripheral structures have a small number of tightly connected

nodes which form the core, with a large number of sparsely connected nodes which form the periphery. Structures representing tiered or core-periphery systems have been used in the past [78, 10, 13, 92, 59]. However, these are constructed using different approaches and are not formally tested for whether or not they can be regarded as tiered.

While previous studies on network topology have focussed on comparing one type of realistic structure with a random network, we address the need for investigating a range of structures (including more than one type of tiered network). We investigate structures that can be easily applied to smaller networks (i.e. networks that are tiered, but not necessarily scale-free), together with liquidity risk and market sentiment.

For the purpose of this study the 'structure' of the network not only refers to the degree distribution of nodes, but the way in which the degree distribution is influenced by the relative asset values of banks. Previous studies considering the structural effect of networks on systemic risk interpret the structure/topology as the level of interconnectedness in the system, or the degree distribution as seen independently of the relative asset values between individual banks (see for example [51, 54, 83]). While the way in which financial institutions are connected to one another plays an important role in the propagation of shocks [89], we argue that it must not be investigated in isolation, but in conjunction with lending preferences[2] and other network characteristics such as capital levels, average interconnectedness etc. In this work, the different structures explicitly take account of the relative asset sizes of banks when determining banks' lending behaviour.

Empirical evidence of the role of lending preferences in network structure differ. For example, [74] find that asset size is not always clearly associated with lending preferences. On the other hand, [41] find that the German core-peripheral network's highly connected core consists of money centre banks which act as intermediaries between other banks and are identified by their size, specialization and balance sheet ratios. Similarly, [30] find evidence that factors such as bank size, sector and type are indicative of a bank's in- and out-degrees.

It should be noted that many of the references contained in this thesis are based on working papers, which is necessitated by the chosen research problem. The working papers used for this thesis are published by reliable sources such as the BIS or reserve banks such as the Bank of England.

[2]The work in [71] takes account of lending preferences between banks, but does this in a dynamic setting where lending preferences change according to expected profitability.

1.3 Purpose and significance of the research

This study makes the distinction between small and large networks for the following reasons:

- Network simulation models of systemic risk offer much flexibility, allowing one to easily include concepts such as liquidity risk and market sentiment. It is not practical to use simulation models for large networks because of the computational demands. By considering a smaller network, it is possible to incorporate a higher level of realism into a model.

- It is noted that not all financial networks are sufficiently small to allow for simulation modelling techniques. Some systems contain thousands of participants [94], which may make a simulation approach too resource intensive. In such cases numerical methods would be preferred. For example, simulation methods may be unfeasible for modelling international systems. Therefore it remains important to consider asymptotic results, even if theoretical complexities limit the flexibility of such models.

The purpose of this research is five-fold. Firstly, we aim to investigate how different network structures can influence the relative systemic risk. This is done by considering both a simulation and an analytical model of systemic risk which account for differences in bank asset sizes. The simulation model introduced here accounts for differences in the maturity of assets and business models of banks, liquidity risk and investor sentiment. While liquidity risk and investor sentiment are difficult mechanisms to incorporate in a realistic manner, we argue that they are too important to ignore. Their importance is illustrated in section 2.3, where it is shown that the effect of network structure is highly dependent on the inclusion of these factors and the parameter values associated with them.

As mentioned in section 1.1, certain banking systems exhibit power law degree distributions. However, not all networks follow a power law distribution [32] and for smaller networks (e.g. the South African, Mexican or Belgian banking systems) this is difficult to test for. We therefore aim to include a broader range of possible network structures when analysing network sensitivities. This analysis is used to answer the questions below, which are discussed further in section 4.1.1.

1. How does the structure of an interbank network influence the risk inherent in a system?

2. Are some structures inherently less risky than others?

3. How do network characteristics (i.e. interconnectedness, asset distribution and size of the system) influence the answer to question 1?

4. How do indirect contagion mechanisms influence the answers to the above questions?

5. Can network properties (e.g. shortest path, clustering, level of tiering) assist in explaining some of the variation in risk levels observed from different structures?

6. What policy suggestions can be made regarding network characteristics?

The second purpose of the research is to assess the usefulness of network models of systemic risk. Despite the growth in empirical analyses of banking networks[3], there is a gap in the literature regarding assessments of whether such models can capture increases in systemic risk during stressed market conditions. To the best of the author's knowledge, this has not been investigated before in a South African context. This research addresses this need by considering systemic risk at different points in time, during which incidents of market stress were experienced.

The third purpose of this work is to contribute to the literature on analytical methods of modelling systemic risk. This is done by building upon the theoretical results from [8] to show that it can be applicable to sequences of networks with random degree distributions and not only to deterministic degree distributions. This is of practical interest as the interbank connections between banks change continuously over time. It bears similarities to [9], who consider similar results for groups of graphs. That study also extends the results from [8] to a more general setting. However, the randomness of that model originates from the probability of a contagious link existing between any two nodes, whereas the randomness considered in this study originates from the probability of one bank lending to another which leads to random degree sequences. This is motivated by empirical evidence from [46] who show that interbank relationships are formed randomly on a daily basis, based on the true underlying structure [91].

Finally, this study formalises the definition of a class of networks that can be used to analyse a wide range of different network structures. The class of stochastic networks defined in this study is based on the concept of multiple interacting networks [28]. The difference between this study and other studies that make use of such networks (for example [17]) is that we do not consider different *types* of network connections (such as different

[3]For example [55] investigated the network properties of the South African interbank system, [65] focused on Asian and Pacific banks, [95] on listed European banks, [94] on the German interbank market, [74] on the Mexican system and [25] on the Austrian interbank market to name but a few.

types of loan), but rather consider banks being grouped together according to some criteria, such as their asset size. We therefore assume the existence of multiple groups of banks (that can each be seen as a network on its own) that interact with one another via interbank links. Banks within any one group are assumed to be similar in size and exhibit similar lending behaviour towards one another and towards banks belonging to other groups.

This study provides two methodologies for addressing the problem of determining how the structure of a network affects relative systemic risk levels. The methods presented and developed here can be useful to regulators and for future research for the following reasons:

- We test a range of probability models to simulate agreements between banks, none of which are overly complex. We show that three of the resulting network structures can be seen as tiered structures when they are sparse, making them appropriate for simulation studies of systemic risk.

- Sensitivity tests show that network structure plays an important role in determining the level of systemic risk.

- Simulation results show that a combination of different network characteristics influence which structures a regulator could encourage to reduce risk.

- The proposed model can be embedded in a dynamic model of systemic risk with a macro-economic environment for future research.

- We show that the network structure should influence the focus of regulation as it can assist in informing which network characteristics influence systemic risk the most.

- The framework presented here can be useful to answer 'what-if' questions that arise in practice and to give insight into what might happen to the system given an appropriate network. The framework in itself can be used to generate a wide range of output, for example one can investigate a range of different risk measures (average capital lost, average proportion of asset value lost by the system etc.), and consider correlations between this and the size of the initially defaulted bank.

- A useful class of networks are formally defined and we illustrate how this can be used to generalize results for Erdős-Rényi (i.e. homogeneous) networks to inhomogeneous networks. This creates the opportunity for existing research on Erdős-Rényi graphs to be applied to a much richer collection of networks.

- We illustrate how this class of networks can be used to model and compare different types of core-peripheral financial networks. These networks can explicitly take account of the lending preferences of banks based on asset sizes or any other network characteristic.

1.4 Structure of the thesis

The study considers small and large systems separately from one another, since different modelling techniques are appropriate depending on the size of the system. The focus of chapter 2 is on a simulation methodology that is appropriate for smaller systems. Section 2.1 explains and motivates the modelling procedure. Section 2.2 presents and discusses a range of structures that are investigated as part of the study. The results of the tests performed on this model are presented in section 2.3. The simulation model is applied to the South African banking system in section 2.4. The work presented in sections 2.1 to 2.3 has been submitted for publication in a shortened form. The work presented in section 2.4 has been accepted for publication in a shortened, alternative form.

Chapter 3 contains the asymptotic results for large networks. Section 3.1 presents the relevant background and notation for this chapter. Section 3.2 discusses the existing results for deterministic networks in depth. These results are generalised in section 3.3 to random networks. Illustrations of the generalised results are included in section 3.4. The work presented in chapter 3 has been submitted for publication in a shortened form, after which feedback was received with the opportunity to resubmit. It has been resubmitted with the revisions incorporated into this thesis.

Chapter 4 serves to conclude the thesis. The implications of the results from chapters 2 and 3 are summarised in section 4.1. The final conclusions of the research are discussed in section 4.2. The shortcomings of the research and avenues for future work are also included in section 4.2.

Chapter 2

Numerical model

This chapter is devoted to modelling systemic risk via a simulation approach, which is appropriate for small networks. The simulation model presented in this chapter is based on the work of [60] and [78]. It includes a novel contagion mechanism that represents the spread of contagion due to a meltdown of trust in the system.

The chapter begins with section 2.1 which describes the mechanics of the simulation model. This includes a description of how banks interact with one another and describes the different contagion mechanisms that spread losses throughout the system. Six different structures are investigated in this chapter. These are defined in section 2.2, where the characteristics (average shortest path, clustering, and level of tiering) of each structure are investigated. Section 2.3 investigates the sensitivities of the simulation model output to changes in parameter values and compares the different structures in terms of this. Section 2.4 presents a practical application of the simulation model which illustrates the use of the general methodology in practice. Here, the modelling procedure is adapted to reflect the fact that losses due to direct counterparty exposures is inappropriate in a South African context.

2.1 Simulation method

The description of the simulation method is broadly divided into two parts. The first is section 2.1.1 which gives a description of the banks' balance sheets and the relevant notation used for the remainder of this chapter. The second part (section 2.1.2) describes how losses propagate through the system upon the default of an initial bank. It describes the three mechanisms used to spread and amplify losses in the system. Finally, the section defines and motivates the systemic risk measure used for this chapter.

2.1.1 Network description

The basics of the model used for investigating the propagation of shocks are similar to those from [60] and [78]. However, the model proposed here can be regarded as a refinement as it incorporates liquidity effects (comparable to [75], but with slightly different dynamics) as well as a proxy for investor/depositor psychology.

Problems can spread very quickly throughout the system [52]. Therefore a bank's failure is assumed to be dependent on its level of Common Equity Tier 1 (CET1) capital (this is similar to the approach from [97], [77] and [39]) since CET1 capital can quickly be converted into cash. Additional Tier 1 capital is excluded since these must first be converted to cash in the event of a crises.

Even though not all countries will have the exact same requirements, it is not uncommon for regulators to step in before insolvency or liquidation occurs. For example, the U.S. Federal Deposit Insurance Corporation (FDIC) has thresholds specified in terms of CET1 capital that triggers corrective action against banks [45]. Regulators will likely also have more than one criteria for corrective action. As these criteria are jurisdiction specific and may be subjective, it is not possible to adopt a methodology that will reflect the conditions for corrective action of all banking systems. That is why we align our methodology to the rules published by the Basel Committee [18]. The BIS requires a minimum of 4.5% CET1 capital to risk-weighted assets, and the FDIC regards a bank as significantly undercapitalised if the ratio is less than 3%. For simplicity, we will assume that a bank is removed from the network when its ratio of CET1 capital to total assets falls below 3% (the model can easily be adapted to reflect different default definitions). This is because corrective action will likely be taken against a bank well before all of its CET1 capital is depleted. As this assumed 3% is 33% lower than the prescribed minimum of 4.5% required by the BIS, it is reasonable to assume that such levels of capital will trigger corrective action to be taken against a bank. For ease of reference we refer to such a bank as being insolvent, even though we acknowledge that in practice there are specific rules for determining whether a bank should be declared insolvent or be placed under liquidation.

The simplified assumption is made that a bank's providers of capital are shareholders, banks, other institutional investors and depositors (in increasing order of seniority). In practice, the seniority hierarchy would likely differ and be more refined. For example certain depositors could lose out before a bank starts to default on its interbank payments as mentioned in section 1.1. However, for the general purpose of this model, this hierarchy will suffice. If the model is to be applied to a specific country, it can easily be adjusted to reflect that country's regulation.

Table 2.1: Illustration of a simplified balance sheet for a bank i.

	Assets		**Equity & Liabilities**
$\sum_j l_{i,j}$	Interbank assets	$a_i\gamma_i$	CET1 Capital
$a_i^{(s)}$	Short-term assets	$\sum_j l_{j,i}$	Interbank liabilities
$a_i^{(m)}$	Medium-term assets	N/A	Other liabilities
$a_i^{(l)}$	Long-term assets		

Therefore, bank counterparties start losing out on their capital after the shareholders' capital (which very roughly makes up the CET1 capital in addition to disclosed reserves) is depleted. This is the basic mechanism with which losses spread through the system — any loss amount that could not be absorbed by a bank's CET1 capital will lead to losses for that bank's lenders. These lenders will then also experience losses in accordance with the amount of their loans that could not be repaid. Apart from this, there are two additional mechanisms by which failures lead to further losses in the system. The different ways in which losses spread through the system are described in detail in section 2.1.2 below.

For each bank a simplified balance sheet is constructed that can be seen as a refinement of the structure illustrated in table 1.2. Here the non-interbank assets are subdivided according to their term. Three categories are chosen for the term namely short, medium and long. Short-term assets are defined as those with a maturity less than one month. The medium-term assets are those with a maturity of less than one year (but more than one month), and the long-term assets include those with maturities longer than one year.

The short-, medium- and long-term assets of a bank i are denoted by $a_i^{(s)}$, $a_i^{(m)}$ and $a_i^{(l)}$ respectively and the total assets of i is a_i. The value of the interbank assets held by institution i as a result of loans granted to j is denoted by $l_{i,j}$. In other words, this can be regarded as i's exposure to j. In a network representation of the system where the banks represent the vertices, the connections formed through interbank lending and borrowing form the edges. If $l_{i,j} > 0$ then it is said that there exists a directed edge with weight $l_{i,j}$ from i to j. The ratio of CET1 capital to total assets held by institution i is denoted by γ_i and will be referred to as a bank's capital ratio for our purposes. The liability and equity side of the balance sheet will consist of CET1 capital, interbank liabilities and other liabilities. The simplified balance sheet structure and the relevant notational conventions adopted by this study is illustrated by table 2.1.

Let N denote the total number of banks in the network. For reasons mentioned in section 1.2, the model considers systemic risk from the perspective of a lender of last resort and hence assumes the absence of bailouts. Even though banks have counterparty expo-

sures outside of the banking system, we consider a closed system as we are interested in the risk arising from interactions between the banks only. In other words, we are interested in the endogenous risk due to a combination of network structure, network characteristics, liquidity risk and investor sentiment. The size of each vertex is determined by its total asset value — banks with higher asset values will represent bigger vertices. The network is therefore represented by a weighted directed graph with N vertices of differing sizes.

2.1.2 Simulation of shock propagation

Initial shock and direct losses

Once the edges connecting the banks are determined (section 2.2 discusses different ways for determining this), each edge can be assigned a corresponding exposure, $l_{i,j}$. This is the amount that bank i has lent to bank j. It is assumed that $l_{i,j} = 0$ whenever there is no edge from i to j. The values of these exposures can be determined via simulation, or they can be based on actual balance sheet data. Section 2.3 explains how the exposures are determined for the purpose of our simulations.

Unless specified otherwise, assume for the remainder of this chapter that for all i, $\gamma_i = \gamma > 0.03$ to ensure that all banks start off solvent and well capitalised. The system is shocked by reducing the capital of a single bank n, presumably because of suffering a loss on its non-interbank assets. For the purpose of this study, such an event is called an 'initial shock' since we assume that it was a significant and unexpected event. This implies that bank n needs to use its CET1 capital, $a_n\gamma_n$, to try and absorb the loss caused by the shock. To standardise the initial shock across banks of different size, it is specified as a fraction, say s of its total assets. The resulting loss amount is subtracted from the capital held by the bank. If the capital of n falls below 3% of its total assets, the bank is deemed significantly undercapitalised and is removed from the network (for ease of reference, we refer to the bank as having failed). If the loss is greater than the capital, the shortfall must be absorbed by the funds held for the interbank liabilities. If this is not enough, the rest of the loss is absorbed by the funds held for the other liabilities. It is assumed that losses absorbed by the non-interbank liabilities do not propagate through the system.

In other words, when bank n is shocked its external assets are reduced by an amount of $S_n = s \cdot \left(a_n^{(s)} + a_n^{(m)} + a_n^{(l)}\right)$. There are three possibilities regarding the spread of this loss through the system. If $S_n < a_n\gamma_n - 0.03a_n$, the loss is absorbed by the capital and no further losses occur. Secondly if $a_n\gamma_n - 0.03a_n < S_n < a_n\gamma_n$, then n is deemed undercapitalised, but there are no direct counterparty losses since the loss is less than n's capital. There will only be liquidity and proximity shocks to the system, which are explained below in this

section.

Finally, if $S_n > a_n\gamma_n$ then n fails and its interbank counterparties suffer a total credit loss of $\min\left\{S_n - a_n\gamma_n, \sum_k l_{n,j}\right\}$ (liquidity and proximity shocks apply as well, and are explained below). In this case if $S_n - a_n\gamma_n < \sum_k l_{n,j}$, then the amount of the loss above the capital is less than the total interbank loans made by bank n. This shortfall then needs to be divided between the counterparties. It is done proportionately according to the amount owed to each counterparty. For example, the amount lost by each counterparty i is equal to

$$L_{i,n}^{(1)} = (S_n - a_n\gamma_n)\,\frac{l_{n,i}}{\sum_k l_{n,k}}. \tag{2.1}$$

Hence, direct losses suffered by a counterparty cannot exceed the amount of the original loan. Of course if $S_n - a_n\gamma_n \geq \sum_k l_{n,k}$, then each counterparty loses the full amount of their loan.

This concludes the direct spread of interbank losses. In practice, losses caused due to indirect contagion effects are highly relevant and should be included in such an interbank model. The spread of these indirect losses through the system is discussed next.

Liquidity effects on the system

Further losses spread through the system in the form of liquidity shocks, which are comparable to the liquidity shocks introduced in [75]. Every time a bank defaults, the non-interbank assets of all other banks receive a shock that reduces the value of their holding in those assets by a specified factor. This is a simplified way of accounting for raised provisions and market liquidity risk. The non-interbank assets are categorised broadly according to term. This is based on the presumption that assets of differing maturity will be affected differently by problems in the system. An important aspect for future work would be a refinement of this classification and to model the reaction of the markets in these assets more thoroughly. Let $g^{(s)}$, $g^{(m)}$ and $g^{(l)}$ be parameters associated with the reduction of value for the short-, medium- and long-term assets respectively. The short-term non-interbank assets of each bank i is reduced from $a_i^{(s)}$ to $a_i^{(s)} \cdot \exp\left(-g^{(s)}\right)$, where the reduced values for the medium- and long-term non-interbank assets are calculated similarly. The loss to each bank i is therefore equal to

$$L_i^{(2)} = \sum_{\eta\in\{s,m,l\}} a_i^{(\eta)}\left[1 - \exp\left(-g^{(\eta)}\right)\right], \tag{2.2}$$

This is a commonly used method in the systemic risk literature for modelling changes in asset prices due to changes in supply and demand [37, 75, 52, 78]. It makes sense in this setting because the factor $\exp\left(-g^{(s)}\right)$ will always be in the range $(0, 1)$, but will never be equal to 0 or 1 (i.e. there will always be *some* loss, but not all of the assets will be lost). This implicitly assumes that all banks in the system hold similar classes of assets, which is generally not the case. Therefore the less stringent assumption is made that banks following the same business model hold similar classes of assets. This is appropriate since the business model of a bank is driven by its business objectives, which in turn affects its funding, asset mix and impairment provisions.

For the purpose of this study, we use the three bank business models empirically identified in [82]. Banks are deemed either retail-funded, wholesale-funded or capital markets-oriented based on balance sheet ratios identified by the authors. The asset mix held by retail and wholesale-funded banks were found to be noticeably similar. Therefore suppose first that the defaulted bank n and another bank i are either both retail-funded or wholesale-funded, or that one bank is retail-funded and the other wholesale-funded. In this case the expression in (2.2) for the reduction in the value of i's non-interbank assets remains the same.

On the other hand, suppose that the defaulted bank n is capital markets-oriented and another bank i is retail-funded or wholesale-funded (or vice versa). Then bank i's short-term non-interbank assets are reduced from $a_i^{(s)}$ to $a_i^{(s)} \cdot \exp\left(-\frac{g^{(s)}}{\nu}\right)$, where $\nu \in (1, \infty)$ and with $g^{(s)}$ replaced by $g^{(m)}$ and $g^{(l)}$ for the medium- and long-term assets respectively. In this case, bank i suffers a loss of

$$L_i^{(2)} = \sum_{\eta \in \{s,m,l\}} a_i^{(\eta)} \left[1 - \exp\left(-\frac{g^{(\eta)}}{\nu}\right)\right]. \tag{2.3}$$

The factor ν in equation (2.3) is used to ensure that banks with asset compositions different to the failing bank receive smaller liquidity shocks than implied by equation (2.2). By adjusting the expression in (2.2) in this way, we achieve two important goals. Firstly, the model is based on less restrictive assumptions as we no longer assume that all banks hold similar classes of assets. Secondly, this is achieved without introducing many new parameters or overcomplicating the model unnecessarily. The parameter ν serves as a measure of the difference between the asset composition of two banks. For the purposes of our illustrations we assume that $\nu = 2$. The implication of this choice for ν depends on the chosen parameter values of $g^{(\eta)}$, $\eta = \{s, m, l\}$ — The greater the values of $g^{(\eta)}$, the greater the difference will be between $\exp\left(-g^{(s)}\right)$ and $\exp\left(-\frac{g^{(\eta)}}{\nu}\right)$."

This type of loss captures reductions in the market values of assets resulting from the fire-sales or inopportune sales of bank n's assets. It is referred to as liquidity losses or liquidity shocks for the remainder of this thesis. The associated parameters are referred to as liquidity reduction parameters. Higher values of these parameters would result in greater reductions in the value of external assets. Each of these parameters represents the expected effect that the insolvency of a bank would have on the non-interbank assets in the system.

A bank should normally have more short-term assets that are available for sale compared to medium- and long-term assets. This should increase the value chosen for $g^{(s)}$ compared to $g^{(m)}$ and $g^{(l)}$. Next, as long-dated assets are inherently less liquid than shorter dated assets, the probability of those assets being affected should be higher than for shorter dated assets. Finally, since longer dated assets have higher discounted mean terms, their anticipated devaluation should be higher than for shorter dated assets.

Each of the parameters $g^{(s)}$, $g^{(m)}$ and $g^{(l)}$ should incorporate a combination of these three factors. Since it is not possible to determine their net effect without conducting an in-depth analysis of the system in question, it is necessary to make a simplified assumption. One of the three effects push up the parameter associated with short-dated assets, and the remaining two increase the parameter associated with long-dated assets. Therefore we assume that the liquidity risk parameters are increasing with term, i.e. $g^{(s)} < g^{(m)} < g^{(l)}$. This is reflected in the benchmark system discussed in section 2.3. While the objective of this study is not to model the values for the reduction parameters, it remains an important consideration for future work.

Loss of confidence through proximity shocks

A third type of shock is aimed at including factors associated with investor confidence. Any bank that is viewed as being 'in close proximity' to the failing bank is penalised as the market might start to doubt its financial soundness. The reaction of the market will typically depend on the circumstances surrounding the default and therefore it is preferred to take a generalised modelling approach to capture the effects of market sentiment. To put this into context, consider the following chain of events: An initial bank fails due to an event that adversely affects its assets. We do not specify what type of event this is — it could be due to unsustainable lending practices, large proportions of its customer base suffering from financial strain or any other event that causes impairments to increase significantly and unexpectedly. For example, following the curatorship of African Bank, some of the larger banks received credit downgrades from Moody's which increased their cost of

borrowing. The reason given for the downgrade was that while the Reserve Bank did mitigate contagion risk by issuing a bailout, some creditors were allowed to suffer losses and hence Moody's was of the view that there is a "lower likelihood of systemic support from South African authorities to fully protect creditors in the event of need"[1]. Other examples include the banking crisis in Greece which saw a run on the banks that negatively affected banks' liquidity positions, and the European sovereign debt crisis which resulted in credit downgrades and increased costs of borrowing. Instead of restricting ourselves to particular scenarios, we consider a wide range of possibilities regarding the spread of distrust in the system. This is done by simulating network paths according to the structures presented in section 2.2, which avoids the need to consider the circumstances surrounding the initial default.

The model proposed by this thesis investigates what could happen if such an event (shock) causes this bank to become insolvent or severely undercapitalised. This could result in fire-sales when the shocked bank has to forcibly sell assets, depressing the value of similar assets held by other banks (this is captured by the liquidity parameters in section 2.1.2 above). Furthermore, the market will react to the news of the initial shock event even if the root cause of the shock is not immediately clear. In fact, it may take several months for audit investigations to be completed before the details of the failure are known to the public. During this time, the market might be uncertain as to whether the initial shock event is self-contained, or whether it affects other banks as well. Depending on the circumstances, the market may suspect other banks of suffering from similar shortcomings as the first bank (perhaps due to similar business models or target markets). This may cause a loss of confidence in those banks. This in turn could lead to problems for those banks in rolling forward their short-term liabilities, causing further losses due to maturity mismatches. In this context, losses due to such a lack of confidence will be called a proximity shock. This avoids confusion with losses due to the fire sales of a defaulted bank.

The factors that determine how close one bank is to another are debatable. In practice it would be related to the initial shock event, the modelling of which is not the purpose of this thesis. The measure of closeness adopted should depend on information that is available to investors, since proximity shocks are based upon the market's view of the remaining banks.

Suppose $d_{i,n} > 0$ denotes the distance from any bank i to a failing bank n, which can be measured in any way that makes sense for the situation to which the model is applied. In order to reflect the shrinkage of a bank i's balance sheet due to increased

[1] `https://www.moodys.com/research/Moodys-downgrades-four-South-African-banks-on-review-for-further--PR_306571`, accessed 2018/09/28

funding costs and any resulting forced sales, the short-term non-interbank assets of bank i become $a_i^{(s)} \cdot \exp\left(-\frac{\delta}{d_{i,n}}\right)$ where the formulae for the medium- and long-term assets are constructed in the same way. The loss suffered by bank i is given by

$$L_{i,n}^{(3)} = \sum_{\eta \in \{s,m,l\}} a_i^{(\eta)} \left[1 - \exp\left(-\frac{\delta}{d_{i,n}}\right)\right].$$

The parameter δ is a reduction factor associated with proximity shocks. It represents the degree to which banks may suffer losses due to the difficulty of raising funding. Similar to the liquidity reduction parameters, different proximity factors should be assigned to different types of assets when applying this model. However, we avoid introducing too many parameters for the purpose of illustration by using the same parameter for all non-interbank assets.

This approach implicitly assumes that distrust in the system is only initiated after a default, and that banks of equal distance to the defaulting bank will experience losses of a similar degree. The appropriateness of the first assumption depends on the circumstances surrounding the default, and whether the market was aware of any friction within the system beforehand. The second assumption is unlikely to hold in practice but is required to keep the model simple and tractable. Finally, the model makes the underlying assumption that the loss of market sentiment due to a bank's failure is independent of the failed bank's size. In practice, it is expected that the failure of bigger banks will affect market sentiment more adversely than that of smaller banks. However, it is not straightforward to determine the magnitude of such differences, and the resulting effects may obscure the network implications that we aim to investigate in this thesis.

The 'closeness' between banks can be measured in a variety of different ways. The below examples serve to demonstrate this.

Example 1: Shortest path Distance from a solvent bank i to a failing bank n can be measured by means of the shortest path from i to n. Note that direction plays an important role here, as such paths should follow the direction of the connecting edges for this to make sense. This is based on the presumption that the market is concerned about the interbank counterparty losses that i experienced. In this case, we have that $d_{i,n} \in \mathbb{N}$, with $d_{i,n} = 1$ if i has an outgoing edge directly connected to n. In this case i suffers a loss of $a_i\,(1 - \exp(-\delta))$, which is the largest proximity loss that it can experience. However, this is not a realistic option, since in practice the market will not have knowledge of banks' bilateral agreements. It is therefore important to look at other measures of closeness as

well.

Example 2: Asset size Relative asset size can be used as a measure of closeness if the market believes that banks of similar size to a failing bank may suffer from similar difficulties. If we let $d_{i,n} = \frac{\max\{a_i,a_n\}}{\min\{a_i,a_n\}}$, then $d_{i,n} \in [1,\infty)$. A bank i is closest to the failing bank n if $a_i = a_n$, in which case i will suffer a loss of $a_i\,(1-\exp(-\delta))$. The greater the difference between a_i and a_n, the smaller the loss $a_i\left(1-\exp\left(-\frac{\delta}{d_{i,n}}\right)\right)$ becomes.

Example 3: Business model Another possibility is to link the distance between banks to differences in the business models that they adopt. Banks that make use of similar business models can be regarded as close, while banks with different business models are regarded as further away from one another. The intuition behind this approach is the notion that banks with similar business models invest in similar asset classes or industries. If one bank fails as a result of a reduction in the value of its external assets, the market may be sceptical about the financial state of other banks that hold similar external assets. Suppose banks are categorised according to the same business models as for the liquidity shocks, namely retail-funded, wholesale-funded or capital markets-oriented. Here, one possibility for $d_{i,n}$ is as follows:

$$d_{i,n} = \begin{cases} 1 & \text{if } i \text{ and } n \text{ employ the same business model} \\ 2 & \text{if } i \text{ is retail-funded and } n \text{ wholesale-funded, or vice versa} \\ 3 & \text{if either } i \text{ or } n \text{ is capital markets-oriented and the other is} \\ & \text{retail-funded or wholesale-funded.} \end{cases} \tag{2.4}$$

For the sample of banks from [82], all three business models differ significantly from one another with respect to their funding profiles. On the other hand, the retail-funded and wholesale-funded business models' asset compositions were noticeably similar. This is why the retail-funded and wholesale-funded categories are considered closer to one another than the capital markets-oriented and retail-funded (or alternatively, the capital markets-oriented and wholesale-funded) categories. This partly informed the choices for the above parameters: If banks i and n employ the same business model, the resulting proximity shock is the greatest. If i is retail-funded and n wholesale-funded (or vice versa), then the resulting proximity shock is reduced slightly. If either i or n is capital markets-oriented and the other is retail-funded or wholesale-funded, then the resulting proximity shock is reduced further. Other than to capture this hierarchy, the choices for the above parameters are arbitrary and intended for illustration purposes, since no empirical evidence is available

with which to make more informed choices.

Example 4: Network approach The final example of how proximity shocks can be incorporated makes use of a networks-on-networks approach. In addition to the network of interbank exposures, we construct a second network where the connections between banks have a different meaning. Here, the market considers two banks as close if they have similar characteristics such as a target industry. For example, a directed edge from i to n could exist if i invests in an industry that is similar to what n invests in. A shortest path approach to bank n can be used here. This will be similar to example 1, except that the edges we consider are based on a second network. An advantage of this approach over example 1 is that it avoids the assumption that the market is aware of banks' interbank counterparties. However, it is difficult to decide on how to construct an appropriate network structure to use. This approach is adopted in section 2.4, where this simulation framework is applied to the South African banking system.

After the default of the initial bank n, the total loss to each remaining bank i is given by $L_{i,n} = L_{i,n}^{(1)} + L_i^{(2)} + L_{i,n}^{(3)}$. This concludes the description of the three ways in which losses spread through the system, i.e. through direct credit losses, liquidity effects and proximity shocks. The section below provides a description of how these losses are combined to produce a default cascade.

Default cascades

Suppose that the initial bank n is shocked and that it fails because of this shock. If the initial loss exceeds the capital held by n, its counterparty creditors will suffer direct losses because of this. Since n failed, there will be a liquidity shock to all banks regardless of whether they were a counterparty to n or not. The extent of this shock will depend on the different business models of the banks, to account for the fact that not all of them hold the same assets. There will also be proximity shocks to the other banks based on how close they are to n.

For each remaining bank, the potential losses due to all three channels of contagion are added together. The total loss for each bank is then subtracted from their respective reserves. Those banks that received losses of more than 3% of their assets fail as a result. This concludes the first round of default. Note that all remaining banks in the system (excluding those that have already defaulted as part of the first round of default) will have their assets adjusted at this point. For example, suppose that a bank k did not default, but

that its assets were reduced as a result of the failure of n. Then its assets are recalculated accordingly, such that its interbank and non-interbank assets all reflect the losses that had to be absorbed. This will affect any future calculations of its threshold $0.03a_k$ used to determine whether it is undercapitalised.

Now there may be more than one bank that had defaulted. The direct counterparty losses resulting from all the defaulted banks are determined. For every bank that had failed, liquidity and proximity losses are assigned to all remaining banks based on the procedures detailed above. Note that banks defaulting at the same time do not affect each others' losses. For example, suppose that banks i and j both default as a result of bank n's failure. Let $S_{n,i}$ be the total loss to i as a result of the failure of n. Then the fact that j has also defaulted because of the failure of bank n will not increase $S_{n,i}$ any further. Similarly, the default of i will not affect $S_{n,j}$, the loss suffered by j. This means that any direct counterparty losses, liquidity shocks and proximity shocks of i and j will not affect each other.

For each remaining bank in the system, the losses experienced by it are added together to determine the next round of default. The above procedure is then repeated for all remaining banks, and the default cascade stops whenever no more banks are failing. The cascade can therefore either stop when all losses in the network have been absorbed, or when no more banks are left in the system. Over all banks in the system, the total loss in capital over all nodes in the system is calculated as a proportion of the system's total capital. In other words if c_i is the capital of bank i before the start of the cascade, and $c_i'(n)$ is i's capital at the end of the cascade which was initiated by the default of bank n, then we calculate the proportion

$$\theta_n = \frac{\sum_{i=1}^{N} \left(c_i - c_i'(n)\right)}{\sum_{i=1}^{N} c_i}. \tag{2.5}$$

Suppose there is a total of N banks in the system. A simplified measure of systemic risk is calculated as follows: The above procedure is repeated N times, and for each repetition a different bank is chosen as the initially defaulted bank. After each repetition of the cascade, the proportion given by equation (2.5) is calculated. Thereafter, the balance sheets are reset to their original states. The average of this ratio over the N repetitions of the cascade is then calculated, i.e. we calculate

$$\bar{\theta} = \frac{1}{N} \sum_{n=1}^{N} \theta_n. \tag{2.6}$$

This whole procedure is repeated m times, where the links between the banks are re-

simulated for each repetition. Due to the various sources of randomness, it is not possible to theoretically determine a suitable value for m. Therefore, the value of m was chosen to be 10 000, which yielded stable results without unduly increasing the processing time required for the simulations. The average of $\bar{\theta}$ over all the simulations[2] is then used as a straightforward measure of systemic risk. This measure reminds of the methodology used to construct the Contagion Index [39] and is called the average *capital reduction ratio (CRR)* for our purpose. The model can easily incorporate more refined measures of systemic risk, which is left for future research. However, this measure is sufficiently intuitive for our purposes since it satisfies the following:

1. It takes account of the extent of losses, i.e. the larger the losses that are suffered by banks, the higher the level of risk that will be measured.

2. The default of a big bank will increase the risk measure by more than the default of a small bank if both banks lose the same percentage of their assets. For example, suppose that both banks initially had a CET1 capital ratio of 5% and that at the end of the cascade both had defaulted by having their CET1 capital ratio reduced to 2%. Since the bigger bank would have had a bigger monetary loss, it would have contributed more to the numerator in equation (2.5) than the smaller bank.

3. It is standardised over systems with different total asset values and equal capital ratios if the fraction of capital left over at the end of the cascade remains the same. This is because equation (2.5) is expressed as a percentage of total capital.

4. Increasing/decreasing the number of banks N in the system will not skew the results unduly, since the denominator in equation (2.5) will increase/decrease along with a potential increase/decrease in the numerator.

2.2 Different network structures

Six network structures were investigated for the purpose of this chapter. The definitions of the structures are presented in section 2.2.1 below, where a standardisation procedure is presented to ensure that the structures are comparable to one another. Three global network measures are used in section 2.2.2 to compare the structures to one another and to gain insight into how relevant the structures are for modelling systemic risk.

[2]It makes sense to consider the average over a number of simulations for the links, since in practice the short-term interbank lending relationships change daily. This means that the network links are continuously changing and therefore by taking the average over a number of simulations, it is possible to get an indication of the systemic risk on a 'typical' day.

2.2.1 Definition and standardisation of structures

Each structure is defined according to the probability that there exists a connecting edge between any two banks. This is motivated by empirical evidence from [46] who show that interbank relationships are formed randomly on a daily basis, based on the true underlying structure [91]. Our probability model approach is therefore appropriate, as we are considering static model as opposed to a dynamic approach where the underlying structure might change over time.

Three of the structures are included in the interest of investigating a wider range of structures, including ones that are widely known in network theory. These are the Erdős-Rényi, assortative and disassortative structures. The remaining three structures improve on others commonly found in the banking network literature by capturing real-world behaviour, namely the presence of tiering in banking networks. Banks included in the upper tier are more connected to the remainder of the network, while the lower tiered banks are less connected amongst themselves than to the upper tiered banks. Such tiered structures have been found to be more representative of actual banking networks than scale-free structures, and have a much clearer economic meaning [41]. When only two tiers are considered, a tiered structure is similar to a core-peripheral structure (where the system consists of a few, highly connected core banks and a larger number of small, less interconnected banks that form the periphery).

For all structures apart from the Erdős-Rényi case, the connection probabilities are based on the total asset values of banks. It is natural to assume that the number of debtors and creditors for small banks will differ from that of large banks, and that the probability of a bank extending (receiving) credit to (from) another bank will differ between banks of differing sizes. Indeed, empirical evidence has shown that banks in the upper tier tend to be large [41]. This behaviour is captured in our three more realistic structures. The formulae for determining the connection probabilities for all structures are now presented.

Let $p(i, j)$ be the probability that bank i has an outgoing edge to bank j. As banks cannot lend to themselves, $p(i, j) = 0$ whenever $i = j$. For the rest of this section, assume that $i, j \in \{1, 2, \ldots, N\}$, with $i \neq j$. The proposed formulae for $p(i, j)$ for structures 2 to 6 are motivated by the fact that they are simple and easy to calculate, while managing to capture the intended lending behaviour as a function of banks' asset values. In order to enable comparisons of the different structures, the $p(i, j)$ probabilities are normalised before applying them. The different network structures are specified below. After this, we discuss the procedure for normalising the the $p(i, j)$ probabilities to enable comparisons between them.

1. Erdős-Rényi, where each edge from i to j exists with a fixed probability $p(i, j) = p$ for all banks i and j where $i \neq j$. This is used as benchmark network structure as well-known early network models of banking systems (e.g. [78, 52, 75]) are based on it due to its simplicity. However, it is not very realistic as banking networks have been found to exhibit different types of structures. The Erdős-Rényi structure is investigated by this study for comparison purposes, even though it does not necessarily have much practical value if the structure of a system plays an important role.

2. Disassortativeness, where banks of similar size are generally unwilling to do business with one another. The opposite holds for banks that differ greatly in size. This structure is investigated for the sake of theoretical interest since this is a well-known structure type in network theory. For this structure, the connection probabilities are given by

$$p(i, j) = \frac{\max\left\{\frac{a_i}{a_j}, \frac{a_j}{a_i}\right\}}{\max\limits_{\substack{k,m\in\{1,\dots,N\}\\ k\neq m}} \left\{\frac{a_k}{a_m}, \frac{a_m}{a_k}\right\}}. \tag{2.7}$$

3. Assortativeness, where banks of similar size are keener to do business with one another than banks that differ in size. As for the disassortative case, this structure is unlikely to be found in actual banking networks. It is included here for theoretical purposes since it is a well-known structure from network theory and the modelling concept is based on ideas from network theory. The connection probabilities before normalisation are given by

$$p(i, j) = \frac{\min\left\{a_i, a_j\right\}}{\max\left\{a_i, a_j\right\}}. \tag{2.8}$$

4. "Attraction to size", where any bank is generally more willing to lend to a bigger bank. In other words, banks with higher asset values are viewed as more creditworthy by lender banks. This type of structure might be more representative of periods of financial distress, as lenders may perceive bigger banks to be better able to absorb financial losses. This can also be viewed as a situation where the bigger banks are regarded as too-big-to-fail, reflecting a general belief that a lender of last resort will step in if such banks run into trouble (even though this is not the case with the chosen model methodology). As the bigger banks are highly interconnected with the rest of the system, this structure may lead to the bigger banks also being too-interconnected-to-fail. A detailed analysis of the consequences of bailouts for such banks is left for

future research as the focus here is on the propagation of risk, and not assessing the cost of bailouts or regulation. For this structure, the formula for the probability of a bank i lending to another bank j is given by

$$p(i,j) = \frac{a_j}{\max\limits_{k\in\{1,\dots,N\}} \{a_k\}}. \tag{2.9}$$

5. Tiered type I, where large banks tend to lend to each other frequently and small banks are reluctant to lend to one another. The probability of smaller and larger banks doing business with one another is in between these two extremes. This structure attempts to provide a more realistic view of potential banking structures. In this case

$$p(i,j) = \frac{a_i + a_j}{\max\limits_{\substack{k,m\in\{1,\dots,N\}\\ k\neq m}} \{a_k + a_m\}}. \tag{2.10}$$

   The denominator is a constant used to ensure that $p(i,j) \leq 1$. The numerator $a_i + a_j$ then ensures that $p(i,j)$ is at its lowest when a_i and a_j are the two lowest asset values, and that $p(i,j)$ is one when a_i and a_j are the two highest asset values in the system. In general, $p(i,j)$ will be low when both a_i and a_j are low, and it will be higher when both a_i and a_j are high.

6. Tiered type II. This structure attempts to create a hierarchical banking network which can also be considered as more realistic. The probability of

   - a small bank lending to another small bank is low,
   - a small bank lending to a large bank is also relatively low
   - a large bank lending to a small bank is high,
   - a large bank lending to another large bank is also high.

   To achieve this structure, the following formula is used:

$$p(i,j) = \frac{a_i + a_j + \max\left\{a_i - a_j, 0\right\}}{3 \cdot \max\limits_{k\in\{1,\dots,N\}} \{a_k\}}. \tag{2.11}$$

   Similar to the tiered type I structure, the denominator is a constant that ensures $p(i,j) \leq 1$ and the behaviour of the structure is captured by the numerator. Part of the numerator is the same as that of the tiered type I structure, namely the quantity

$a_i + a_j$. The term $\max\left\{a_i - a_j, 0\right\}$ is used to refine this structure further — if the lender bank i is larger than the borrower bank j, then the probability $p(i, j)$ is increased.

Figure 2.1 illustrates the behaviour of each structure. The size of a node is indicative of the bank's total asset value and the transparency of a line suggests the probability of a directed edge existing. For example, a solid line from a larger node to a smaller node indicates that the probability of a directed edge existing from a large bank to a small bank is high. A dashed, transparent line indicates a lower probability of an edge existing.

This concludes the description of the standard connection probabilities between nodes in the system. Note that for a given set of asset values, the average number of edges will differ between the different structures. For comparative purposes, it is important to work with similar levels of connectedness between the different structures. Otherwise differences in systemic risk levels between structures may largely be driven by differences in connectivity as opposed to differences in the network structure alone. In addition to this, it is important to have a systematic way of varying the level of connectedness in the network to investigate its effect on systemic risk.

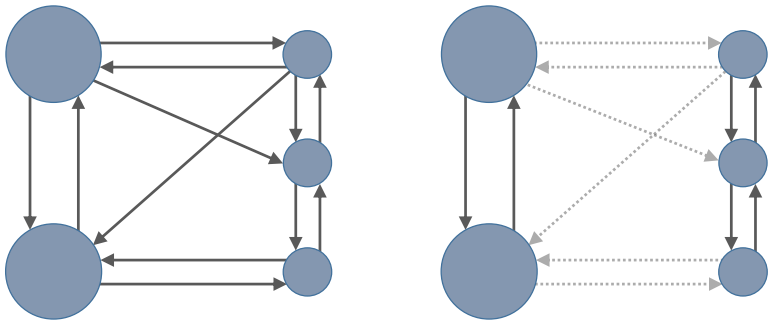
For these reasons, a method for scaling the probabilities is required. Such a method will need to provide scaled probabilities that remain in the range $[0, 1]$. It must further allow one to standardise the different structures to represent the same level of connectivity between them. We will measure the level of connectivity by means of the average probability of an edge existing in the system (i.e. the average probability of one bank lending to another).

Let $\bar{p}_0$ be the average probability of an edge existing in the system based on any one of the above six structures. Then

$$\bar{p}_0 = \frac{\sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} p(i, j)}{N(N-1)}. \tag{2.12}$$

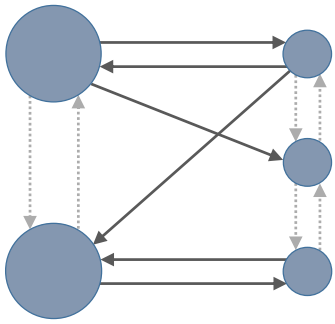
Note that $p(i, j) = 0$ whenever $i = j$. These entries are not included in the calculation of the above average and they are to remain zero after scaling. The aim is now to find a way of scaling the $p(i, j)$ probabilities in such a way that a new average, say $\bar{p}$, is obtained. The new average $\bar{p}$ is specified beforehand and the $p(i, j)$ probabilities need to be scaled accordingly. Here, the probabilities $p(i, j)$ should be adjusted in a consistent way such that the new probabilities remain in the range $[0, 1]$ and that the ranking of the probabilities remain the same. In other words, if two probabilities $p(k, l)$ and $p(m, n)$ satisfy $p(k, l) < p(m, n)$ before scaling, then the same inequality must hold after the probabilities have been scaled.

Scaling all the $p(i, j)$ probabilities by the same factor will not suffice. If this were

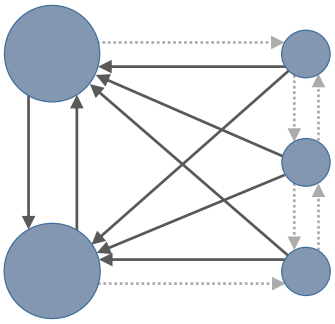


(a) Erdős-Rényi

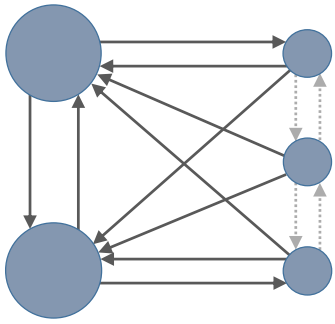
(b) Assortativeness



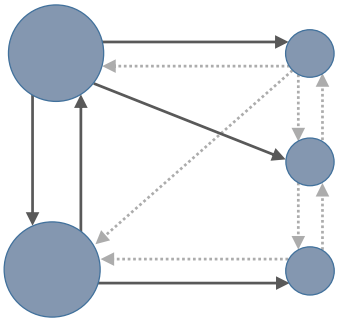
(c) Disassortativeness



(d) Attraction to size



(e) Tiered type I



(f) Tiered type II

Figure 2.1: Illustration of the different network structures considered for application of this simulation model.

to be done, it would not be guaranteed that the probabilities will remain smaller than 1. Therefore one needs to consider the distances of $\bar{p}_0$ and $\bar{p}$ to 1 and to 0. Suppose $\bar{p}_0 > \bar{p}$. Then the $p(i, j)$ probabilities need to be scaled downward. Let the new probabilities be

given by

$$\tilde{p}(i,j) = p(i,j)\,\frac{\bar{p}}{\bar{p}_0}. \tag{2.13}$$

The new probabilities (where $i \neq j$) will now have the required average:

$$\frac{\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\tilde{p}(i,j)}{N(N-1)} = \frac{\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}p(i,j)\,\bar{p}}{N(N-1)\,\bar{p}_0} = \bar{p}. \tag{2.14}$$

Suppose now that $\bar{p}_0 < \bar{p}$. The probabilities need to be scaled upward closer toward one. The new probabilities are then given by

$$\tilde{p}(i,j) = 1 - (1 - p(i,j))\,\frac{1-\bar{p}}{1-\bar{p}_0}, \tag{2.15}$$

and they have the required average since

$$\begin{aligned}
\frac{\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\tilde{p}(i,j)}{N(N-1)} &= \frac{1}{N(N-1)}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\left(1-(1-p(i,j))\,\frac{1-\bar{p}}{1-\bar{p}_0}\right)\\
&= 1-\left(\frac{1-\bar{p}}{1-\bar{p}_0}\right)\left(1-\frac{\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}p(i,j)}{N(N-1)}\right)\\
&= 1-\left(\frac{1-\bar{p}}{1-\bar{p}_0}\right)(1-\bar{p}_0)\\
&= \bar{p}.
\end{aligned} \tag{2.16}$$

2.2.2 Characteristics of the proposed structures

The purpose of this section is to investigate meaningful network characteristics of the structures above. This is done firstly to facilitate interpretation of the results in section 2.3 and to determine whether such characteristics can be used to explain differences in systemic risk. Secondly, the results are used to determine whether the attraction to size and tiered type I and II structures are more appropriate for systemic risk modelling compared to the other structures defined in section 2.2.1. Finally, we use the results presented here to infer an appropriate choice of $\bar{p}$ to use as a base parameter for the analyses in section 2.3.

The following global network characteristics are considered in this section:

1. **Average shortest path:** The Floyd-Warshall algorithm [47] is used to calculate the shortest paths for all nodes. Note that this average does not need to be calculated

separately for the out-degrees and the in-degrees, since one node's in-coming links are others' out-going links.

2. **Weighted clustering coefficient:** This takes account of both the direction and the weight of the edges. The measure and its explanation below is based on [79].

3. **Tiering error:** This is based on a method of measuring how far a given network is from a perfectly tiered network [41]. The adopted definition of a tiered network is specifically chosen to make sense in a banking context.

We restrict our attention to the above measures, as other measures of network characteristics do not provide additional meaningful information in our context. For example, the level of tiering provides similar, but more meaningful, information compared to centrality measures since it specifically requires larger nodes to be more central to the network. Other measures such as the diameter of the network do not provide any additional information over the chosen measures and cannot be used to test for important characteristics such as the small-world property.

The simulations done for the purpose of this section are based on an example network of $N = 35$ banks, where the total assets are evenly distributed over a bounded Pareto distribution[3] with parameters $\alpha = 0.001$, $x_{min} = 1\,000$ and $x_{max} = 140\,000\,000$. The Pareto distribution is appropriate as it has been used in the past to model the distribution of wealth [81] and empirical evidence has found that assets follow a heavy tailed distribution. Note however that if the unbounded Pareto distribution is chosen, it typically results in only a single large bank that overshadows the rest due to the sub-exponential property, while it is of interest here to have more than one large bank. The bounded Pareto distribution results in a few large banks and various smaller banks, which is more representative of actual banking systems than a system containing a single large bank. The minimum and maximum values chosen are roughly within the range of asset values observed in the South African banking system (if the asset values are given in R'0 000s), which is of similar size as our benchmark system. A graphical representation of this system for the tiered type II structure with $\bar{p} = 0.1$ is given in figure 2.2 for illustration purposes. Here the nodes will, on average, have 3.4 outgoing edges and 3.4 incoming edges. It can be seen that the smallest nodes on the perimeter of the network generally have two to four edges in total connected to them. That is, the outgoing and incoming edges together add up to between two and four edges as opposed to the average total of 6.8. This means that the largest

[3]The version of the bounded Pareto that we consider has probability density function $f(x) = \frac{\alpha\, x_{min}^{\alpha}\, x^{-\alpha-1}}{1-\left(\frac{x_{min}}{x_{max}}\right)^{\alpha}}$, where $x_{min} \leq x \leq x_{max}$.

nodes have many more edges connected to them to compensate for the numerous small nodes with few edges.

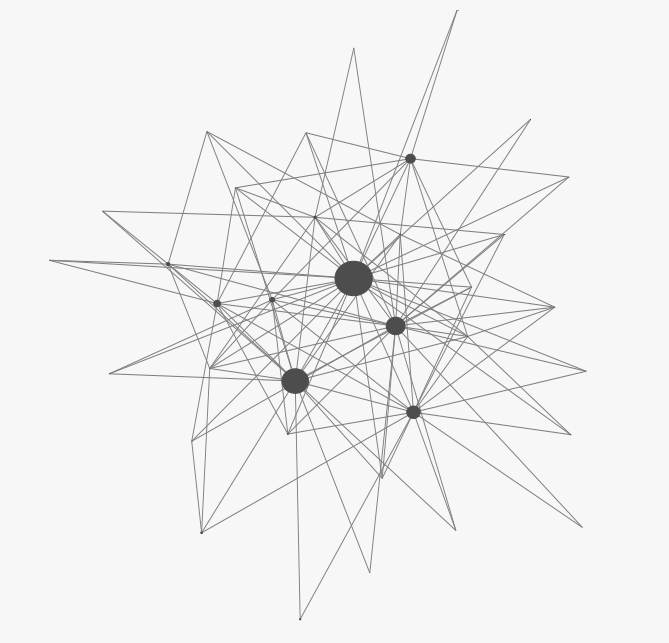


Figure 2.2: Graphical representation of a hypothetical tiered type II banking network with average connection probability $\bar{p} = 0.1$.

Average shortest path

The average shortest path from one node to another in the network provides information on how close banks are to one another in terms of their direct credit exposures. This is closely related to the average level of interconnectedness of the network.

A low average shortest path may mean that contagion in the network reaches the remaining banks in the system more quickly. In network theory, low average path lengths can be an indication of efficiency, since information is passed on easily between nodes. Since we consider the spread of losses, it is ideal in our setting to prevent the spread of 'information' between nodes. When more than one bank fails, the other banks are more likely to receive direct counterparty losses. This is because they are more likely to be directly connected to the failing banks.

On the other hand, the typical shortest path is an average taken over all nodes. Therefore a low shortest path could mean that there are many disconnected banks in the system, with some nodes having lots of edges connected to them. If this were the case, then some banks won't suffer any direct counterparty losses and others are on average likely to distribute its

interbank assets across many banks, meaning that losses are spread more thinly between counterparties than otherwise. This may prevent the remaining banks from experiencing large losses, thereby preventing the spread of contagion.

Figure 2.3 shows a comparison of the average shortest paths obtained from each network structure for varying levels of interconnectedness, $\bar{p}$. As expected, figure 2.3 shows a general decline in average shortest paths as $\bar{p}$ increases. For the attraction to size structure, there is a sudden increase in the average shortest path at approximately $\bar{p} = 0.1$. This is due to the fact that the original average level of interconnectedness, $\bar{p}_0$, for the attraction to size structure as given by equation (2.9) is equal to 0.0991 for our choice of base parameters. Suppose bank k is the largest bank. Then before any scaling of the probabilities $p(i, j)$ (i.e. before applying equation (2.13) or (2.14)), we have that $p(i, k) = 1$ for all $i \neq k$. Therefore if the required average level of connectedness satisfies $\bar{p} < \bar{p}_0$, then there will be a non-zero probability of having at least one bank with no in-degrees or out-degrees. But if $\bar{p} \geq \bar{p}_0$, then all banks $i \neq k$ in the system will always be connected to k. This is what gives rise to the jump in shortest path at approximately $\bar{p} = 0.1$.

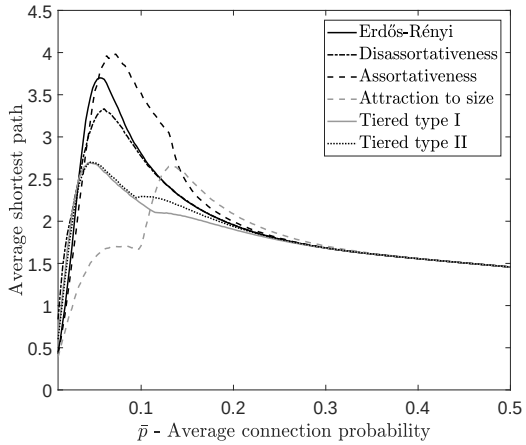


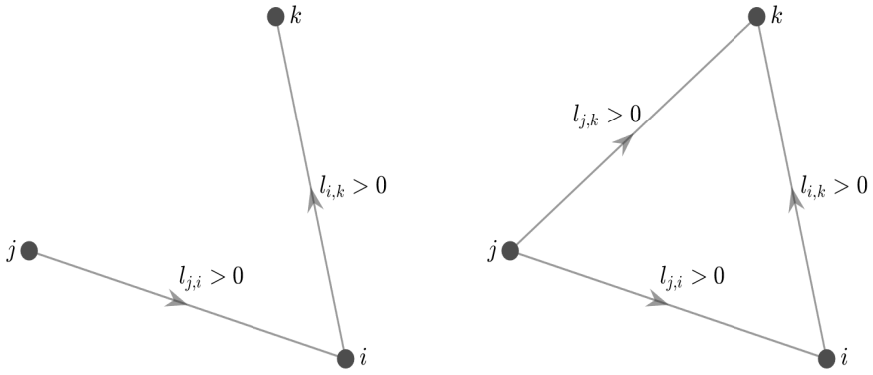
Figure 2.3: The average shortest path in the system for each structure as a function of $\bar{p}$, the level of interconnectedness in the system.

In addition to the attraction to size structure, the assortative, tiered type I and tiered type II structures also have kinks in their respective lines in figure 2.3. This is due to the way that the scaling is done in equations (2.13) and (2.14). Since the formula used for scaling changes when $\bar{p}$ is increased beyond $\bar{p}_0$, the resulting rate of increase for each individual $p(i, j)$ changes as well. The point at which this happens differs between structures, since the value of $\bar{p}_0$ differs between them. The disassortative structure however, does not visibly display the same behaviour at the relatively small original average of $\bar{p}_0 = 0.02$.

Global clustering coefficient

In social networks, clustering can be used to describe the likelihood that two acquaintances of a person are also acquainted. In other words, suppose a node i is connected to two other nodes j and k. Then the global clustering coefficient provides information on the likelihood that j is also connected to k [79].

Three nodes i, j and k form a triplet whenever they are connected by two edges such that there is a one-directional flow from the first node to the last. For example a triple centred on node i will satisfy either $l(j,i) > 0$ and $l(i,k) > 0$ or $l(k,i) > 0$ and $l(i,j) > 0$. A triplet is transitive or closed whenever there is a directed edge from the first node in the triplet chain to the last. Otherwise it is called intransitive. This is illustrated in figure 2.4. In this setting the weights of the edges are given by the exposure amounts $l(\alpha,\beta)$. This is used to determine the weight assigned to a triplet, which is given by $(l(j,i) + l(i,k))/2$ in figure 2.4.



(a) Illustration of an intransitive triplet (b) Illustration of a transitive triplet

Figure 2.4: An illustration of an intransitive and transitive triplet centred on node i. The exposures between the banks are the weights of the edges and both triplets' weights are equal to $\frac{l(j,i)+l(i,k)}{2}$.

The global weighted, directed clustering coefficient of a graph $\mathcal{G}$ is then given by

$$C(\mathcal{G}) = \frac{\text{Total weight of transitive triplets}}{\text{Total weight of all triplets}}. \tag{2.17}$$

A high clustering coefficient indicates that a large number of triplets in the system are transitive. To investigate what this implies for a banking system, consider what happens to each triplet in figure 2.4 in the event that bank k fails. Let l_m denote the interbank assets of any bank m and assume for simplicity that the triplet is disjoint from the rest of the system. This amount needs to be divided amongst m's counterparties. Therefore the more

Table 2.2: Comparison of the potential direct counterparty losses for an intransitive and transitive triplet, given the failure of k or/and i.

Direct counterparty losses		
	Intransitive triplet	Transitive triplet
If only k fails:	l_i	$l_i + \frac{l_j}{2}$
If only i fails:	l_j	$\frac{l_j}{2}$
If k and i fails:	$l_i + l_j$	$l_i + l_j$

counterparties that m has, the smaller the loan amount allocated to each counterparty. For the purpose of this illustration, we assume without loss of generality that interbank assets are divided equally amongst counterparties, failing banks default on the full amount of their interbank loans and that the effect of liquidity and proximity risk parameters can be ignored.

Assume first that the triplet centred on bank i is intransitive, so that it is represented by figure 2.4a. If k fails, then i loses l_i due to direct counterparty losses. If i also fails, then j loses l_j in direct counterparty losses. Note that we do not consider the potential failure of j, because a triplet does not require any directed edges from i or k to j. Suppose on the other hand that the triplet centred on i is transitive (figure 2.4b). Then if k fails, i loses an amount of l_i and j loses an amount of $\frac{l_j}{2}$. If i fails, then j loses a further $\frac{l_j}{2}$.

These possibilities are summarised in table 2.2, from which it can be seen that neither column always contains losses less than or equal to the other column. If only bank k fails, then the total capital lost is greater for a transitive triplet. The opposite holds true if it were i that defaulted. Consider now the event that k defaulted first and that this caused i to default as well. Then the total loss would be equal, but the loss to node j would occur earlier in the default cascade for a transitive triplet. This is because j would lose an amount of $\frac{l_j}{2}$ during the first round of default already, which is not the case for the intransitive triplet. Therefore, even if i did not default in the case of a transitive triplet, j would still suffer a loss which may consequently lead to its failure. This suggests that the effect of clustering on the risk in the system is not straightforward, but it may be possible that a large number of transitive triplets can lead to more defaults if bank j as illustrated in figure 2.4 does not hold enough capital to withstand the event that bank k fails. This discussed further in section 2.3.

Figure 2.5 contains the average clustering coefficient for the different structures for a range of interconnection levels. As expected, the clustering in the system generally increases for increasing levels of interconnectedness. Similar to figure 2.3, there is a sudden

change in the slope of each structure's line where $\bar{p} = \bar{p}_0$ in figure 2.4. This is again due to the difference in the scaling formulae used, depending on whether $\bar{p} < \bar{p}_0$ or $\bar{p} > \bar{p}_0$ (see equations (2.13) and (2.14)).

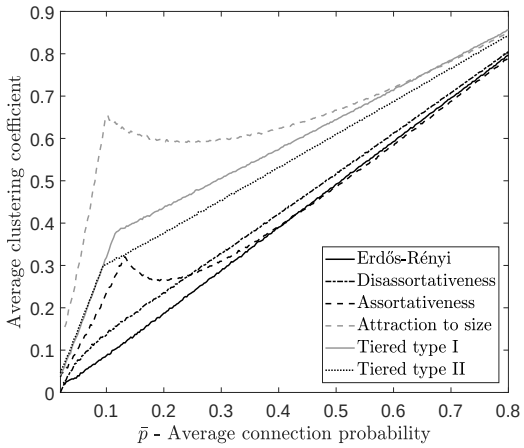


Figure 2.5: The average weighted, directed clustering coefficient for each network structure as a function of $\bar{p}$, the level of interconnectedness in the system.

The attraction to size structure exhibits a significantly higher level of clustering compared to the rest of the structures, while the two tiered structures also exhibit higher clustering levels compared to the the Erdős-Rényi network. This is suitable, since empirical evidence has found that real banking systems tend to exhibit high clustering [66]. Even though some empirical studies find that disassortative structures are representative of real-world banking systems [66], we argue that a tiered approach is more appropriate since it is more refined and carries more economic meaning [41]. This is supported by the fact that the disassortative structure shows clustering levels that are too close to a purely random network (i.e. to the Erdős-Rényi network) in figure 2.5.

Tiering error

In a perfectly tiered network, the following requirements are satisfied:

(a) Top-tier banks always lend to one another;

(b) lower-tier banks never lend to one another;

(c) each top-tier bank must lend to at least one lower-tier bank; and

(d) each top-tier bank must borrow from at least one lower-tier bank.

The tiering error measures how far a given network is from a perfectly tiered network. The definition of a perfectly tiered network bears similarities to our proposed structures 4 to 6. One would therefore expect these to have a lower tiering error than the other structures.

Suppose there are n banks in the system, c of which are top-tier banks. For all $i, j \in \{1, 2, \ldots, n\}$, let $L_{i,j} = 1$ if $l_{i,j} > 0$ and $L_{i,j} = 0$ otherwise. Then the tiering error of a graph $\mathcal{G}$ is given by

$$E(\mathcal{G}) = \frac{E_{CC} + E_{PP} + E_{CP} + E_{PC}}{\sum_{i=1}^{n} \sum_{j=1}^{n} L_{i,j}}, \tag{2.18}$$

where E_{CC}, E_{PP}, E_{CP} and E_{PC} respectively denote the components of the error measure that are associated with each of the tiering requirements (a) through (d). Note that in order to calculate this measure, each node in the network $\mathcal{G}$ must be categorised as either top-tier or lower-tier. The combination of top-tier and lower-tier banks must be chosen such that the tiering error is minimised. Simulated annealing can be used to find the optimal combination of top-tier and lower-tier nodes that minimises the tiering error. Simulated annealing is an optimisation technique that can be used to find global maxima or minima of a given function. The terminology used to explain simulated annealing come from thermodynamics, where the process of slowly lowering a substance's temperature is called annealing [27]. The process of performing simulated annealing is described by [27] as follows:

1. Choose a starting temperature $T = T_0$ and a decreasing function that is used to decrease the temperature at each iteration of the procedure. The values that are assigned to the temperature T at the various iterations of the procedure is called the annealing schedule. There are no concrete rules for determining T_0 and the annealing schedule. For this application it was found that $T_0 = 1$ and a temperature function $f(k) = 1 - \frac{k}{100}$, $k \in \{0, 1, \ldots, 99\}$ could be used to find the optimal core.

2. Choose a starting set of initial estimates. This choice was inferred by the discussion in [41], where it is indicated that a bank is part of the initial periphery if it either has no borrowers or no lenders. Otherwise it forms part of the initial core. Our initial estimate is therefore the set initial core banks.

3. Randomly select a 'neighbour' of the initial estimate. Following [41], a neighbour is found by randomly picking any node from either the core or the periphery and moving it to another category. In other words, if the chosen node was part of the core, it is moved to the periphery and vice versa. The new resulting set of core banks is then the neighbour.

This step is subject to the constraint that a node in the core must borrow from, or lend to at least one node in the periphery. Hence, nodes that cannot be moved to the core cannot be selected to determine the neighbour estimate.

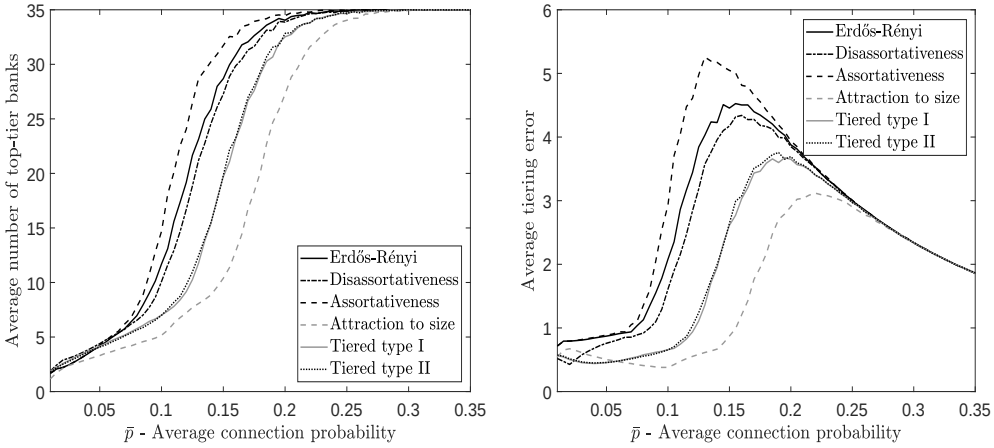
4. Compare the function values based on the initial estimate and its randomly chosen neighbour. The function is this application is the tiering error given by equation (2.18). Let θ_{old} and $\theta_{neighbour}$ be the tiering errors based on the initial estimate and the neighbour respectively. Then if $\theta_{neigbour} < \theta_{old}$, the neighbour estimate is chosen as the new estimate for the core. Otherwise, select the neighbour estimate with probability $\exp\left\{\frac{\theta_{neigbour}-\theta_{old}}{T}\right\}$ and keep the old estimate with probability $1 - \exp\left\{\frac{\theta_{neigbour}-\theta_{old}}{T}\right\}$.

5. Repeat steps 3 and 4 until an equilibrium is reached. Following [41], equilibrium is achieved when the error score cannot be reduced any longer. The equilibrium is taken as the set of core banks that produced the least tiering error after repeating steps 3 and 4 100 times.

6. Lower the temperature according to the annealing schedule and repeat steps 2–5, using the equilibrium from the previous step as the new initial estimate. After the temperature has reached 0, the algorithm is stopped and the final equilibrium is the approximation for the global minimum.

Given the refined definition of the tiering error, its consequences for systemic risk is not straightforward. This is discussed further in section 2.3 below. As illustrated by figure 2.6, the level of tiering and the number of banks belonging to each tier will vary with the interconnectedness of a network. This is because a network with no interconnections should be perfectly tiered with all nodes belonging to the lower tier. Similarly, a fully connected system will be perfectly tiered with all banks belonging to the top tier. This is illustrated by figure 2.6a where the average number of top-tier banks is an increasing function of $\bar{p}$, starting close to zero and increasing to 35 (i.e. the total number of banks in the system). Between these two extremes it is expected that networks will rarely be perfectly tiered. This is supported by figure 2.6b, which shows that the error is relatively small for sparse networks, but never very close to zero. As $\bar{p}$ increases, the tiering error increases and then shows a decreasing trend as the network becomes more interconnected.

The three network structures that we consider to be the most realistic exhibit the lowest average tiering error for most values of $\bar{p}$. This is to be expected given the similarities between the adopted definition of a perfectly tiered system and the formulae used to determine the probabilities $p(i, j)$. It is interesting to see that the tiered types I and II structures



(a) Average number of banks in the top tier as a function of $\bar{p}$

(b) Average tiering error as a function of $\bar{p}$

Figure 2.6: Comparison of the tiering error and the average number of banks in the top tier between all structures. These are expressed as functions of the average connection probability $\bar{p}$.

exhibit the same average tiering error throughout, even though they are not identical. Furthermore, the less sophisticated attraction to size structure matches a tiered structure more closely than tiered types I and II. If $E(\mathcal{G}) \geq 1$, then it is not worth fitting a tiering model at all [41]. This shows that the choice of $\bar{p}$ is an important factor to consider when deciding whether a structure can be regarded as tiered or not. Although figure 2.6b suggests that the attraction to size and tiered type I and II structures are significantly more tiered than a random network, we test this formally in appendix A.1.

Implications of network characteristics

The results from this section alone are not enough to fully explain differences in risk levels for the network structures. This is because a single number used to characterise a whole system does not contain enough information to make any inferences. The average shortest path and the clustering coefficient may be useful in explaining some of the results from section 2.3 below, but the tiering error is more useful for assessing the appropriateness of different structures.

Figures 2.3, 2.5 and 2.6 can, however, be useful in assessing whether the proposed network structures satisfy empirically observed characteristics. As real-world banking systems tend to satisfy the small-world property, it is preferable for structures to have smaller average shortest paths and higher clustering coefficients compared to an Erdős-

Rényi network. Furthermore, such structures need to be tiered and sparse. Sparseness can be achieved for any structure by choosing a low enough value for $\bar{p}$. This needs to be balanced against the fact that $\bar{p}$ influences the average shortest path, clustering and level of tiering of a network. Figures 2.3, 2.5 and 2.6 show that the attraction to size, tiered type I and tiered type II structures are good candidates for appropriate network structures. They generally exhibit low average shortest paths (except for the attraction to size structure when $\bar{p} > 0.11$), high clustering coefficients and low tiering errors compared to an Erdős-Rényi network.

The attraction to size and tiered type I and II structures can therefore be used to infer a default value for $\bar{p}$ for use in section 2.3, as these most closely resemble real-world systems. The largest banks should ideally form the top tier of the network. From figure 2.2 it can be seen that between four and seven banks in the network are significantly larger than the others. The formulae for the connection probabilities $p(i, j)$ for the three realistic structures generally lead to higher levels of interconnectedness for the largest banks. Therefore the largest banks will generally be part of the top tier. Based on these observations, figure 2.6a implies that $\bar{p} = 0.1$ is a sensible choice, since the core consists of approximately five to 6.5 banks on average. This is supported by figure 2.6b, which shows that the average tiering errors are relatively low (and less than one) for the realistic structures when $\bar{p} = 0.1$. Finally, since this choice of $\bar{p}$ will result in sparse networks, we conclude that it is an appropriate choice for illustration purposes.

This concludes the discussion regarding the network properties of the different structures. In section 2.3 below, the simulation model described in section 2.1 is applied to these structures to investigate their effect on systemic risk.

2.3 Sensitivities and the effect of different network characteristics

This section presents the sensitivity tests performed on the simulation model investigated in this chapter. The benchmark system used for the sensitivity tests are described in section 2.3.1. Thereafter, section 2.3.2 presents the effects on systemic risk levels when the model parameters are varied.

2.3.1 Constructing the benchmark system

For a single run of the model, the links are simulated once and the system is shocked N times. For each time i that the system is initially shocked, bank i's total assets are reduced by $s = 0.2$ of their original value[4]. The average proportion of capital lost by all banks over the N cascades is then calculated (see equations (2.5) and (2.6) on page 30 for the definition of the CRR, the capital reduction ratio). For each time n that the system is shocked, the balance sheet entries of the banks are reset to their original positions, and the default cascade starts with bank n receiving the initial loss. For each network structure, the model is then run 10 000 times to obtain an average CRR over a number of simulations.

To ensure comparability with the results of section 2.2.2, the illustrations in this section are based on a system of $N = 35$ banks with total assets evenly distributed over a bounded Pareto distribution with the same parameters as before ($\alpha = 0.001$, $x_{min} = 1\,000$ and $x_{max} = 140\,000\,000$). The average probability of an edge existing between any two nodes i and j ($i \neq j$) is given by $\bar{p} = 0.1$. This is in line with our results from section 2.2.2 and ensures that we obtain sparse networks that exhibit empirically observed characteristics for our three most realistic structures.

The CET1 ratio for all banks is $\gamma = 0.05$, which is higher than the Basel III risk weighted minimum of 4.5% (note that the minimum CET1 ratio is lower than the minimum Tier I capital ratio of 8%). The difference between our CET1 ratio and the Basel minimum CET1 ratio is sensible since regulators generally set their own minimum capital requirements which are often higher than the Basel minimum requirements. Furthermore, the ratio used as a benchmark here is not directly comparable to that of Basel III since the former ratio is unweighted.

The benchmark system based on the above parameters is investigated with and without liquidity and proximity effects. When liquidity and proximity reduction factors are included, the base values are given by $g^{(s)} = 0.05$, $g^{(m)} = 0.06$, $g^{(l)} = 0.07$ and $\delta = 0.06$ respectively. All of the above base parameters are summarised in table 2.3.

In order to apply the method outlined in section 2.1.2 for determining the liquidity shocks, it is necessary to assign a business model to each bank. The assignment of banks to retail-funded, wholesale-funded or capital markets-oriented business models is done in line with the empirical findings of [82]. Therefore we assume 25 retail-funded, 6 wholesale-funded and 4 capital markets-oriented banks which is, on average, in line with their latest

[4]It is noted that in practice not all external assets lose their value in the same way, for example banks may hold mostly covered bonds or mostly sovereign bonds. However, for the purpose of this illustration the mechanism that leads to the loss is not specified to avoid over-complicating the model and distracting from the research objectives.

Table 2.3: Parameter values for the base model.

	Parameter	Value
N	Number of banks	35
α	Bounded Pareto shape parameter	0.001
x_{min}	Bounded Pareto scale parameter	1 000
x_{max}	Bounded Pareto scale parameter	140 000 000
γ	Capital ratio	0.05
$g^{(s)}$	Short-term liquidity reduction factor	0.05
$g^{(m)}$	Medium-term liquidity reduction factor	0.06
$g^{(l)}$	Long-term liquidity reduction factor	0.07
δ	Proximity reduction factor	0.06
$\bar{p}$	Average probability of an edge existing between nodes $i \neq j$	0.1

(2013) sample of banks. For each simulation of the links, the banks will be randomly assigned a business model according to these proportions. It is important to note that the typical proportions of a banking system that belong to each business model category is not the same for different countries. This observation is supported by the empirical findings of [14]. Therefore it will be necessary to adjust these proportions if this model is to be applied to a specific banking system.

It now remains to describe how the asset side of the balance sheet is constructed for our illustrative banking system. The proportions of interbank assets to total assets are chosen to be consistent with the business models assigned to the banks. Based on [82], retail-funded banks are assigned a ratio of 9% interbank assets to total assets. Wholesale-funded banks and capital markets-oriented banks are assigned ratios of 8% and 22% respectively. Note that if a bank does not have any outgoing edges emanating from it, the ratio is automatically set equal to 0% since it should then have no interbank assets. The ratios for the interbank assets of retail-funded and wholesale-funded banks are relatively low compared to the capital markets-oriented banks. This is reasonable since deposits and wholesale debt make up more than 70% of retail-funded and wholesale-funded banks' liabilities [82] compared to 50% for capital markets-oriented banks. Hence, it is expected that retail-funded and wholesale-funded banks will hold a large proportion of assets in the form of loans to retail and wholesale customers (as opposed to other banks), with capital markets-oriented banks holding a slightly lower proportion."

Now each bank's interbank assets are divided proportionally amongst its counterparties

according to their relative asset sizes, i.e. we have that

$$l_{i,j} = \frac{a_j}{\sum_{\{k:\, l_{i,k}\neq 0\}} a_k}. \tag{2.19}$$

Even though this is still a simplification, it is more realistic and practical than dividing the interbank assets equally amongst counterparties.

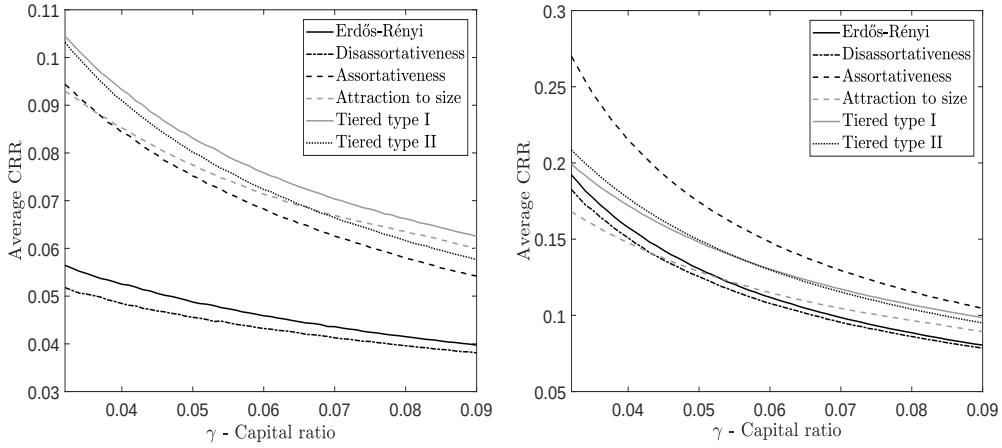
Note that the division of non-interbank assets according to term will be system-specific. For simplicity, our base model will be roughly based on the South-African system, since it is approximately the same size as our benchmark system. The short-, medium- and long-term assets will therefore comprise 20%, 35% and 45% of the non-interbank assets.

This system described above is taken as a benchmark for the different sensitivity tests conducted in the remainder of this section. Unless specified otherwise, the parameters therefore remain the same throughout. For ease of reference, the effect of liquidity losses and market sentiment on the spread of systemic risk will be referred to as indirect risk for the remainder of the thesis.

2.3.2 The effect of varying model parameters

Consider first the sensitivity of the average CRR to the capital ratio that is chosen. The capital ratio is varied from 0.032 to 0.09 and the resulting average CRR is calculated over all simulations. The results for all network structures considered in section 2.2 are illustrated in figure 2.7a (without indirect risk) and figure 2.7b (including indirect risk). As expected, the average CRR always decreases for increasing capital ratios. This shows that the proposed model of financial contagion behaves as expected.

For our choice of base parameters, the disassortative structure exhibits the lowest risk, except for very low capital ratios in figure 2.7b. Apart from the fact that figure 2.7a exhibits lower CRR averages (which is to be expected), there are three noteworthy observations to be made from figure 2.7 alone. Firstly, the ranking of the structures in terms of systemic risk changes when indirect risk parameters are included. While the attraction to size, tiered type I and tiered type II structures exhibit the most risk in figure 2.7a, this is not necessarily the case in figure 2.7b. Secondly, the lower CRR averages for the Erdős-Rényi structure compared to the attraction to size and tiered type I and II structures show that simulation studies that assume an Erdős-Rényi structure for simplicity may underestimate the level systemic risk. Finally, the CRR averages in figure 2.7b decline more quickly with increasing capital levels than in figure 2.7a. This is likely because for every default in the system, the indirect risk parameters increase the amount of capital lost over and above any losses



(a) Without indirect risk (b) With indirect risk

Figure 2.7: Sensitivity of the average capital reduction ratio (CRR) to the capital ratio, with and without the inclusion of indirect risk.

due to direct exposures. Therefore for every default prevented by holding higher capital, the reduction in losses is greater when indirect risk parameters are included.

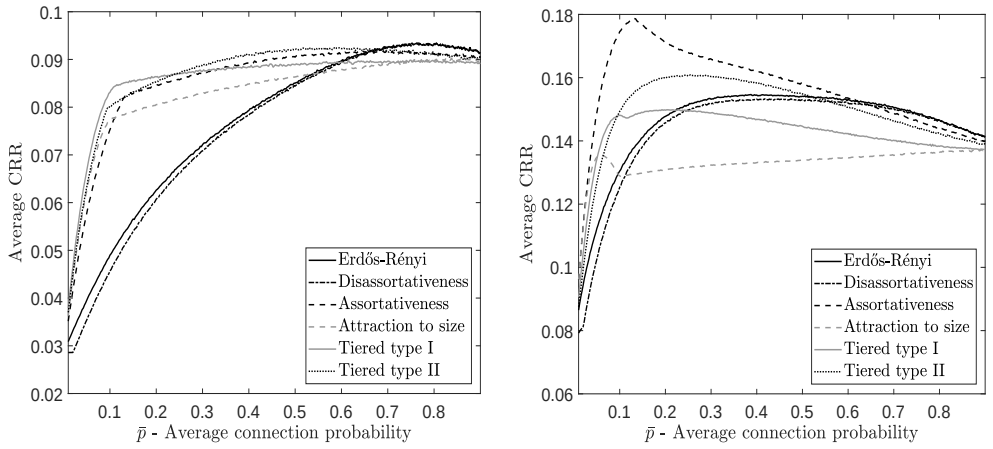
Since the ranking of the structures changes depending on the capital ratio and the inclusion of indirect risk parameters, the network properties discussed in section 2.2.2 will not be able to predict the relative riskiness of the structures under all circumstances. It remains interesting to see whether the characteristics can be used to explain some of the differences observed between the structures.

The ranking of the average shortest paths and clustering coefficients do not correspond to the CRR averages for any level of capital, with or without indirect risk parameters. A general observation that can be made is that the structures with low average shortest paths and high clustering coefficients tend to exhibit higher risk, with the exception of the assortative structure. First consider the effect of the average shortest path. Since the average number of disconnected banks for these structures is less than 1 at $\bar{p} = 0.1$, this is not the reason for the low average shortest path lengths. The traditional interpretation of the average shortest path should then imply that the spread of contagion is generally facilitated by the short path lengths. It is important to note that shorter average path lengths do not always lead to higher risk. The disassortative and Erdős-Rényi structures exhibit a very similar average shortest path at $\bar{p} = 0.1$ in figure 2.3, but do not exhibit the same level of risk in figure 2.7. Furthermore, the assortative structure exhibits the highest average path length, but does not exhibit the lowest level of risk. This may be because banks are in general better able to absorb losses from other banks that are significantly smaller than

themselves (depending on the number of counterparties their assets are spread amongst). Since only similar-sized banks tend to lend to one another, this may deter the contagion cascade from stopping when it reaches small banks.

The assortative, attraction to size and tiered types I and II structures all show higher average CRR levels in figure 2.7a and clustering coefficients in figure 2.5 at $\bar{p} = 0.1$ compared to the Erdős-Rényi and disassortative structures. This does not mean that structures with higher clustering levels always result in more risk compared to those with lower levels. Even though the attraction to size structure exhibits the highest average clustering coefficient, it does not have the highest level of risk. This shows that on a system level, the average shortest path and clustering coefficient may be useful in explaining some of the behaviour of different structures when indirect risk parameters are excluded. They do not contain enough information to determine the relative riskiness of different structures. When indirect risk parameters are included, their interpretation becomes less straightforward, even though different structures still exhibit differences in risk trends.

Consider now the behaviour of the average CRR as the connectedness of the system is varied. We use the average probability $\bar{p}$ of any two nodes $i \neq j$ being connected to represent the connectedness of the system. The parameter $\bar{p}$ is varied from 0.01 to 0.9, and the results are shown in figures 2.8a and 2.8b (with and without indirect effects respectively).



(a) Without indirect risk

(b) With indirect risk

Figure 2.8: Sensitivity of the average CRR to the average connection probability.

Figure 2.8a shows that the trends resulting from varying interconnectedness levels differ both between different structures and between the two different graphs. Whether the system benefits from an increase in $\bar{p}$ or not therefore depends on the network structure,

indirect risk parameters and the current value of $\bar{p}$. Some banking networks may benefit from an increase in connectedness while other networks do not, even if they have similar levels of interconnectedness to begin with. Differences between all structures apart from the Erdős-Rényi and disassortative structures are exasperated when indirect risk is included, further motivating why such factors are important to consider when comparing risk levels under different structures.

The CRR averages for all structures but the Erdős-Rényi and disassortative structures exhibit a notable change in slope around $\bar{p} = 0.1$. Recall from section 2.2.2 that the average shortest paths and the clustering coefficient also had notable differences in slope around $\bar{p} = 0.1$. This change is again related to the change in the scaling formula given by equations (2.13) and (2.14). As the average connection probability approaches 1, the average CRR for the structures converge. This is to be expected since each structure starts to represent a fully connected system.

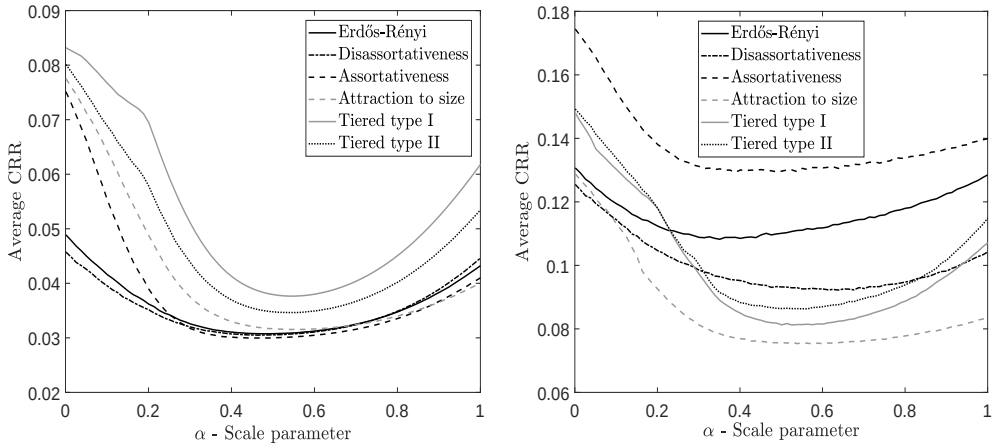
The risk in the system generally increases in figure 2.8a as the shortest paths decrease. For very small values of $\bar{p}$, the shortest paths increase while the risk in the system also increases sharply at the same time. This may be because the number of disconnected banks decrease sharply at this point, opening up new channels of contagion in the system. However, it is more likely due to the scaling of the $p(i, j)$ probabilities, since the sudden changes in slope correspond to the points at which $\bar{p} = \bar{p}_0$ and the disassortative structure (for which $\bar{p}_0$ is too small to feature on the graph) and the Erdős-Rényi structure (for which $\bar{p}$ is irrelevant) do not display sudden changes in slope. It is interesting to note that for $\bar{p} > 0.3$, the shortest paths of the structures are almost identical, but the levels of risk in figure 2.8a clearly differ between the structures. (Note that this observation will likely change when considering different base parameter choices.) Differences in risk levels are even more prominent in figure 2.8b, where indirect risk is included. This again shows that on a system-wide level, differences in the risk levels of network structures are too intricate to be captured with simple network characteristics.

Where the assortativeness, attraction to size and tiered type I and II structures have a steep increase in clustering coefficient (see figure 2.5), the average risk levels in figure 2.8a also show significant increases. However, it is again likely to be due to changes in the scaling formula of the $p(i, j)$'s rather than changes in the clustering coefficient. This is motivated by the fact that the assortative structure's clustering coefficient decreases sharply after approximately $\bar{p} = 0.12$, whereas its average CRR does not.

It is interesting to note that the tiered type I and II structures exhibit almost identical tiering errors in figure 2.6b, but do not exhibit the same level of risk in figure 2.8. This shows that while two structures may exhibit identical characteristics in some ways, they

may still exhibit differences in risk levels. They may also respond differently to changes in certain network parameters as shown in figure 2.8b.

Varying the shape parameter of the Pareto distribution shows how the risk levels respond to changes in how heavy tailed the distribution of assets is. This incorporates two components, namely differences in balance sheet sizes and differences in the connection probabilities $p(i, j)$ which are functions of the asset sizes (except for the Erdős-Rényi structure). The lower the scale parameter, the greater the differences between the larger and smaller banks' asset sizes and connection probabilities. The effect of varying the scale parameter α on the risk levels is shown in figure 2.9 with and without indirect risk.



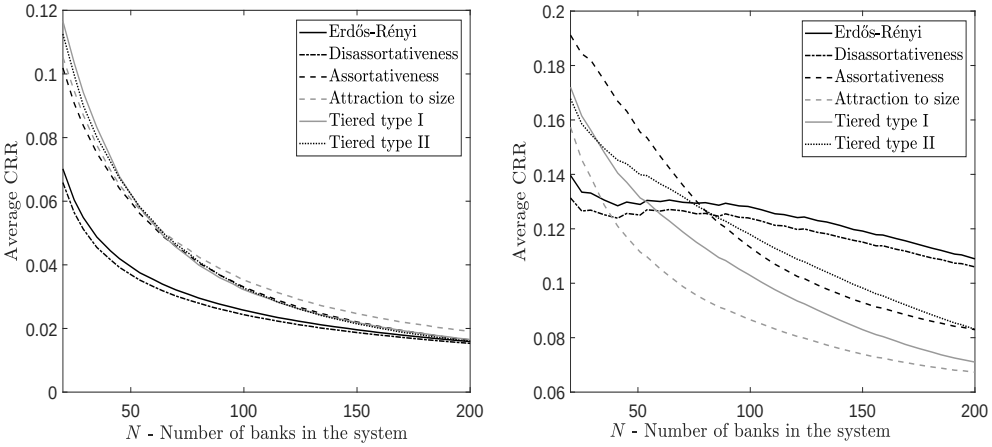
(a) Without indirect risk (b) With indirect risk

Figure 2.9: Sensitivity of the average CRR to the distribution of assets as measured by the shape parameter of the bounded Pareto distribution.

Figure 2.9 shows that the relationship between the Pareto shape parameter and the average CRR is not linear. The average default fractions first reduce as the distribution becomes more light tailed, and then starts to increase. Both figures show that the average CRR is at its highest for very heavy tailed distributions of assets. As α approaches 1, the risk starts to increase again. This shows that while a very heavy tailed distribution of assets leads to high risk levels, a light tailed distribution of assets is also not ideal. If the asset values follow a Pareto distribution, large differences between the largest and smallest banks' assets can be beneficial to the system, provided that the differences are not too large. Figure 2.9b shows once again that the inclusion of indirect risk parameters can have a significant effect on the relative risk levels between the structures. Due to this, network characteristics that do not take such factors into account cannot be used to compare the risk posed by different network structures.

The final model parameter that is considered is N, the number of banks in the network (figure 2.10). The average connection probability remains fixed as the size is increased from 20 banks to 200. To ensure consistency with the base network where $N = 35$, the proportions of retail-funded, wholesale-funded and capital market-oriented banks in the system are kept fixed.

Figures 2.10a (without indirect risk) and 2.10b (with indirect risk) contain the corresponding average default fractions. It can be seen that the distinction between different network structures becomes greater when liquidity losses and market sentiment are taken into account. As before, the inclusion of indirect risk parameters result in changes to the slopes and the ranking of the structures in terms of risk. The average CRR generally decreases as the size of the network increases. This does not mean that larger networks are always safer than smaller ones. For example, if the effects of liquidity losses and loss of market sentiment (as measured by the indirect risk parameters) are high in a very interconnected system, it is possible that larger networks are more prone to collapse than smaller networks.



(a) Without indirect risk (b) With indirect risk

Figure 2.10: Sensitivity of the average CRR to the size of the system.

This is illustrated in figure 2.11, where the following adjusted parameters were used: $\bar{p} = 0.02$ and $g^{(s)} = g^{(m)} = g^{(l)} = \delta = 0.175$. Here it can be seen that the average CRR does not necessarily decrease for larger network sizes. This highlights the importance of looking at all the network characteristics together instead of in isolation. It also shows that a change in network characteristics that may make one financial system safer might not necessarily be good for another system.

From figure 2.10 alone it appears that larger banking networks are safer compared

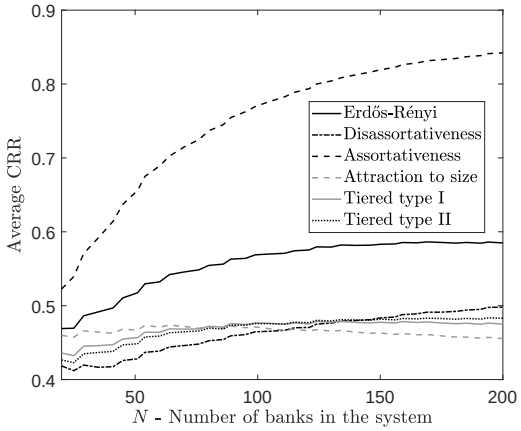


Figure 2.11: Sensitivity of the average CRR to the size of the system for a more interconnected system with a higher level of indirect risk in the system.

to smaller ones. However, when adjusting the network parameters to represent a more interconnected system with higher indirect risk parameters, this is no longer the case. In this case, increasing the network size leads to a larger number of links to be added to the system. This shows why increasing stability in banking networks cannot follow a 'one size fits all' approach, and that network characteristics should be considered as a package instead of in isolation. Therefore it is very important for regulators to take account of the network structure together with liquidity losses, market sentiment and all other network characteristics when making decisions.

Now that the model has been considered on a theoretical level, it is of interest to apply it to real-world data. This will firstly show how it can be used by regulators and help determine it usefulness in real world applications. This is done in section 2.4 below.

2.4 Application to the South African system

This section applies the methodology presented in section 2.1 to the South African banking system using standardised bank balance sheet data. The methodology presented in section 2.1 is adapted for the South African banking environment. The relevant justifications and details for this are presented in section 2.4.1 below. The data used for the application is described in section 2.4.2, where the construction of the simplified bank balance sheets is described as well. The adjusted modelling procedure is described in more detail in section 2.4.3, and the results are presented in section 2.4.4. Additional analyses are presented in section 2.4.5. The work presented in this section has been accepted for publication in the

South African Actuarial Journal. The pre-print can be viewed on arXiv.org[5].

2.4.1 Overview and adjustments to model

Systemic risk and the spread of financial contagion are important considerations for regulators tasked with overseeing stability of banking systems. While the literature on systemic risk has become vast over the past decade, the complex nature of banking systems remains difficult to replicate and model precisely. It is therefore of interest to find simplified models capable of detecting changes in systemic risk. We contribute to this by showing that network models of systemic risk can satisfy this requirement to a large extent. We illustrate how such a model can be used, by applying it to real-world balance sheet data and showing that changes in risk are detected under times of market stress for various network structures.

It is common for network models of systemic risk to have the edges between the nodes represent interbank assets and liabilities. Such models assume that whenever a bank in the system defaults, it cannot honour its commitments to its creditors, and hence defaults on its interbank commitments. We note that in the South African banking environment, the hierarchy of interbank loans compared to other unsecured debt is not well-defined. Therefore, in the event of a bank's default, other unsecured liabilities may be subjected to bail-in before interbank liabilities. That is why we take a different approach in this study, and do not model the interbank lending relationships. Instead, we consider contagion mechanisms applicable to any jurisdiction, namely how a loss of trust in the market may spill over to other banks, creating uncertainty and a difficulty to roll forward short-term debt.

When one or more banks experience solvency or liquidity problems, other banks may be reluctant to provide liquid funding to one another (similar to the approach taken in section 2.1). This hoarding of liquid instruments in the market, coupled with any shift in the yield curve as a result of distress selling shrink the balance sheets of banks in the system. This is still explicitly accounted for in this application by again including the factors representing market liquidity risk for short-, medium- and long-term assets. As before, short-term assets are defined to have a maturity of less than a month, and medium-term assets a maturity of more than a month and less than a year. Assets with a maturity of more than a year are deemed long-term. The simplified balance sheet used for this application is illustrated in table 2.4.

The way that banks are connected to one another via links (edges) in the network is

[5] http://arxiv.org/abs/1811.04223

Table 2.4: Illustration of a simplified balance sheet used for applying the simulation model of section 2.1 to the South African banking system.

Assets	**Equity & Liabilities**
Short-term assets	CET1 Capital
Medium-term assets	Other equity & liabilities
Long-term assets	

still referred to as the structure of the network. One difficulty arising from the way that this application is approached is the specification of the links between banks, since it is not possible to determine the paths through which losses will spread. Empirical network studies that focus on interbank lending as the direct contagion channel can use maximum entropy estimation techniques to estimate connections between banks [94]. However, this estimation method can lead to inaccuracies when assessing systemic risk [77, 11] and makes use of each bank's total interbank assets and loans. It is therefore not appropriate to use this here.

The process of deciding which banks to connect to one another should be dependent consider the definition of the edges in the network. For this to make sense, the formation of edges should ideally be consistent with the event that initiated the contagion in the first place. For example, banks that are heavily exposed to the mining sector may experience a loss of investor sentiment following the default of a bank with mine workers as its target market if the default is a result of sudden mine closures. From a modelling perspective, it is impractical to try to account for all possible contagion events, seeing that this would need to cover a wide range of risks such as market and operational risks. For this reason, we consider the same range network structures as in section 2.3. The details of the adjusted modelling procedure is discussed in section 2.4.3.

From previous research [51, 54, 70, 83] and the results of section 2.3, we note that different network structures may exhibit different levels of risk and the effect of changes to network characteristics is dependent on the chosen structure. As the true network structure is unknown, it is important to investigate how changes in the structure can affect the results of a real-world application such as this.

The results of this section are used to assess whether the model can be used to monitor systemic risk levels by capturing increases in systemic risk during stressed market conditions. This is done by considering systemic risk at different points in time during which incidents occurred that adversely affected the local economy. To summarise, this section aims to make the following contributions:

1. A top-down network approach is used to model systemic risk over time in the South African banking system. We investigate whether this model is capable of monitoring systemic risk by detecting instances of market turmoil.

2. A novel contagion mechanism is introduced that focuses on market sentiment.

3. The effect of the network structure on the results is investigated. Since the actual network structure is unknown, it is important to obtain insight into how sensitive the results are to the choice of network structure.

The remainder of section 2.4 is structured as follows: Section 2.4.2 describes the data and explains how the data is used to construct the balance sheets. Section 2.4.3 discusses the adjusted modelling procedure, and section 2.4.4 presents the results obtained by applying the model to the South African system. Finally, section considers additional analyses that can be performed by means of the modelling framework presented here.

2.4.2 Data and balance sheet construction

In this section, we describe the data and the construction of the balance sheets used for the application. Standardised monthly bank balance sheet data[6] of South African banks are used from April 2015 to March 2017. The BA900 returns are not granular enough to allow the extraction of CET1 capital data, which was instead obtained from banks' annual statements, Pillar III capital disclosures and the Orbis Bank Focus database[7]. It is not sensible to use a period dating back further since the capital data becomes too scarce. Capital data before 2015 is difficult to obtain for all banks since numerous small banks either did not exist, or their capital data for earlier periods is simply not available in the public domain.

As at March 2017 there were ten locally controlled banks, three mutual banks, six foreign controlled banks and fifteen branches of foreign banks, making up a total of 34 registered banks. For the purpose of this investigation, we do not consider the parent companies of the foreign branches. Firstly, subsidiaries may not be supported by the parent company. Secondly, while the assumption may under- or overestimate systemic risk in the local banking sector if subsidiaries were supported by the parent (depending on the solvency position of the parent company), it is necessary in order to keep the system closed.

[6] https://www.resbank.co.za/RegulationAndSupervision/BankSupervision/Banking%20sector%20ata/Pages/Banks-BA900-Returns.aspx., accessed 2017/07/31

[7] https://orbisbanks.bvdinfo.com/version-2017713/home.serv?product=OrbisBanks, accessed2017/07/31

In other words, to ensure that risk levels within the system are not influenced by external market players, any actions that they may take or any regulations that may apply to them.

This banking system can be considered as a typical candidate for a core-peripheral structure, as it consists of five large, 'core' banks and 29 smaller banks. To illustrate this, the total asset values of the banks are shown graphically in appendix B.1, figure B.1. For this study, eight banks are excluded because of a lack of capital data (see section 2.4.2 below for more detail), leaving 26 that are included in the analysis. The process for composing the simplified balance sheets as illustrated in table 2.4 is explained below.

On the asset side, items are categorised according to whether they have a short, medium or long time to maturity at inception of the asset. Recall that short-term assets have a maturity of less than a month, medium-term assets have a maturity of more than a month and less than a year, and long-term assets have a maturity of more than a year. Assets that do not have a contractual maturity date are categorised according to their expected holding period, for example remittances in transit which are categorised as medium-term. Not all balance sheet items fall distinctly into only one category. Most of these items are placed into the category in which most individual assets are expected to fall. For example, the local Treasury Bills can have maturities ranging from one day to twelve months, but normally have an unexpired maturity of 91 days or 182 days. Therefore, these are categorised as medium-term assets for our purpose. There are two exceptions to this rule:

1. Marketable government stock on the BA900 forms (line item 198) is only given with a maturity of up to three years, and a maturity of over three years. Marketable government stock with a maturity of over three years are included in the long-term asset category. Marketable government bonds with a maturity of up to three years are assumed to be evenly distributed across short-, medium- and long-term assets as all three of these maturity categories are included in this line item.

2. Derivatives with non-banking counterparties are divided according to term on the liability side of the BA900 forms, but not on the asset side. The assumption is therefore made that on the asset side of each bank, the proportion of short-term non-bank derivatives to total non-bank derivatives is the same as on the liability side. The same assumption holds for the medium- and long-term derivative instruments with non-banking counterparties. If there are no derivatives on the liability side, the derivatives on the asset side are divided equally among the short-, medium- and long-term assets.

Derivative exposures constitute an important source of systemic risk because increased

margining requirements during stress scenarios can place excessive strain on banks' liquidity positions. While this can be modelled using a network approach (see e.g. [73]), we do not explicitly model these exposures, but include such effects indirectly via the trust mechanism. This is because counterparty relationships are not publicly available, and it avoids complicating a model that is meant to remain simple.

Credit impairments with respect to loans and advances are deducted from the medium-term assets. This is because private sector loans and advances (that are categorised as medium-term assets) generally make up a large portion of total loans and advances and should also contain the majority of impaired accounts. Any impairments in respect of investments are deducted from the long-term assets, since investments are generally regarded as long-term assets. The categorisation of assets is illustrated in appendix B.2, table B.1.

Note that with more granular data the categorisation of assets according to maturity can be done more precisely. In this case it is necessary to aggregate the balance sheet items at this level since the available detail does not allow for a finer categorisation according to term. Regulators with more detailed information could use a larger number of categories so that assets can be grouped according to more time horizons and other characteristics as well.

Banks' CET1 capital represents the capital part of the balance sheet, similar to the approach used in section 2.1. The equity side of the BA900 balance sheets is not sufficiently granular to allow for calculation of the banks' CET1 capital. For this reason, data from financial statements, published Pillar III capital disclosures and Orbis Bank Focus is used to supplement the primary balance sheet information. However, the data obtained via these sources are quarterly at best (in some cases only annually) and not all banks publish these on the same dates. Furthermore, some banks publish only risk weighted CET1 ratios and do not necessarily include a monetary amount for this type of capital. The available data for this part of the balance sheet must therefore be used to estimate the missing data points where possible.

To estimate a bank's monthly CET1 capital, the available CET1 amounts are divided by the corresponding total asset values of the respective bank at the available points in time. This gives an unweighted ratio of CET1 to total assets at selected points in time. There are two main reasons for using this ratio. It firstly strips out any inflationary effects over time and secondly removes the effect of significant increases or decreases in banks' growth rates. Where available, the unweighted ratio of CET1 to total assets are very stable for all banks over the period considered (the maximum variance for this ratio for over all banks is 0.00344). Therefore, for most banks the missing unweighted CET1 ratios could easily be estimated.

Table B.2 in appendix B.2 shows all the unweighted CET1 ratios for registered local banks that could be obtained from the available data. For each bank that has at least three CET1 data points available between May 2017 and February 2015, the remaining ratios are estimated for the outstanding months. Banks that have less than three CET1 data points are excluded from the analysis, reducing the total number of banks from 34 to 26. The total assets of all excluded banks make up less than 3% of all banks' assets as at May 2017.

For the remaining banks, the available unweighted CET1 capital ratios are used to estimate the unknown CET1 ratios as follows:

- Where missing data points fall in-between two known data points, linear interpolation between the two known data points is used to estimate the missing values. For example, if the ratios C_τ and $C_{\tau+3}$ are available for months τ and $\tau+3$, but ratios for months $\tau+1$ and $\tau+2$ are not, we use the estimates $\hat{C}_{\tau+k} = C_\tau + \frac{k(C_{\tau+3}-C_\tau)}{3}$ for $k = 1, 2$.
- Where a missing data point does not lie between two known data points, the average unweighted CET1 ratio for the associated bank is taken. For example, if no CET1 data is available for month $t = 1$, then $\hat{C}_1 = \frac{1}{m}\sum_\tau C_\tau$, where m is the number of months τ for which C_τ is available and the sum is taken over all available ratios C_τ.

This simple method of choosing the estimates is chosen for two reasons. Firstly, the CET1 ratios were very stable as has been noted above. Therefore, there is no need for more sophisticated techniques or for using more data points in calculating an estimate. Secondly, slight trends in the data are accounted for by using linear interpolation as described in the first point above. Once the estimates $\hat{C}_\tau$ for the CET1 capital are determined, all the required balance sheet entries are known. The next step is then to specify the interactions between the banks that are represented by the edges, where different assumptions regarding these interactions lead to different network structures.

2.4.3 Adjusted modelling procedure

Suppose the system consists of N banks, where each bank i's total assets is denoted by a_i. The short-, medium- and long-term assets of a bank i are again denoted by $a_i^{(s)}$, $a_i^{(m)}$ and $a_i^{(l)}$ respectively. The only difference between these and the asset variables specified in section 2.1.2 is that we now include direct interbank exposures in the asset categories. Bank i's CET1 capital is again denoted by c_i. For ease of reference, the terms CET1 capital and capital will once again be used interchangeably.

As before, we choose an initial bank n and suppose that it suffers an initial loss. In this event bank n loses a fraction, say s, of its assets. In section 2.1, a bank is deemed undercapi-

talised if the loss is greater than 3% of the bank's assets, i.e. if $s \cdot \left(a_n^{(s)} + a_n^{(m)} + a_n^{(l)}\right) > 0.03 a_n$. This threshold of 3% is based on regulations by the FDIC. It is not necessarily appropriate for the South African environment, as it is aimed at U.S. institutions. Furthermore, the CET1 ratios differ greatly between South African banks (see table B.2 in appendix B.1) and it is therefore difficult to assign a sensible threshold for our purposes. South Africa does not have such a published threshold because the South African banking regulations follow the Basel requirements closely, and the BIS has not published such ratios. In South Africa, it is determined on a case-by-case basis whether corrective action should be taken against a bank. Therefore for our purposes it is more appropriate to use a threshold that is expressed in terms of the total amount of capital that is lost. For the purpose of this illustration, we assume that a bank is deemed undercapitalised whenever it loses 80% or more of its CET1 capital. If $s \cdot \left(a_n^{(s)} + a_n^{(m)} + a_n^{(l)}\right) > 0.8 c_n$, then it is assumed to have failed as in section 2.1.2.

As mentioned in section 2.4.1, the hierarchy of interbank loans compared to other unsecured debt is not well-defined in the South African banking environment.In some situations the regulator may require other banks to assist with capitalisation in order to limit the spread of losses to other parts of the economy by making whole the retail and institutional creditors' unsecured loans. We assume that a proportion, say u, of this shortfall must be covered by the remaining banks. The remaining proportion $1 - u$ is absorbed by the Total Loss Absorbing Capacity (TLAC) part of the troubled bank's balance sheet, after which unsecured creditors bear the loss. The resulting funding requirement is spread over all banks in the system in proportion to their assets sizes. In other words, if bank n experiences an initial loss event, its capital is reduced by $S_n = s \cdot a_n$. If $S_n > 0.8 c_n$, then n defaults and each bank i suffers a loss of

$$L_{i,n}^{(1)} = u\,(S_n - c_n)\,\frac{a_i}{\sum_{k=1}^{n} a_k}.$$

As before, we include losses due to raised provisions and mark-to-market effects. Each remaining bank i in the system suffers a loss of

$$L_i^{(2)} = \sum_{\eta \in \{s,m,l\}} a_i^{(\eta)} \left[1 - \exp\left(-g^{(\eta)}\right)\right].$$

Finally, we include losses due to a deterioration in market sentiment. The perceived exposure of other banks to the problems faced by bank n determines the edges in the network. An edge starting at a bank i and pointing towards bank n means that the market believes i may be exposed to similar difficulties as n, or may be adversely affected by the

default of n. Recall that the shortest distance between two nodes is the smallest number of edges that can be used to travel from the one to the other. The shortest distance $d_{i,n}$ in the network from any bank i to the failing bank n determines the degree to which i is affected by a loss of trust. Small values of $d_{i,n}$ indicate 'closeness' in the network, which represents a perceived tendency for a bank i to experience similar problems as n.

In order to reflect the shrinkage of each bank i's balance sheet, each asset class of i is reduced by a factor $\exp\left(-\frac{\delta}{d_{i,n}}\right)$. Therefore, each remaining bank i in the system experiences a further loss of

$$L_{i,n}^{(3)} = \sum_{\eta\in\{s,m,l\}} a_i^{(\eta)}\left[1 - \exp\left(-\frac{\delta}{d_{i,n}}\right)\right], \tag{2.20}$$

where δ is the associated reduction factor. This means that the network structure is not used to determine direct losses any more. Instead, the network structure is only used to model the spread of indirect losses due to a lack of market trust similar to example 4, page 29. Note that different proximity factors should be assigned to different types of assets when applying this model in practice. However, we avoid introducing too many parameters for the purpose of illustration by using the same parameter for all asset classes.

As before, this type of loss is called a proximity shock. The way that proximity shocks are modelled in the network accounts for the fact that some banks will experience a worse loss of confidence than others. From equation (2.20) it is seen that banks with smaller shortest distances to the failing bank will experience worse losses than those with greater shortest distances. This is illustrated in figure 2.12, where the failing bank is indicated by the cross. The darker nodes experience greater losses than the lighter node, since they have a smaller shortest distance to the failing bank. The edges in the network is directed to take account of a wide range of possibilities without overcomplicating the model. In some circumstances the default of one bank may lead to distrust in another bank, but not the other way around. For example, the default of a large, systemically important bank may affect small banks, but the default of a small bank may not necessarily affect the financial positions of much larger banks. Note that the direction of the edges does not represent the direction in which the losses spread but rather represents the similarity between banks. For example, if a directed edge exists from bank i to bank j, the interpretation is that bank i is similar to bank j in the sense that the market perceives i to be exposed to similar difficulties as bank j in the event of bank j's default.

For simplicity we ignore the effect of different business models on the liquidity shocks in the system. In other words, we do not adjust any of the formulae for $L_i^{(2)}$ or $L_{i,n}^{(3)}$ to account for differences in business models as has been done in 2.1.2. Therefore the total loss to each remaining bank i following the default of bank n is given by $L_{i,n} = L_{i,n}^{(1)} + L_i^{(2)} + L_{i,n}^{(3)}$.

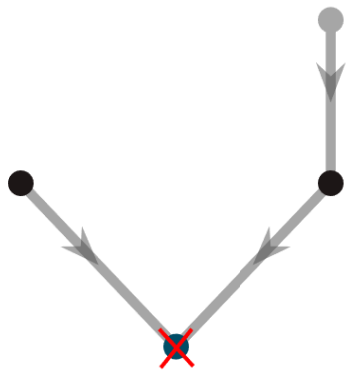


Figure 2.12: Illustration of a proximity shock following the default of a bank. The bank marked with a cross represents the failing bank. The black nodes represent the banks closest to it, and these will receive the largest proximity shock.

2.4.4 Results for the South African system

Recall that the network structure is determined by the way that banks are connected to one another in the network. We construct a network of trust deterioration, where the edges represent paths through which market trust is lost in the system. As it is not possible to know beforehand which banks will be perceived as being affected by another bank's failure, the edges in the network are again assumed to be random and the same range of network structures are considered as before. Even though some structures may be unrealistic, it is of interest to include them to consider a wider range of outcomes. This allows for a better understanding of the relevance of network structure in a network model based on trust deterioration.

The different network structures discussed in section 2.2.1 are compared to one another. Here, the structures have different interpretations as before. One would expect the default of large banks to have a more significant effect on the confidence in the banking system compared to smaller banks. Therefore the attraction to size is likely to be more realistic in this application compared to most of the other structures (figure 2.1 on page 36 contains illustrations of the structures). This is because the large banks have higher probabilities of having incoming edges connected to them compared to small banks. This can be interpreted as follows: whenever a small bank fails, the probability that it will affect the perceived financial strength of other banks. On the other hand, the failure of a large bank will lead to a loss of market trust for any other bank with a higher probability. Here,

the market assumed before the shock event that the bigger banks were the most financially sound. The failure of a big bank therefore causes widespread panic, affecting most other banks in the system.

For consistency with the previous sections, we include graphs for the average shortest path, global clustering coefficient and the tiering error in appendix B.3. The weights for the global clustering coefficient have to be adjusted to account for the fact that we no longer use interbank exposures. Instead, the weight of an edge from i to j is taken as the asset value of bank i, since i's loss in the event of j's default is proportional to the asset value of i. From this, it can be seen that if $\bar{p}$ is 0.15, then the attraction to size, tiered type If and tiered type II structures theoretically result in tiered networks. It is noted that since the meaning of the network differs from before, it is not necessarily required that these networks be tiered. However, for consistency we adopt the parameter $\bar{p} = 0.15$ in our low interconnectedness scenarios below.

The combined effect of network structure, the system's interconnectedness and the consequences of liquidity shortages and a deterioration of market sentiment on systemic risk is investigated. For each case considered, an initial shock of 40% is applied to the system. In other words, the bank that suffers the initial loss experiences a loss equal to 40% of its total asset value. Whenever a loss causes a bank's capital to be depleted, it is assumed that a proportion $u = 0.3$ of the capital shortfall must be absorbed by the banking system.

We illustrate how the systemic risk changes over time by calculating a point in time measure at each month during the investigation period. At each time interval, four scenarios are considered regarding the interconnectedness and the effect of indirect risk factors on systemic risk:

- Low indirect risk parameters ($g^{(s)} = 0.005$, $g^{(m)} = 0.0075$, $g^{(l)} = 0.01$ and $\delta = 0.0075$) and a low level of interconnectedness ($\bar{p} = 0.15$).
- Low indirect risk parameters ($g^{(s)} = 0.005$, $g^{(m)} = 0.0075$, $g^{(l)} = 0.01$ and $\delta = 0.0075$) with a moderate level of interconnectedness[8] ($\bar{p} = 0.5$).
- High indirect loss parameters ($g^{(s)} = 0.015$, $g^{(m)} = 0.0225$, $g^{(l)} = 0.03$ and $\delta = 0.0225$) with a low level of interconnectedness ($\bar{p} = 0.15$).
- High indirect risk parameters ($g^{(s)} = 0.015$, $g^{(m)} = 0.0225$, $g^{(l)} = 0.03$ and $\delta = 0.0225$) and a moderate level of interconnectedness ($\bar{p} = 0.5$).

[8]Higher levels of interconnectedness are not considered for this study, as each network structure's unique characteristics become lost when the system is too interconnected.

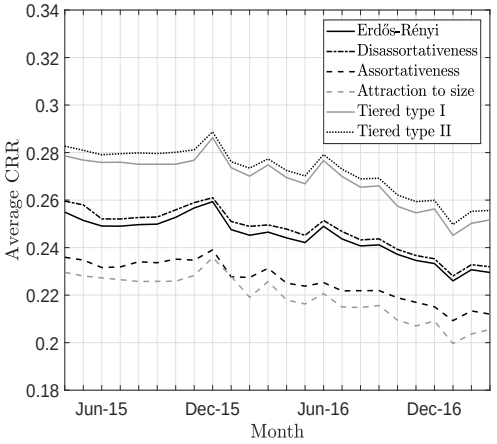
The values for $g^{(s)}$, $g^{(m)}$, $g^{(l)}$ and δ for the high indirect risk scenarios are scaled versions of the corresponding low indirect risk parameters. It is important to note that one may obtain different results with different combinations of these parameters. However, it is impractical to consider an arbitrary number of combinations without more information regarding realistic values. Therefore only a few combinations are considered in this study. It is noted that this way of modelling the risk over time assumes that all parameters remain constant over time. While this is not necessarily the case in practice, it is a reasonable assumption to make in the absence of additional information.

For the low risk scenario, the chosen parameter values imply that after each default, the short-term assets of banks are decreased by $1 - \exp\{-0.005\} = 0.5\%$, the medium-term assets by 0.75% and the long-term assets by 1%. The proximity shock parameter reduces all asset values of banks with a shortest distance of one to the failing bank by 0.75%. Banks with a shortest distance of two to the failing bank have their assets reduced by 0.37%. For the high risk scenarios, the short-term assets are reduced by 1.49%, the medium-term assets by 2.22% and the long-term assets by 2.96%. The proximity shocks decrease all assets of banks directly connected to the failing bank by 2.22%.

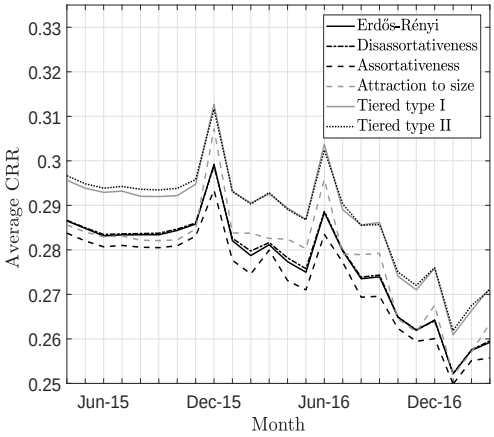
Figure 2.13 shows the relative levels of systemic risk over time for all network structures considered in section 2.2. It is noted that the CRR averages shown by the graphs are based on hypothetical values of the parameters associated with indirect risks, and are not necessarily accurate. This is because the focus of this study is on the relative risk levels associated with different network structures, and not to calculate actual proportions for these events.

As expected, higher levels of interconnectedness result in lower levels of discrimination between the different structures. This is because the higher value of $\bar{p}$ pushes the probabilities $p(i, j)$ towards one for all structures start and hence they become more representative of fully connected systems. The levels of risk over time generally show a slight increase in risk when the interconnectedness is increased for low indirect risk parameters (see figures 2.13a and 2.13b). In that case, the additional connections serve as additional channels of contagion and increase the risk. The spikes and dips seen in figure 2.13a are accentuated when interconnectedness is increased, although the overall shapes of the graph are preserved.

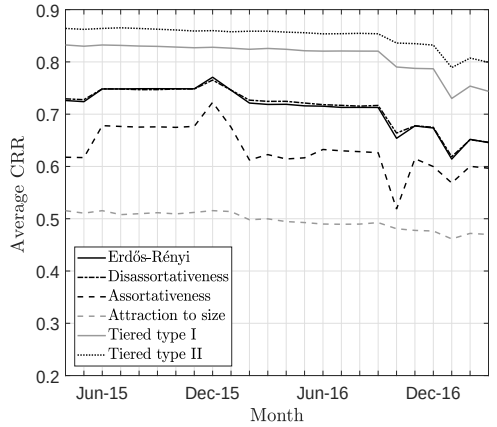
When indirect risk is high (see figures 2.13c and 2.13d), increased interconnectedness increases systemic risk for some structures and decreases it for others. Once again, increasing the interconnectedness results in accentuating the dips in systemic risk levels. Here, the additional connections serve to stabilise the system, leading to greater reductions in systemic risk. The spikes in systemic risk are not accentuated as much, which is likely due



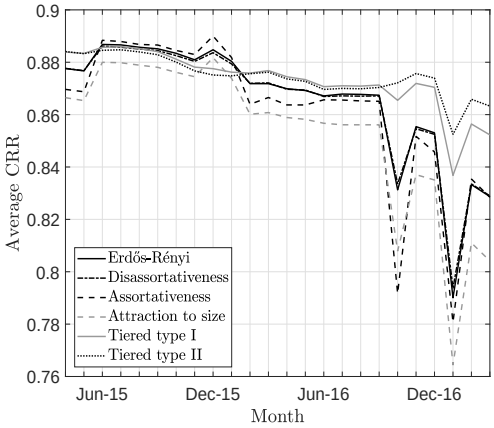
(a) Low indirect risk, low interconnectedness



(b) Low indirect risk, moderate interconnectedness



(c) High indirect risk, low interconnectedness



(d) High indirect risk, moderate interconnectedness

Figure 2.13: Comparing the CRR resulting from different network structures over time, under the assumption of no regulatory intervention.

to the fact that the risk levels are already close to their upper bound of one. This shows that the structure of the network and the level of indirect risk influences which events lead to an increase in systemic risk.

From all four scenarios, it is seen that there is a spike in systemic risk around December 2015. This corresponds to the month during which former South African finance Minister Nhlanhla Nene was replaced, which was an unexpected and controversial political event in South Africa. The local financial market reacted negatively, and the local currency depreciated significantly during that period. Only the tiered type I and II structures do not

exhibit this behaviour in figures 2.13c and 2.13d where the indirect risk parameters are high.

A second spike in systemic risk is observed around June 2016 for figures 2.13a and 2.13b (where lower risk parameters are used). This was also a time during which the Rand depreciated steeply against the US Dollar. This was due a combination of factors, namely a weak economic growth outlook, rumours that the former finance Minister was to be arrested and an approaching credit review by Standard and Poor to decide whether they will downgrade South Africa's sovereign rating to junk status.

During March 2017, former finance Minister Pravin Gordhan was also replaced during another controversial political event. However, this time there were numerous media reports beforehand suggesting that it was likely to happen, and therefore the markets expected it. Since there are no prominent spikes in systemic risk during that time, it shows that either the banks made sufficient use of publicly available information to protect themselves, or that the model is not able to detect the increase in systemic risk.

The prominence of the December 2015 spike may be explained by looking at the average balance sheet items over time (see figure B.2 in appendix B.1). At December 2015, the relative increase in the average for the short-term and long-term assets are much greater compared to the CET1 capital. This could, on average, lead to relatively larger losses for the initially shocked bank (because this is defined as a proportion of assets) that need to be absorbed by the capital. However, the average asset values at June 2016 and March 2017 do not show the same extreme behaviour. This could explain why the June 2016 spike is slightly less prominent in figure 2.13a and absent in figures 2.13c and 2.13d. This may also explain why we do not see a significant spike at March 2017 for any of the figures.

In general, it appears that the importance of the network structure is to a large extent influenced by the interconnectedness of the system and the values chosen for the risk parameters. For example, for moderate interconnectedness and high indirect risk parameters (figure 2.13d) at months when systemic risk levels are high, the network structures exhibit small differences. On the other hand, for low interconnectedness and high indirect risk parameters, there are significant differences between the risk exhibited by some of the structures. For low interconnectedness (figures 2.13a and 2.13c), the network structure generally plays the greatest role in the level of systemic risk. An exception to this is figure 2.13d, where large differences can be observed at October 2016 and December 2016.

In most cases the attraction to size structure exhibits the least risk, and the tiered type II structure the most risk. However, this is not always the case, since the assortative structure exhibits the least risk in figure 2.13b and the most risk during June 2015 to December 2015 in figure 2.13d. The differences between the risk levels for structures that are tiered when

$\bar{p} = 0.15$ (i.e. attraction to size, tiered type I and tiered type II) are noteworthy. While the attraction to size structure exhibits significantly less risk compared to a completely random (Erdős-Rényi) network, the opposite is true for the tiered type I and II networks. This shows that even for fixed levels of connectivity and indirect risk, different definitions of tiered structures can lead to different conclusions.

From all four scenarios it is seen that the structures are mostly consistent regarding directional changes, i.e. the structures' risk levels move in the same direction at each time step, albeit at different rates. Furthermore, for a given set of parameters the ranking of the structures in terms of risk remain relatively consistent over time (with the exception of June 2015 to December 2015 in figure 2.13d).

The above results show that the indirect risk parameters have a significant effect on the behaviour of network structure. To further illustrate this point, we consider the effect of changing the relative values of the indirect risk parameters. A base parameter value of 0.0075 is used for all indirect risk parameters. These parameter values will be referred to as the base parameters for the remainder of the section. The resultant graph of systemic risk over time is shown in figure 2.14.

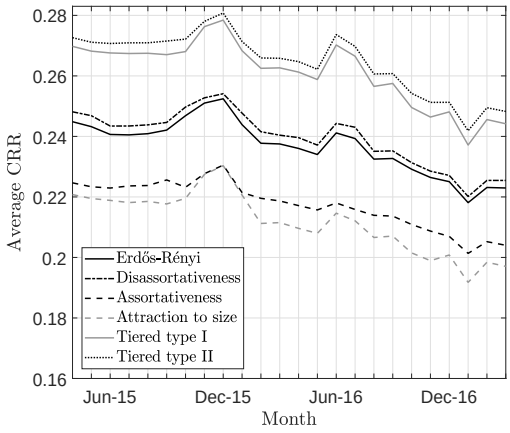
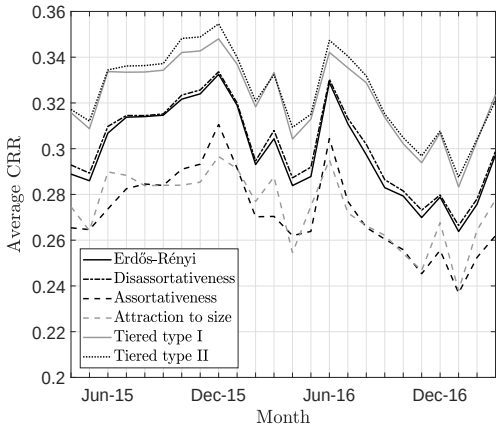
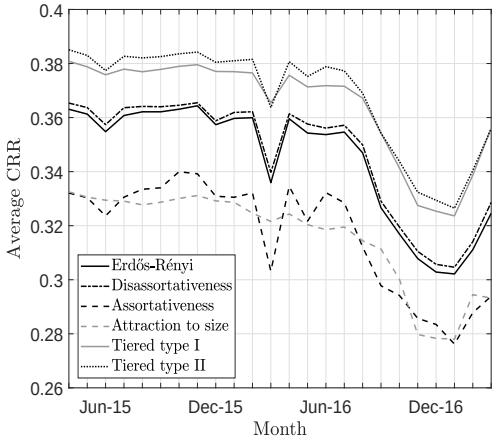


Figure 2.14: Systemic risk over time for base parameter values of $g^{(s)} = g^{(m)} = g^{(l)} = \delta = 0.0075$.

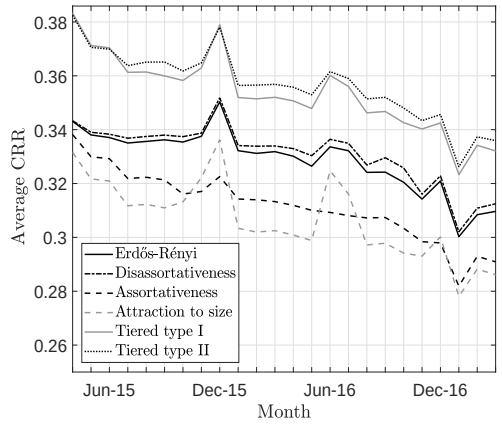
The effect of increasing any one risk parameter to $0.0075 \cdot 3 = 0.0225$ is considered. The level of interconnectedness is kept at 0.15, since it was seen from figure 2.13 that differences between the structures are accentuated when connectivity is low. Figure 2.15a shows what happens if only the parameter associated with the short-term liquidity losses is increased from 0.0075 to 0.0225. Figures 2.15b to 2.15d show the same results for the medium-term, long-term and proximity shock parameters, respectively.



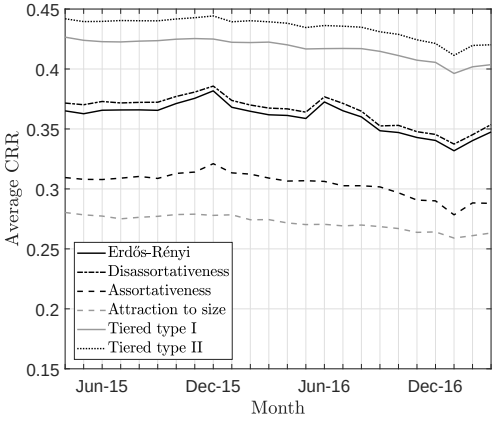
(a) Systemic risk over time for base parameters, but with $g^{(s)} = 0.0225$



(b) Systemic risk over time for base parameters, but with $g^{(m)} = 0.0225$



(c) Systemic risk over time for base parameters, but with $g^{(l)} = 0.0225$



(d) Systemic risk over time for base parameters, but with $\delta = 0.0225$

Figure 2.15: Comparing different network structures over time and increasing each indirect risk parameter in turn.

From figure 2.15 it is seen that the parameters don't have the same effect on systemic risk. By increasing only the short-term liquidity parameter in figure 2.15a, the peaks in systemic risk are more pronounced than in figure 2.14 for the base parameters. Differences between the network structures remain more or less the same. Compared to figures 2.15b to 2.15d, the general level of risk is lower when only the short-term liquidity parameter is increased.

Increasing only the medium-term liquidity parameter (figure 2.15b) flattens out the graph to such an extent that the December 2015 and June 2016 spikes are no longer prominently visible. Only the December 2016 decrease in risk is preserved and accentuated.

Once again, the differences between the network structures remain more or less the same when compared to figure 2.14.

When only the long-term liquidity risk parameter is increased in figure 2.15c, the general shape of graph remains largely unaltered. Exceptions to this is that for figure 2.15c, the attraction to size structure has much more pronounced peaks at December 2015 and June 2016, and the peaks at December 2015 are slightly more pronounced for the remaining structures.

By increasing only the parameter associated with market sentiment, the differences between the risk levels become large compared to figures 2.15a to 2.15c. This may be because this parameter is directly related to the network structure, and increasing it accentuates differences between the structures. At the same time, increasing the parameter does not influence the ranking of the structures in terms of risk. The lines on the graph become significantly flatter compared to figure 2.14 for the base parameters. The December 2015 and June 2015 peaks are only easily identifiable for the Erdős-Rényi and disassortative structures.

The differences between figures 2.15a to 2.15c are likely because of differences in asset values for different maturities between banks and within each bank. This is because the three liquidity parameters enter the model in the same way via reductions in the associated asset values. Therefore, networks derived from different countries' banking systems will likely differ in the way that they react to changes in network structure and liquidity risk parameters. For regulators, it is important to note that conclusions reached for one banking system will not necessarily hold for another.

The results show that both network structure and indirect risk are important in determining the level of risk present in the system. Network structure can significantly affect the level of systemic risk. Determining parameters associated with interconnectedness, liquidity risk for different asset types and market sentiment is important for network models of systemic risk, since these can influence the degree to which market disturbances fuel systemic risk. It is important to stress that different observations may be made depending on the network model is applied. For example, figures B.6 to B.8 in appendix B.4 show how the figures of this section would look if the average fraction of defaulted banks were used as a risk measure. This measure is calculated as follows: Let α_n be the proportion of banks (including bank n) that have defaulted after bank n has received the initial shock. The average of these proportions over m simulations is denoted by $\bar{\alpha}_n$. The average defaulted fraction is then defined by $\bar{\alpha} := \frac{1}{N}\sum_{i=1}^{N} \bar{\alpha}_n$.

The implications of the results from this section are discussed further in chapter 4, where all the results from this thesis are discussed together.

2.4.5 Additional analyses

To understand how the network structure affects how the different banks in the system contribute to systemic risk, it is useful to consider how different risk measures vary with the asset value of the initially shocked bank n. It is reasonable to expect the default of larger banks to have a greater knock-on effect on the system compared to smaller banks and hence measures of risk are expected to be higher for larger banks. This is confirmed by the results of this section.

The objective here is to illustrate different ways in which this modelling framework can be used to analyse the spread of contagion in the network. This is done by analysing the following quantities:

- The CRR (Capital Reduction Ratio);
- The fraction of defaulted banks ($\bar{\alpha}_n$ as defined above); and
- the number of contagion rounds that the network experiences following a shock before no more defaults occur (either because all banks have failed or because the remaining banks have successfully absorbed all losses). For a description of contagion rounds, see the discussion of default cascades starting on page 29.

The following analyses were performed, the results of which are contained in the subsections below:

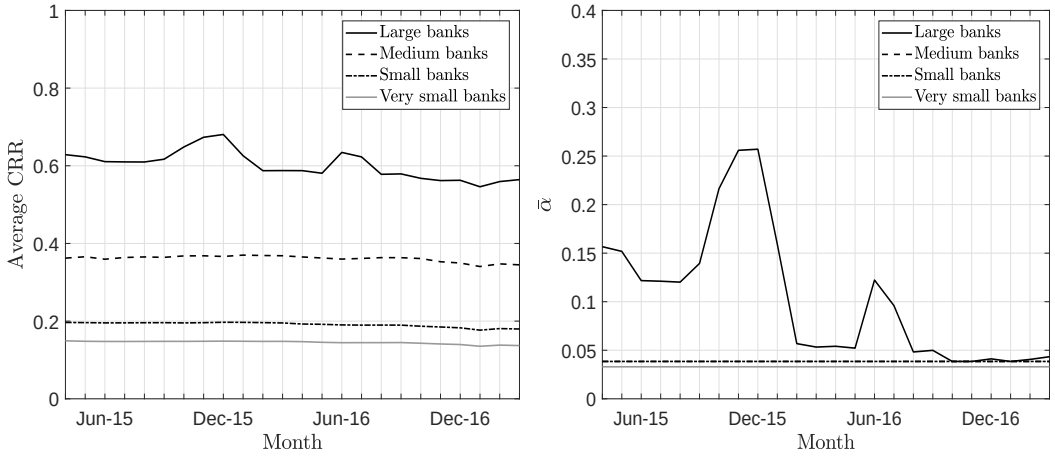
1. The banks were divided into four groups according to size, and within each group the average of each quantity mentioned above was calculated from April 2015 to March 2017. Figure B.1 in appendix B.1 was used to determine the groups. The four largest banks in the system are included in the 'large' group, the fifth largest bank in the 'medium' group, the sixth to thirteenth largest banks (Capitec Bank to African Bank) in the 'small' group, and the remainder in the 'very small' group.

2. For each initially shocked bank, the resulting quantity was compared against the asset value of that bank. This was done for December 2015 (where systemic risk levels are generally high), January 2017 (where systemic risk levels are generally low) and March 2017 (where systemic risk levels are generally at intermediate levels).

3. For December 2015, January 2017 and March 2017, the empirical distribution was determined for the average of each quantity, where the average is taken over all initially shocked banks. For example, if θ_n denotes the CRR after initially shocking bank n (see equation (2.5)), then we consider the distribution of $\bar{\theta} = \frac{1}{N}\sum_{n=1}^{N}\theta_n$ (see equation (2.6)).

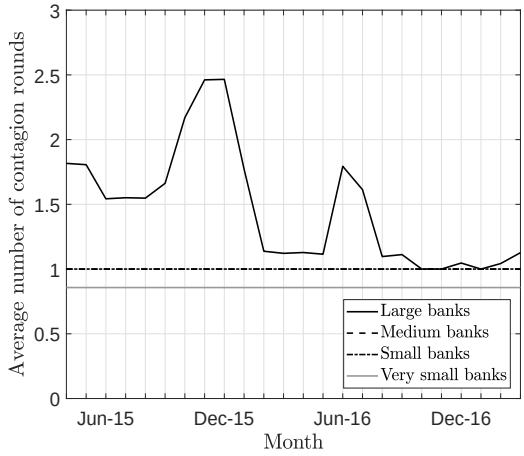
1. Risk quantities over time by size category

Figure 2.16 below considers the Erdős-Rényi structure and how banks of different sizes affect each of the three risk quantities listed above. The figures for the remaining structures display the same patterns, and hence those are included in appendix B.5.1. Figures B.9, B.10 and B.11 contain all structures' figures for the CRR, the average defaulted fraction and the number of default rounds respectively.



(a) Capital Reduction Ratio

(b) Average fraction of defaulted banks



(c) Average number of contagion rounds

Figure 2.16: Plotting three different risk quantities for the Erdős-Rényi structure over time for very small, small, medium and large banks.

Interestingly, all three structures exhibit increases and decreases in risk at the same points in time. This suggests that the modelling framework is relatively insensitive to changes in the risk measure as far as general trends over time are concerned. As expected,

the average of any risk quantity is highest when the initially shocked bank is large. From the three graphs it can be seen that the average of any risk quantity over time is more variable for large banks. This shows that the variation of risk observed over time can be attributed to the large banks for this choice of parameters. For higher values of the indirect risk parameters, it is possible that the smaller banks may also exhibit variation over time.

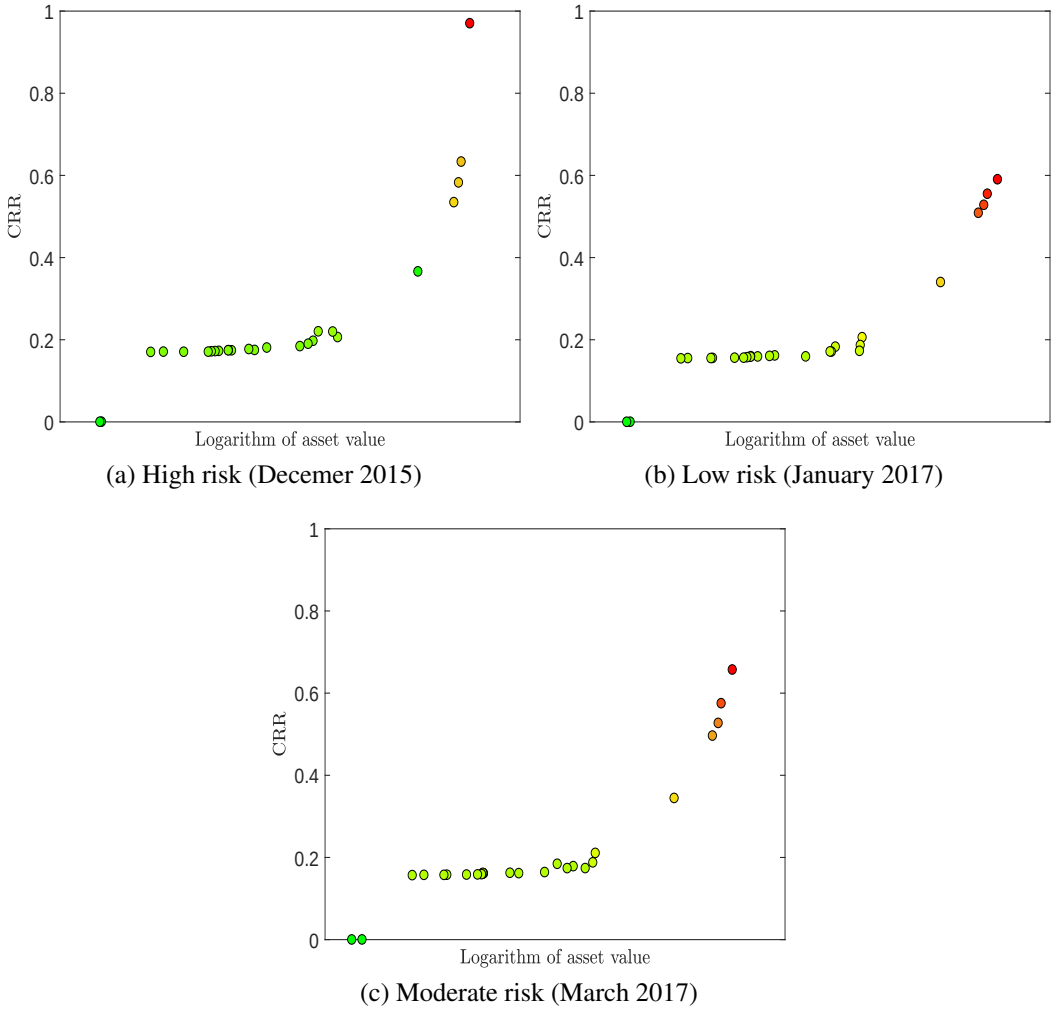
During times of relatively high systemic risk, the difference between the large banks and the rest of the banks are much more pronounced for the average fraction of defaulted banks (figure 2.16b) and the number of contagion rounds (figure 2.16c) than for the CRR (2.16a). At the same time, the differences between the medium, small and very small groups are less pronounced for the average fraction of defaulted banks and the number of contagion rounds than for the CRR. The CRR is the only one of the risk measures where the medium and small banks can be distinguished from one another on the graph. This is due to the fact that the CRR takes into account the amounts of the losses suffered in the system, and not just just the number of losses.

2. Risk quantities vs. asset values

Here, the three different risk quantities are compared to the logarithm of the initially shocked bank's asset value by means of scatter plots. Hence, the x axis in the figures represents the logarithm of the intially failed bank's asset value, ordered from smallest to largest. This is done for three different scenarios. The first scenario is one where systemic risk levels were high (December 2015). The second scenario considers a month when systemic risk levels were relatively low (January 2017). The third scenario is an intermediate one, which considers a month which had intermediate levels of systemic risk (March 2017). Only the figures for the Erdős-Rényi structure are included here, since the figures for the remaining structures follow similar trends. The figures for all structures are included in appendix B.5.2. Figure B.12 to B.14 contain all the graphs for the CRR measure, figures B.15 to B.17 contain the graphs for the fraction of defaults and figures B.18 to B.20 contain the graphs for the number of contagion rounds.

Figure 2.17 shows the Erdős-Rényi structure scatter plots for the CRR for the three scenarios under consideration. For all scenarios, the CRR increases along with the asset value of the intially shocked bank. For very small asset values, there is a sudden jump in the CRR after a certain point, after which the graphs show an exponentially increasing trend of the CRR with the logarithm of the initially shocked bank's asset value. The rate of increase observed in the graphs is higher in figure 2.17a (high risk scenario) than in figures 2.17b and 2.17c. This is consistent with 2.16, which showed varying levels of risk over

time for the large banks, but not for the other banks.



(a) High risk (Decemer 2015)

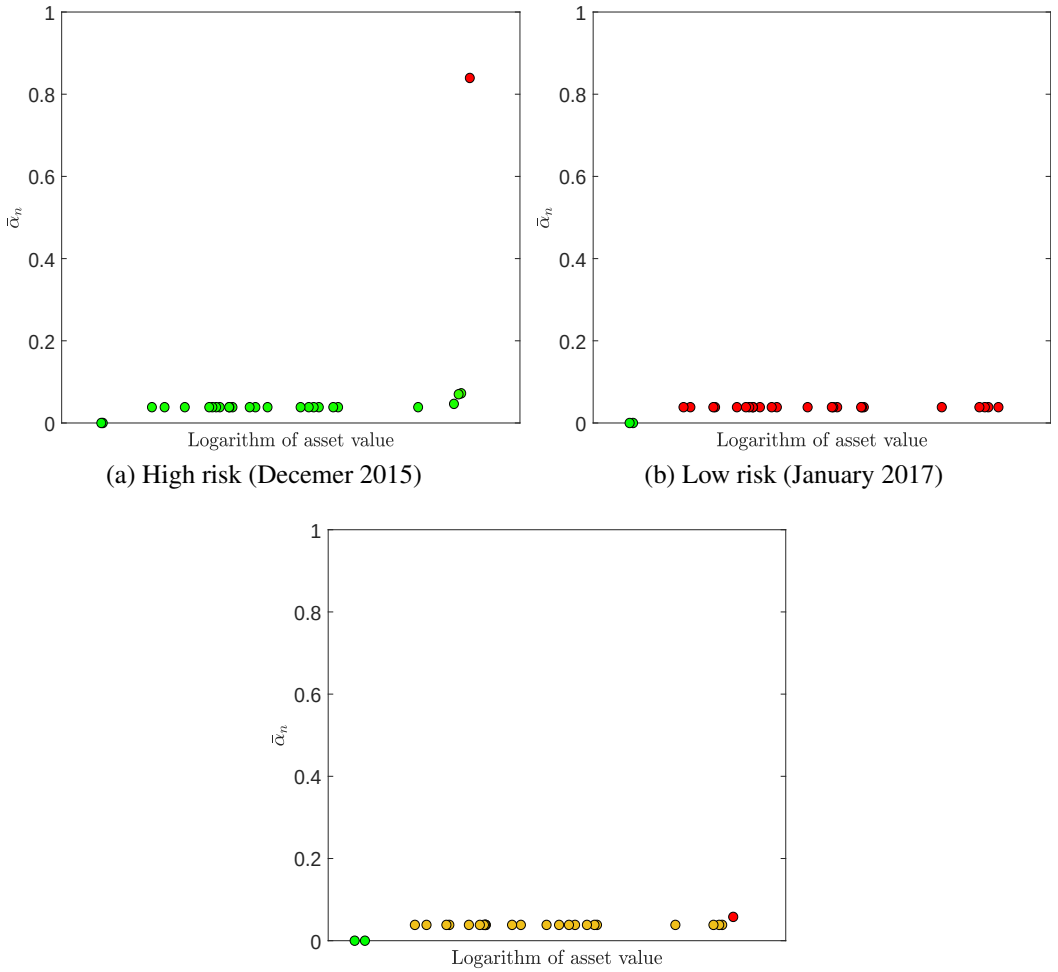
(b) Low risk (January 2017)

(c) Moderate risk (March 2017)

Figure 2.17: Comparing the CRR for each bank to its asset value for three different risk scenarios based on the Erdős-Rényi structure.

Figure 2.18 contains the scatter plots for the fraction of defaults for the Erdős-Rényi structure. Here, the graphs show a non-decreasing trend as opposed to a strictly increasing trend. The fraction of defaults is zero for the two smallest banks, and the default fraction is the same for the remaining small and medium banks. This is consistent with figure 2.16b, where the lines for the small and medium banks are indistinguishable from one another and the line representing the very small banks is just below. The default of the largest bank leads to a significant jump in the fraction of defaults in figure 2.18a for high levels of systemic risk. In figure 2.18c for moderate systemic risk levels, the jump in the fraction of defaults in less pronounced. For low systemic risk levels, there is no visible jump in

systemic risk. This is consistent with figure 2.16b, where the lines associated with the large, medium and small banks are indistinguishable at January 2017.



(a) High risk (Decemer 2015)

(b) Low risk (January 2017)

(c) Moderate risk (March 2017)

Figure 2.18: Comparing the average fraction of defaulted banks for each bank to its asset value for three different risk scenarios based on the Erdős-Rényi structure.

Since the defaulted fraction is zero for the two smallest banks, it means that the initial shock was not enough to make them default in the first place. The majority of banks have a defaulted fraction of 0.4 in figure 2.18. As there are 26 banks in the system, it means that contagion does not spread further in the majority of cases, since the system only experiences one default (the initially shocked bank). Only the largest of banks can result in more defaults when general systemic risk levels are higher and figure 2.18a shows that the largest bank can make more than 80% of the system fail on average when systemic risk levels are at their highest.

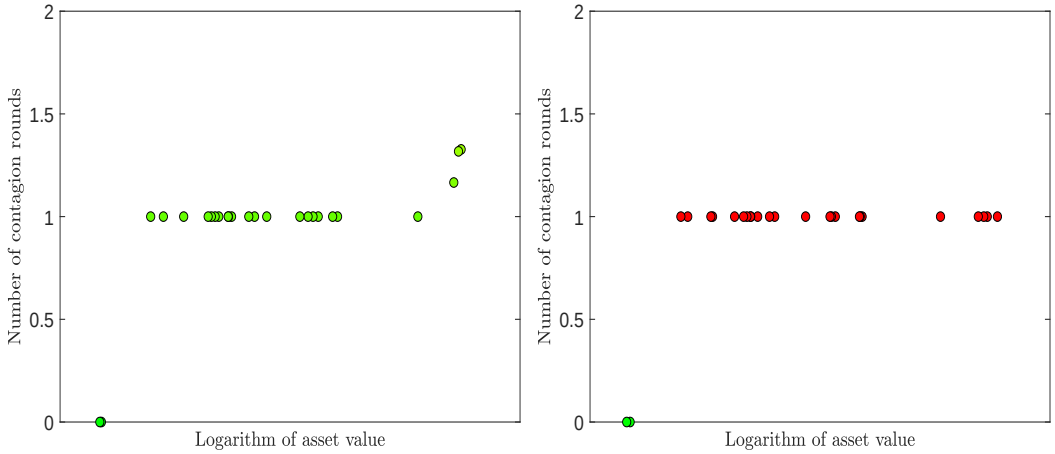
A comparison of figures 2.17a and 2.18a show that when analysing systemic risk, it is important to consider both the number of defaults and the severity of losses in the system. For example, consider the second largest bank in the system. The average defaulted fraction as a result of shocking this bank is smaller than 10%, where the average fraction of capital lost in the system is above 60%. This is largely due the capital lost by the bank itself, but shows that measures of systemic risk need to consider the severity of losses.

The scatter plots of the number of contagion rounds against the logarithm of the asset values for the Erdős-Rényi structure is contained in figure 2.19. As for the fraction of defaults, a non-increasing trend is seen in the graphs. This again illustrates that the CRR is better able to capture differences due to bank size than the defaulted fraction and the number of contagion rounds, since it accounts for the amount of losses and not just the number of losses.

It shows that for the smallest two banks, the average number of contagion rounds is zero regardless of the overall level of systemic risk in the system. As noted above, the initial shock does not result in the default of these two banks. For the majority of the banks, only one round of contagion takes place. This is also consistent with the above observation that only the initially shocked bank defaults in those cases, after which contagion does not spread further.

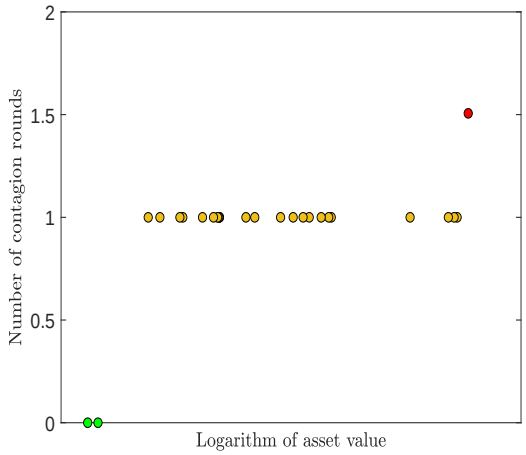
For larger banks, contagion does spread further for some simulations and therefore the average number of contagion rounds is higher than one in figures 2.19a and 2.19c. It is interesting to see that the average number of contagion rounds resulting from applying the initial shock to the largest bank is slightly lower for high systemic risk levels than for moderate levels. However, this observation does not necessarily hold for all network structures (see appendix B.5.2, figures B.18 and B.20). This means that in some months and for some network structures, the default of large banks may lead to additional defaults less frequently than for other months, but when they do lead to additional defaults, the damage to the system is worse.

A comparison of figures 2.17a, 2.18a and 2.19a shows a typical illustration of the 'robust-yet-fragile' property of financial networks [52]: For most initially shocked banks, contagion does not spread further even during times of high systemic risk. Therefore, the probability of contagion is low. However, by considering the largest bank it can be shown that contagion can be devastating to the system if it manages to spread. The average number of contagion rounds is less than 1.5 when shocking the largest bank, but the average ratio of capital lost in the system is close to 1 and the average fraction of defaults is over 80%. This shows that in the unlikely event that contagion spreads just beyond the initial bank, the damage to the system can be severe.



(a) High risk (Decemer 2015)

(b) Low risk (January 2017)



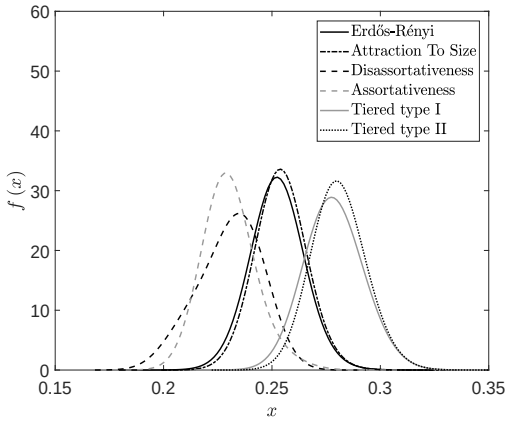
(c) Moderate risk (March 2017)

Figure 2.19: Comparing the average number of default rounds for each bank to its asset value for three different risk scenarios based on the Erdős-Rényi structure.

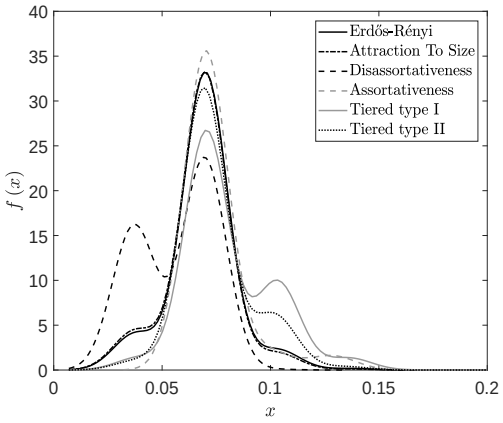
3. Risk quantity distributions

Finally, we consider the empirical distributions of the three different risk quantities contained in figure 2.20. That is, we consider the empirical distributions of the CRR, the fraction of defaulted banks $\bar{\alpha}_n$ and the number of contagion rounds that the network experiences following a shock before no more defaults occur. Figure 2.20a contains the empirical distribution of the average CRR for all network structures when systemic risk levels are high. Appendix B.5.3, figure B.21 contains the empirical distributions for the CRR for low and moderate risk levels. The general shapes of the CRR empirical distributions are the same for all three scenarios, which is why only one scenario's graph is displayed here. The

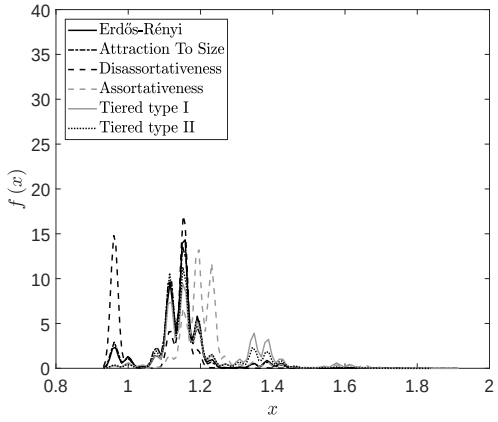
only difference is that the domain of the distributions shift to the left (as expected) when the risk levels are increased.



(a) Average Capital Reduction Ratio



(b) Average fraction of defaults



(c) Average number of contagion rounds

Figure 2.20: The empirical distribution of three risk quantities for a scenario where systemic risk levels are high (December 2015).

Figures 2.20b and 2.20c contain the empirical distributions for the average fraction of defaults and the average number of contagion rounds for all structures for high systemic risk levels. For the defaulted fraction, the graphs for low and moderate systemic risk levels are included in appendix B.5.3 as all structures' lines are indistinguishable from one another and both graphs lie over the same domain (see figure B.22b and B.22c). For the number of default rounds, the graph for low systemic risk levels is included in the appendix, since the distributions for all graphs are indistinguishable from one another and lie on a very small range (see B.23b). The graph for moderate systemic risk levels is included in

the appendix (figure B.23c) since the distributions for the structures follow similar shapes as for the high systemic risk levels (figure 2.20c), but with the former graph having the distributions over a much narrower range. Figure 2.20c containing the scenario with high systemic risk levels distinguish better between the different structures than those figures containing the other scenarios. Furthermore, the domain for the distributions are wider for high systemic risk levels compared to low or medium levels. This is likely due to the few larger banks, which were the only banks that were able to cause a spread of contagion upon default during months with high or moderate risk systemic risk levels.

It is interesting to see that for high systemic risk levels, the distribution of the average CRR tends to be symmetrical, with only one peak. However, the distributions of the average fraction of defaults and the average number of contagion rounds have more than one peak, and the distributions for the different structures have different shapes.

Chapter 3

Asymptotic results

This chapter is devoted to a network model of systemic risk for large networks. Here, we generalise results from [8], formally define a class of random networks, and show that the generalised results hold for this class. The chapter begins with section 3.1 presenting and discussing the background to the analytical network model that forms the basis of our results. The construction of the banking network, relevant notation and important definitions are presented there. The section ends by describing how the default cascades are modelled.

Section 3.2 is devoted to understanding and presenting existing results. It provides in-depth discussions of relevant functions before presenting the main results from [8]. Preliminary results for standard Erdős-Rényi networks are presented in section 3.3. Thereafter a general class of random networks are defined and we show that results similar to theorem 3.2.3 hold for this class of random networks. Section 3.4 illustrates the main theorem of this thesis and uses it to compare the systemic risk inherent in different network structures.

3.1 Background and notation

As mentioned above, this chapter is devoted to understanding and generalising the concepts and results from [8]. As a result, the notation and definitions used here correspond to those used by that article. The purpose of the research presented in [8] is to examine the systemic risk and stability inherent in large financial networks. The main result quantifies the asymptotic fraction of defaults expected in a network of financial entities as the network size grows large. For ease of reference, it is assumed that the network consists of banks, although the results could be used for networks of other entities as well. There is a set of initially defaulted nodes which spread losses to their credit counterparties. These

Table 3.1: Illustration of a simplified bank balance sheet.

Assets		**Equity and liabilities**	
$\sum_{j=1}^{n} e(i,j)$	Interbank Assets	$\gamma(i)\sum_{j=1}^{n} e(i,j)$	Capital/Net worth
N/A	Other assets	$\sum_{j=1}^{n} e(j,i)$	Interbank liabilities
		N/A	Other liabilities

counterparties may then also default as a result, thereby spreading losses throughout the system.

3.1.1 Network description

Similar to chapter 2, each bank in the network has a simplified balance sheet structure which determines the bank's initial financial position (i.e. before any defaults have occurred). Due to the complexity of asymptotic network results, the balance sheets are less refined compared to chapter 2, and liquidity effects are not considered here. An illustration of the assumed balance sheet structure, together with the notation discussed below is given in table 3.1.

For the results of this chapter, only the capital and the interbank assets are of interest. It is assumed that a bank's capital is used to absorb losses and that a bank defaults whenever its capital is depleted. The interbank assets are subdivided between a bank's counterparties. It therefore represents the exposure to other banks in the system and determines a bank's incoming and outgoing edges.

Suppose there are n banks in the system. For now, n is kept fixed and therefore the notation introduced below does not contain any reference to n[1]. Let $\gamma(i)$ be bank i's ratio of capital to interbank assets. Note that this is defined differently here compared to chapter 2, since the capital ratio is no longer specified as a percentage of total assets. This avoids the need to specify the total assets of each bank, leading to a more parsimonious and flexible model. Now let $e(i,j)$ be the amount that bank i has lent to bank j in a network of size n (i.e. bank i's exposure to j). Bank i's total interbank assets is therefore given by $\sum_{j=1}^{n} e(i,j)$, where $e(i,k) = 0$ for those banks k to which i has no exposure to. Bank i's total capital is then given by $\gamma(i)\sum_{j=1}^{n} e(i,j)$.

All n banks' exposures can be represented by a matrix $\mathbf{e}$ where the $(i,j)^{\text{th}}$ entry is the exposure of i to j. From this matrix, the links between the banks can be determined, as

[1]From section 3.2 onwards, the notation makes reference to n since networks of different sizes are considered there.

well as the weight of those links. This is because the non-zero entries in $\mathbf{e}$ indicate the exposures between banks in the system, whereas the zero entries in the matrix indicate which banks are not exposed each other. The system can therefore be represented by a weighted directed graph with n vertices, whose edges can be derived from $\mathbf{e}$. This matrix indicates the presence, direction and weight of all edges in the network. Note that for any exposure matrix, the diagonal entries $e(i, i)$ must be zero, whereas the rest of the entries must be non-negative.

Let $\vec{\gamma}$ be the vector containing the capital ratios of all n banks. A network of banks is then characterised by its exposure matrix $\mathbf{e}$ and its vector of capital ratios $\vec{\gamma}$. In other words, two networks are considered to be identical if and only if they have the same exposure matrix and capital ratio vector. This leads to the following definition [8, Definition 2.1]:

Definition 3.1.1 (Financial network)**.** Let $\mathbf{e}$ be an exposure matrix for a network consisting of n banks, and let $\vec{\gamma}$ be the associated vector of capital ratios. The pair $(\mathbf{e}, \vec{\gamma})$ is then called a *financial network*.

A node's out-degree is the number of outgoing edges connected to it (in other words the number of banks that it has lent money to). The out-degree of a node i is defined by $d^+(i) = \#\{j \in \{1, \dots, n\} \mid e(i, j) > 0\}$. Similarly, the in-degree of a node i is the number of incoming edges connected to it and represents the number of banks that it has borrowed money from. It is defined by $d^-(i) = \#\{j \in \{1, \dots, n\} \mid e(j, i) > 0\}$. For a network of size n, $\vec{d^+} = (d^+(1), \dots, d^+(n))$ and $\vec{d^-} = (d^-(1), \dots, d^-(n))$ are the vectors containing the out-degrees and in-degrees of the nodes in the system respectively.

3.1.2 Random financial networks

In practice a bank's counterparties change on a daily basis. Therefore it would be preferable to choose the counterparties of a bank i uniformly over a range of possibilities. Consider the case where the number of counterparties to each bank is known, but the identities of those counterparties are unknown. The counterparties of bank i are then chosen uniformly over all possibilities where the total number of its counterparties remains the same.

As for the exposure sizes, it is assumed that there is a set of $d^+(i)$ exposure amounts, randomly assigned to each of i's chosen counterparties. This concept is formalised as follows (see [8, Definition 2.4]):

Definition 3.1.2 (Random financial network)**.** Suppose $(\mathbf{e}, \vec{\gamma})$ is a financial network of size n with fixed degree vectors $\vec{d^+}$ and $\vec{d^-}$ as determined by $\mathbf{e}$. The set of exposures for a fixed bank i in this financial network is given by $\{e(i, j) \mid e(i, j) > 0, j \in \{1, \dots, n\}\}$.

Let $\mathcal{G}(\mathbf{e})$ be the set of all possible exposure matrices of size n that

1. imply the same degree vectors $\vec{d}^+$ and $\vec{d}^-$ as $\mathbf{e}$ and
2. for each node i, have the same set of exposure sizes as $\mathbf{e}$ (each of which need not be assigned to the same counterparty as in the original financial network, as long as the node i which lent the money remains the same as in the original financial network).

Note that $\mathcal{G}(\mathbf{e})$ is not empty, since the original exposure matrix $\mathbf{e}$ is an element of $\mathcal{G}(\mathbf{e})$.

Define $\Omega(\mathbf{e}, \vec{\gamma})$ to be the set of all possible financial networks with degree vectors $\vec{d}^+$ and $\vec{d}^-$, where the set of exposures for a bank i is the same as for $\mathbf{e}$, i.e. it is $\{e(i,j) \mid e(i,j) > 0, j \in \{1, \ldots, n\}\}$. Let $\mathcal{S}$ be the field consisting of all subsets of $\Omega(\mathbf{e}, \vec{\gamma})$. Note that the sets $\mathcal{G}(\mathbf{e})$ and $\Omega(\mathbf{e}, \vec{\gamma})$ are quite similar, the difference being that the elements of $\mathcal{G}(\mathbf{e})$ are exposure matrices, whereas the elements of $\Omega(\mathbf{e}, \vec{\gamma})$ are financial networks (exposure matrices with capital ratios).

For each financial network $(\mathbf{e}, \vec{\gamma})$, let $(\Omega(\mathbf{e}, \vec{\gamma}), \mathcal{S}, \mathbb{P})$ be the probability space where $\mathbb{P}$ is the counting measure on $\Omega(\mathbf{e}, \vec{\gamma})$. For this probability space, define $\mathbf{E}\colon \Omega(\mathbf{e}, \vec{\gamma}) \to \mathcal{G}(\mathbf{e})$ as a random exposure matrix, uniformly distributed on $\mathcal{G}(\mathbf{e})$. The nodes of $\mathbf{E}$ are endowed with the capital ratios $\vec{\gamma}$, and the resulting financial network $(\mathbf{E}, \vec{\gamma})$ is called a *random financial network*.

Example illustrating how $\mathcal{G}(\mathbf{e})$ is determined Suppose we have the following exposure matrix for a network consisting of $n = 4$ nodes:

$$\mathbf{e} = \begin{bmatrix} 0 & 1 & 2 & 0 \\ 3 & 0 & 0 & 4 \\ 5 & 0 & 0 & 6 \\ 0 & 7 & 0 & 0 \end{bmatrix} \tag{3.1}$$

Here the degree vectors are given by $\vec{d}^+ = (2, 2, 2, 1)$ and $\vec{d}^- = (2, 2, 1, 2)$. The set $\mathcal{G}(\mathbf{e})$ contains (for example) the following matrices as elements:

$$\begin{bmatrix} 0 & 1 & 2 & 0 \\ 3 & 0 & 0 & 4 \\ 5 & 0 & 0 & 6 \\ 0 & 7 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 2 & 1 & 0 \\ 4 & 0 & 0 & 3 \\ 6 & 0 & 0 & 5 \\ 0 & 7 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 1 & 0 & 2 \\ 4 & 0 & 3 & 0 \\ 0 & 5 & 0 & 6 \\ 7 & 0 & 0 & 0 \end{bmatrix} \tag{3.2}$$

The set $\mathcal{G}(\mathbf{e})$ does not, for example, contain the matrices

$$\begin{bmatrix} 1 & 0 & 2 & 0 \\ 3 & 0 & 0 & 4 \\ 5 & 0 & 0 & 6 \\ 0 & 7 & 0 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & 7 & 2 & 0 \\ 3 & 0 & 0 & 4 \\ 5 & 0 & 0 & 6 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \tag{3.3}$$

The first matrix is not in $\mathcal{G}(\mathbf{e})$ since it does not have only zeros along the diagonal and therefore is not a valid exposure matrix. The second matrix is not contained in $\mathcal{G}(\mathbf{e})$ since only the rows of the original matrix may be shuffled, and not the columns. In other words, $\mathcal{G}(\mathbf{e})$ contains all matrices where the entries in each row of the original matrix $\mathbf{e}$ are shuffled around, subject to the requirements that the in-degree vector must remain the same and that the diagonal entries must remain zero (so that the matrix remains a valid exposure matrix, since banks cannot lend to themselves).

3.1.3 Default and contagion mechanics

The set $\mathbb{D}_0(\mathbf{e}, \vec{\gamma})$ is defined to be the set of initial defaults in the financial system $(\mathbf{e}, \vec{\gamma})$. As mentioned in section 3.1.1, an institution defaults whenever its capital is depleted. In order to assess the effect of a shock to the financial system, there needs to be one or more initial defaults. Banks whose capital ratios are zero therefore form the set of initial defaults so that

$$\mathbb{D}_0(\mathbf{e}, \vec{\gamma}) = \{i \in \{1, \ldots, n\} \mid \vec{\gamma}(i) = 0\}. \tag{3.4}$$

These are then the institutions that may cause a default cascade in the system. Once the initial defaults have occurred, losses are spread through the defaulted nodes' incoming edges. This is because the direction of edges imply the direction of lending in the system. A bank $i \in \mathbb{D}_0(\mathbf{e}, \vec{\gamma})$ will cause a loss of $(1 - R(i))\, e(j, i)$ to each of lenders j, where $R(i)$ is the recovery rate associated with the defaulted node i.

The losses caused by the nodes in $\mathbb{D}_0(\mathbf{e}, \vec{\gamma})$ might lead to additional defaults in the system. The set of nodes that have defaulted up to this point is given by

$$\mathbb{D}_1(\mathbf{e}, \vec{\gamma}) = \left\{ i \in \{1, \ldots, n\} \mid \gamma(i) \sum_{j=1}^{n} e(i, j) \leq \sum_{j \in \mathbb{D}_0(\mathbf{e}, \vec{\gamma})} (1 - R(j))\, e(i, j) \right\}. \tag{3.5}$$

These losses form the first 'round' of default, as these are the first nodes to have defaulted

as a result of the initial set of defaults. For subsequent rounds of default we have that

$$\mathbb{D}_k(\mathbf{e}, \vec{\gamma}) = \left\{ i \in \{1, \ldots, n\} \mid \gamma(i) \sum_{j=1}^{n} e(i,j) \leq \sum_{j \in \mathbb{D}_{k-1}(\mathbf{e}, \vec{\gamma})} (1 - R(j))\, e(i,j) \right\}, \tag{3.6}$$

for $k \geq 1$.

Note that $\mathbb{D}_0(\mathbf{e}, \vec{\gamma}) \subseteq \mathbb{D}_1(\mathbf{e}, \vec{\gamma})$, since the latter set also includes the initial defaults. This is because $\vec{\gamma}(i) \sum_{j=1}^{n} e(i,j) = 0$ for all initially defaulted nodes, so that they satisfy $\gamma(i) \sum_{j=1}^{n} e(i,j) \leq \sum_{j \in \mathbb{D}_0(\mathbf{e}, \gamma)} (1 - R(j))\, e(i,j)$ regardless of whether they are connected to one another or not.

Similarly, $\mathbb{D}_1(\mathbf{e}, \vec{\gamma}) \subseteq \mathbb{D}_2(\mathbf{e}, \vec{\gamma})$ because $\mathbb{D}_2(\mathbf{e}, \vec{\gamma})$ is determined by $\mathbb{D}_1(\mathbf{e}, \vec{\gamma})$, and the latter set contains the initial defaults. Any defaults caused directly by the initially defaulted nodes will therefore also be included in $\mathbb{D}_2(\mathbf{e}, \vec{\gamma})$.

By following this argument for all $k \geq 1$, it is seen that the sequence $\mathbb{D}_0(\mathbf{e}, \vec{\gamma}) \subseteq \mathbb{D}_1(\mathbf{e}, \vec{\gamma}) \subseteq \cdots \subseteq \mathbb{D}_{n-1}(\mathbf{e}, \vec{\gamma})$ is nested. Here $\mathbb{D}_{n-1}(\mathbf{e}, \vec{\gamma}) \subseteq \{1, \ldots, n\}$ because there can be at most n defaults and there can be at most $n - 1$ rounds of default. There cannot be more than $n - 1$ rounds because there must be at least one node in the set $\mathbb{D}_0(\mathbf{e}, \vec{\gamma})$ for there to be any subsequent rounds of default. If the default cascade stops when there have only been $k_0 < n - 1$ rounds of default, then $\mathbb{D}_k(\mathbf{e}, \vec{\gamma}) = \mathbb{D}_{k+1}(\mathbf{e}, \vec{\gamma})$ for all $k = k_0, \ldots, n - 2$.

From the defaults $\mathbb{D}_{n-1}(\mathbf{e}, \vec{\gamma})$ caused by nodes in the set $\mathbb{D}_0(\mathbf{e}, \vec{\gamma})$, we can find the final/-total fraction $\alpha_n(\mathbf{e}, \vec{\gamma})$ of defaults in a financial network $(\mathbf{e}, \vec{\gamma})$ given by

$$\alpha_n(\mathbf{e}, \vec{\gamma}) \coloneqq \frac{\mathbb{D}_{n-1}(\mathbf{e}, \vec{\gamma})}{n}. \tag{3.7}$$

For some of the subsequent results it is necessary to keep track of the order in which a node i's counterparties default. For this purpose, let $\Sigma(i)$ be the set of all combinations of orders in which i's counterparties can default. In other words for a node i with $d^+(i)$ outgoing edges, $\Sigma(i)$ will be the set of all permutations of the numbers $1, \ldots, d^+(i)$, so that it will be a set of $d^+(i)!$ vectors, each of size $d^+(i)$.

Once an order of default is chosen for a node's outgoing edges, the minimum number of counterparty defaults that can cause it to become insolvent can be determined. This is called a node's default threshold, and is dependent on the order of default chosen. Let $\tau \in \Sigma(i)$ for a fixed node i. Then i's default threshold $\Theta(i, \tau)$ as determined by τ is given by

$$\Theta(i, \tau) = \min \left\{ k \geq 0 \mid \gamma(i) \sum_{j=1}^{n} e(i,j) \leq \sum_{j=1}^{k} (1 - R(\tau(j)))\, e(i, \tau(j)) \right\}. \tag{3.8}$$

Suppose that the orders of default for all nodes have been determined. Then for notational convenience, it is said that a node is of type (j, k, θ) if it has j out-degrees, k in-degrees and a default threshold of θ. Similarly, if a node has j out-degrees and k in-degrees, it is said to be of degree (j, k).

3.2 Existing results for deterministic networks

3.2.1 Discussion of relevant functions

The purpose of this section is to discuss functions that are required in order to present and discuss the results of this chapter. Each function is formally defined and the intuitive meaning behind it is discussed. From this point onwards, the notation is adapted to indicate whenever a variable is dependent on n, the number of nodes in the network.

Proportion of nodes with degree (j, k)

The exposure matrix $\mathbf{e}_n$ of a financial network of size n determines the degrees of all the nodes. The resulting distribution of degrees μ_n is given by

$$\mu_n(j,k) = \frac{1}{n}\#\left\{i \in \{1,\dots,n\} \mid d_n^+(i) = j, d_n^-(i) = k\right\}. \tag{3.9}$$

In other words this is the proportion of nodes with degree (j, k), where there are n nodes in total.

Now let $(\mathbf{e}_n, \vec{\gamma}_n)_{n\geq 1}$ be a sequence of financial networks. For each $n = 1, 2, \dots$ the exposure matrix $\mathbf{e}_n$ determines degree vectors $\vec{d}_n^+$ and $\vec{d}_n^-$ and degree distribution $\mu_n(j, k)$. The degree vectors and degree distributions obtained from the sequence $(\mathbf{e}_n, \vec{\gamma}_n)_{n\geq 1}$ must satisfy the following assumptions [8, Assumption 3.1]:

Assumption 3.1. There exists a probability distribution μ on $\mathbb{N}^2$ which is independent of n, has finite mean $\lambda = \sum_{j,k} j\mu(j,k) = \sum_{j,k} k\mu(j,k)$ and satisfies the following:

1. $\mu_n(j,k) \to \mu(j,k)$ as $n \to \infty$ for all $j, k \geq 0$ and
2. $\sum_{i=1}^n \left[(d_n^+(i))^2 + (d_n^-(i))^2\right] = O(n)$.[a]

As $\mu(j, k)$ has mean λ, this means that as $n \to \infty$, nodes will, on average, have λ in-degrees. This is the same as saying that nodes will have λ out-degrees on average.

[a] Let x_n and y_n be non-negative sequences. Then $x_n = O(y_n)$ if there exists a non-negative integer N and a $C > 0$ such that $x_n \leq Cy_n$ for all $n \geq N$.

Since μ_n is determined by $\mathbf{e}_n$, it is of interest to investigate the different types of exposure matrices that would result in realised degree distributions satisfying assumption 3.1. This is done in section 3.3, where a range of possibilities are discussed.

Proportion of nodes with degree (j,k), insolvent after θ counterparty defaults

For each n, $p_n(j,k,\theta)$ represents the fraction of nodes with degree (j,k) that become insolvent after θ of their counterparties default in the random network $\mathbf{E}_n$. It is defined as follows:

$$p_n(j,k,\theta) = \frac{\#\left\{(i,\tau_n) \mid d_n^+(i) = j, d_n^-(i) = k, \tau_n \in \Sigma_n(i)\,, \Theta_n(i,\tau_n) = \theta\right\}}{n\mu_n(j,k)\; j!}. \tag{3.10}$$

The numerator represents the number of pairs (i,τ_n) where the node i has degree (j,k) and the permutation τ_n leads to i having default threshold θ. In the denominator, $n\mu_n(j,k)$ represents the total number of nodes with degree (j,k). The $j!$ term represents the number of ways in which the j counterparties of a node with degree (j,k) can default. The denominator therefore represents the total number of pairs (i,τ_n) where the node has degree (j,k). This definition of p_n can therefore be replaced by

$$p_n(j,k,\theta) = \frac{\#\left\{(i,\tau_n) \mid d_n^+(i) = j, d_n^-(i) = k, \tau_n \in \Sigma_n(i)\,, \Theta_n(i,\tau_n) = \theta\right\}}{\#\left\{(i,\tau_n) \mid d_n^+(i) = j, d_n^-(i) = k, \tau_n \in \Sigma_n(i)\right\}}. \tag{3.11}$$

If the permutations are chosen uniformly, then this represents the expected fraction of (j,k)-nodes that will default after θ of their counterparties default.

The function $p\colon \mathbb{N}^3 \to [0,1]$ must satisfy the following [8, Assumption 3.4]:

$$p_n(j,k,\theta) \to p(j,k,\theta) \text{ as } n \to \infty \tag{3.12}$$

for all $(j,k,\theta) \in \mathbb{N}^3$ where $\theta \leq j$. As the network size becomes large, $p(j,k,\theta)$ represents the expected fraction of (j,k)-nodes to default after θ counterparty defaults. Therefore $p(j,k,0)$ is the asymptotic fraction of all nodes with degree (j,k) that form part of the set of initially defaulted nodes, as these will have defaulted after zero counterparty defaults.

For notational convenience, the survival function of a binomial random variable $X \sim \text{Bin}(j,\pi)$ is defined as follows:

$$\bar{B}(j,\pi,\theta) = \mathbb{P}(X \geq \theta) = \sum_{l \geq \theta}^{j} \binom{j}{l} \pi^l \,(1-\pi)^{j-l}\,. \tag{3.13}$$

Suppose that for any node, its counterparties each default with probability π. The above expression is equivalent to the probability that a node with j outgoing links experiences θ or more counterparty defaults. Therefore for a node i with j counterparties and a default threshold of θ, $\bar{B}(j, \pi, \theta)$ represents the probability that the node's default threshold will be reached, hence leading to the default of i.

The function $I\colon\ [0, 1] \to [0, 1]$ is defined by

$$I(\pi) = \sum_{j,k} \frac{\mu(j, k)\, k}{\lambda} \sum_{\theta=0}^{j} p(j, k, \theta)\, \bar{B}(j, \pi, \theta)\,. \tag{3.14}$$

The terms making up $I(\pi)$ can be interpreted as follows:

- $p(j, k, \theta)\, \bar{B}(j, \pi, \theta)$ is the expected proportion of nodes with degree (j, k) and default threshold θ, who will default after one round if any one of their counterparties had defaulted with probability π. This is because $p(j, k, \theta)$ can be approximated by

  $$\frac{\#\text{ nodes of type }\ (j, k, \theta)}{\#\text{ nodes in the system}}, \tag{3.15}$$

  where the size of the network is large. Similarly if π represents the probability that a counterparty of a node defaults, then $\bar{B}(j, \pi, \theta)$ can be approximated by

  $$\frac{\#\text{ defaulting nodes with } j \text{ counterparties and default threshold } \theta}{\#\text{ nodes with } j \text{ counterparties and default threshold } \theta}. \tag{3.16}$$

  Since $\bar{B}$ is independent of the number of incoming links to a node, it can also be approximated by

  $$\frac{\#\text{ defaulting nodes of type }\ (j, k, \theta)}{\#\text{ nodes of type }\ (j, k, \theta)}. \tag{3.17}$$

  Therefore $p(j, k, \theta)\, \bar{B}(j, \pi, \theta)$ can be approximated by

  $$\frac{\#\text{ defaulting nodes of type }\ (j, k, \theta)}{\#\text{ nodes in the system}}, \tag{3.18}$$

  which can be interpreted as the fraction of all nodes that have defaulted and are of type (j, k, θ).

- $\sum_{\theta=0}^{j} p(j, k, \theta)\, \bar{B}(j, \pi, \theta)$ is the expected fraction of all nodes that have degree (j, k) and will default after one round if any one of their counterparties had defaulted with probability π.

- For large n, $\frac{\mu_n(j,k)\,k}{\lambda}$ can be used to approximate $\frac{\mu(j,k)\,k}{\lambda}$. The former can be rewritten as follows:

$$\frac{\mu_n(j,k)\,k}{\lambda} = \frac{\left(\#\left\{i \mid d_n^+(i) = j, d_n^-(i) = k\right\}\right)\,k}{n\lambda}. \tag{3.19}$$

The numerator represents the number of in-degrees used by nodes of degree (j,k). Regarding the denominator, λ represents the average number of in-degrees[2] for a node when the network size tends to infinity. Therefore for large n, $n\lambda$ represents the total expected number of in-degrees in the system. The term then represents the fraction of all in-degrees in the system used by nodes of degree (j,k).

To illustrate this, consider the simple network shown in figure 3.1. The degrees of the nodes are shown as labels on the graph and the network has a total of nine edges.

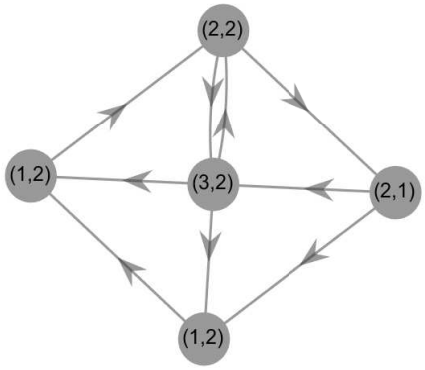


Figure 3.1: A simple network example to illustrate the meaning of $\frac{\mu_n(j,k)k}{\lambda}$. The labels inside the nodes represent their respective degrees.

For this example, we have that λ is given by

$$\begin{aligned}\lambda &= 3\mu_n(3,2) + 2\mu_n(2,2) + 2\mu_n(2,1) + \mu_n(1,2) + \mu_n(1,2)\\ &= 3\,(0.2) + 2\,(0.2) + 2\,(0.2) + (0.4) + (0.4)\\ &= \frac{9}{5}.\end{aligned} \tag{3.20}$$

Note that here, λ was calculated as the average in-degree of nodes in the system. The same answer would be obtained by calculating the average out-degree of nodes in the system.

Consider nodes of degree $(1,2)$. In this case the fraction of incoming edges used by nodes of this type is $\frac{4}{9}$, as there are four incoming edges used by these nodes and the

[2]This is also the average number of out-degrees.

system has nine edges in total. Now since $\mu_n(1,2) = 0.4$ and $\lambda = \frac{9}{5}$, then

$$\frac{\mu_n(1,2) \cdot 2}{\lambda} = \frac{4}{5} \cdot \frac{5}{9} = \frac{4}{9}, \tag{3.21}$$

which is the same as the proportion calculated above.

- Based on equation (3.19), the function I can be rewritten as follows:

$$I(\pi) = \sum_{j,k} \frac{\left(\#\left\{i \mid d_n^+(i) = j, d_n^-(i) = k\right\}\right) k}{n\lambda} \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j,\pi,\theta)\,. \tag{3.22}$$

  The term $\left(\#\left\{i \mid d_n^+(i) = j, d_n^-(i) = k\right\}\right) k \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j,\pi,\theta)$ represents the number of in-degrees used by (j,k) nodes multiplied by the proportion of those nodes that will default as determined by π. Therefore the term represents the number of in-degrees from the (j,k) nodes that will spread losses during the next round of defaults.

- The sum $\sum_{j,k} \left(\#\left\{i \mid d_n^+(i) = j, d_n^-(i) = k\right\}\right) k \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j,\pi,\theta)$ is then the number of in-degrees from all nodes that will spread losses during the next round of defaults.

Therefore $I(\pi)$ is the proportion of all in-degrees that will spread losses during the next round of default, if each counterparty of a node defaults with probability π. In other words at the end of a default round, this is the typical fraction of counterparty defaults that a node can expect, given that at the start of that round, a node would have expected a fraction of π counterparty defaults. Therefore if a node had j counterparties at the start of the round, then at the end of the round it would be expected to have $j(1-\pi)$ counterparties left, of which $j(1-p)\, I(\pi)$ would have defaulted.

Using this interpretation, the term $I(0)$ represents the fraction of counterparty defaults that a node can expect after one round of default, given that at the start of that round, a node would have had no counterparty defaults. The interpretation makes sense if one considers $I(0)$ to be the fraction of initially defaulting nodes, since those nodes wouldn't have had counterparty defaults themselves.

Next, define π^* to be the smallest fixed point[3] of I:

$$\pi^* = \inf\{\pi \in [0,1] \mid I(\pi) = \pi\}\,. \tag{3.23}$$

A fixed point π_0 of I (provided it exists) has the following intuitive explanation: If nodes experienced a proportion π_0 of their counterparties defaulting at the start of a round,

[3]Proposition 3.2.2 below establishes the existence of π^*.

then at the end of that round a randomly chosen node will have the same proportion of its remaining counterparties defaulting.

Let $I^k(\pi)$ denote the k^{th} composition of I with itself. Note that the fraction of initially insolvent counterparties to a node is given by $I(0)$. Therefore $I^k(0)$ is the fraction of counterparty defaults that a node can expect at the end of the k^{th} round. For large enough k, $I^k(0)$ is the fraction of counterparty defaults that a node can expect at the end of the contagion process.

To relate this to the definition of π^*, Kleene's fixed point theorem is used. The statement of the theorem and the definitions used in relation to it are based on [16, 88].

Let $(P, \leq)$ be a partially ordered set. Then $(P, \leq)$ is said to be ω-complete if every increasing sequence of elements in P have a supremum in P. The function $f\colon P \to P$ is said to be ω-continuous if, for every increasing sequence $(x_n)_{n\in\mathbb{N}}$ of elements in P, the following hold:

1. If $\sup\{x_n\}$ exists in P, then $\sup\{f(x_n)\}$ exists in P as well, and

2. $$f(\sup\{x_n\}) = \sup\{f(x_n)\}\,. \tag{3.24}$$

Kleene's fixed point theorem is given as follows:

Theorem 3.2.1. *Let $(P, \leq)$ be an ω-complete partially ordered set with a least element x_0 and let $f\colon P \to P$ be ω-continuous. If $x \in P$ satisfies $x \leq f(x)$, then $x^* = \sup\{f^n(x_0)\}$ is a fixed point of f. In particular, it is the least fixed point of f in P.*

If Kleene's fixed point theorem can be applied with $x = 0$, then it will show that the least fixed point of I represents the probability that a randomly chosen edge ends at a defaulted node at the end of the contagion process.

Proposition 3.2.2. *Kleene's fixed point theorem can be applied to $P = [0, 1]$, $f = I$ and $x = 0$.*

Proof. The set $P = [0, 1]$ together with the usual meaning of $\leq$ defines a partial ordering. It is ω-complete because $[0, 1]$ is closed. Now it is proved that I is ω-continuous.

Let $(\pi_n)_{n\in\mathbb{N}}$ be any increasing sequence in $[0, 1]$, so that $\sup\{\pi_n\}$ exists because $([0, 1], \leq)$ is ω-complete. Then $\sup\{I(\pi_n)\}$ is contained in the range $[0, 1]$ since it is closed. Now let $a = \sup\{\pi_n\}$. Since $\{\pi_n\}$ is an increasing sequence, then $\lim_{n\to\infty} \pi_n = a$. It needs to be shown that $\sup\{I(\pi_n)\} = I(a)$. To do this, it is first shown that $\lim_{n\to\infty} I(\pi_n) = I(a)$ and then that I is a non-decreasing function.

For any $n \in \mathbb{N}$,

$$
\begin{aligned}
&|I(\pi_n) - I(a)| \\
&= \left| \sum_{j,k} \frac{\mu(j,k)\,k}{\lambda} \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j,\pi_n,\theta) - \sum_{j,k} \frac{\mu(j,k)\,k}{\lambda} \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j,a,\theta) \right| \\
&= \left| \sum_{j,k} \frac{\mu(j,k)\,k}{\lambda} \sum_{\theta=0}^{j} p(j,k,\theta) \left(\bar{B}(j,\pi_n,\theta) - \bar{B}(j,a,\theta) \right) \right| \\
&= \left| \sum_{j,k} \frac{\mu(j,k)\,k}{\lambda} \sum_{\theta=0}^{j} p(j,k,\theta) \sum_{l \geq \theta}^{j} \binom{j}{l} \left(\pi_n^l\, (1-\pi_n)^{j-l} - a^l\, (1-a)^{j-l} \right) \right|. \qquad (3.25)
\end{aligned}
$$

Therefore if $n \to \infty$, then $\pi_n \to a$ and hence $|I(\pi_n) - I(a)| \to 0$. This shows that $\lim_{n\to\infty} I(\pi_n) = I(a)$.

Furthermore, I is a non-decreasing function of π because all the terms of which I consist are non-negative and $\bar{B}(j,\pi,\theta)$ is a non-decreasing function of π for all j and all θ. This means that $\{I(x_n)\}$ is a non-decreasing sequence because $\{x_n\}$ is an increasing sequence. Therefore $\lim_{n\to\infty} I(x_n) = \sup\{I(x_n)\}$. Since $\lim_{n\to\infty} I(\pi_n) = I(a)$, the function I on $[0,1]$ is ω-complete.

Since I is non-decreasing, $0 \leq I(0)$, so that the assumptions underlying Kleene's fixed point theorem are satisfied. □

To conclude, proposition 3.2.2 therefore shows that $\pi^* = \lim_{k\to\infty} I^k(0)$, where $I(0)$ is the fraction of initially defaulted nodes.

3.2.2 Asymptotic fraction of total defaults

The theorem below is a restatement of theorem 3.8 in [8].

Theorem 3.2.3. *Suppose the sequence $(\mathbf{e}_n, \vec{\gamma}_n)$ of networks satisfies assumptions 3.1 and 3.4 as in [8] and that $(\mathbf{E}_n)_{n\geq 1}$ is the corresponding sequence of random matrices on $(\Omega, \mathcal{A}, \mathbb{P})$.*

1. *If $\pi^* = 1$, i.e. if $I(\pi) > \pi$ for all $\pi \in [0,1)$, then*

$$\alpha_n(\mathbf{E}_n, \vec{\gamma}_n) \to 1 \qquad (3.26)$$

weakly as $n \to \infty$. In other words, almost all nodes default as the network size goes to infinity.

2. *If $\pi^* < 1$, and π^* is a stable fixed point of I (i.e. $I'(\pi^*) < 1$), then*

$$\alpha_n(\mathbf{E}_n, \vec{\gamma}_n) \to \sum_{j,k} \mu(j,k) \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j,\pi^*,\theta) \tag{3.27}$$

weakly as $n \to \infty$. This is then the asymptotic fraction of defaults as the network size goes to infinity.

Remark 3.1. Theorem 3.2.3 part 1 shows that if the proportion of in-degrees spreading losses during each consecutive round increases as the cascade continues, then eventually all nodes will default. Part 2 gives an expression for the asymptotic fraction of defaults for the case where the proportion of in-degrees spreading losses during each consecutive round stabilises.

Theorem 3.2.3 and the assumptions underlying it are essentially the focus of the remainder of this chapter. Note that generalising the remaining results from [8] are beyond the scope of this study. Appendix C.1 contains an overview of those results. Theorem 3.2.3 is based on a sequence of networks with deterministic degree sequences. Every node in a random financial network retains its original number of debtors and creditors, as well as the monetary amounts of the interbank loans that it granted. The element of randomness in a network should reflect the daily change in interbank relationship because of the high frequency of this change. The definition of a random financial network (see definition 3.1.2, page 84) does not completely capture this daily change in interbank relationships since degree vectors and the exposure amounts are assumed to remain the same.

It is for this reason that section 3.3 below generalises theorem 3.2.3 to apply to stochastic exposure matrices instead. Hence we let $\mathbf{e}_n$ be determined via a random graph. A second reason for using a random graph is that the results for the default Erdős-Rényi case can be extended to the case where there are multiple Erdős-Rényi networks that interact with one another. This allows us to apply the results to a very flexible network and thus examine a range of different network structures in section 3.4 based on bank lending behaviour.

3.3 Results for stochastic networks

In this section, we consider different sequences of networks with a view to applying the results of theorem 3.2.3 to them. The assumptions underlying this theorem in its original form apply to sequences of networks with deterministic numbers of degrees. The goal of this section is to show that assumptions similar to assumption 3.1 (see page 88) hold

for certain sequences of networks with random numbers of degrees. The focus will be on randomising the links between the banks, since assumption 3.1 deals with the degree distributions of the networks implied by the sequence $(\mathbf{e}_n, \vec{\gamma}_n)$.

Section 3.3.1 contains results for the Erdős-Rényi case that are used as a building block for further work. Thereafter, section 3.3.2 uses the results of section 3.3.1 to prove a version of theorem 3.2.3 that applies to a stochastic, inhomogeneous class of networks.

3.3.1 Results for Erdős-Rényi networks

Preliminary definitions and explanations

For this section we will consider an Erdős-Rényi graph of size n, where the probability that a directed edge exists between two nodes is denoted by q. The initial results of the section are applicable to standard Erdős-Rényi graphs. These are then extended to an arbitrary number of connected groups of Erdős-Rényi networks in section 3.3.2.

A randomly chosen node will, on average, have $(n-1)q$ incoming edges. This is the same as the average number of outgoing edges. Therefore if q is kept fixed when letting $n \to \infty$, then μ_n cannot converge to a probability mass function with finite mean. For this reason the value of q must depend on n. If we specify beforehand that we want the average number of incoming edges to converge to a finite amount $\lambda < \infty$, then for a given n it must hold that $q_n = q = \frac{\lambda}{n-1}$.

Such an Erdős-Rényi graph of size n with connection probability $q_n = \frac{\lambda}{n-1}$ will be denoted by $K_{\lambda,n}$. Accordingly we let $\kappa_{\lambda,n}$ denote a realization of the random network $K_{\lambda,n}$, chosen uniformly over all the possible networks that $K_{\lambda,n}$ can result in.

We let h_n^+ and h_n^- be the probability mass functions of $D_n^+(i)$ and $D_n^-(i)$ respectively. Note that D_n^+ and D_n^- are independent of i, but may depend on n. This is because the edges of nodes in an Erdős-Rényi graph all have the same probability of being present, and this probability is a function of n. Now if h_n is the two-dimensional probability mass function whose marginals are h_n^+ and h_n^-, then $h_n(j,k)$ is the probability that a node i has degree (j,k) (i.e. that $D_n^+(i) = j$ and $D_n^-(i) = k$.). For the random network $K_{\lambda,n}$ we will have that

$$h_n(j,k) = \binom{n-1}{j} q_n^j (1-q_n)^{n-1-j} \binom{n-1}{k} q_n^k (1-q_n)^{n-1-k}, \tag{3.28}$$

since each node has $n-1$ other nodes to which it can be connected via incoming and/or outgoing edges.

The function h_n converges to a probability mass function h with finite mean λ. It is

known from the Poisson limit theorem that if $n \to \infty$ and $q_n \to 0$ such that $nq_n \to \lambda$, then

$$\binom{n}{j} q_n^j (1-q_n)^{n-j} \to e^{-\lambda} \frac{\lambda^j}{j!} \tag{3.29}$$

for $j > 0$. If $j = 0$, then $\binom{n}{j} q_n^j (1-q_n)^{n-j} = (1-q_n)^n$. Since $nq_n \to \lambda$, then $(1-q_n)^n \to \left(1 - \frac{\lambda}{n}\right)^n$. Using the fact that $\left(1 + \frac{x}{n}\right)^n \to e^x$, we know that $\left(1 - \frac{\lambda}{n}\right)^n \to e^{-\lambda}$. Therefore if $j = 0$, then $\binom{n}{j} q_n^j (1-q_n)^{n-j} \to e^{-\lambda}$ as $n \to \infty$.

Now if $\lambda < \infty$ is chosen beforehand, then $q_n = \frac{\lambda}{n-1}$ is chosen as the Erdős-Rényi probability. Now $q_n \to 0$ as $n \to \infty$ and $(n-1)q_n = \lambda$ remains constant. Therefore when letting $n \to \infty$ in equation (3.28), it is seen that

$$h_n(j,k) \to e^{-\lambda} \frac{\lambda^j}{j!} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-2\lambda} \frac{\lambda^{j+k}}{j!k!} = h(j,k)\,, \tag{3.30}$$

with the appropriate adjustments if $j = 0$ and/or $k = 0$.

Asymptotic results for Erdős-Rényi networks

Suppose $\lambda \in (0, \infty)$ is fixed and consider the Erdős-Rényi graph $K_{\lambda,n}$. Proposition 3.3.1 below serves as a preliminary result. Theorem 3.3.2 uses a procedure similar to the proof of the law of large numbers to prove a result similar to the first part of assumption 3.1.

Proposition 3.3.1. *Let $\lambda \in (0, \infty)$ be the expected number of incoming (or outgoing) edges in an Erdős-Rényi network, where n is the number of nodes in the network (λ being independent of n). For fixed $j, k, n \in \mathbb{N}$, define the random variables $X_i^{(n,j,k)}$, $i = 1, 2, \ldots, n$ to indicate whether a node i is of type (j,k) or not. Then $cov\left(X_i^{(n,j,k)}, X_l^{(n,j,k)}\right) \to 0$ when $n \to \infty$ for $i \neq l$.*

Proof. Fix any given $j, k \in \mathbb{N}$. Note that $X_i^{(n,j,k)} \sim \text{Ber}(h_n(j,k))$ for each $n \in \mathbb{N}$ and each $i \in \{1, 2, \ldots, n\}$, where

$$h_n(j,k) = \binom{n-1}{j} q_n^j (1-q_n)^{n-1-j} \binom{n-1}{k} q_n^k (1-q_n)^{n-1-k} \tag{3.31}$$

and $q_n = \lambda/n-1$. Let $b(n, q_n, j) = \binom{n}{j} q_n^j (1-q_n)^{n-j}$ denote the binomial probability mass

function. Then we have that

$$
\begin{aligned}
&E\left[X_i^{(n,j,k)} X_l^{(n,j,k)}\right] \\
=&P\left(X_i^{(n,j,k)} = 1, X_l^{(n,j,k)} = 1\right) && (3.32)\\
=&q_n^2 \left[b(n-2, q_n, j-1)\, b(n-2, q_n, k-1)\right]^2 && (3.33)\\
+&2q_n (1-q_n)\, b(n-2, q_n, j-1)\, b(n-2, q_n, k)\, b(n-2, q_n, j)\, b(n-2, q_n, k-1) && (3.34)\\
+&(1-q_n)^2 \left[b(n-2, q_n, j)\, b(n-2, q_n, k)\right]^2 . && (3.35)
\end{aligned}
$$

The first step (3.32) follows from the fact that $X_i^{(n,j,k)} X_l^{(n,j,k)}$ can only be non-zero if both random variables are equal to one. Therefore both nodes i and l must be of type (j,k). The final step above captures the relationship between the degrees of the two nodes. The following possible cases must be considered here:

1. Both directed edges from i to l and from l to i exist. This happens with probability q_n^2. For nodes i and l to be of degree (j,k), they must both have an additional $j-1$ and $k-1$ outgoing and incoming edges respectively which are connected to any of the remaining $n-2$ nodes in the network. The probability of this happening is given by (3.33).

2. Either one of the directed edges from i to l or from l to i exists, while the other one does not. This happens with probability $2q_n(q_n-1)$. The node whose outgoing edge to the other node exists, must have $j-1$ more outgoing edges connected to the remaining $n-2$ nodes, and k incoming edges connected to them. The other node must have j outgoing edges and $k-1$ incoming edges connected to the remaining $n-2$ nodes in the system. This is given by (3.34).

3. Neither of the two possible directed edges between nodes i and l exist. This happens with probability $(q_n-1)^2$. Here both nodes must have j outgoing and k incoming edges connected to the remaining $n-2$ nodes in the system. This probability is given by (3.35).

Since $q_n = \frac{\lambda}{n-1} \to 0$ and $(n-2)\,q_n = \frac{(n-2)\lambda}{n-1} \to \lambda$ as $n \to \infty$, all of the $b(n-2, q_n, \cdot)$ terms in (3.33) – (3.35) approach a Poisson probability mass function as $n \to \infty$. Therefore q_n^2 and $q_n(1-q_n)$ factors from (3.33) and (3.34) respectively, make the two terms approach zero. The $(1-q_n)^2$ factor from (3.35) approaches 1, and therefore the whole of (3.35)

approaches the product of Poisson probability mass functions, namely

$$\left(e^{-\lambda}\frac{\lambda^j}{j!}e^{-\lambda}\frac{\lambda^k}{k!}\right)^2 = \left[h(j,k)\right]^2, \tag{3.36}$$

where $h(j,k)$ is defined as in equation (3.30). Therefore we have that

$$E\left[X_i^{(n,j,k)}X_l^{(n,j,k)}\right] \to \left[h(j,k)\right]^2 \text{ as } n \to \infty. \tag{3.37}$$

From this, it follows that

$$\begin{aligned} cov\left(X_i^{(n,j,k)}, X_l^{(n,j,k)}\right) =& E\left[X_i^{(n,j,k)}X_l^{(n,j,k)}\right] - E\left[X_i^{(n,j,k)}\right]E\left[X_l^{(n,j,k)}\right] \\ =& E\left[X_i^{(n,j,k)}X_l^{(n,j,k)}\right] - \left[h_n(j,k)\right]^2 \\ \underset{\infty}{\overset{n}{\to}}& \left[h(j,k)\right]^2 - \left[h(j,k)\right]^2 = 0, \end{aligned} \tag{3.38}$$

which proves the proposition. □

Theorem 3.3.2. *Let $K_{\lambda,n}$ be an Erdős-Rényi network with n nodes, each with average degree $\lambda \in (0,\infty)$, and fix the integers $j,k \in \mathbb{N}_0$. If $\tilde{\mu}_n(j,k)$ is the sample proportion of nodes with degree (j,k) of the network $K_{\lambda,n}$ and $h(j,k) = e^{-2\lambda}\frac{\lambda^{j+k}}{j!k!}$, then for any $\epsilon > 0$*

$$P\left(|\tilde{\mu}_n(j,k) - h(j,k)| > \epsilon\right) \underset{\infty}{\overset{n}{\to}} 0. \tag{3.39}$$

Proof. Let $\epsilon > 0$ be given and fix the integers $j,k \in \mathbb{N}_0$. Further let $h_n(j,k)$ be defined as in equation (3.28). By the triangle inequality

$$\begin{aligned} &|\tilde{\mu}_n(j,k) - h(j,k)| \\ \leq\ &|\tilde{\mu}_n(j,k) - h_n(j,k)| + |h_n(j,k) - h(j,k)|. \end{aligned} \tag{3.40}$$

From (3.30) we know that $|h_n(j,k) - h(j,k)| \underset{\infty}{\overset{n}{\to}} 0$. Hence we can choose $N \in \mathbb{N}$ large enough so that $|h_n(j,k) - h(j,k)| < \frac{\epsilon}{2}$ for all $n \geq N$. Therefore for all $n \geq N$ we have that

$$\begin{aligned} &|\tilde{\mu}_n(j,k) - h(j,k)| \\ \leq\ &|\tilde{\mu}_n(j,k) - h_n(j,k)| + \frac{\epsilon}{2}. \end{aligned} \tag{3.41}$$

Using this inequality, it is seen that

$$P(|\tilde{\mu}_n(j,k) - h(j,k)| > \epsilon)$$

$$\leq P\left(|\tilde{\mu}_n(j,k) - h_n(j,k)| + \frac{\epsilon}{2} > \epsilon\right)$$
$$= P\left(|\tilde{\mu}_n(j,k) - h_n(j,k)| > \frac{\epsilon}{2}\right). \tag{3.42}$$

The next step is to show that $P\left(|\tilde{\mu}_n(j,k) - h_n(j,k)| > \frac{\epsilon}{2}\right) \to 0$ when $n \to \infty$. The proof of this is similar to the proof of the weak law of large numbers, with a slight modification to deal with the fact that $h_n(j,k)$ depends on n.

We know that $\tilde{\mu}_n(j,k)$ is the proportion of nodes with degree (j,k), where there are n nodes in total. Therefore it can be seen as the average of n Bernoulli trial outcomes, where the probability of success is $h_n(j,k)$. The probability of success is equal to $h_n(j,k)$ since this is the probability that a node will be of degree (j,k) when there are n nodes in the system. It then also follows that $E[\tilde{\mu}_n(j,k)] = h_n(j,k)$.

Let $X_1^{(n,j,k)}, X_2^{(n,j,k)}, \ldots, X_n^{(n,j,k)}$ be n Ber($h_n(j,k)$) random variables (note that the distribution of the random variables depends on n). The $X_i^{(n,j,k)}$'s are not independent of each other, since the degrees of one node affect the degrees of the nodes connected to it. They are, however, asymptotically mutually uncorrelated from proposition 3.3.1. Therefore $cov\left(X_i^{(n,j,k)}, X_l^{(n,j,k)}\right) \to 0$ for all $i \neq l$ when $n \to \infty$. and hence

$$\begin{aligned}
&var[\tilde{\mu}_n(j,k)] \\
=&var\left[\frac{1}{n}\sum_{i=1}^{n} X_i^{(n,j,k)}\right] \\
=&\frac{1}{n^2}\left(n\, var\left[X_i^{(n,j,k)}\right] + \left(n^2 - n\right) cov\left(X_i^{(n,j,k)}, X_l^{(n,j,k)}\right)\right),\ i \neq l \text{ since } X_i\text{'s are identically distributed} \\
=&\frac{1}{n}\left(var[X_i] + (n-1)\, cov\left(X_i^{(n,j,k)}, X_l^{(n,j,k)}\right)\right) \\
=&\frac{1}{n}\left(h_n(j,k) - (h_n(j,k))^2 + (n-1)\, cov\left(X_i^{(n,j,k)}, X_l^{(n,j,k)}\right)\right).
\end{aligned} \tag{3.43}$$

From Chebyshev's inequality, we have that

$$\begin{aligned}
&P\left(|\tilde{\mu}_n(j,k) - h_n(j,k)| > \frac{\epsilon}{2}\right) \\
\leq&\frac{4}{\epsilon^2} var[\mu_n(j,k)] \\
=&\frac{4}{\epsilon^2 n}\left(h_n(j,k) - (h_n(j,k))^2 + (n-1)\, cov\left(X_i^{(n,j,k)}, X_l^{(n,j,k)}\right)\right) \\
=&\frac{4}{\epsilon^2 n}\left(h_n(j,k) - (h_n(j,k))^2\right) + \frac{(n-1)}{n}\frac{4\, cov\left(X_i^{(n,j,k)}, X_l^{(n,j,k)}\right)}{\epsilon^2}.
\end{aligned} \tag{3.44}$$

If $n \to \infty$, then $h_n(j,k) \to h(j,k) = e^{-2\lambda}\frac{\lambda^{j+k}}{j!k!} < \infty$. Then since $\frac{4}{\epsilon^2 n} \to 0$ when $n \to \infty$, the first term above tends to 0. For the second term, $\frac{(n-1)}{n} \to 1$ and $\frac{4\,cov\left(X_i^{(n,j,k)}, X_l^{(n,j,k)}\right)}{\epsilon^2} \to 0$. Therefore

$$P\left(|\tilde{\mu}_n(j,k) - h_n(j,k)| > \frac{\epsilon}{2}\right) \underset{\infty}{\overset{n}{\to}} 0. \tag{3.45}$$

Since $\epsilon > 0$ was arbitrary, this shows that $\tilde{\mu}_n(j,k) \underset{\infty}{\overset{n}{\to}} h(j,k)$ in probability. □

Theorem 3.3.2 serves as a building block for similar results for a variant of the Erdős-Rényi graph where there are groups of connected Erdős-Rényi graphs. This will allow us to apply the results considered in this chapter to a very versatile type of network. This can be used to compare the risk implied by different types of network structure. Section 3.3.2 below defines the type of network that this study is focused on, and presents a result analogous to theorem 3.3.2.

3.3.2 Semi-heterogeneous Erdős-Rényi graphs

Suppose now that there are d groups of Erdős-Rényi networks that all interact with one another to form a new network of size n. Each group α comprises n_α nodes, so that $\sum_{\alpha=1}^{d} n_\alpha = n$. The probabilities of edges existing between any two nodes are predetermined based on the groups to which the nodes belong. Suppose a node i is in group α and a node j is in group β. The probability that the edge from i to j exists is denoted by $q_{\alpha\beta}^{(n)}$. The n in the superscript is included to indicate that the value of this probability will be dependent on the size of the network. It is shown below that this structure satisfies equation (3.39).

For each group α, let $w_\alpha = \frac{n_\alpha}{n}$ and assume that these remain constant when $n \to \infty$. Let $\lambda_{\alpha\beta}$ be the expected number of edges from any node in group α to nodes in group β. If $\alpha = \beta$ then a node in group α can connect to $n_\alpha - 1$ other nodes in the same group, this means that $q_{\alpha\alpha}^{(n)} = \frac{\lambda_{\alpha\alpha}}{n_\alpha - 1}$. If $\alpha \neq \beta$ then a node in group α can connect to n_β nodes in group β and therefore $q_{\alpha\beta}^{(n)} = \frac{\lambda_{\alpha\beta}}{n_\beta}$.

For the purpose of this research we will call this a semi-heterogeneous Erdős-Rényi graph. To formalize this, we have the following definitions:

Definition 3.3.1 (Average connection matrix)**.** For a group of $d \geq 1$ Erdős-Rényi graphs $K_{\lambda_1,n_1}, K_{\lambda_2,n_2}, \ldots, K_{\lambda_d,n_d}$ where there exists edges linking nodes between different graphs, the average connection matrix λ is defined to be the $d \times d$ matrix whose elements $\lambda_{\alpha\beta}$, $\alpha, \beta = 1, 2, \ldots, d$ represent the expected number of edges from any node in graph α to

nodes in graph β. The diagonal entries are given by $\lambda_{\alpha\alpha} = \lambda_\alpha$, $\alpha = 1, 2, \ldots, d$. An average connection matrix is said to be positive if $\lambda_{\alpha\beta} > 0$ for all α, β.

Definition 3.3.2 (Semi-heterogeneous Erdős-Rényi graph)**.** Let $d \in \mathbb{N}$ and consider the set $\{K_{\lambda_{11},n_1}, \ldots, K_{\lambda_{dd},n_d}\}$ of d Erdős-Rényi graphs with positive average connection matrix λ. If $\vec{n} = (n_1, \ldots, n_d)$ and the nodes from these graphs may be connected to one another such that $\lambda_{\alpha\beta}$ is the expected number of edges from any node in group α to nodes in group β, then we call the resulting graph a semi-heterogeneous Erdős-Rényi graph. This graph will be denoted by $K^d_{\lambda,\vec{n}}$.

Recall that $D^+_n(i)$ and $D^-_n(i)$ are the random variables representing the number of out- and in-degrees of a randomly chosen node i in the network. Now let $D^{+,\alpha}_n(i)$ and $D^{-,\alpha}_n(i)$ be the random variables representing the number of out- and in-degrees of any node i in group α. Similar to h^+_n and h^-_n which denote the respective probability mass functions of $D^+_n(i)$ and $D^-_n(i)$, we let $h^{+,\alpha}_n$ and $h^{-,\alpha}_n$ denote the probability mass functions of $D^{+,\alpha}_n(i)$ and $D^{-,\alpha}_n(i)$ respectively.

Now let h^α_n be the joint probability mass function of $h^{+,\alpha}_n$ and $h^{-,\alpha}_n$. Then $h^\alpha_n(j,k)$ is the probability that a randomly chosen node in group α (where the total network size is n) has j and k outgoing and incoming edges connected to it respectively. Then $h_n(j,k)$, the probability that any node i has degree (j,k), is given by

$$\begin{aligned} h_n(j,k) &= \sum_{\alpha=1}^{d} \frac{n_\alpha}{n} h^\alpha_n(j,k) \\ &= \sum_{\alpha=1}^{d} \frac{n_\alpha}{n} h^{+,\alpha}_n(j)\, h^{-,\alpha}_n(k)\,. \end{aligned} \tag{3.46}$$

Preliminary discussions

For $d = 2$ Erdős-Rényi networks First consider a semi-heterogeneous Erdős-Rényi graph $K^d_{\lambda,\vec{n}}$ where $d = 2$. For a preliminary discussion we will consider the expressions for $h_n(j,k)$ and its limit. When $d = 2$ we have four connection probabilities, namely $q^{(n)}_{11}$, $q^{(n)}_{12}$, $q^{(n)}_{21}$ and $q^{(n)}_{22}$.

For $d = 2$ groups, equation (3.46) can be written as follows:

$$\begin{aligned} h_n(j,k) &= \frac{n_1}{n}\, h^1_n(j,k) + \frac{n_2}{n}\, h^2_n(j,k) \\ &= \frac{n_1}{n}\, h^{+,1}_n(j)\, h^{-,1}_n(k) + \frac{n_2}{n}\, h^{+,2}_n(j)\, h^{-,2}_n(k)\,. \end{aligned} \tag{3.47}$$

To find an expression for $h_n^{+,1}(j)$, note that there are $n_1 - 1$ other nodes in group 1 to which a node in group 1 can connect. However there are n_2 nodes in group 2 to which a node in group 1 can connect. There can either be 0 of the j edges connected to group 1 nodes and j to group 2 nodes, or 1 edge connected to group 1 nodes and $j-1$ to group 2 nodes, etc. Therefore

$$\begin{aligned}
h_n^{+,1}(j) = &\binom{n_1-1}{0}\left(q_{11}^{(n)}\right)^0\left(1-q_{11}^{(n)}\right)^{n_1-1}\binom{n_2}{j}\left(q_{12}^{(n)}\right)^j\left(1-q_{12}^{(n)}\right)^{n_2-j} \\
&+\binom{n_1-1}{1}\left(q_{11}^{(n)}\right)^1\left(1-q_{11}^{(n)}\right)^{n_1-1-1}\binom{n_2}{j-1}\left(q_{12}^{(n)}\right)^{j-1}\left(1-q_{12}^{(n)}\right)^{n_2-(j-1)} \\
&+\cdots \\
&+\binom{n_1-1}{j}\left(q_{11}^{(n)}\right)^j\left(1-q_{11}^{(n)}\right)^{n_1-1-j}\binom{n_2}{0}\left(q_{12}^{(n)}\right)^0\left(1-q_{12}^{(n)}\right)^{n_2} \\
= &\sum_{l=0}^{j}\binom{n_1-1}{l}\left(q_{11}^{(n)}\right)^l\left(1-q_{11}^{(n)}\right)^{n_1-1-l}\binom{n_2}{j-l}\left(q_{12}^{(n)}\right)^{j-l}\left(1-q_{12}^{(n)}\right)^{n_2-(j-l)}. \qquad (3.48)
\end{aligned}$$

For notational convenience, let the binomial probability mass function be denoted by

$$b(n,q,l) = \binom{n}{l}q^l(1-q)^{n-l}. \qquad (3.49)$$

Then

$$h_n^{+,1}(j) = \sum_{l=0}^{j} b\left(n_1-1, q_{11}^{(n)}, l\right) b\left(n_2, q_{12}^{(n)}, j-1\right) \qquad (3.50)$$

and similarly,

$$\begin{aligned}
h_n^{-,1}(k) &= \sum_{l=0}^{k}\binom{n_1-1}{l}\left(q_{11}^{(n)}\right)^l\left(1-q_{11}^{(n)}\right)^{n_1-1-l}\binom{n_2}{k-l}\left(q_{21}^{(n)}\right)^{k-l}\left(1-q_{21}^{(n)}\right)^{n_2-(k-l)} \\
&= \sum_{l=0}^{k} b\left(n_1-1, q_{11}^{(n)}, l\right) b\left(n_2, q_{21}^{(n)}, k-l\right), \qquad (3.51) \\
h_n^{+,2}(j) &= \sum_{l=0}^{j}\binom{n_2-1}{l}\left(q_{22}^{(n)}\right)^l\left(1-q_{22}^{(n)}\right)^{n_2-1-l}\binom{n_1}{j-l}\left(q_{21}^{(n)}\right)^{j-l}\left(1-q_{21}^{(n)}\right)^{n_1-(j-l)} \\
&= \sum_{l=0}^{j} b\left(n_2-1, q_{22}^{(n)}, l\right) b\left(n_1, q_{21}^{(n)}, j-l\right) \qquad (3.52)
\end{aligned}$$

and

$$h_n^{-,2}(k) = \sum_{l=0}^{k} \binom{n_2-1}{l} \left(q_{22}^{(n)}\right)^l \left(1-q_{22}^{(n)}\right)^{n_2-1-l} \binom{n_1}{k-l} \left(q_{12}^{(n)}\right)^{k-l} \left(1-q_{12}^{(n)}\right)^{n_1-(k-l)}$$
$$= \sum_{l=0}^{k} b\left(n_2-1, q_{22}^{(n)}, l\right) b\left(n_1, q_{12}^{(n)}, k-l\right). \tag{3.53}$$

These expressions can then be substituted into the expression for $h_n(j,k)$ given by equation (3.47). For the expressions to make sense it is assumed, without loss of generality, that $j,k < n_1$ and $j,k < n_2$.

Note that since w_1 and w_2 remain constant as $n \to \infty$, then n_1 and n_2 tend to infinity at the same rate as n. This means that all of the $q_{\alpha\beta}^{(n)}$ probabilities converge to 0 as $n \to \infty$. Next it is shown that all of the factors making up equation (3.47) converge.

Consider first the expression for $h_n^{+,1}(j)$ given by (3.50). Since $(n_1-1)q_{11}^{(n)} = \lambda_{11}$ stays constant when n increases,

$$\binom{n_1-1}{l} \left(q_{11}^{(n)}\right)^l \left(1-q_{11}^{(n)}\right)^{n_1-1-l} \underset{\infty}{\overset{n}{\to}} e^{-\lambda_{11}} \frac{\lambda_{11}^l}{l!} \tag{3.54}$$

for $l = 1, 2, \ldots, j$ and

$$\binom{n_1-1}{l} \left(q_{11}^{(n)}\right)^l \left(1-q_{11}^{(n)}\right)^{n_1-1-l} \underset{\infty}{\overset{n}{\to}} e^{-\lambda_{11}} \tag{3.55}$$

if $l = 0$. The second factor in the summation for $h_n^{+,1}(j)$ converges to a similar expression since $n_2 q_{12}^{(n)} = \lambda_{12}$ is constant. In this case

$$\binom{n_2}{j-l} \left(q_{12}^{(n)}\right)^{j-l} \left(1-q_{12}^{(n)}\right)^{n_2-(j-l)} \underset{\infty}{\overset{n}{\to}} e^{-\lambda_{12}} \frac{\lambda_{12}^{j-l}}{(j-l)!} \tag{3.56}$$

for $l = 0, 1, \ldots, j-1$. If $l = j$, then

$$\binom{n_2}{j-l} \left(q_{12}^{(n)}\right)^{j-l} \left(1-q_{12}^{(n)}\right)^{n_2-(j-l)} \underset{\infty}{\overset{n}{\to}} e^{-\lambda_{12}}. \tag{3.57}$$

Since all of the above limits are finite, and the sum from equation (3.50) consists of a finite number of terms,

$$h_n^{+,1}(j) \underset{\infty}{\overset{n}{\to}} \sum_{l=0}^{j} e^{-\lambda_{11}} \frac{\lambda_{11}^l}{l!} e^{-\lambda_{12}} \frac{\lambda_{12}^{j-l}}{(j-l)!}$$

$$= e^{-\lambda_{11}-\lambda_{12}} \sum_{l=0}^{j} \frac{\lambda_{11}^{l}\lambda_{12}^{j-l}}{l!\,(j-l)!} \tag{3.58}$$

with the understanding that $\frac{\lambda_{11}^{l}}{l!}$ is replaced by 1 when $l = 0$ and $\frac{\lambda_{12}^{j-l}}{(j-l)!}$ is replaced by 1 when $l = j$.

Now consider equation (3.51) for $h_n^{-,1}(k)$. The convergence of the first factor in the summation has already been established. For the second factor, note that $n_2 q_{21}^{(n)} = n_2 \frac{\lambda_{21}}{n_1} = \frac{w_2}{w_1}\lambda_{21}$ is constant. Therefore

$$\binom{n_2}{k-l}\left(q_{21}^{(n)}\right)^{k-l}\left(1-q_{21}^{(n)}\right)^{n_2-(k-l)} \underset{\infty}{\overset{n}{\to}} e^{-\frac{w_2}{w_1}\lambda_{21}} \frac{\left(\frac{w_2}{w_1}\lambda_{21}\right)^{k-l}}{(k-l)!} \tag{3.59}$$

for $l = 0, 1, \ldots, k-1$, with $\frac{\left(\frac{w_2}{w_1}\lambda_{21}\right)^{k-l}}{(k-l)!}$ replaced by 1 if $l = k$. Therefore

$$\begin{aligned} h_n^{-,1}(k) &\underset{\infty}{\overset{n}{\to}} \sum_{l=0}^{k} e^{-\lambda_{11}} \frac{\lambda_{11}^{l}}{l!} e^{-\frac{w_2}{w_1}\lambda_{21}} \frac{\left(\frac{w_2}{w_1}\lambda_{21}\right)^{k-l}}{(k-l)!} \\ &= e^{-\lambda_{11}-\frac{w_2}{w_1}\lambda_{21}} \sum_{l=0}^{k} \frac{\lambda_{11}^{l}\left(\frac{w_2}{w_1}\lambda_{21}\right)^{k-l}}{l!\,(k-l)!} \end{aligned} \tag{3.60}$$

with the understanding that $\frac{\lambda_{11}^{l}}{l!}$ is replaced by 1 when $l = 0$ and $\frac{\left(\frac{w_2}{w_1}\lambda_{21}\right)^{k-l}}{(k-l)!}$ is replaced by 1 when $l = k$.

For $h_n^{+,2}(k)$ given by equation (3.52), we have the following:

$$\binom{n_2-1}{l}\left(q_{22}^{(n)}\right)^{l}\left(1-q_{22}^{(n)}\right)^{n_2-1-l} \underset{\infty}{\overset{n}{\to}} e^{-\lambda_{22}} \frac{\lambda_{22}^{l}}{l!} \tag{3.61}$$

and

$$\binom{n_1}{j-l}\left(q_{21}^{(n)}\right)^{j-l}\left(1-q_{21}^{(n)}\right)^{n_1-(j-l)} \underset{\infty}{\overset{n}{\to}} e^{-\lambda_{21}} \frac{\lambda_{21}^{j-l}}{(j-l)!}. \tag{3.62}$$

Therefore

$$\begin{aligned} h_n^{+,2}(k) &\underset{\infty}{\overset{n}{\to}} \sum_{l=0}^{j} e^{-\lambda_{22}} \frac{\lambda_{22}^{l}}{l!} e^{-\lambda_{21}} \frac{\lambda_{21}^{j-l}}{(j-l)!} \\ &= e^{-\lambda_{22}-\lambda_{21}} \sum_{l=0}^{j} \frac{\lambda_{22}^{l}\lambda_{21}^{j-l}}{l!\,(j-l)!}, \end{aligned} \tag{3.63}$$

with the appropriate adjustments in the summation when $l = 0$ and $l = j$. Finally, for equation (3.53) we have that

$$\binom{n_1}{k-l}\left(q_{12}^{(n)}\right)^{k-l}\left(1-q_{12}^{(n)}\right)^{n_1-(k-l)} \underset{\infty}{\overset{n}{\to}} e^{-\frac{w_1}{w_2}\lambda_{12}}\frac{\left(\frac{w_1}{w_2}\lambda_{12}\right)^{k-l}}{(k-l)!}, \tag{3.64}$$

and hence

$$\begin{aligned} h_n^{-,2}(k) &\underset{\infty}{\overset{n}{\to}} \sum_{l=0}^{k} e^{-\lambda_{22}}\frac{\lambda_{22}^{l}}{l!}e^{-\frac{w_1}{w_2}\lambda_{12}}\frac{\left(\frac{w_1}{w_2}\lambda_{12}\right)^{k-l}}{(k-l)!} \\ &= e^{-\lambda_{22}-\frac{w_1}{w_2}\lambda_{12}}\sum_{l=0}^{k}\frac{\lambda_{22}^{l}\left(\frac{w_1}{w_2}\lambda_{12}\right)^{k-l}}{l!\,(k-l)!}. \end{aligned} \tag{3.65}$$

Equations (3.58), (3.60), (3.63) and (3.65) can now be used to find $\lim_{n\to\infty} h_n(j,k)$ based on equation (3.47). In addition to this, $\frac{n_1}{n}$ and $\frac{n_1}{n}$ is replaced with w_1 and w_2 respectively, so that

$$\begin{aligned} \lim_{n\to\infty} h_n(j,k) =& h(j,k) \\ =& w_1\left(e^{-\lambda_{11}-\lambda_{12}}\sum_{l_1=0}^{j}\frac{\lambda_{11}^{l_1}\lambda_{12}^{j-l_1}}{l_1!\,(j-l_1)!}\right)\left(e^{-\lambda_{11}-\frac{w_2}{w_1}\lambda_{21}}\sum_{l_2=0}^{k}\frac{\lambda_{11}^{l_2}\left(\frac{w_2}{w_1}\lambda_{21}\right)^{k-l_2}}{l_2!\,(k-l_2)!}\right) \\ &+ w_2\left(e^{-\lambda_{22}-\lambda_{21}}\sum_{l_3=0}^{j}\frac{\lambda_{22}^{l_3}\lambda_{21}^{j-l_3}}{l_3!\,(j-l_3)!}\right)\left(e^{-\lambda_{22}-\frac{w_1}{w_2}\lambda_{12}}\sum_{l_4=0}^{k}\frac{\lambda_{22}^{l_4}\left(\frac{w_1}{w_2}\lambda_{12}\right)^{k-l_4}}{l_4!\,(k-l_4)!}\right) \\ =& w_1 e^{-2\lambda_{11}-\lambda_{12}-\frac{w_2}{w_1}\lambda_{21}}\left(\sum_{l_1=0}^{j}\frac{\lambda_{11}^{l_1}\lambda_{12}^{j-l_1}}{l_1!\,(j-l_1)!}\right)\left(\sum_{l_2=0}^{k}\frac{\lambda_{11}^{l_2}\left(\frac{w_2}{w_1}\lambda_{21}\right)^{k-l_2}}{l_2!\,(k-l_2)!}\right) \\ &+ w_2 e^{-2\lambda_{22}-\lambda_{21}-\frac{w_1}{w_2}\lambda_{12}}\left(\sum_{l_3=0}^{j}\frac{\lambda_{22}^{l_3}\lambda_{21}^{j-l_3}}{l_3!\,(j-l_3)!}\right)\left(\sum_{l_4=0}^{k}\frac{\lambda_{22}^{l_4}\left(\frac{w_1}{w_2}\lambda_{12}\right)^{k-l_4}}{l_4!\,(k-l_4)!}\right). \end{aligned} \tag{3.66}$$

This probability mass function has a finite mean, since the average number of outgoing edges connected to a randomly chosen node in the system is given by

$$\begin{aligned} &\frac{n_1}{n}\left[(n_1-1)\,q_{11}^{(n)}+n_2 q_{12}^{(n)}\right]+\frac{n_2}{n}\left[n_1 q_{21}^{(n)}+(n_2-1)\,q_{22}^{(n)}\right] \\ =& w_1\left(\lambda_{11}+\lambda_{12}\right)+w_2\left(\lambda_{21}+\lambda_{22}\right). \end{aligned} \tag{3.67}$$

For d Erdős-Rényi networks in general In this case we have d^2 connection probabilities $q_{\alpha\beta}^{(n)}$, where $\alpha, \beta = 1, 2, \ldots, d$. Recall that for d groups, $h_n(j,k)$ can be expressed as follows:

$$h_n(j,k) = \sum_{\alpha=1}^{d} w_\alpha h_n^{+,\alpha}(j)\, h_n^{-,\alpha}(k)\,. \tag{3.68}$$

In order to find general expressions for $h_n^{+,\alpha}(j)$ and $h_n^{-,\alpha}(k)$, we look at the different ways in which a node in group α can have j out-degrees and k in-degrees. Without loss of generality we assume that $n_\beta > j, k$ for $\beta = 1, 2, \ldots, d$. This is a reasonable assumption, since for $\beta = 1, 2, \ldots, d$ we have that $n_\beta \to \infty$ as $n \to \infty$. Then we have that for $d > 2$ and $\alpha < d$

$$\begin{aligned} h_n^{+,\alpha}(j) = \sum_{m_1=0}^{j} \sum_{m_2=0}^{j-m_1} \cdots \sum_{m_\alpha=0}^{j-m_1-\cdots-m_{\alpha-1}} \cdots \sum_{m_{d-1}=0}^{j-m_1-\cdots-m_{d-1}} & b\Big(n_1, q_{\alpha 1}^{(n)}, m_1\Big)\, b\Big(n_2, q_{\alpha 2}^{(n)}, m_2\Big) \\ & \cdots b\Big(n_{\alpha-1}, q_{\alpha,\alpha-1}^{(n)}, m_{\alpha-1}\Big)\, b\Big(n_\alpha - 1, q_{\alpha\alpha}^{(n)}, m_\alpha\Big)\, b\Big(n_{\alpha+1}, q_{\alpha,\alpha+1}^{(n)}, m_{\alpha+1}\Big) \\ & \cdots b\Big(n_{d-1}, q_{\alpha,d-1}^{(n)}, m_{d-1}\Big)\, b\Big(n_d, q_{\alpha d}^{(n)}, j - m_1 - \cdots - m_{d-1}\Big)\,, \end{aligned} \tag{3.69}$$

with a similar expression for $h_n^{-,\alpha}(k)$ and for the case $\alpha = d$.

Each of the terms above is a binomial probability mass function and will converge to a Poisson probability mass function as $n \to \infty$. Therefore $h_n^{+,\alpha}(j)$ and $h_n^{-,\alpha}(k)$ will converge for all groups α. Hence there exists an h such that $h_n(j,k) \xrightarrow[\infty]{n} h(j,k)$ for all $j, k \in \mathbb{N}$.

Preliminary results

We can now show that equation (3.39) is satisfied when we have any finite number of Erdős-Rényi graphs that interact with one another. Similar to proposition 3.3.1 we have the following result:

Proposition 3.3.3. *Suppose there are $d \geq 1$ groups of random Erdős-Rényi graphs, and let $q_{\alpha\beta}^{(n)}$, w_α, h_n^α, h^α and h be defined as before. For any given $j, k, n \in \mathbb{N}$, let $X_i^{(n,j,k)}$, $i = 1, 2, \ldots, n$ be the indicator random variable which is equal to 1 when node i is of type (j,k). Then $cov\left(X_i^{(n,j,k)}, X_l^{(n,j,k)}\right) \to 0$ when $n \to \infty$ for all nodes $i \neq l$.*

Proof. Let j, k and d be given. Recall that $h_n^\alpha(j,k)$ is the probability that a node in group α is of type (j,k), where the system is of size n. Analogous to this we let $h_n^\alpha(j,k,\beta)$ be the probability that a node in group α is of type (j,k), where one node in group β is disregarded,

and the system is treated as if it has $n-1$ nodes. Then for any two nodes $i \neq l$ we have that

$$\begin{aligned}
&E\left[X_i^{(n,j,k)} X_l^{(n,j,k)}\right] \\
&= \sum_{\alpha=1}^{d}\sum_{\beta=1}^{d} w_\alpha w_\beta P\Big(X_i^{(n,j,k)} = 1, X_l^{(n,j,k)} = 1 \mid \{\text{node } i \text{ is in group } \alpha\} \cap \{\text{node } l \text{ is in group } \beta\}\Big) \\
&= \sum_{\alpha=1}^{d}\sum_{\beta=1}^{d} w_\alpha w_\beta \Big[q_{\alpha\beta}^{(n)} q_{\beta\alpha}^{(n)} h_n^\alpha(j-1,k-1,\beta)\, h_n^\beta(j-1,k-1,\alpha) \\
&\quad + q_{\alpha\beta}^{(n)}\left(1-q_{\beta\alpha}^{(n)}\right) h_n^\alpha(j-1,k,\beta)\, h_n^\beta(j,k-1,\alpha) + \left(1-q_{\alpha\beta}^{(n)}\right) q_{\beta\alpha}^{(n)} h_n^\alpha(j,k-1,\beta)\, h_n^\beta(j-1,k,\alpha) \\
&\quad + \left(1-q_{\alpha\beta}^{(n)}\right)\left(1-q_{\beta\alpha}^{(n)}\right) h_n^\alpha(j,k,\beta)\, h_n^\beta(j,k,\alpha)\Big]. \qquad (3.70)
\end{aligned}$$

If $n \to \infty$ then for all $\alpha, \beta = 1, 2, \ldots, d$ we have that $q_{\alpha\beta}^{(n)} \to 0$ and that w_α remains constant. Therefore

$$\begin{aligned}
E\left[X_i^{(n,j,k)} X_l^{(n,j,k)}\right] &\to \sum_{\alpha=1}^{d}\sum_{\beta=1}^{d} w_\alpha w_\beta h^\alpha(j,k)\, h^\beta(j,k) \\
&= \left[h(j,k)\right]^2 \\
&= E\left[X_i^{(n,j,k)}\right] E\left[X_l^{(n,j,k)}\right], \qquad (3.71)
\end{aligned}$$

which proves the theorem. □

The following theorem can be deduced based on proposition 3.3.3:

Theorem 3.3.4. *Let $K_{\lambda,\vec{n}}^d$ be a semi-heterogeneous Erdős-Rényi graph. Let $n = \sum_{i=1}^d n_i$ and for $\alpha, \beta = 1, 2, \ldots, d$, let $q_{\alpha\beta}^{(n)}$, w_α, h_n^α, h^α and h be defined as before. If $\tilde{\mu}(j,k)$ is the proportion of nodes with degree (j,k), then for every $j,k \in \mathbb{N}_0$ and any $\epsilon > 0$*

$$P\left(|\tilde{\mu}_n(j,k) - h(j,k)| < \epsilon\right) \xrightarrow[\infty]{n} 0 \quad . \qquad (3.72)$$

In order to show that $\tilde{\mu}_n(j,k) \xrightarrow[\infty]{n} h(j,k)$ for all $j,k \in \mathbb{N}$, the same steps as for theorem 3.3.2 can be used in conjunction with proposition 3.3.3. In order to avoid repetitiveness, the proof is not written out again.

Assume now that after the links for any graph $\kappa_{\lambda,\vec{n}_i}^d$ have been determined, the exposure amounts of any node in group α are identically distributed random variables with distribution function F_α. It is further assumed that the non-zero exposure amounts of any two nodes are independent. Suppose that the fraction of initial defaults is π_0 (chosen uniformly over all the nodes), that these nodes have capital ratios equal to zero and that all other nodes

have capital ratios equal to $c > 0$. Then for all nodes i

$$\gamma(i) = \begin{cases} c \text{ with probability } 1 - \pi_0 \\ 0 \text{ with probability } \pi_0. \end{cases}$$

Recall that the function $p_n(j, k, \theta)$ represents the expected fraction of nodes with degree (j, k) that has a default threshold of θ. Therefore $p_n(j, k, 0) = \pi_0$. Similarly for $\alpha = 1, 2, \ldots, d$, let $p_n^\alpha(j, k, \theta)$ denote the expected fraction of nodes in group α with degree (j, k) that default after θ counterparties have defaulted. Note that $p_n(j, k, 0) = p_n^\alpha(j, k, 0) = \pi_0$.

Suppose therefore that $\theta > 0$. The fact that the order of default has not yet been determined can be ignored, as the exposures of nodes are i.i.d. within each group. Therefore $p_n^\alpha(j, k, \theta)$ is simply the probability that a node in group α with degree (j, k) will have total capital that is more than the total loss suffered on $\theta - 1$ of its exposures, but less than the loss suffered on θ exposures.

For a fixed $j \in \mathbb{N}$, let $X_1^\alpha, X_2^\alpha, \ldots, X_j^\alpha$, $\alpha = 1, 2, \ldots, d$ be d sequences of i.i.d. random variables, where F_α is the distribution function of the α^{th} sequence's random variables. Let $LGD = 1 - R$ denote the loss given default for any counterparty. Note that a node i can only have a default threshold greater than zero if $\gamma(i) = c < LGD$. Therefore since the capital ratios are independent of the exposures,

$$p_n^\alpha(j, k, \theta) = (1 - \pi_0)\, P\left(LGD\, X_\theta^\alpha > c \sum_{l=1}^{j} X_l^\alpha - LGD \sum_{m=1}^{\theta-1} X_m^\alpha > 0 \right), \tag{3.73}$$

with the appropriate adjustments whenever $\theta = 1$ and/or $j = 1$. By using Bayes' theorem we then have that

$$p_n(j, k, \theta) = \sum_{\alpha=1}^{d} p_n^\alpha(j, k, \theta)\, \frac{h^\alpha(j, k)\, w_\alpha}{h(j, k)}, \tag{3.74}$$

where $h^\alpha(j, k)$ and $h(j, k)$ are defined as before.

The term $c \sum_{l=1}^{j} X_l^\alpha$ in (3.73) is the total capital held by a node in group α with degree (j, k), as c represents the ratio of capital to total interbank assets (which is i.i.d. within each group for a fixed j). The total loss suffered by the default of $\theta - 1$ counterparties is given by $LGD \sum_{m=1}^{\theta-1} X_m^\alpha$. The event that a node in group α of degree (j, k) has not defaulted after $\theta - 1$ counterparty defaults is given by $\left\{ c \sum_{l=1}^{j} X_l^\alpha - LGD \sum_{m=1}^{\theta-1} X_m^\alpha > 0 \right\}$. For such a node to have default threshold θ, it must default after the next counterparty default. Therefore the remaining capital after $\theta - 1$ counterparty defaults must be less than the loss suffered on the θ^{th} default. This is the event $\left\{ LGD\, X_\theta^\alpha > c \sum_{l=1}^{j} X_l^\alpha - LGD \sum_{m=1}^{\theta-1} X_m^\alpha \right\}$. Equation (3.73)

depends on j and through the joint distribution of $X_1^\alpha, X_2^\alpha, \ldots, X_j^\alpha$, but does not depend on n and therefore $p(j,k,\theta) = p_n(j,k,\theta)$.

Asymptotic results for semi-heterogeneous Erdős-Rényi networks

For a network of size n and fixed j,k, let the random variable $\tilde{\mu}_n(j,k)$ be defined on the probability space $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ and assume that the capital ratios $\vec{\gamma}_n$ are given for each n. Define the mapping $\mathcal{H}_n : \Omega_n \to M_n$, where M_n is the set of all exposure matrices of size n and $\mathcal{H}_n(\omega_n) = \mathbf{e}_n$.

As in section 3.2, we let π^* be the smallest fixed point of the function $I\colon\ [0,1] \to [0,1]$, where

$$I(\pi) = \sum_{j,k} \frac{h(j,k)\,k}{\bar{\lambda}} \sum_{\theta=0}^{j} p(j,k,\theta)\,\bar{B}(j,\pi,\theta)\,, \tag{3.75}$$

and where $\bar{B}(j,\pi,\theta) = P(X \geq \theta) = \sum_{l\geq\theta}^{j} \binom{j}{l} \pi^l (1-\pi)^{j-l}$ denotes the survival function of a binomial random variable. It was shown in section 3.2 that since I is non-decreasing, then $\pi^* = \lim_{k\to\infty} I^k(0)$ where $I(0)$ is the fraction of initially defaulted nodes.

Theorem 3.3.5 below, which constructs a measure on the product space and is a special case of a theorem in [90], is used together with theorem 3.8 in [8] in order to prove theorem 3.3.6. Theorem 3.3.6 makes the results in [8] (which are based on deterministic in- and out-degree sequences) applicable to semi-heterogeneous Erdős-Rényi graphs where the in- and out-degree sequences are random. It shows that it is possible to find a subsequence of semi-heterogeneous Erdős-Rényi graphs for which the results in [8] hold almost surely.

Theorem 3.3.5 (Special case of theorem 2.4.4 in [90])**.** *For each $n \in \mathbf{N}$, let $(X_n, \mathcal{A}_n, P_n)$ be a probability space where X_n is a locally compact, σ-compact metric space with Borel σ-algebra $\mathcal{A}_n$. Then there exists a unique probability measure $P = \prod_{i=1}^{\infty} P_n$ on $(X, \mathcal{A}) := (\prod_{n=1}^{\infty} X_n, \prod_{n=1}^{\infty} \mathcal{A}_n)$ with the property that*

$$P\left(\prod_{n=1}^{\infty} U_n\right) = \prod_{n=1}^{\infty} P_n(U_n) \tag{3.76}$$

whenever $U_n \in \mathcal{A}_n$ for each $n \in \mathbf{N}$ and one has $U_n = X_n$ for all but finitely many of the n.

Theorem 3.3.6. *For each $n \in \mathbb{N}$ and $j,k \in \mathbb{N}_0$, let $\tilde{\mu}_n(j,k)$ denote the fraction of nodes with degree (j,k) in the semi-heterogeneous Erdős-Rényi graph $K^d_{\lambda,\vec{n}}$, where $\lambda = (\lambda_{il})$, $\bar{\lambda} = \sum_{i,l} w_i \lambda_{il}$ and $n = \sum_{i=1}^{d} n_i$. Then there exists a sequence $(n_m)_{m\geq 1}$ in $\mathbb{N}$ such that for any $\omega = (\omega_{n_m})_{m\geq 1}$ in the product space $\prod_{m=1}^{\infty} \Omega_{n_m}$, the corresponding sequence of exposure matrices $(\mathcal{H}_{n_m}(\omega_{n_m}))_{m\geq 1} = (\mathbf{e}_{n_m})_{m\geq 1}$ will satisfy the following with probability one:*

1. *If $\pi^* = 1$, i.e. if $I(\pi) > \pi$ for all $\pi \in [0, 1)$, then*

$$\alpha_{n_m}(\mathbf{E}_{n_m}, \vec{\gamma}_{n_m}) \to 1 \tag{3.77}$$

weakly as $m \to \infty$. In other words, almost all nodes default as the network size goes to infinity.

2. *If $\pi^* < 1$, and π^* is a stable fixed point of I (i.e. $I'(\pi^*) < 1$), then*

$$\alpha_{n_m}(\mathbf{E}_{n_m}, \vec{\gamma}_{n_m}) \to \sum_{j,k} h(j,k) \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j, \pi^*, \theta) \tag{3.78}$$

weakly as $m \to \infty$. This is then the asymptotic fraction of defaults as the network size tends to infinity.

Proof. Fix $j, k \in \mathbb{N}$ and let $(\Omega, \mathcal{F}) = (\prod_{i=1}^{\infty} \Omega_i, \prod_{i=1}^{\infty} \mathcal{A}_i)$. Then for each $i \in \mathbb{N}$, define the projection $\Pi_i \colon \Omega \to \Omega_i$ by $(x_1, x_2, \dots) \mapsto x_i$. From theorem 3.3.5, there exists a unique probability measure $\mathbb{P}$ on Ω such that if $\tilde{\mu}(j,k) = (\tilde{\mu_1}(j,k), \tilde{\mu_2}(j,k), \dots)$ is a random variable on $(\Omega, \mathcal{F})$, then for all $\epsilon > 0$

$$\mathbb{P}(|\Pi_n \tilde{\mu}(j,k) - h(j,k)| > \epsilon) = \mathbb{P}_n(|\tilde{\mu}_n(j,k) - h(j,k)| > \epsilon) \to 0. \tag{3.79}$$

The left-hand side of equation (3.79) follows from theorem 3.3.5, and the convergence from Proposition 3.3.3. This shows that $\Pi_n \tilde{\mu}(j,k) \to h(j,k)$ in probability. Therefore there exists a subsequence $n_1, n_2, \dots$ such that $\Pi_{n_k} \tilde{\mu}(j,k) \to h(j,k)$ almost surely.

Now let χ_C denote the indicator function of the set C. In a system of size n, the number of nodes with degree (j,k) can be expressed as

$$\frac{\sum_{i=1}^{n} \left[(D_n^+(i))^2 + (D_n^-(i))^2 \right] \chi_{\{D_n^+(i)=j\}} \chi_{\{D_n^-(i)=k\}}}{j^2 + k^2}, \tag{3.80}$$

where D_n^+ and D_n^- are defined as before. Hence

$$\tilde{\mu}_n(j,k) = \frac{1}{n} \frac{\sum_{i=1}^{n} \left[(D_n^+(i))^2 + (D_n^-(i))^2 \right] \chi_{\{D_n^+(i)=j\}} \chi_{\{D_n^-(i)=k\}}}{j^2 + k^2}. \tag{3.81}$$

Therefore

$$\frac{1}{n_m} \sum_{i=1}^{n_m} \left[\left(D_{n_m}^+(i) \right)^2 + \left(D_{n_m}^-(i) \right)^2 \right] = \sum_{j,k} \tilde{\mu}_{n_m}(j,k) \left(j^2 + k^2 \right)$$

$$\to \sum_{j,k} h(j,k)\left(j^2 + k^2\right) < \infty \tag{3.82}$$

almost surely, since h is the joint probability mass function of two random variables with finite second moments. Now we have that

$$\sum_{i=1}^{n_m} \left[\left(D^+_{n_m}(i) \right)^2 + \left(D^-_{n_m}(i) \right)^2 \right] = O(n_m) \tag{3.83}$$

almost surely. By theorem 3.8 in [8] we now have that for any $(\omega_{n_1}, \omega_{n_2}, \dots) \in \prod_{m=1}^{\infty} \Omega_{n_m}$ the corresponding sequence of exposure matrices $(\mathcal{H}_{n_m}(\omega_{n_m}))_{m\geq 1} = (\mathbf{e}_{n_m})_{m\geq 1}$ satisfies equation (3.77) and (3.78).

□

Section 3.4 now deals with illustrating theorem 3.3.6 and shows how semi-heterogeneous Erdős-Rényi graphs can be used to compare different types of network structures.

3.4 Application to stochastic, heterogeneous financial networks

3.4.1 Illustration of theoretical results

A simple Erdős-Rényi structure is used for this section as the computational inefficiencies of evaluating large networks are exasperated when dealing with multiple Erdős-Rényi networks that interact with one another. Two cases are considered regarding the non-zero exposures of each bank. The first case is where all exposures are assumed to be equal, and the second is where the positive exposures are assumed to be exponentially distributed with parameter η. This keeps the function $p(j,k,\theta)$ mathematically tractable while ensuring that counterparty exposures remain positive.

Calculating the analytical weak limits

Recall that theorem 3.3.6 provides weak limits for $\alpha(\mathbf{E}_n, \vec{\gamma}_n)$, the final fraction of defaults under different circumstances. The first step in determining the weak limit is to find the value of π^*, so that the appropriate case can be chosen above. Recall that π^* is defined by

$\pi^* = \inf\{\pi \in [0,1] \mid I(\pi) = \pi\}$, where $I(\pi)$ is given by equation (3.14) as

$$I(\pi) = \sum_{j,k} \frac{\mu(j,k)\,k}{\lambda} \sum_{\theta=0}^{j} p(j,k,\theta)\,\bar{B}(j,\pi,\theta)\,. \tag{3.84}$$

It was shown through proposition 3.2.2 that in fact $\pi^* = \lim_{k\to\infty} I^k(0)$, where $I(0) = \pi_0$ represents the fraction of initially defaulted nodes.

The calculation of $I(\pi)$ and the limit for the second case of theorem 3.3.6 requires computation of the functions $\bar{B}(j,\pi,\theta)$, $\mu(j,k)$ and $p(j,k,\theta)$. Recall that the $\bar{B}$ function is merely the binomial distribution survival function (see equation (3.13)), and hence is easily calculated. Consider now the functional form of $\mu(j,k)$.

The average in- and out-degrees are kept fixed at $\lambda < \infty$ for different network sizes. Now let j and k be fixed. From theorem 3.3.2 we know that the proportion of nodes that have degree (j,k) satisfies $P\left(|\tilde{\mu}_n(j,k) - h(j,k)| > \epsilon\right) \xrightarrow[\infty]{n} 0$, where $h(j,k) = e^{-2\lambda}\frac{\lambda^{j+k}}{j!k!}$.

Finally, consider the functional form of $p(j,k,\theta)$. Assume that the exposure sizes of all nodes (irrespective of their in- and out-degrees) follow an exponential distribution with parameter η. Further assume that any node i is included in the set of initially defaulted nodes with probability π_0. Initially defaulted nodes will have capital ratios of $\gamma(i) = 0$ and the rest will have capital ratios of, say, c where $c > 0$. Therefore $p(j,k,0) = \pi_0$ and for $\theta > 0$, equation (3.73) shows that

$$p(j,k,\theta) = (1-\pi_0)\,\mathbb{P}\left(LGD\,X_\theta > c\sum_{l=1}^{j} X_l - LGD\sum_{m=1}^{\theta-1} X_m > 0\right), \tag{3.85}$$

where $X_1, X_2, \ldots, X_j$ are i.i.d. $\exp(\eta)$ random variables and LGD is the loss given default ratio for all counterparties. The term $LGD\sum_{m=1}^{\theta-1} X_m$ is omitted whenever $\theta = 1$. Note that we must have $LGD > c$, since otherwise no node will ever default apart from the set of initially defaulted nodes.

Note that the right-hand factor in equation (3.85) is difficult to calculate directly since the summations are not independent. In order to address this, we need to consider different values that j and θ can take. The different cases considered below are $j > \theta > 1$; $j > 1$ and $\theta = 1$; and $j = 1$ with $\theta = 0$ or $\theta = 1$.

Assume that $j > \theta > 1$:

The probability on the right hand side of equation (3.85) can then be rewritten as fol-

lows:

$$\mathbb{P}\left(LGD\,X_\theta > c\sum_{l=1}^{j} X_l - LGD\sum_{m=1}^{\theta-1} X_m > 0\right) \tag{3.86}$$

$$=\mathbb{P}\left(c\sum_{l=1}^{j} X_l - LGD\sum_{m=1}^{\theta-1} X_m < LGD\,X_\theta\right) - \mathbb{P}\left(c\sum_{l=1}^{j} X_l - LGD\sum_{m=1}^{\theta-1} X_m \leq 0\right)$$

$$=\mathbb{P}\left(c\sum_{l=1}^{j} X_l - LGD\sum_{m=1}^{\theta} X_m < 0\right) - \mathbb{P}\left(c\sum_{l=1}^{j} X_l - LGD\sum_{m=1}^{\theta-1} X_m \leq 0\right)$$

$$=\mathbb{P}\left(c\sum_{l=\theta+1}^{j} X_l - (LGD-c)\sum_{m=1}^{\theta} X_m < 0\right) \tag{3.87}$$

$$-\,\mathbb{P}\left(c\sum_{l=\theta}^{j} X_l - (LGD-c)\sum_{m=1}^{\theta-1} X_m \leq 0\right). \tag{3.88}$$

For each of the above probabilities (3.87) and (3.88), the two summations are now independent of one another. Furthermore since the X_i's are independent exp(η) distributed random variables, the summations above will have gamma distributions.

First we consider (3.87) above and determine the distribution of each of the two summations. The random variable cX_l has an $\exp\left(\frac{\eta}{c}\right)$ distribution. Therefore $\sum_{l=\theta+1}^{j} cX_l$ will have a $\text{gamma}\left(j-\theta, \frac{c}{\eta}\right)$ distribution. Similarly $(LGD-c)\sum_{m=1}^{\theta} X_m$ will have a $\text{gamma}\left(\theta, \frac{LGD-c}{\eta}\right)$ distribution. Regarding the summations in (3.88), we have that $c\sum_{l=\theta}^{j} X_l$ has a $\text{gamma}\left(j-\theta+1, \frac{c}{\eta}\right)$ distribution and $(LGD-c)\sum_{m=1}^{\theta-1} X_m$ has a $\text{gamma}\left(\theta-1, \frac{LGD-c}{\eta}\right)$ distribution. Note that $LGD - c > 0$ and therefore the parameters for the gamma distributions will not become negative.

Since the summation within each of the probabilities (3.87) and (3.88) independent, they can be calculated by means of integration. To illustrate this procedure, suppose that Y and Z are independent random variables with probability mass functions f_Y and f_Z respectively. Then

$$\mathbb{P}(Z - Y \leq 0) = \int_{-\infty}^{\infty}\int_{-\infty}^{y} f_Z(z)\, f_Y(y)\, dz\, dy. \tag{3.89}$$

To calculate (3.87), we let Z and Y have $\text{gamma}\left(j-\theta, \frac{c}{\eta}\right)$ and $\text{gamma}\left(\theta, \frac{LGD-c}{\eta}\right)$ distributions respectively. For (3.88), Z and Y will have $\text{gamma}\left(j-\theta+1, \frac{c}{\eta}\right)$ and $\text{gamma}\left(\theta-1, \frac{LGD-c}{\eta}\right)$ distributions respectively, which then allows one to calculate $p(j,k,\theta)$ for $j > \theta > 1$.

Assume that $\theta = 1$ and $j > 1$:

For a node to have a default threshold of one, it must default when its very first coun-

terparty defaults. Therefore (3.86) will become

$$\mathbb{P}\left(LGD\, X_1 - c\sum_{l=1}^{j} X_l > 0\right) = \mathbb{P}\left((LGD - c)\; X_1 - c\sum_{l=2}^{j} X_l > 0\right)$$
$$= \mathbb{P}\left(c\sum_{l=2}^{j} X_l - (LGD - c)\; X_1 < 0\right), \tag{3.90}$$

where $c\sum_{l=2}^{j} X_l \sim \text{gamma}\left(j-1, \frac{c}{\eta}\right)$ and $(LGD - c)\, X_1 \sim \exp\left(\frac{\eta}{LGD-c}\right)$. The probability can then be calculated using (3.89).

Assume that $j = 1$:

If $j = 1$, then either $\theta = 0$ or $\theta = 1$. The case $\theta = 0$ has already been discussed (recall that $p(j, k, 0) = \pi_0$) and therefore $p\,(1, k, 1) = 1 - \pi_0$.

Suppose now that $j = \theta > 1$. Equation (3.86) then becomes

$$\begin{aligned}
&\mathbb{P}\left(LGD\, X_j > c\sum_{l=1}^{j} X_l - LGD\sum_{m=1}^{j-1} X_m > 0\right)\\
=&\mathbb{P}\left(c\sum_{l=1}^{j} X_l - LGD\sum_{m=1}^{j-1} X_m < LGD\, X_j\right) - \mathbb{P}\left(c\sum_{l=1}^{j} X_l - LGD\sum_{m=1}^{j-1} X_m \leq 0\right)\\
=&\mathbb{P}\left((c - LGD)\sum_{l=1}^{j} X_l < 0\right) - \mathbb{P}\left(cX_j - (LGD - c)\sum_{l=1}^{j-1} X_l \leq 0\right)\\
=&1 - \mathbb{P}\left(cX_j - (LGD - c)\sum_{l=1}^{j-1} X_l \leq 0\right),
\end{aligned} \tag{3.91}$$

where $cX_j \sim \exp\left(\frac{\eta}{c}\right)$ and $(LGD - c)\sum_{l=1}^{j-1} X_l \sim \text{gamma}\left(j-1, \frac{LGD-c}{\eta}\right)$.

Now that it is possible to calculate $\mu(j, k)$ and $p(j, k, \theta)$, it is possible to calculate the weak limit of $\alpha(\mathbf{E}_n, \vec{\gamma}_n)$ as given by theorem 3.3.6. This allows us to compute the analytical part of the computations. What is left is to describe the steps taken to produce the simulated results, including the chosen parameter values.

Description of the simulation method

This section broadly describes the coding algorithm used to obtain the simulated results. For a given system, let N denote the size of the system, i.e. the number of nodes in the system. Since the results relate to the weak convergence of random variables as $N \to \infty$, the simulations need to be performed for increasing values of N. For each value of N, the

following steps are performed:

1. The $N \times N$ matrix containing the probabilities of any node being connected to another is determined[4]. The entry (i, j) corresponds to the probability that a directed edge exists from node i to node j. In this case, all the non-diagonal entries are equal to $\frac{\lambda}{N-1}$, and all the diagonal entries are zero.

2. Now the edges of the network are determined randomly based on the probabilities determined in the previous step. The existence or otherwise of the possible edges in the system are determined independently from one another. This gives the adjacency matrix consisting of ones and zeros, where the entry (i, j) is equal to one if there is a directed edge from i to j and zero otherwise.

   The rest of the steps below are now repeated for each new simulation of the adjacency matrix.

3. Now that the connections between the banks have been determined, the amounts of all the interbank loans (i.e. the weights of the edges) can be determined. The amount of each interbank loan is independently drawn from the exponential distribution (note that all parameter values, including the exponential distribution's parameter, are given further along in this section.

4. Recall from section 3.3.2 that we assumed that any node i is included in the set of initially defaulted nodes with probability π_0. In this step the set of initially insolvent nodes is determined by randomly selecting a set of $N \cdot \pi^*$ nodes, rounded to the nearest integer.

5. The capital of all nodes can now be determined. The set of initially insolvent nodes start off with capital values of zero. The remaining nodes will have capital values determined by the total monetary amount of interbank loans that they have issued. This is a direct consequence of the fact that capital ratios are kept constant, and that the capital value is determined by multiplying the total interbank assets by the capital ratio. When applying this model, the capital levels should be determined in a more realistic way. However for the purpose of this section, this capital ratio assumption will suffice.

[4]For the Erdős-Rényi case it is not necessary to use a matrix, as the connection probabilities between two nodes $i \neq j$ are equal. A matrix is used here to make the rest of the code easily adaptable for cases where the probabilities may differ.

At this stage it is important to note that not all banks will necessarily always have interbank exposures. These banks have to be given artificial capital values, since otherwise their capital will be wrongly set to zero without them having defaulted. The code assigns the average capital value between the non-defaulted banks with positive interbank exposures to them. Note that the assumption regarding these banks' capital values will not influence the results.

6. The counterparty losses resulting from the initially insolvent nodes can be determined now. For each loan issued to an insolvent node, the issuer loses the amount of the loan multiplied by the assumed *LGD* level. Any negative capital values are set equal to zero. Now the program finds all nodes with depleted capital, but which has not defaulted before. In other words, it finds all the newly defaulted nodes.

7. While the set of newly defaulted nodes is not empty, the same procedure as in step 6 is repeated for all the newly defaulted nodes.

   When there are no more newly defaulted nodes, calculate the average number of defaulted nodes which includes the set of initially insolvent nodes.

8. Based on the initial exposure matrix, a new random exposure matrix needs to be obtained in line with definition 3.1.2. The new exposure matrix must yield the same in- and out-degrees for all the nodes, but each node's existing exposures must be shuffled. In order words, each row of the exposure matrix must be shuffled randomly in such a way that the diagonals remain zero and the number of non-zero entries in each column remain the same as before[5].

   For large systems, it is not computationally feasible to obtain an enumeration of all possible exposure matrices that satisfy the above requirement. The number of possible combinations of rows simply becomes too large. For this reason the following procedure was followed:

   (a) Remove the diagonals from the current exposure matrix so that all entries in each row can be shuffled.

   (b) Within each row, shuffle the entries randomly and afterwards include the diagonal entries again.

[5]The number of non-zero entries in each row will clearly remain the same, since they are just a re-shuffling of each original row.

(c) Calculate the resulting in-degrees of the nodes and compare them with the in-degrees of the original exposure matrix. If the in-degrees don't match, do the following:

i. Find the columns of the new exposure matrix that correspond to in-degrees that are greater than those required by the original exposure matrix. Similarly, find the columns of thee new exposure matrix that correspond to in-degrees that are less than those based on the original exposure matrix.

ii. Now do the following for each column j whose resulting in-degrees are too many:

A. Find all the non-zero entries of column j. Some of these will need to be moved to other columns (within the same row) in order to reduce the in-degrees of bank j. First determine the number of entries that need to be moved so that bank j will have the correct in-degree, and then randomly choose the non-zero entries in column j that will be moved.

For each of the entries that were chosen to be moved, do the following:

B. Determine the set of possible places to which the entry may be moved. It must be moved to a column that has an in-degree that is too low, must remain in the same row, may not be moved to a diagonal position and cannot move to a position where there is already a non-zero entry. If there are no possible places to which an entry may be moved, do nothing for now. Otherwise choose a random position out of the eligible positions to move the entry to, and update the random exposure matrix accordingly. Now calculate the in-degrees of the updated matrix and redetermine which columns have in-degrees that are too few. This is to ensure that a column which initially had too few non-zero entries do not receive too many non-zero entries later on.

iii. Once steps (i) and (ii) have been performed for all columns with too many non-zero entries, test whether the in-degrees of the resulting random exposure matrix are correct, and break the loop if it is. Otherwise repeat steps (d) and (e) until the in-degrees are correct.

(d) If these steps cannot produce the correct in-degrees, begin from step (a) again. Repeat this until a suitable exposure matrix is found.

9. Use the new random exposure matrix and repeat the wholes process from step 5

again. Note that the original set of initially insolvent nodes is retained.

10. Up to now the steps have all been based on one sample of an Erdős-Rényi graph. In order to determine whether the results hold in probability for a random Erdős-Rényi graph, the whole procedure from step 2 needs to be repeated for different samples of the random graph.

Once the above procedure has been determined for increasing values of N, we can calculate the following probability empirically:

$$\mathbb{P}\left(\left|\alpha_n(\mathbf{E}_n, \vec{\gamma}_n) - \alpha_0\right| < \epsilon\right) \tag{3.92}$$

where $\epsilon > 0$ and $\alpha_0 = \sum_{j,k} h(j,k) \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j, \pi^*, \theta)$. The value of $\alpha_n(\mathbf{E}_n, \vec{\gamma}_n)$ is determined via simulation for increasing values of n, and α_0 is determined analytically.

Comparison of the analytical and simulation results

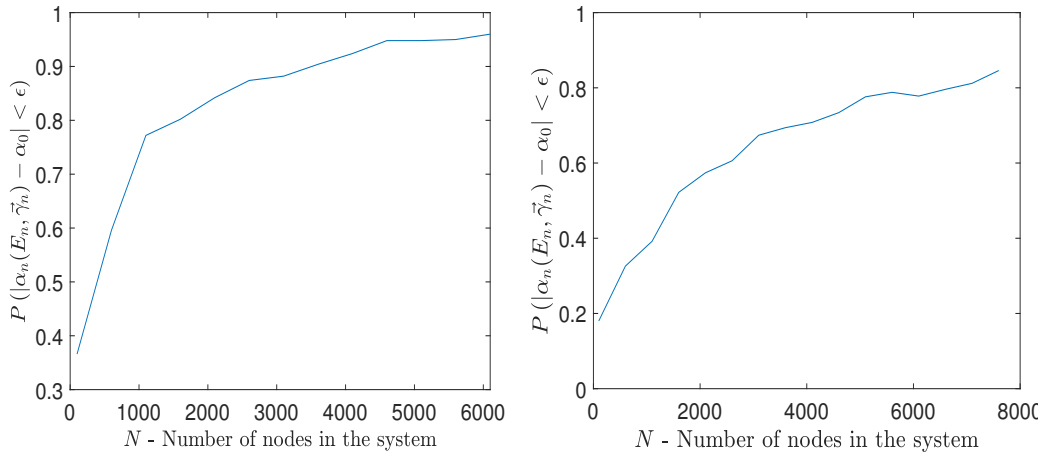
Table 3.2 contains the parameters used for the purpose of this illustration.

Table 3.2: The parameter values used for illustrating the convergence given by theorem 3.3.6.

Parameter	Description	Parameter value
γ	Ratio of interbank assets to capital	0.4
π_0	Initial fraction of defaults	0.05
$\bar{\lambda}$	Average out-degree/in-degree of the system	4
η	Mean exposure amount	1
ϵ	Error term used for evaluating equation (3.92)	0.025

Figure 3.2 shows how equation (3.92) moves closer to one for increasing values of N, which supports the conclusion of theorem 3.3.6. For equal exposures (figure 3.2a), convergence is achieved much faster than for random exposure amounts (figure 3.2b). This is expected, since there is less variation between nodes in the network. Similarly, convergence is expected to be slower when groups of interacting Erdős-Rényi graphs are considered.

Note that the fraction of defaults based on simulation results only start to converge to the theoretical quantity α_0 for very large values of n, which may not be attained in a practical setting. For example, the German banking system had approximately 1,800 banks as at 2014 [21]. In this case figure 3.2 implies that the theoretical results would be close to the observed fraction of defaults approximately 80% of the time for equal exposures, and



(a) Equal exposures

(b) Exponentially distributed i.i.d. exposures

Figure 3.2: Illustrating the convergence as given by theorem 3.3.6.

approximately 40% of the time for exponential i.i.d. exposures. However, computationally it is still much more efficient than simulation methods to assess sensitivities resulting from changes to combinations of network and bank characteristics. Furthermore, analytical results are often useful tools for understanding complex systems as these can assist in understanding the underlying components before carrying out simulations. Results for large financial networks are also of interest in cases where multiple countries and/or multiple types of financial institutions are considered. In these cases the number of nodes in the network can increase significantly, making an asymptotic approach appropriate.

3.4.2 Applying the results to different network structures

For this section we consider a semi-heterogeneous Erdős-Rényi graph $K^d_{\lambda,\vec{n}}$ with $d = 2$ groups of connected Erdős-Rényi graphs, even though the theory presented in this section can deal with any finite number of interacting graphs. This will be used to compare different network structures based on the matrix λ. In the setting discussed in section 3.3.2, the groups can be determined in any way as long as the exposure amounts satisfy requirements lined out in [8]. In our case we assume that banks are grouped according to size in order to relate this section to banking systems commonly found in practice. Hence we let group one consist of a small fraction w_1 of large banks and group two of a larger fraction w_2 of small banks.

We will assume that the non-zero exposure amounts follow an exponential distribution with means η_1 and η_2 for groups one and two respectively. The exponential distribution

is used because of its analytical tractability which facilitates the calculation of $p(j,k,\theta)$. However, in practice it would make sense to use a truncated distribution for the exposures since a bounded support is more realistic for balance sheet figures.

Recall now that the total asset value of each bank does not feature in any of the results that this study considers. Therefore in this stylized setting we will assume that banks that generally have large counterparty exposures have high asset values and vice versa for banks with lower exposure amounts. This is equivalent to assuming that loans granted by large banks are generally larger than any loans granted by small banks, which is a reasonable assumption to make. Therefore we must have that $\eta_1 > \eta_2$ and these parameters will be used to differentiate between banks of different size.

The matrix λ will in turn be used to differentiate between different network structures by varying the level of interconnectedness between the different groups and within each group. In order to make the structures comparable it is assumed that the average out-degree (or equivalently the in-degree) of a randomly chosen node in each type of network is a fixed quantity $\bar{\lambda}$. Three network structures will be compared to one another based on the structures introduced in section 2.2.1. These structures together with their connection probabilities in the case of a finite network of size n are as follows:

(i) Standard Erdős-Rényi graph, with $q_{11}^{(n)} = q_{12}^{(n)} = q_{21}^{(n)} = q_{22}^{(n)} = \frac{\bar{\lambda}}{n-1}$.

(ii) Tiered type I - Large banks are the most likely to be exposed to one another and small banks less likely to be exposed to one another. The probability that small banks and large banks are exposed to one another is in between the former two probabilities. The probabilities are given by $q_{ij}^{(n)} = \frac{\eta_i+\eta_j}{2\eta_1} L_2^{(n)}$.

(iii) Tiered type II - Large banks have a relatively high probability of lending to any other bank, small banks have a smaller probability of lending to large banks and the probability of small banks lending to one another is the least. Here we have that $q_{ij}^{(n)} = \frac{\eta_i+\eta_j+\max\{\eta_i-\eta_j,0\}}{3\eta_1} L_3^{(n)}$.

The $L_m^{(n)}$ quantities in the formulae above are adjustment factors that ensure that the structures exhibit the required average out-degree $\bar{\lambda}$ and that the connection probabilities are functions of n that tend to zero as $n \to \infty$. It is noted that the structures and formulae chosen are used for illustrative purposes and for investigating how network structure may affect systemic risk. Hence they are not necessarily the most realistic structures for banking systems. However, the second and third structures are representative of core-peripheral networks which explicitly place larger banks in the tightly connected core, and therefore contain elements of structures found in practice.

It now remains to determine the matrix $\lambda = \left(\lambda_{ij}\right)$ for each network structure based on the above probabilities so that theorem 3.3.6 can be applied. Note that in the case of a finite network of size n, the average out-degree (or in-degree) of a node in the network would be given by

$$\bar{\lambda} = w_1 (n_1 - 1) q_{11}^{(n)} + w_1 w_2 n q_{12}^{(n)} + w_1 w_2 n q_{21}^{(n)} + w_2 (n_2 - 1) q_{22}^{(n)}. \tag{3.93}$$

This equation will be used to determine the functional form of $L_m^{(n)}$, $m = 1, 2, 3$, so that we can find $\lambda_{ii} = \lim_{n\to\infty} (n_i - 1) q_{ii}^{(n)}$, $i = 1, 2$ and $\lambda_{ij} = \lim_{n\to\infty} w_j n q_{ij}^{(n)}$, $i \neq j$.

1. Erdős-Rényi This structure is straightforward, since $\lambda_{ij} = \bar{\lambda} w_j$ for $i, j = 1, 2$.

2. Tiered type I For this structure we have that $q_{11}^{(n)} = L_2^{(n)}$, $q_{12}^{(n)} = q_{21}^{(n)} = \frac{\eta_1+\eta_2}{2\eta_1} L_2^{(n)}$ and $q_{22}^{(n)} = \frac{\eta_2}{\eta_1} L_2^{(n)}$. Using equation (3.93) it can be seen that

$$\begin{aligned}\bar{\lambda} &= w_1 (n_1 - 1) q_{11}^{(n)} + 2 w_1 w_2 n q_{12}^{(n)} + w_2 (n_2 - 1) q_{22}^{(n)} \\ &= L_2^{(n)} \left[w_1 (n_1 - 1) + w_1 w_2 n \frac{\eta_1 + \eta_2}{\eta_1} + w_2 (n_2 - 1) \frac{\eta_2}{\eta_1} \right]\end{aligned} \tag{3.94}$$

so that

$$L_2^{(n)} = \bar{\lambda} \frac{\eta_1}{\eta_1 w_1 (w_1 n - 1) + w_1 w_2 n (\eta_1 + \eta_2) + w_2 \eta_2 (w_2 n - 1)}. \tag{3.95}$$

Hence

$$\begin{aligned}\lambda_{11} &= \lim_{n\to\infty} (n_1 - 1) L_2^{(n)} \\ &= \lim_{n\to\infty} \bar{\lambda} \frac{\eta_1 (n_1 - 1)}{\eta_1 w_1 (w_1 n - 1) + w_1 w_2 n (\eta_1 + \eta_2) + 2 w_2 \eta_2 (w_2 n - 1)} \\ &= \bar{\lambda} \frac{\eta_1 w_1}{\eta_1 w_1^2 + w_1 w_2 (\eta_1 + \eta_2) + \eta_2 w_2^2},\end{aligned} \tag{3.96}$$

and in general $\lambda_{ij} = \bar{\lambda} \frac{(\eta_i + \eta_j) w_j}{2\eta_1 w_1^2 + 2 w_1 w_2 (\eta_1 + \eta_2) + 2\eta_2 w_2^2}$.

3. Tiered type II In this case $q_{11}^{(n)} = q_{12}^{(n)} = \frac{2}{3} L_3^{(n)}$, $q_{21}^{(n)} = \frac{\eta_1+\eta_2}{3\eta_1} L_3^{(n)}$ and $q_{22}^{(n)} = \frac{2\eta_2}{3\eta_1} L_3^{(n)}$. Based on equation (3.93) we now have

$$\begin{aligned}\bar{\lambda} &= w_1 (n - 1) q_{11}^{(n)} + w_1 w_2 n q_{21}^{(n)} + w_2 (n_2 - 1) q_{22}^{(n)} \\ &= \frac{L_3^{(n)}}{3} \left(2 w_1 (n - 1) + w_1 w_2 n \frac{\eta_1 + \eta_2}{\eta_1} + w_2 (n_2 - 1) \frac{2\eta_2}{\eta_1} \right),\end{aligned} \tag{3.97}$$

Table 3.3: The default parameter values and their respective ranges used for comparing network structures.

Parameter	Description	Default value	Range
γ	Ratio of interbank assets to capital	0.4	$[0.4, 0.6]$
π_0	Initial fraction of defaults	0.05	
$\bar{\lambda}$	Average out-degree as given by equation (3.93)	4	$[1, 7]$
η_1, η_2	Mean exposure amounts for groups one and two	$4, 1$	$[1, 11]$
w_1	Weight for group one	0.15	

$$L_3^{(n)} = \bar{\lambda}\frac{3\eta_1}{2\eta_1 w_1 (n-1) + w_1 w_2 n (\eta_1 + \eta_2) + 2\eta_2 w_2 (w_2 n - 1)} \tag{3.98}$$

and

$$\begin{aligned}\lambda_{11} &= \lim_{n\to\infty} \bar{\lambda}\frac{2\eta_1 w_1 (w_1 n - 1)}{2\eta_1 w_1 (n-1) + w_1 w_2 n (\eta_1 + \eta_2) + 2\eta_2 w_2 (w_2 n - 1)} \\ &= \bar{\lambda}\frac{2\eta_1 w_1}{2\eta_1 w_1 + w_1 w_2 (\eta_1 + \eta_2) + 2\eta_2 w_2^2}.\end{aligned} \tag{3.99}$$

Furthermore $\lambda_{ij} = \frac{(\eta_k + \eta_i) w_j}{2\eta_1 w_1 + w_1 w_2 (\eta_1 + \eta_2) + 2\eta_2 w_2^2}$, where $k = \min\{i, j\}$.

For all of the structures above, the expressions for λ_{ij} then satisfies the identity $\bar{\lambda} = w_1 (\lambda_{11} + \lambda_{12}) + w_2 (\lambda_{21} + \lambda_{22})$. Table 3.3 now shows the parameter values that were chosen for this analysis. Parameters either have the default value as indicated by the table or are varied within the range given in the final column.

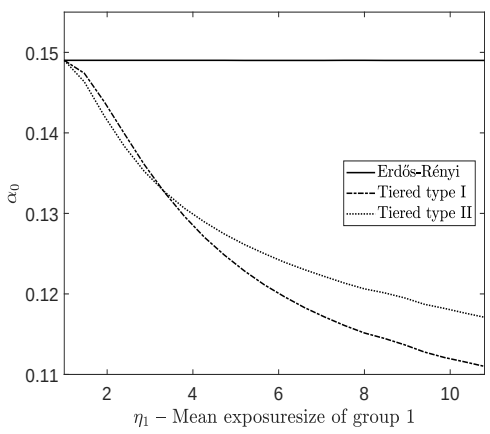
Consider first the variation in the final fraction of defaults as the relative sizes of the two groups are changed. The mean exposure amount of group one is varied from one to 11, whereas the mean exposure amount of group two is kept fixed at one. Therefore when $\eta_1 = 1$, we have a completely homogeneous network where all banks are of the same size. The three structures should therefore yield precisely the same fraction of final defaults in this case, since the discriminatory factor is eliminated when $\eta_1 = \eta_2$. This is illustrated in figure 3.3a, where the graph starts out with all three lines indistinguishable from one another. As the heterogeneity between the banks is increased along with η_1, the structures begin to discriminate between banks of different size.

It is interesting to see that both Tiered structures immediately start to deviate from the standard Erdős-Rényi case. These two structures exhibit decreasing risk for increasing heterogeneity between the groups of banks. This suggests that systems may benefit from

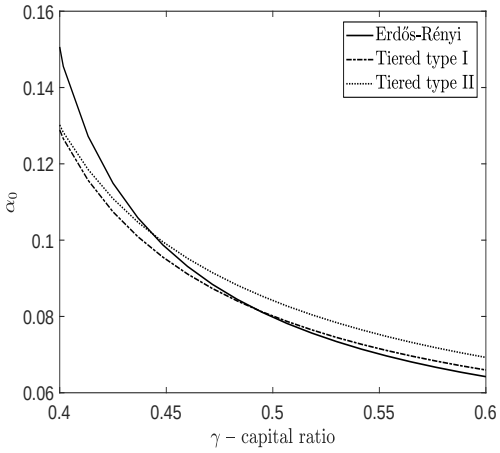
having a core-peripheral structure, with lending preferences that depend thereon.

Consider now the effect of varying the capital ratio from 0.4 to 0.6. Since this is a ratio of capital to interbank assets, then if interbank exposures consist of roughly 20% [8] of total capital, it corresponds to a range of 0.08 to 0.12 of capital to total assets. The results are given by figure 3.3b. As expected, the final fraction of defaults declines for all structures, though the final fraction of default declines more steeply for the Erdős-Rényi structure compared to the tiered type I and tiered type II structures in figure 3.3b.

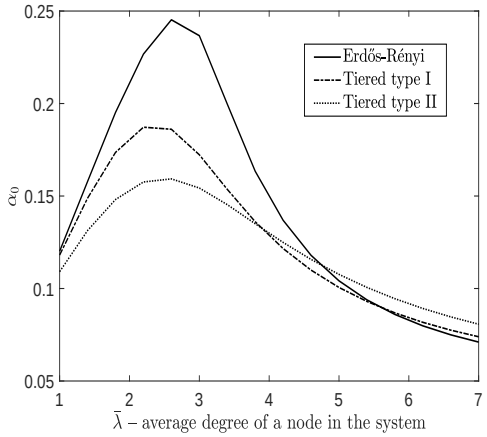
The average degree of the system is considered in figure 3.3c. This parameter is varied from one (an extremely sparse network) to seven. It can be seen that for a very sparse network, the different structures do not result in significantly different levels of default fractions. When the average out-degree is increased, the additional links in the system facilitate the spread of contagion for all structures. When the average out-degrees is just over 2.5, the default fractions start do decline when the additional links in the system serve as a safety mechanism. The peak at $\bar{\lambda} \approx 2.5$ and subsequent decline is more pronounced for the Erdős-Rényi case than for the other structures. This indicates a higher sensitivity to the level of interconnectedness in the system for our base parameters. It shows that conclusions regarding the optimal network structure can be highly dependent on network characteristics.



(a) Varying the mean asset value of group 1



(b) Varying the capital ratio



(c) Varying the average degree per node

Figure 3.3: Illustrating parameter sensitivities of different semi-heterogeneous Erdős-Rényi graphs, where $\alpha_0 = \sum_{j,k} h(j,k) \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j,\pi^*,\theta)$.

Chapter 4

Conclusion

This chapter serves to conclude the study by considering the results obtained from all parts of the study. First, the implications of the results are discussed in section 4.1. This is split into three parts as follows: Section 4.1.1 discusses the implications of the sensitivity tests performed on the numerical model of chapter 2. Section 4.1.2 deals with the application of this model to the South African system. Section 4.1.3 discusses the implications of the asymptotic results of chapter 3. Finally, section 4.2 serves to present the final conclusions of all results, in addition to shortcomings and avenues for future work.

4.1 Policy and theoretical implications of results

4.1.1 Network structure sensitivities

This section summarises the implications of the sensitivity tests from section 2.3 by discussing how these results answer the questions set out in section 1.3. Each question is discussed in turn below, where the findings are compared to those from exiting literature.

1. How does the structure of an interbank network influence the risk inherent in a system? This is illustrated by figures 2.7 through 2.10 where the risk in the system generally differs between different structures. In some cases, the risk between some structures are relatively close to one another (e.g. figure 2.10a), but this is not always the case (e.g. figure 2.9a). This shows that determining the correct structure of a network is important when modelling systemic risk. Methods that attempt to derive the most likely network structure in the absence of knowledge about the true structure need to bear this in mind. For example, maximum entropy techniques can underestimate [11] or overestimate [77]

systemic risk. Therefore it may not be appropriate since the resulting network structure will likely be incorrect, thereby leading to incorrect conclusions.

2. Are some structures inherently less risky than others? In contrast to previous work [54], we find that one type of structure is not necessarily more stable than another. From figure 2.7a alone it seems as if the disassortative and Erdős-Rényi structures are significantly less conductive of contagion compared to the other structures. This is supported by previous work [71], who find that random networks can be more resilient than scale-free networks. However, from our other figures (e.g. figure 2.8b) it can be seen that this is not necessarily the case under all circumstances. This is supported by the results in [83], where it is found that there is no single structure that performs the best under all circumstances.

3. How do network characteristics (i.e. interconnectedness, asset distribution and size of the system) influence the answer to question 1? In some situations the network characteristics influence the extent that the network structures differ from one another in terms of risk levels. Figures 2.8a, 2.9a and 2.10a show that for some structures, there are only small differences in the relative risk levels when the network characteristics are changed. This is the case when the assortative, attraction to size and tiered type I and II structures or the disassortative and Erdős-Rényi structures are compared to one another without indirect risk. Therefore in these cases the distinction between the different structures is not very important. This is, however, not always the case. When, for example, the disassortative structure is compared to the attraction to size structure, the relative risk levels change significantly for different combinations of network characteristics regardless of whether indirect risk is included or not. This is one aspect in which our results differ from [83], who find that network topology only matters when illiquidity is taken into account. While the differences between our network structures are much more pronounced when indirect risk is accounted for (see figures 2.8b, 2.9b, and 2.10b), network characteristics can still influence relative risk levels without indirect risk.

4. How do indirect contagion mechanisms influence the answers to the above questions? Indirect mechanisms have a significant influence over the extent to which network structure influences the risk in the system (question 1). This can be seen from figures 2.8 through 2.10, where the differences between the structures are generally much more pronounced when indirect risk is included. It also influences the answer to question 2, since the disassortative structure is the most resilient in almost all cases when indirect risk is excluded, but this is no longer the case when it is included. Finally, it influences the extent

to which network characteristics in turn influence the relative risk levels (question 3). For example, figure 2.10a shows almost identical risk levels for the structures as N approaches 200, which is not the case in figure 2.10b.

5. Can network properties assist in explaining some of the variation in risk levels observed from different structures? Not necessarily. In some instances when indirect risk is excluded, there are similarities between the average shortest paths, the clustering coefficient and the risk levels of the different structures. However these similarities are not seen under all circumstances, especially if indirect risk is included. Previous research has found that tiered banking systems are more resilient than otherwise [70]. While our definition of a tiered system is not exactly the same, our results show that this is not always the case. Indirect risk and other network characteristics such as the interconnectedness in the system has a significant influence on whether this holds true or not. While network properties may be useful at individual banks level, we conclude that such network properties do not contain sufficient information on a system-wide level to explain differences in risk levels under all circumstances.

6. What policy suggestions can be made regarding network characteristics? From figures 2.8 to 2.10 it is clear that systems with different network structures and different levels of indirect risk respond differently to changes in network characteristics. For example, while the risk in the system increases in most cases when the interconnectedness of the system is increased (as found in [83] for example), this is not necessarily the case under all circumstances. When indirect risk is included, the risk in the system eventually starts to decrease for increasing levels of interconnectedness. Depending on the initial level of interconnectedness, it may therefore be beneficial in some cases to increase interconnectedness rather than to decrease it. It is clear from this research that the combination of all aspects of a banking system is too complex to suggest any 'one size fits all' approach. This complexity may be a reason why different authors arrive at different conclusions regarding systemic risk modelling techniques and policy implications. For example while one study finds that maximum entropy techniques underestimate systemic risk [11], another finds that it overestimates it [77]. While in [19] it is found that intermediate levels of connectivity are safer, the opposite is found in [52], and in [83] it is found that lower levels of connectivity are advisable.

4.1.2 Real-world application

Recall from section 1.3 that the purpose of the real-world application is to assess the usefulness of network models of systemic risk in a South African context. The network structures behaved similarly over time for most of the cases considered here. The levels of risk were generally similar for higher levels of connectivity. The differences in risk levels between the structures were generally consistent over time, with the exception of one scenario. Indirect risk served to increase differences in risk levels for low interconnectedness. In contrast to this, an increase in indirect risk reduced the differences in risk levels for higher levels of interconnectedness with the exception of the period starting at October 2016. The structures mostly followed similar trends over time. This suggests that the changes in systemic risk detected by the model is not highly dependent on the network structure. This observations have the following implications:

1. The materiality of network structure is firstly influenced by the objective of the network model. If the objective is to accurately determine the level of risk in the system, then the network structure may not make a significant difference for highly interconnected systems. For lower levels of interconnectedness, differences between the risk levels are more prominent in general. This observation is, however, dependent on the indirect risk parameters and the dates under consideration.

2. If the objective is to detect changes in point-in-time measures of systemic risk, the materiality of network structure decreases. This means that the uncertainty around the initial shock to the system (and hence the resulting path through which losses spread) is less problematic.

The results show that the risk parameters significantly influence how the risk levels in the system change over time. The liquidity risk mechanism employed by the model is directly dependent on the asset values of all banks. Small changes in the liquidity risk parameters have a non-trivial influence on the risk levels over time, which cannot trivially be explained by changes in asset values.

Furthermore, each risk parameter influences the results in its own way. For example, increasing the short-term liquidity parameter emphasised the December 2015 and June 2016 increases in risk, whereas the medium-term liquidity parameters reduced the significance of these spikes. This either means that the model is not able to detect increases in risk for certain parameter values, or that the system does not experience a significant increase in systemic risk during times of market turmoil for some liquidity scenarios. For example, the high liquidity risk scenarios (see figures 2.13c and 2.13d) might increase the risk

levels during all months to such an extent that the effect of weak economic conditions are diminished.

The above observations have the following implications for the modelling of systemic risk using a network approach:

1. Empirical studies that aim to determine the level of systemic risk should take care to calibrate the liquidity risk parameters to levels appropriate for the system being considered. To do this, bank balance sheet data will be required for past events where banks have failed. By analysing changes in asset values for the remaining banks in the system, it may be possible to determine appropriate ranges or distributions for these parameters.

2. A finer division of assets is recommended. The fact that the liquidity risk parameters each had a different effect on the model output suggests that the classification of assets can be a material aspect of such a study.

3. Since the model showed increases in systemic risk during times of market turmoil, it shows that network models of systemic risk may be valuable modelling tools. A great advantage of this is that publicly available balance sheet information can be used to model systemic risk, thereby avoiding the need to obtain confidential trading information. It may be by chance that the model detected increases in systemic risk due to balance sheet fluctuations. This warrants further investigation to determine with greater certainty whether the model can accurately identify potential crises.

Additional analysis shows that the framework presented in chapter 2 can be useful tools for understanding how systemic risk can spread through a banking system. It can be used to understand how and when an initial shock leads to contagion within a network. For example, for the cases considered here, contagion is only spread by the default of the largest banks during times of market turmoil. The analyses further showed that the effect of an initial shock to a system is best quantified by measures that take account of the severity of losses, and not just the number of defaults. Finally, the analyses illustrated that if an initial shock leads to additional defaults, the consequences to the system can be severe. This is important for regulators since it shows that even if an initially shocked bank is allowed to default, it is important to prevent further banks from defaulting as this can quickly cause the whole (or at least most of the) system to collapse.

4.1.3 Theory and structures for large networks

This part of the study was largely theoretical. Therefore, these implications are focused on the effect that the theory has on systemic risk research, as opposed to sections 4.1.1 and 4.1.2 that were more focussed on practical implications. The purpose of this part of the study is to contribute to the literature on large banking networks and to formally define a versatile class of networks that can be useful for theoretical studies on banking networks. The theoretical results from this chapter firstly shows how the fraction of defaults in a financial system can be approximated for large, random networks. A sequence of random financial networks of increasing size that satisfies certain limiting conditions is considered. It is shown that there will be a subsequence for which the fraction of defaults following an initial shock can be determined.

A class of inhomogeneous graphs is defined and it is illustrated how results that apply to Erdős-Rényi networks may be generalized to apply to this class. This is done by considering the theoretical results developed in this chapter, which can be applied to Erdős-Rényi networks. It follows naturally that the results can be applied to the proposed class of inhomogeneous networks. Potential uses of such a class of networks include modelling the interaction between different types of financial entities (e.g. between banks, investment companies and insurance companies) or between the financial systems of different countries.

As a simple illustration of the versatility of the proposed class of networks, three different structures that comply with the definition are compared to one another. The first is the standard Erdős-Rényi graph, where lending behavior is independent of relative asset sizes. The remaining structures assume different kinds of lending behavior for banks based on their relative asset sizes. In other words, banks' preferred creditors and debtors are determined by their asset sizes, although the framework allows for other characteristics to infer such preferences instead. While the illustration may not be based on entirely realistic structures, the second and third structures do account for a hierarchical formation of edges based on bank size.

The illustration considered here suggests that for large systems, the sensitivity of systemic risk to network characteristics is dependent on the network structure. For example, where one structure may show a significant change in systemic risk when the interconnectedness is varied, another structure may only show a modest change. It is noted that the realism of these observations are influenced by the shocks being transmitted via direct exposures, and that indirect mechanisms such as liquidity risk may well serve to increase the importance of network structure even further [57].

4.2 Final conclusions

On the simulation side, we propose a new network model of systemic risk, accounting for liquidity risk, investor confidence, heterogeneity of banks, division of assets according to term and differences in network structure. The liquidity and proximity shock parameters capture the extent that liquidity shortfalls and loss of market trust respectively impact the stability of the system.

For smaller networks, six different network structures are defined and compared to one another in terms of risk under different circumstances. Three of these structures are novel contribution to systemic risk simulation research, as they can satisfy empirically observed network characteristics, namely tiering and the small world property (i.e. a small average shortest path and high clustering). For larger networks, we define a general, versatile class of stochastic networks and show how structures similar to our tiered structures can be modelled. Many theoretical results are based on Erdős-Rényi networks because of its mathematical tractability. We illustrate how results that hold for Erdős-Rényi networks can be generalised to semi-heterogeneous Erdős-Rényi networks. This may be useful for future research on theoretical models of systemic risk, as assumptions that hold for Erdős-Rényi networks may also hold for semi-heterogeneous Erdős-Rényi networks. This is useful as the latter class of network is able to account for differences in lending behaviour based on assets size.

Our results suggest that the resilience of network structures is dependent on network characteristics such as interconnectivity, distribution of assets and the level of capital. Furthermore, the effect that the capital, level of interconnectedness, the asset distribution and the size of the system can have on the risk posed by different structures depends on the presence of liquidity risk and loss of market trust. This shows that the interaction between these network characteristics and the inclusion of indirect losses in banking network models greatly influences the comparison between different network structures. In line with the results of [57], it further shows the importance of including liquidity losses and market sentiment when investigating network structure in systemic risk models.

The presence of tiers in the banking system which dictate lending behaviour can have a significant effect on the risk in the system. While systemic risk can be lower for tiered structures compared to non-tiered structures (this is supported by e.g. [92]), network characteristics such as heterogeneity between banks, capital ratios and interconnectedness influence whether this is indeed the case. Therefore, whether or not the level of tiering in a network serves to strengthen or weaken the system potentially depends on a combination of network characteristics. It further shows that the precise definition of a tiered struc-

ture is important, as structures with similar levels of tiering may exhibit different levels of risk. Therefore, lending preferences (i.e. banks' preferred creditors and/or debtors based on characteristics such as bank size, sector, type etc.) which may differ between jurisdictions may influence the optimal course of action to reduce the risk of default contagion.

The results of this study are further relevant to regulators, since they could be interested in optimising the network structure and characteristics to result in safer banking systems. This could be achieved by providing incentives for banks to encourage or discourage certain types of lending behaviour. For example, if a regulator wanted to encourage or discourage an attraction to size structure, it could create incentives to influence the ratio of the number of 'small' vs. 'large' banks that other banks lend to. Similarly, incentives or restrictions could be imposed to encourage the system to become either more or less interconnected if this would result in a safer system. Regulators that want to optimise the characteristics and the structure of a system must keep in mind the current characteristics and the effect of liquidity risk and the risk of market trust losses.

Despite the problems associated with determining the correct network structure and liquidity risk parameters, such network models of systemic risk can be useful. These models are simple, easy to understand and makes use of publicly available balance sheet data. During the time frame considered for the South African application, the network model detected increases in systemic risk at times when the economy experienced unexpected market disturbances. An important avenue for future research is to determine whether this is by chance, or whether the model accurately determines the probability that a crisis can occur. Furthermore, it is important to note that the model does not forecast times of distress, but instead provides a proxy for the level of risk at a point in time. In other words, the true proportion of capital lost following a shock to the system is not determined, but rather a value that increases or decreases along with it.

The methodologies for small and large networks introduced by this study may enable a regulator to test the effect of a range of different interventions for given structures. We show that network sensitivities can behave differently under different network structures. This shows that the network structure can help inform the regulator which reporting information to focus on, since it is possible to determine the most important drivers of systemic risk. It is important to note that these drivers will likely differ between different banking networks.

Our results emphasize the importance of network characteristics and indirect losses such as losses due to liquidity problems and a deterioration of market sentiment when considering network structures from a regulator's point of view. There are various other factors to consider that is beyond the scope of this study, for example the cost of any proposed regulatory action, specific incentives that can be introduced to influence the network

structure, measuring the systemic importance of individual banks and the cost to the economy of failing banks. For a regulator to implement the proposed model, refinements to the model are required. The model should be embedded in a macroeconomic model with a lender of last resort. This can be used to investigate the effects of policy decisions based on the network structure and characteristics. This includes an investigation of the costs vs. benefits of policy decisions. It is important to note that such refinements are much more difficult to include in a large network setting compared to a small network setting where simulation is appropriate.

Network models can be refined to consider different types of interbank lending, e.g. considering different maturities for interbank loans [12]. A more granular division of external assets will become important when incorporating the macro-economic environment. Our results for the South African application suggest that this is an important avenue for future research. This goes hand in hand with the need to determine appropriate liquidity reduction parameters associated with each asset class.

Common shocks as opposed to idiosyncratic shocks should be considered, as this can have a significant effect on the spread of contagion [43]. For example, comparing the loss resulting from the failure of one big bank vs. the failure of smaller banks with the same combined asset value as the bigger bank. Furthermore, by changing the way that systemic risk is measured, it possible that different conclusions may be reached. It is therefore an important consideration for future research to look at different ways of measuring systemic risk when investigating implications of policy changes.

For further future work, it is also important to acknowledge that a country's banking system generally forms part of a larger international network of networks. The study from [53] on the international banking network and studies on asymptotic network results (e.g. [8]) could provide excellent groundwork for this.

Appendix A

Numerical model

A.1 Formal tiering test results

To test whether these structures can be considered as tiered, we employ a similar procedure as in [41]. For each value of $\bar{p}$, 1000 simulations are done for each structure. The lower 1% tiering error value of the Erdős-Rényi network, say $e_{0.01}$, is used to test whether the other networks are significantly more tiered than a random network. For each of the simulations done for the non-random networks, it is determined whether the error is smaller than $e_{0.01}$. The proportion of simulations which result in a tiered structure is shown in figure A.1.

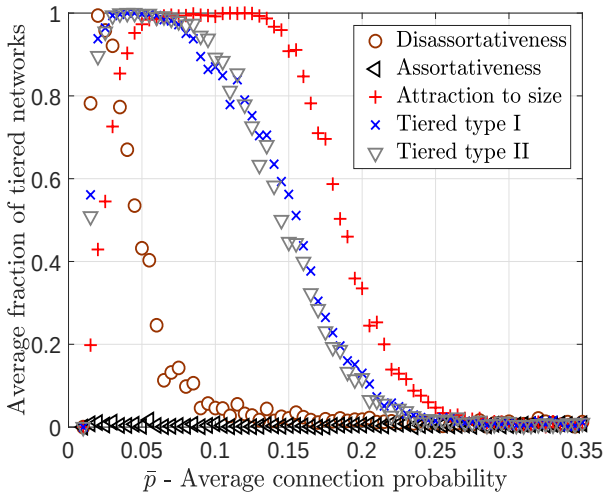


Figure A.1: The average fraction of simulations for each structure that can be regarded as a tiered network.

The attraction to size structure yields tiered networks at least 96% of the time when

$\bar{p} \in [0.05, 0.14]$. The tiered type I and II structures formally result in tiered networks for at least 98% of the simulations when $\bar{p} \in (0.03, 0.07)$. For $\bar{p} = 0.1$ the tiered type I and II structures result in tiered networks for 89% and 87% of the simulations respectively, which is significantly more than the other structures. The disassortative structure only manages to yield tiered networks for a large number of simulations when $\bar{p} \in [0.02, 0.03]$, and the assortative structure does not result in tiered networks for most of the simulations. This shows that the attraction to size and tiered type I and II structures can result in tiered networks sufficiently often, provided that $\bar{p}$ is in the correct range.

Appendix B

Additional information for South African application

B.1 Additional figures for South African banks' assets

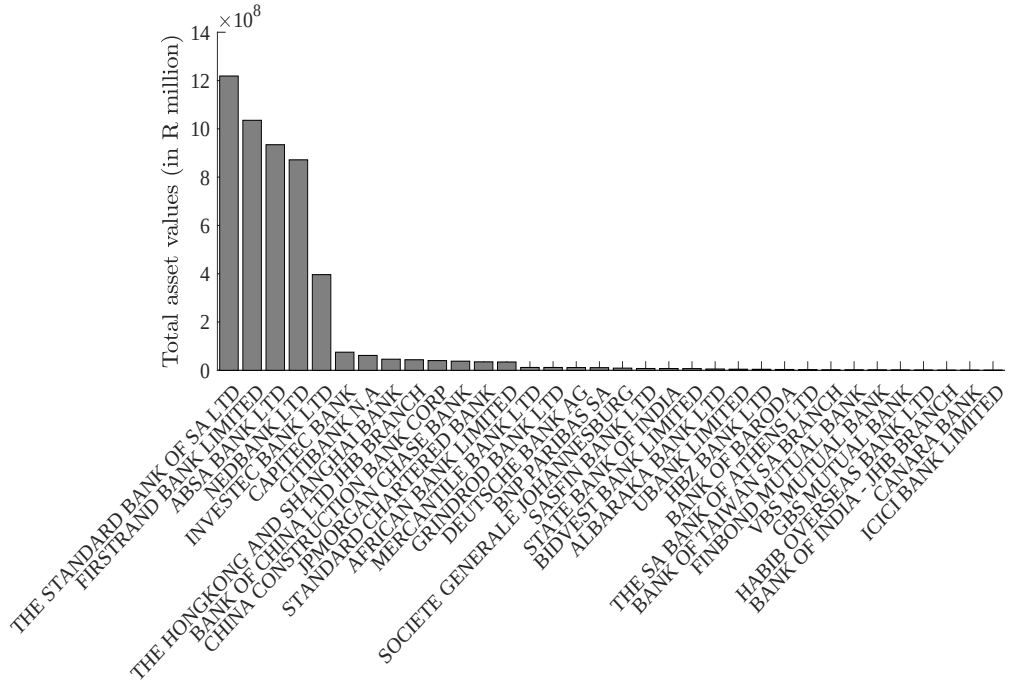


Figure B.1: The distribution of assets in the South African banking sector as at 31 March 2017.

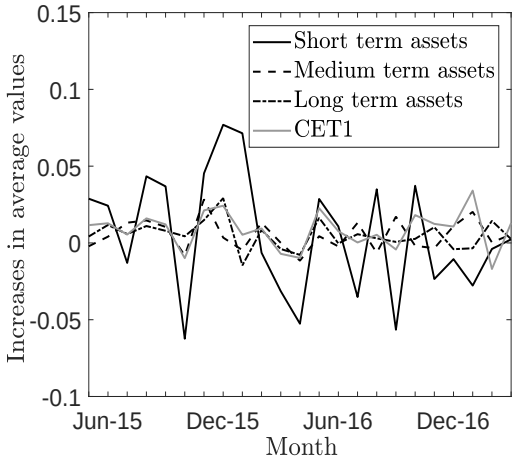


Figure B.2: Relative increases of the average balance sheet items in the system.

B.2 Balance sheet information

Table B.1: Division of the non-interbank assets according to term.

Short Term Assets
– Central bank money and gold.
– Deposits with, and loans and advances to banks.
– Loans granted to the SARB and other non-banking institutions under resale agreements.
– Foreign currency deposits, loans and advances.
– One third of marketable government stock that have an unexpired maturity of less than 3 years.
– Derivative instruments assigned to the short-term assets according to the rules in section 2.4.2.
Medium Term Assets
– Instalment sales.
– Credit-card debtors.
– Overdrafts, loans and advances to the private sector.
– Bankers' acceptances (Treasury bills, SARB bills, promissory notes, commercial paper and Land Bank bills).
– Clients' liabilities per contra.
– Remittances in transit.
– Current income tax receivables and deferred income tax assets.
– One third of marketable government stock that have an unexpired maturity of less than 3 years.
– Derivative instruments assigned to the medium-term assets according to rules in section 2.4.2.
Long Term Assets
– Redeemable preference shares.
– Leasing transactions.
– Mortgage advances.
– Overdrafts, loans and advances to the public sector.
– Non-marketable government stock.
– All marketable government stock excluding two thirds of those stock that have an unexpired maturity of less than 3 years.
– Debentures and other interest-bearing security investments of private sector.
– All equity investments.
– Derivative instruments assigned to the long-term assets according to rules in section 2.4.2.
– Other investments.
– Non-financial assets.
– Retirement benefit assets.
– Assets acquired or bought to protect an advance or investment.

Table B.2: Available unweighted ratios of CET1 to total assets at month-end for all registered banks.

											Reporting month	
Bank Name	**03/2017**	**02/2017**	**12/2016**	**09/2016**	**06/2016**	**03/2016**	**02/2016**	**12/2015**	**09/2015**	**06/2015**	**03/2015**	**02/2015**
ABSA Bank Ltd	0.0642		0.0656	0.0561	0.0540	0.0525		0.0538	0.0525	0.0489	0.0495	
African Bank Ltd	0.2387		0.2259	0.2178								
Albaraka Bank Ltd			0.1057	0.1070	0.1094	0.1111		0.1083		0.1075	0.1089	
Bank of Baroda	0.0459		0.0578		0.0410							
Bank of China LTD Jhb Branch												
Bank of India - Jhb Branch			0.6626	0.7247	0.6865	0.6471		0.5647	0.6692	0.6947	0.6733	
Bank of Taiwan SA Branch												
Bidvest Bank Ltd	0.2483		0.2439	0.2544	0.1745	0.1895		0.1929	0.2051	0.1664	0.2009	
BNP Paribas SA								0.0439				
Canara Bank	0.8010		0.7844	0.7367	0.6515	0.6810		0.6468	0.6796	0.7160	0.7774	
Capitec Bank		0.1990					0.1989					0.1956
China Construction Bank			0.1063		0.0549	0.0618		0.0407	0.0485	0.0482		
Citibank N.A	0.0945		0.0887	0.0838	0.0981	0.0960		0.0568	0.0701	0.0794	0.0926	
Deutsche Bank AG	0.1087		0.1204	0.1007	0.0907	0.0836		0.0603	0.0658	0.0719	0.0589	
Finbond Mutual Bank		0.2085					0.2407					0.2425
Firstrand Bank Ltd	0.0704		0.0672	0.0693	0.0688	0.0663		0.0674	0.0688	0.0681	0.0642	
GBS Mutual Bank												
Grindrod Bank Ltd	0.0665		0.0506	0.0654	0.0640	0.0702		0.0571	0.0749	0.0725	0.0719	
Habib Overseas Bank Ltd	0.1146		0.0994	0.0965	0.0917	0.0862		0.0826	0.0779	0.0787	0.0840	
HBZ Bank Ltd	0.0749		0.0785	0.0865	0.0741	0.0710		0.0662	0.0646	0.0696		
Icici Bank Ltd												
Investec Bank Ltd	0.0854		0.0857	0.0827		0.0813			0.0836		0.0878	
JPMorgan Chase Bank	0.1313		0.1136	0.0840	0.0721	0.0656		0.0568	0.0849	0.0961	0.0785	
Mercantile Bank Ltd	0.1714		0.1649	0.1729	0.1789	0.1837		0.1903	0.1964	0.1971	0.2039	
Nedbank Ltd	0.0594		0.0575	0.0544	0.0556	0.0545		0.0547	0.0524	0.0542	0.0572	
Sasfin Bank Ltd	0.1491		0.1671	0.1836	0.1976	0.2070		0.2090	0.1975	0.1975	0.1944	
Societe Generale Jhb												
Standard Chartered Bank				0.0911	0.0874	0.0829		0.0789	0.0957	0.0972	0.0830	
State Bank of India	0.1722					0.1252		0.1265	0.1323	0.1357	0.1394	
The Hongkong and Shanghai Bank												
The SA Bank of Athens Ltd	0.0851		0.0844	0.0889	0.0767	0.0803		0.0802	0.0928	0.1005	0.1053	
The Standard Bank of SA Ltd	0.0570		0.0552	0.0556	0.0570	0.0546		0.0550	0.0536	0.0549	0.0518	
Ubank Ltd		0.0857					0.0832					0.0877
VBS Mutual Bank						0.0878					0.0973	

B.3 Network properties of the South African system

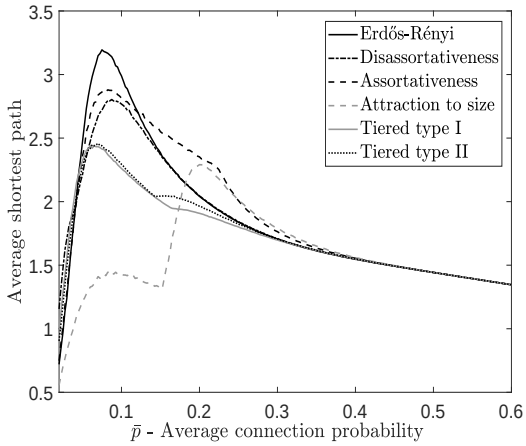


Figure B.3: The average shortest path in the system for each structure as a function of $\bar{p}$, the level of interconnectedness in the system.

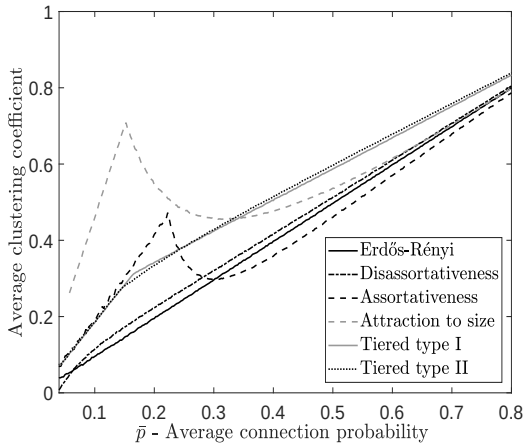
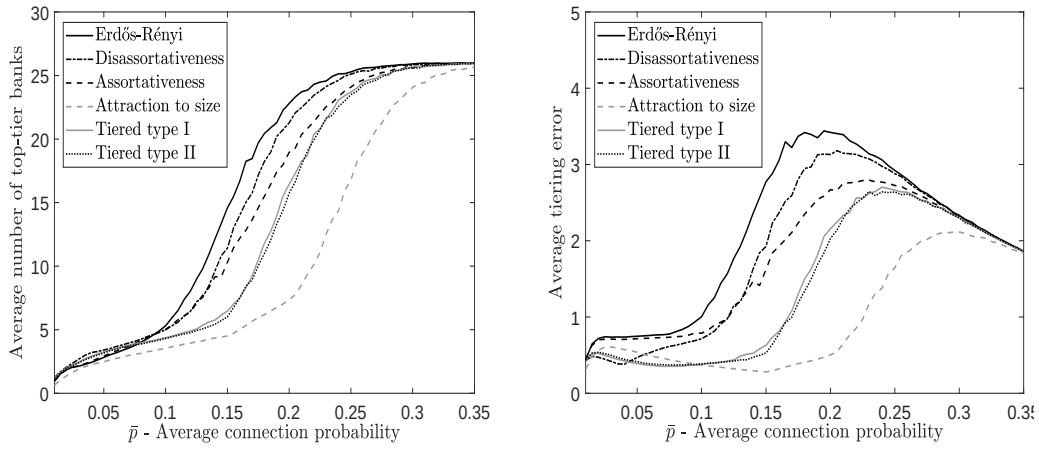


Figure B.4: The average weighted, directed clustering coefficient for each network structure as a function of $\bar{p}$, the level of interconnectedness in the system.

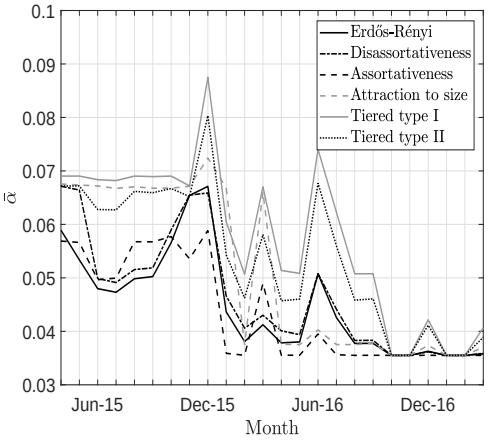


(a) Average number of banks in the top tier as a function of $\bar{p}$

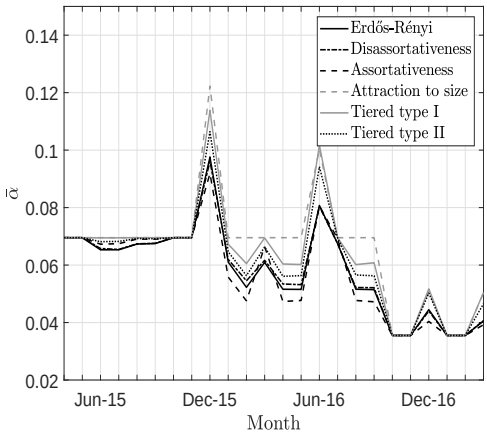
(b) Average tiering error as a function of $\bar{p}$

Figure B.5: Comparison of the tiering error and the average number of banks in the top tier between all structures, expressed as functions of the average connection probability $\bar{p}$.

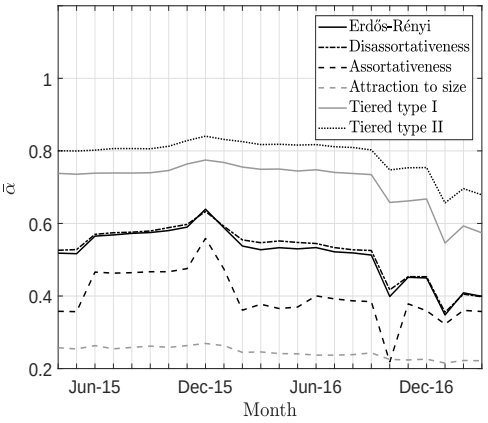
B.4 Changing the risk measure



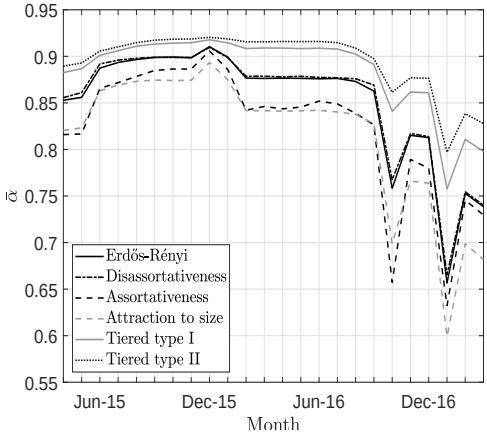
(a) Low indirect risk, low interconnectedness



(b) Low indirect risk, moderate interconnectedness



(c) High indirect risk, low interconnectedness



(d) High indirect risk, moderate interconnectedness

Figure B.6: Comparing the average fraction of defaulted banks resulting from different network structures over time.

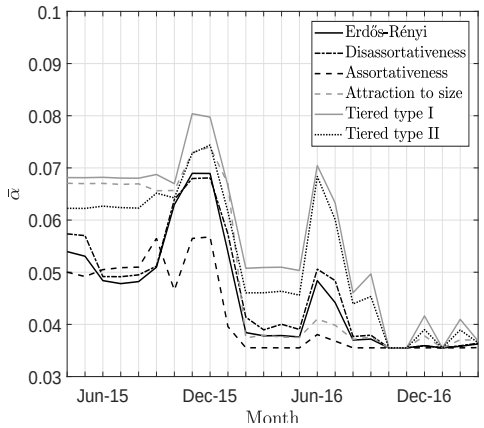
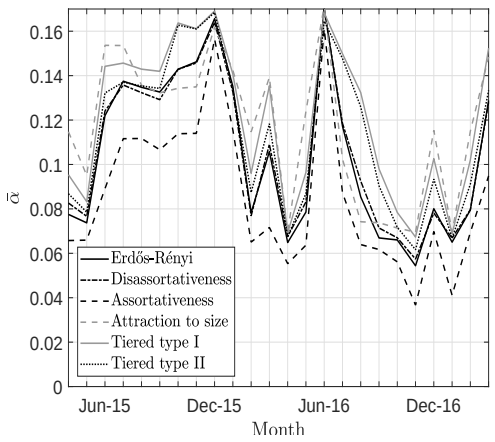
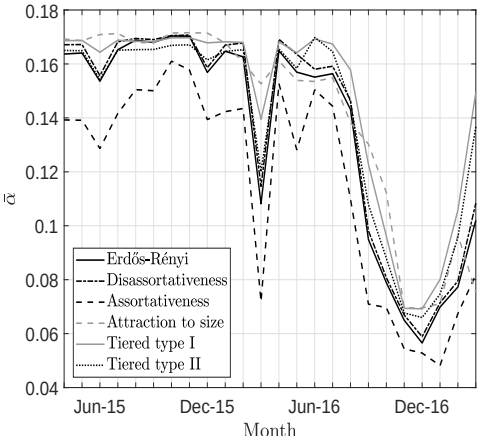


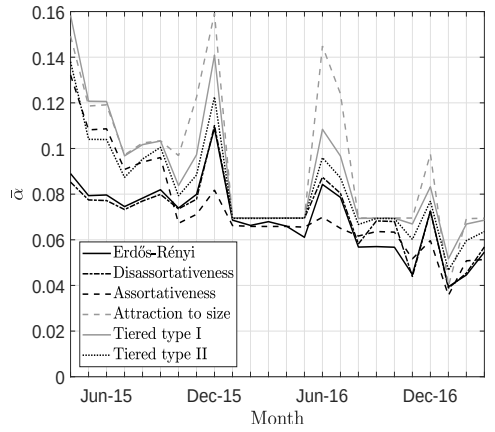
Figure B.7: Systemic risk over time for base parameter values of $g^{(s)} = g^{(m)} = g^{(l)} = \delta = 0.0075$.



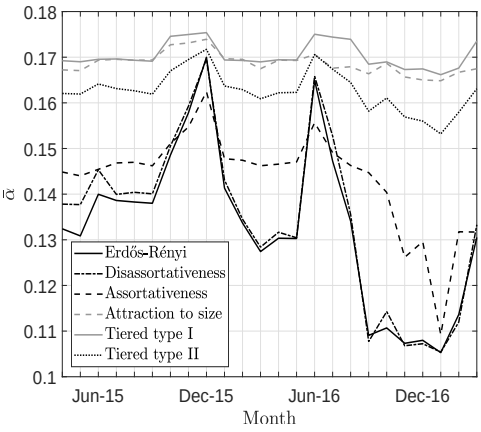
(a) Systemic risk over time for base parameters, but with $g^{(s)} = 0.0225$



(b) Systemic risk over time for base parameters, but with $g^{(m)} = 0.0225$



(c) Systemic risk over time for base parameters, but with $g^{(l)} = 0.0225$



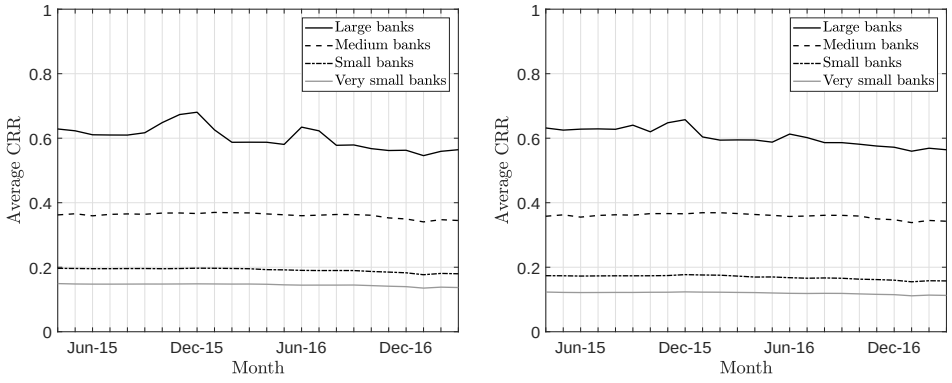
(d) Systemic risk over time for base parameters, but with $\delta = 0.0225$

Figure B.8: Comparing the average defaulted fraction of different network structures over time and increasing each indirect risk parameter in turn.

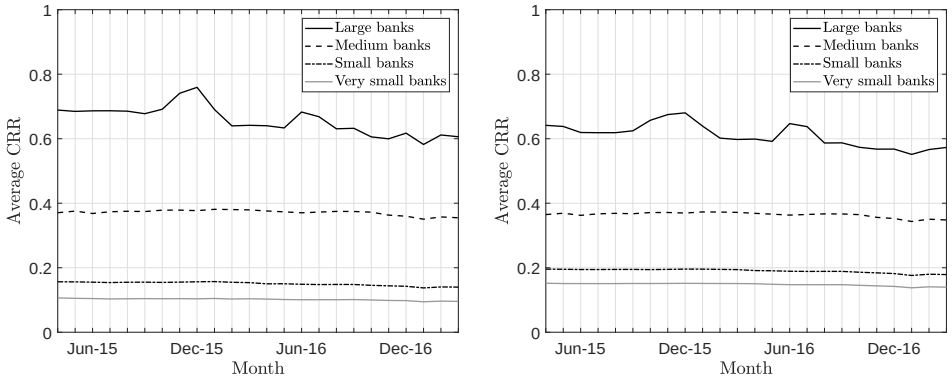
B.5 Additional analyses figures

B.5.1 Risk quantities over time by size category

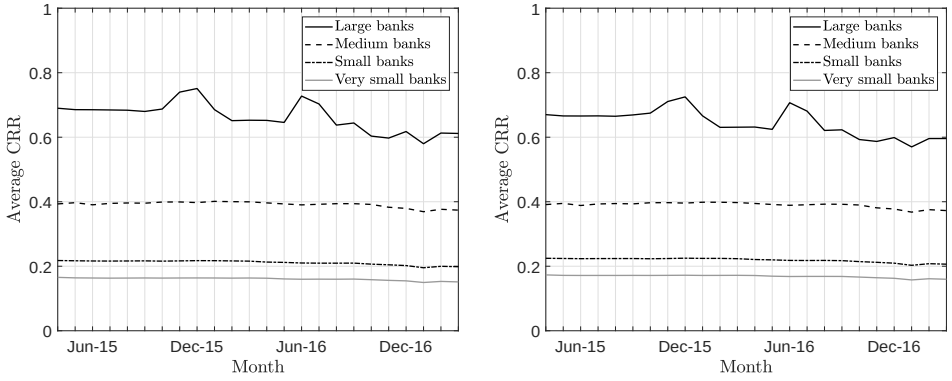
CRR



(a) Erdős-Rényi (b) Disassortative



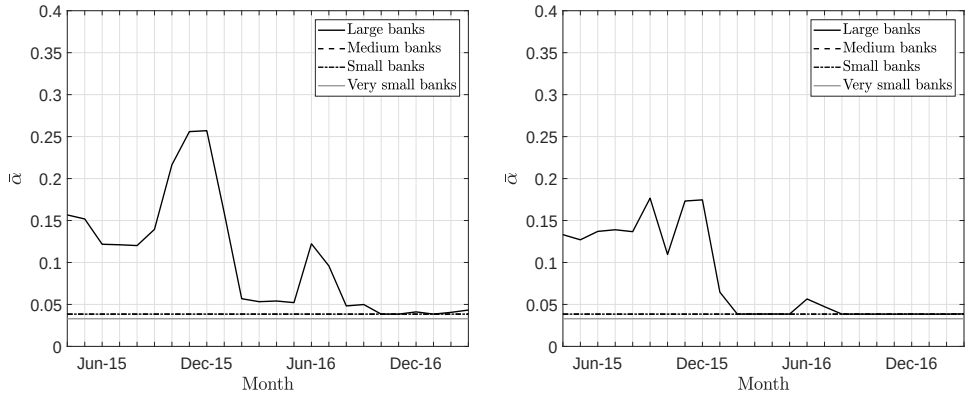
(c) Assortative (d) Attraction to size



(e) Tiered Type I (f) Tiered Type II

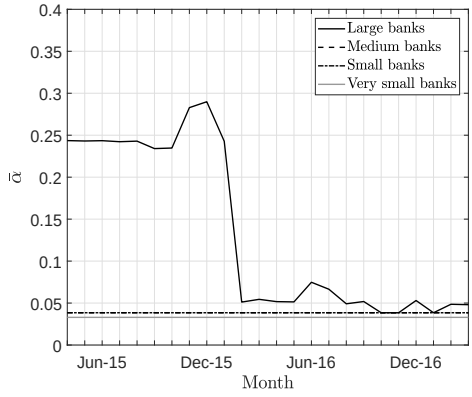
Figure B.9: Plotting the average CRR over time for very small, small, medium and large banks.

Average fraction of defaulted banks

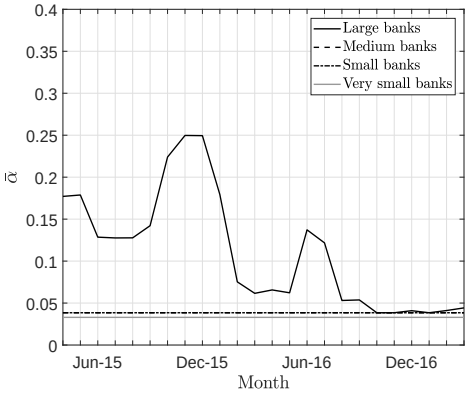


(a) Erdős-Rényi

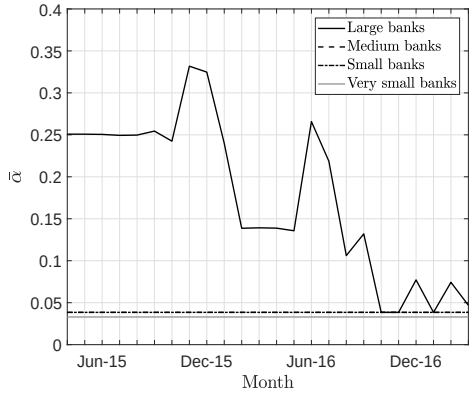
(b) Disassortative



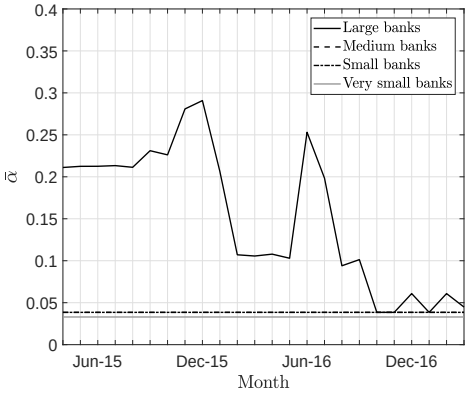
(c) Assortative



(d) Attraction to size



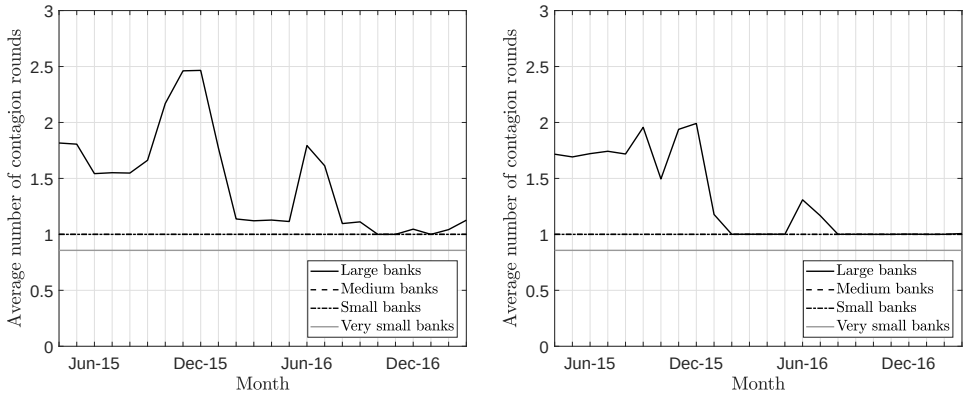
(e) Tiered Type I



(f) Tiered Type II

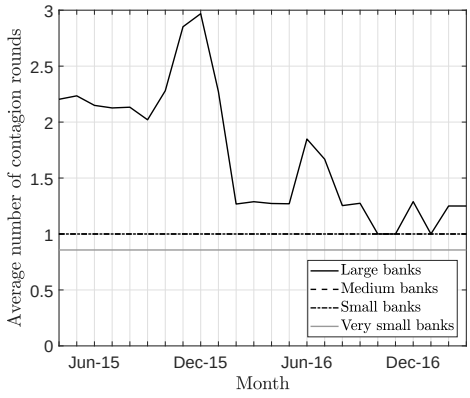
Figure B.10: Plotting the average defaulted fraction of banks over time for very small, small, medium and large banks.

Default rounds

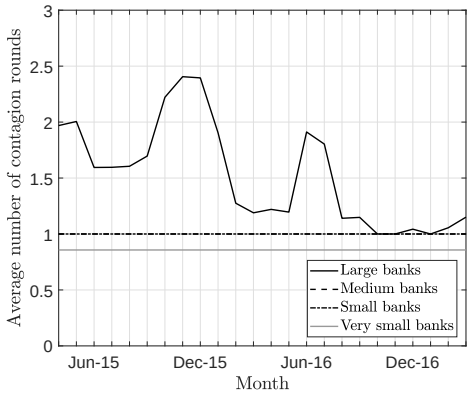


(a) Erdős-Rényi

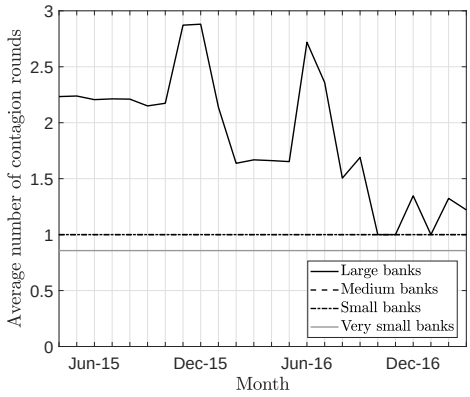
(b) Disassortative



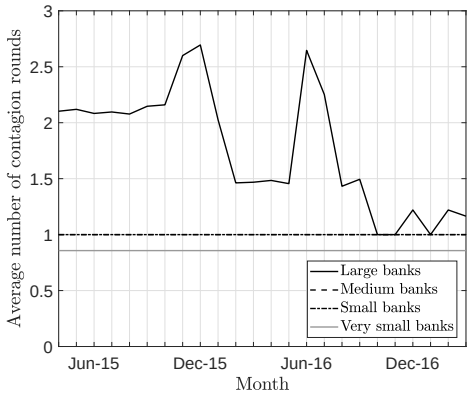
(c) Assortative



(d) Attraction to size



(e) Tiered Type I

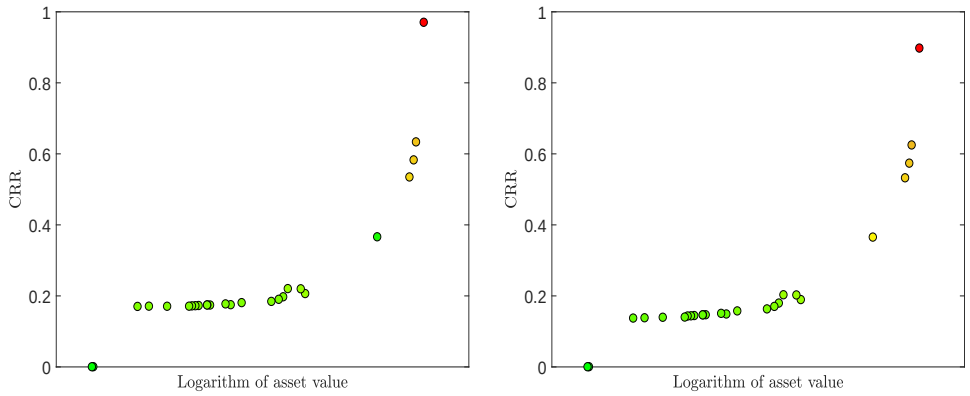


(f) Tiered Type II

Figure B.11: Plotting the average number of default rounds over time for very small, small, medium and large banks.

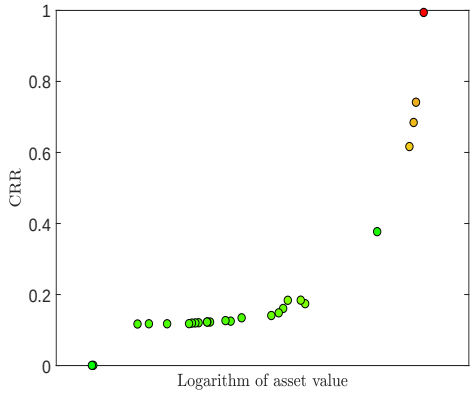
B.5.2 Risk quantities vs. asset values

CRR

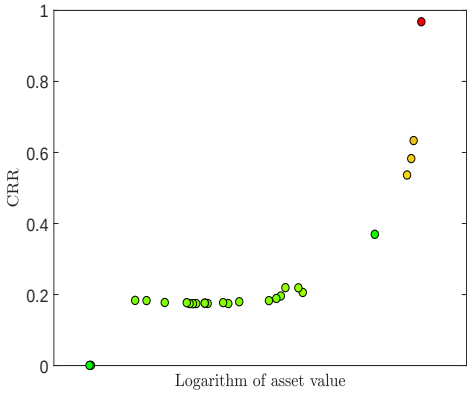


(a) Erdős-Rényi

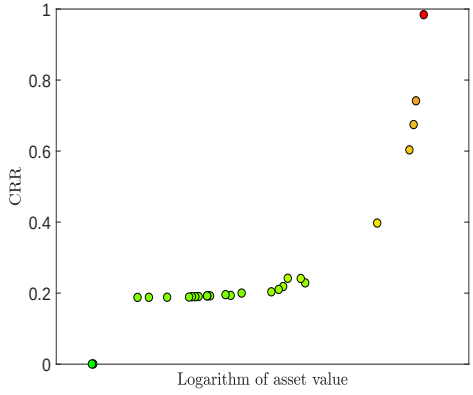
(b) Disassortative



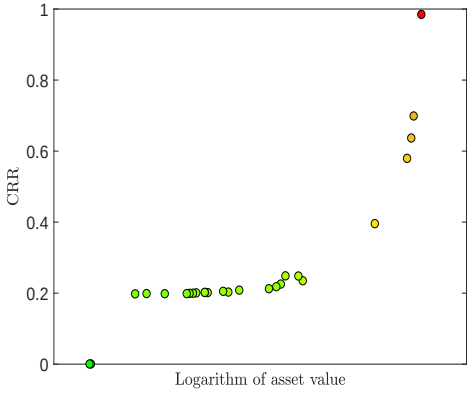
(c) Assortative



(d) Attraction to size

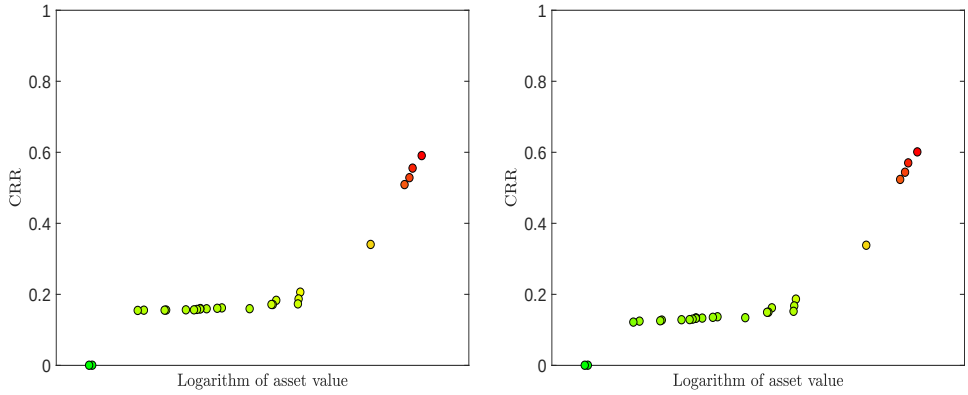


(e) Tiered Type I



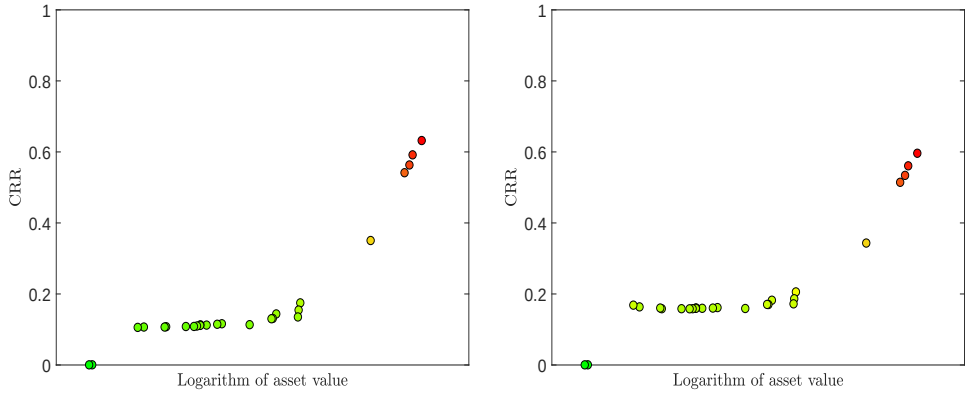
(f) Tiered Type II

Figure B.12: Comparing the CRR for each bank to its asset value for a high-risk scenario (December 2015).



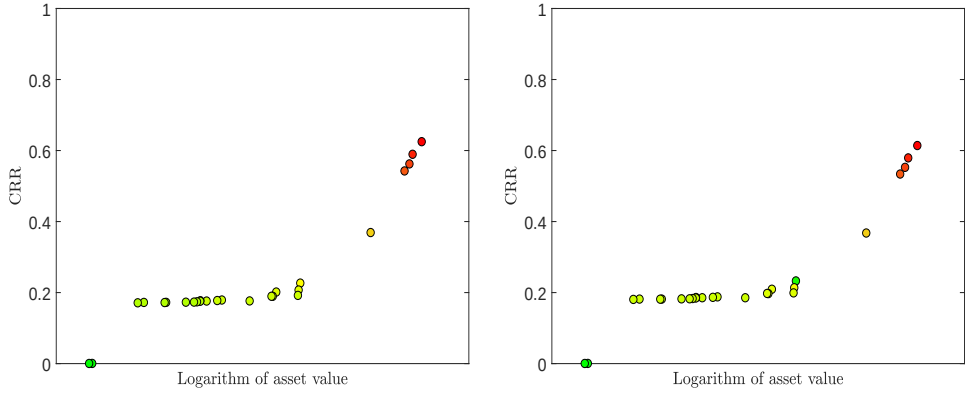
(a) Erdős-Rényi

(b) Disassortative



(c) Assortative

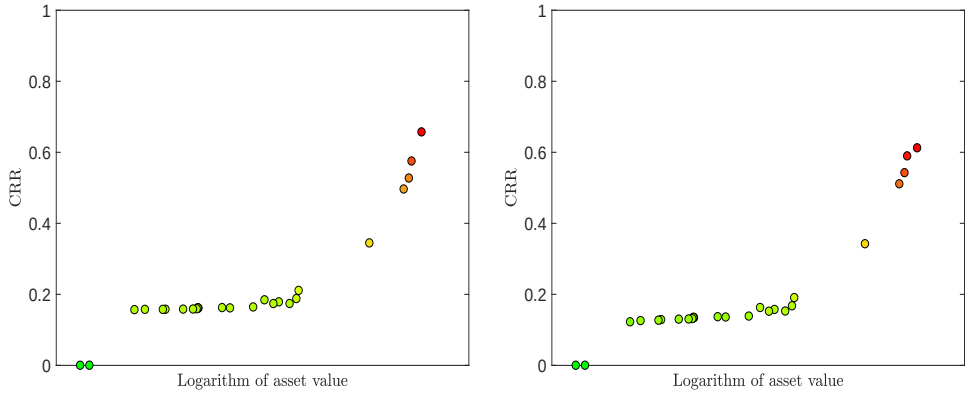
(d) Attraction to size



(e) Tiered Type I

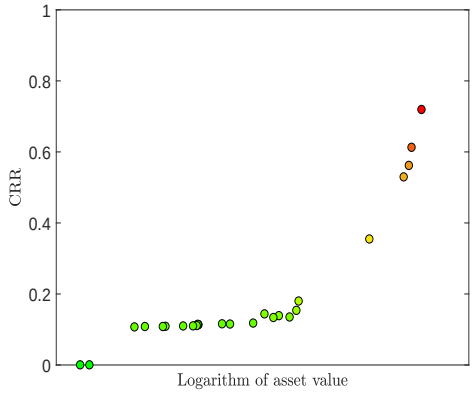
(f) Tiered Type II

Figure B.13: Comparing the CRR for each bank to its asset value for a low-risk scenario (January 2017).

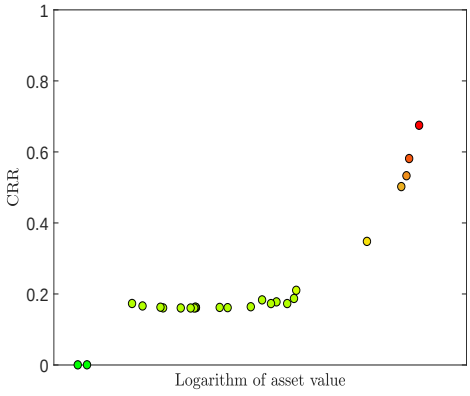


(a) Erdős-Rényi

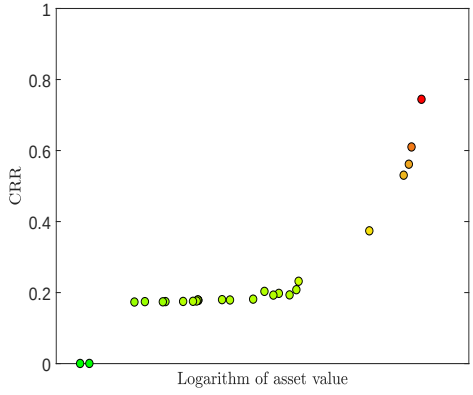
(b) Disassortative



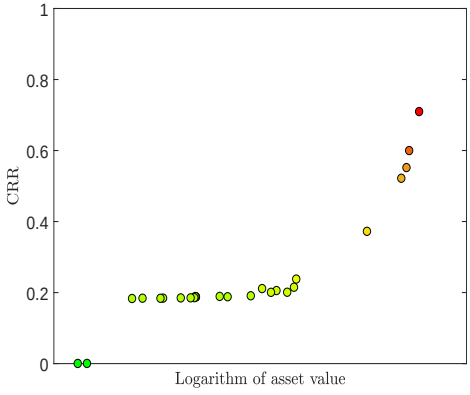
(c) Assortative



(d) Attraction to size



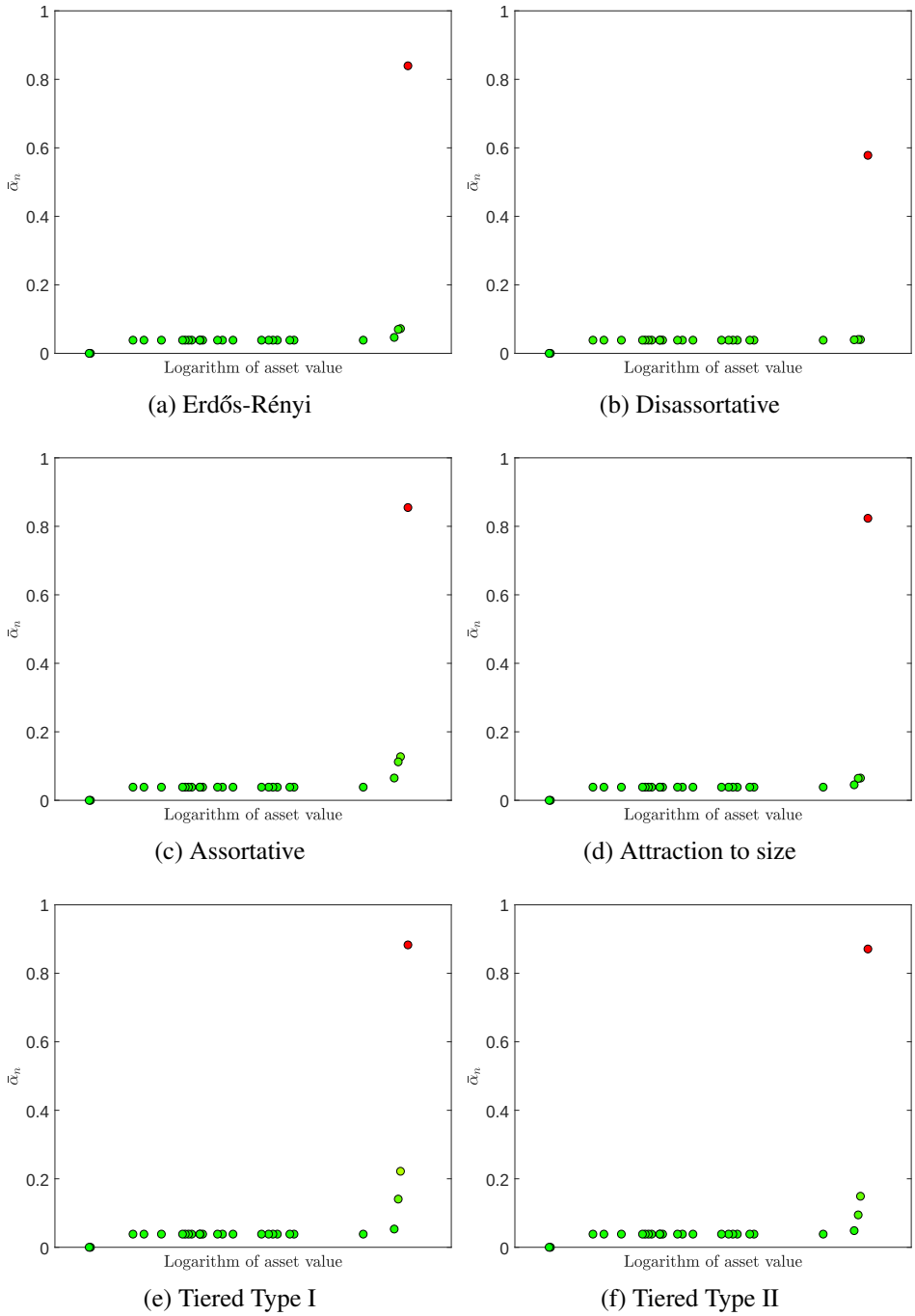
(e) Tiered Type I



(f) Tiered Type II

Figure B.14: Comparing the CRR for each bank to its asset value for a moderate-risk scenario (March 2017).

Average fraction of defaulted banks

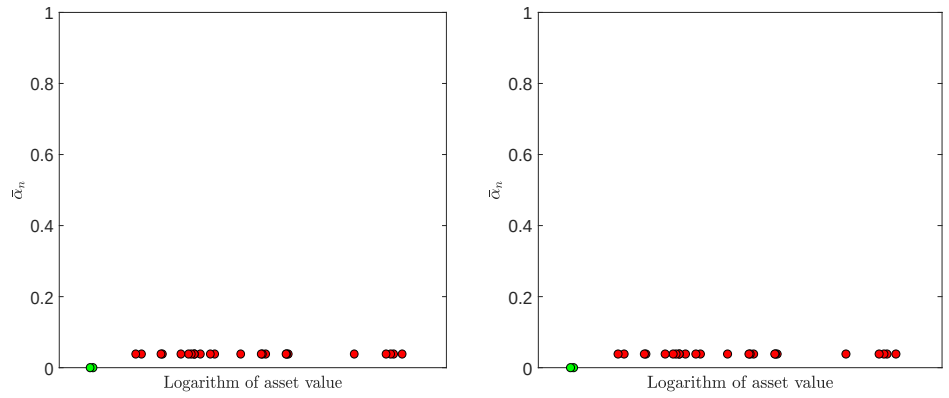


(a) Erdős-Rényi (b) Disassortative

(c) Assortative (d) Attraction to size

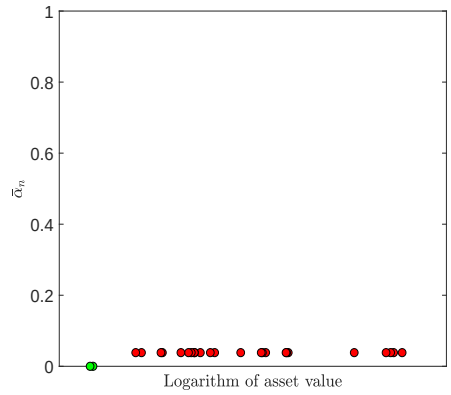
(e) Tiered Type I (f) Tiered Type II

Figure B.15: Comparing the fraction of defaulted banks for each bank to its asset value for a high-risk scenario (December 2015).

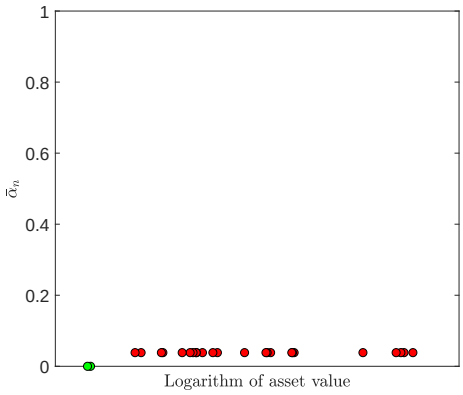


(a) Erdős-Rényi

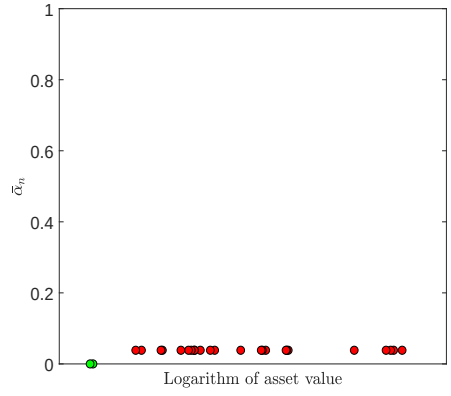
(b) Disassortative



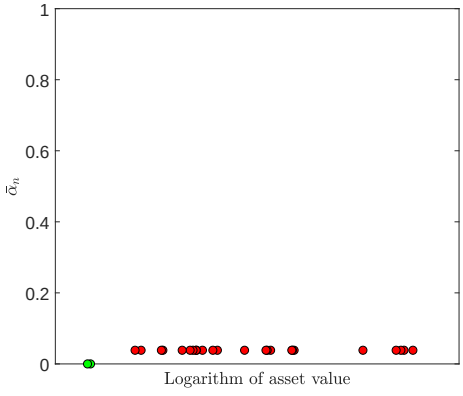
(c) Assortative



(d) Attraction to size

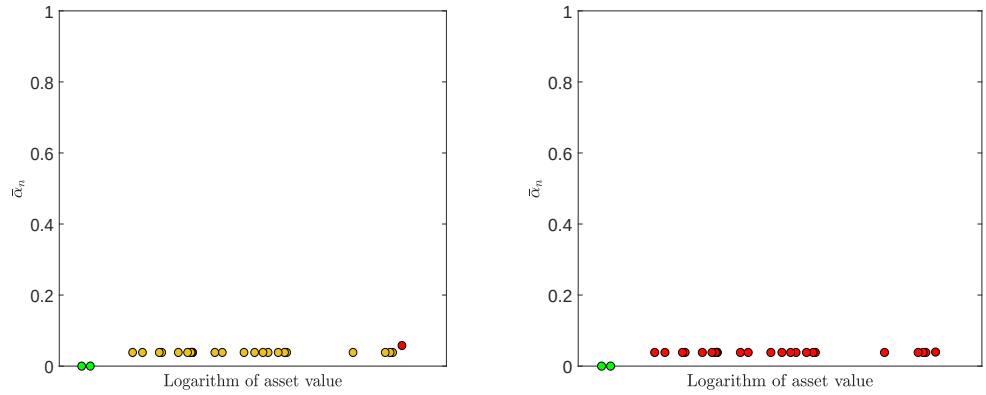


(e) Tiered Type I



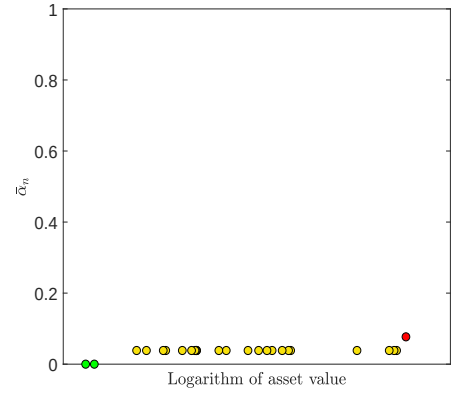
(f) Tiered Type II

Figure B.16: Comparing the fraction of defaults for each bank to its asset value for a low-risk scenario (January 2017).

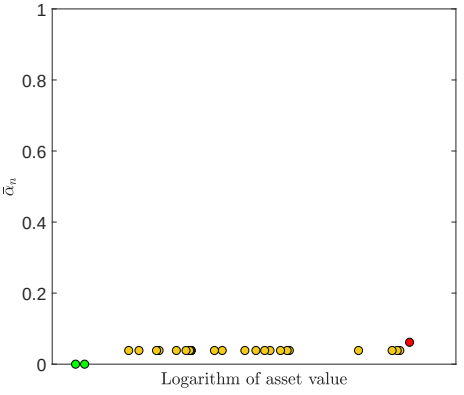


(a) Erdős-Rényi

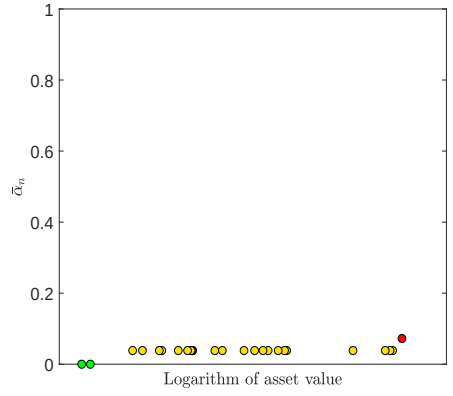
(b) Disassortative



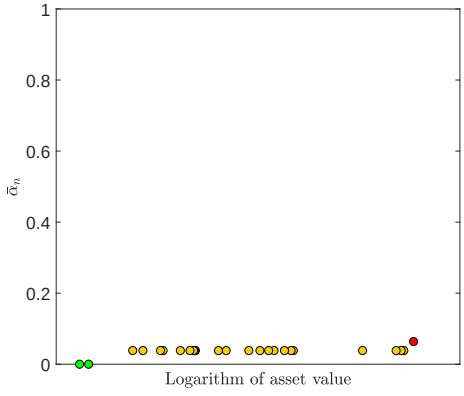
(c) Assortative



(d) Attraction to size



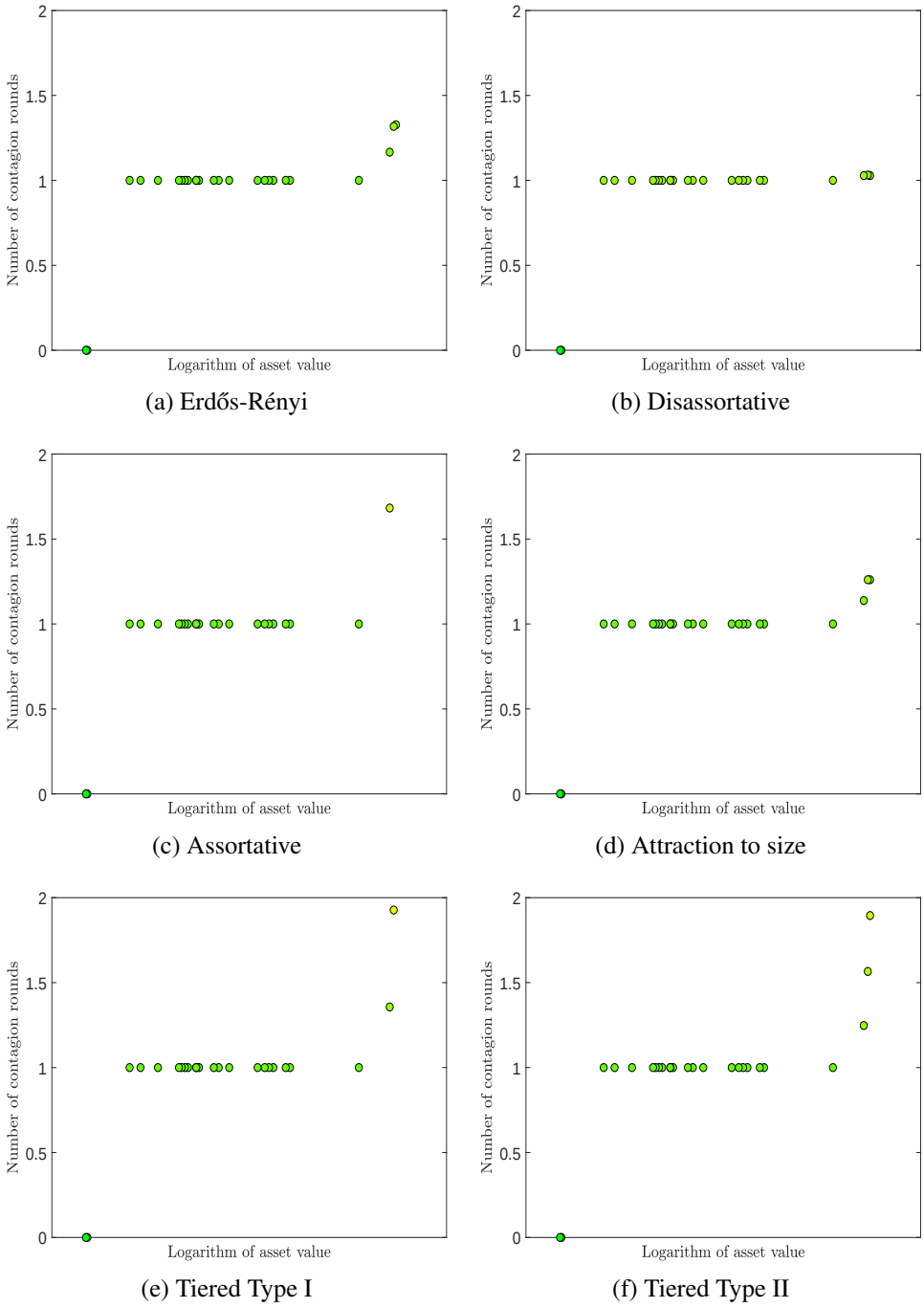
(e) Tiered Type I



(f) Tiered Type II

Figure B.17: Comparing the fraction of defaults for each bank to its asset value for a moderate-risk scenario (March 2017).

Default rounds

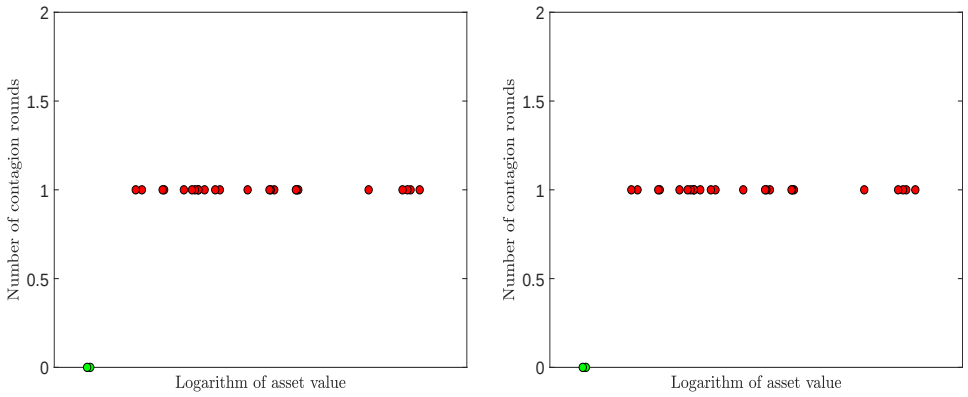


(a) Erdős-Rényi (b) Disassortative

(c) Assortative (d) Attraction to size

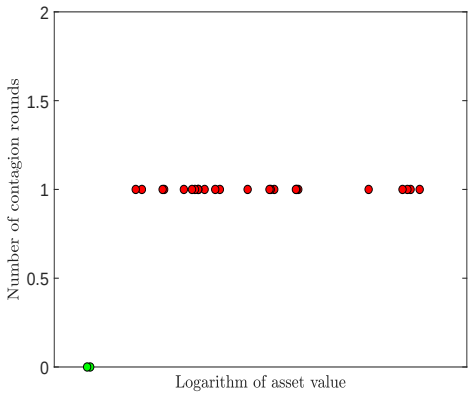
(e) Tiered Type I (f) Tiered Type II

Figure B.18: Comparing the average number of default rounds for each bank to its asset value for a high-risk scenario (December 2015).

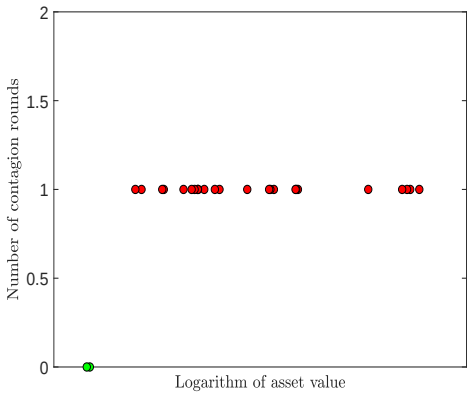


(a) Erdős-Rényi

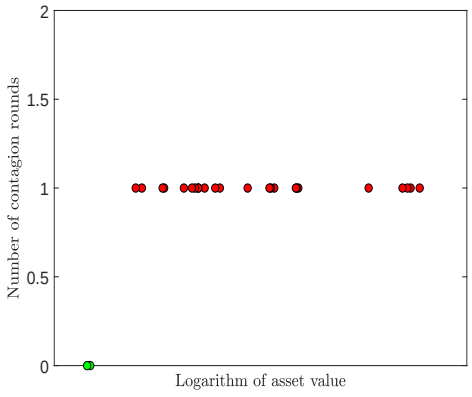
(b) Disassortative



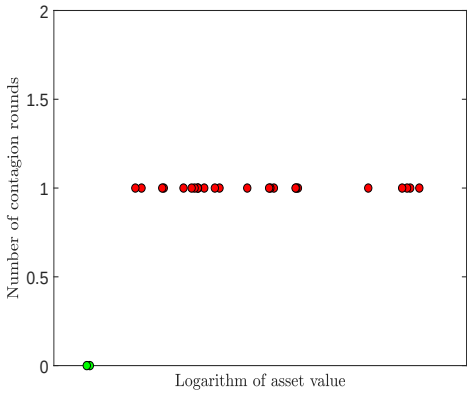
(c) Assortative



(d) Attraction to size

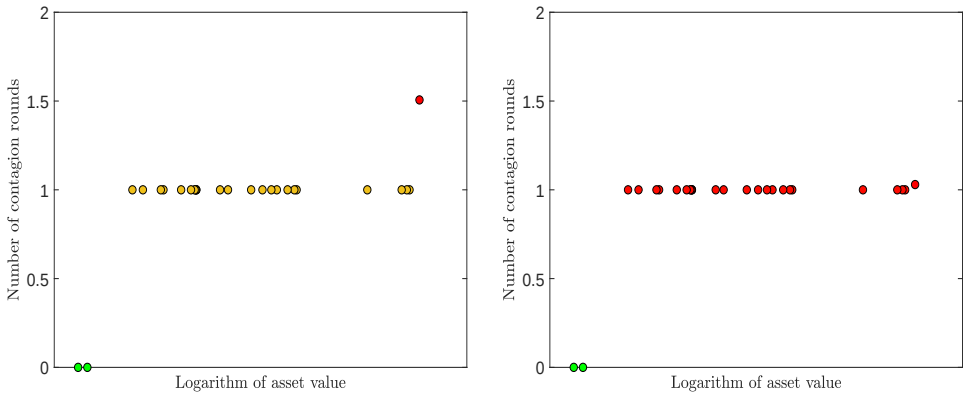


(e) Tiered Type I



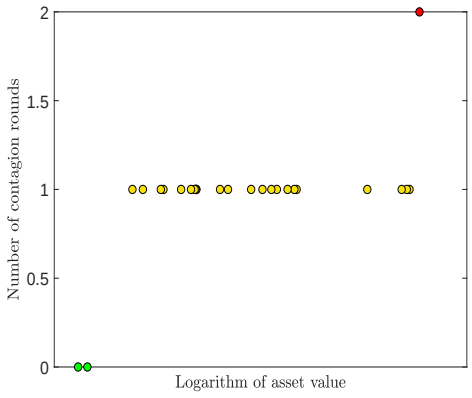
(f) Tiered Type II

Figure B.19: Comparing the average number of default rounds for each bank to its asset value for a low-risk scenario (January 2017).

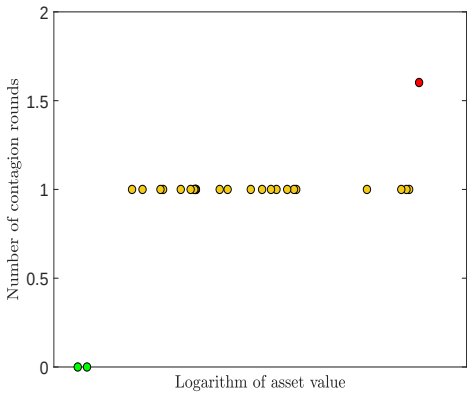


(a) Erdős-Rényi

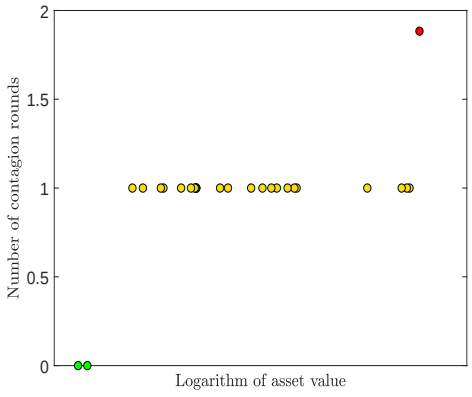
(b) Disassortative



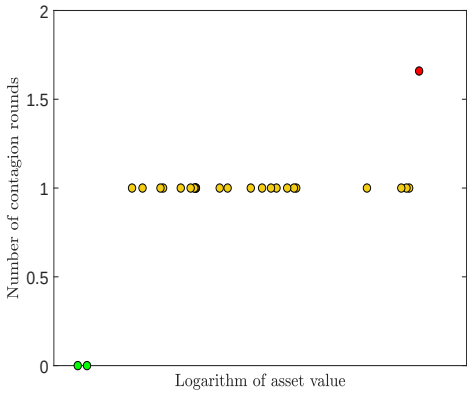
(c) Assortative



(d) Attraction to size



(e) Tiered Type I

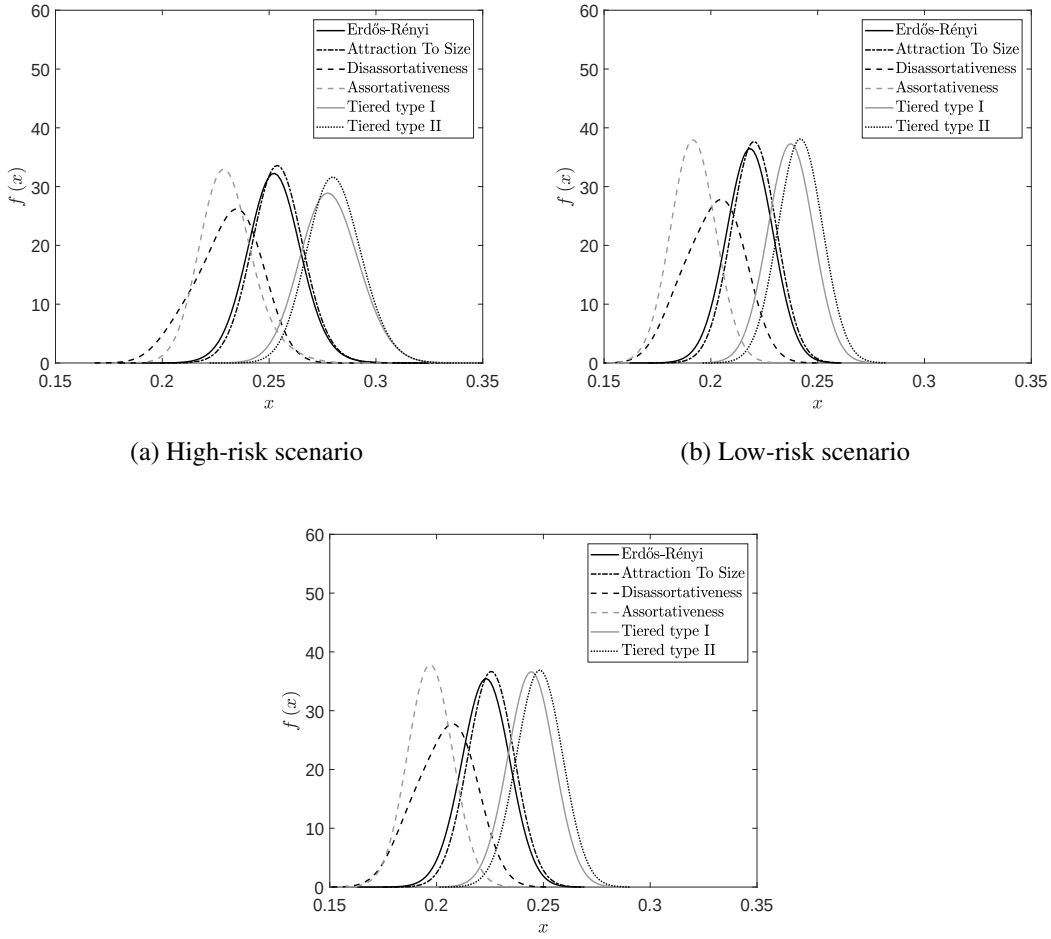


(f) Tiered Type II

Figure B.20: Comparing the average number of default rounds for each bank to its asset value for a moderate-risk scenario (March 2017).

B.5.3 Risk quantity distributions

CRR



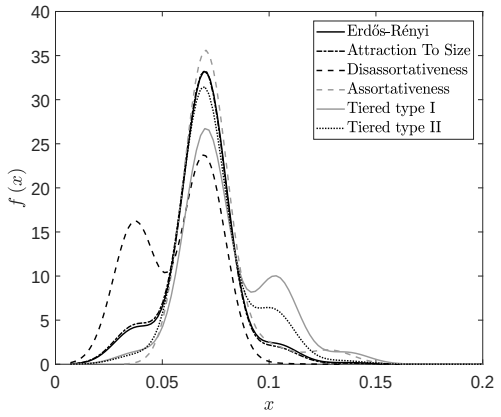
(a) High-risk scenario

(b) Low-risk scenario

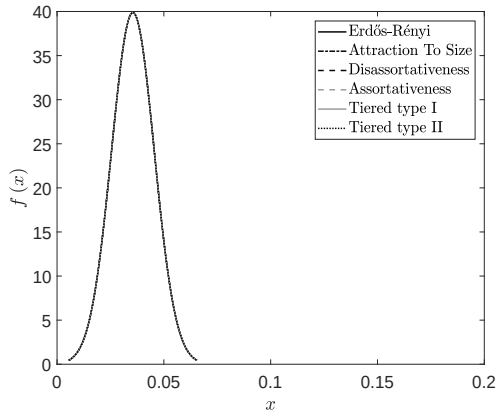
(c) Moderate-risk scenario

Figure B.21: The empirical distribution of the CRR for a high-risk scenario (December 2015), low-risk scenario (January 2017) and moderate-risk scenario (March 2017).

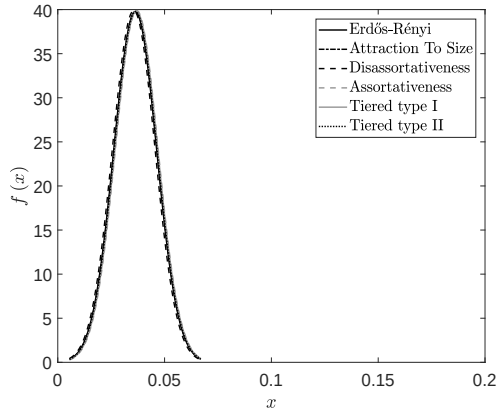
Average fraction of defaulted banks



(a) High-risk scenario



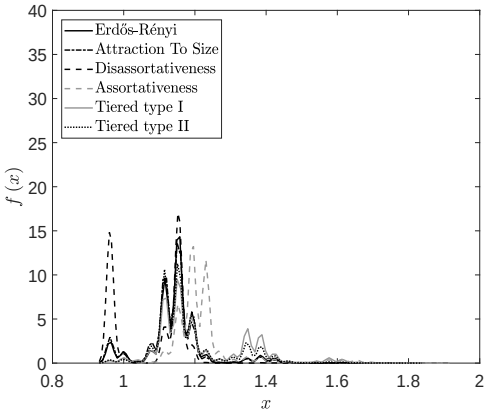
(b) Low-risk scenario



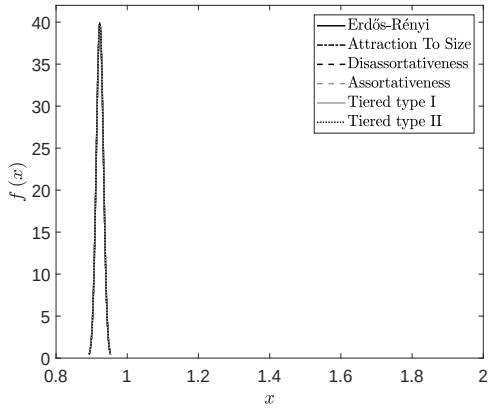
(c) Moderate-risk scenario

Figure B.22: The empirical distribution of the average defaulted fraction for a high-risk scenario (December 2015), low-risk scenario (January 2017) and moderate-risk scenario (March 2017).

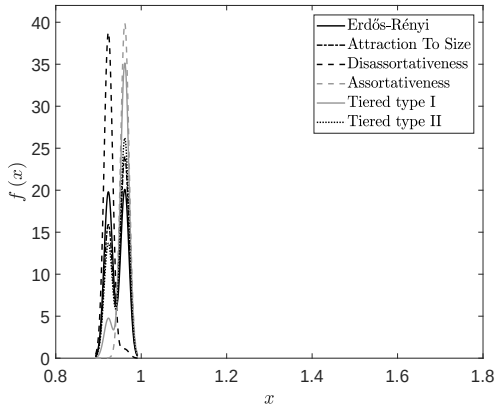
Default rounds



(a) High-risk scenario



(b) Low-risk scenario



(c) Moderate-risk scenario

Figure B.23: The empirical distribution of the number of default rounds for a high-risk scenario (December 2015), low-risk scenario (January 2017) and moderate-risk scenario (March 2017).

Appendix C

Asymptotic results

C.1 Network resilience and susceptibility

Another property of the network to be investigated is its ability to absorb and withstand small shocks. This ability is assessed by means of the *network resilience measure* [8, Definition 3.9], defined as follows:

Definition C.1 (Network resilience measure)**.** Suppose that $(\mathbf{e}_n, \vec{\gamma}_n)_{n\geq 1}$ is a sequence of financial networks satisfying assumptions 3.1 and 3.4 in [8]. Then the *network resilience measure* is defined as

$$1 - \sum_{j,k} \frac{jk}{\lambda} \mu(j,k)\, p(j,k,1) \in (-\infty, 1]\,. \tag{C.1}$$

For a random network, the variable $p(j,k,1)$ represents the expected fraction of (j,k) nodes that default after one counterparty default. Therefore $jp(j,k,1)$ represents the expected number of links that will cause contagion to spread after one round of default.

Each of these nodes will cause a shock (not necessarily leading to default) through k other links. This is because those (j,k) nodes that have defaulted after one round will each spread losses to k other nodes, where the latter nodes are not necessarily unique (in other words, two nodes might spread losses to the same creditor). Therefore $kjp(j,k,1)$ represents the expected number of shocks to the system due to (j,k) nodes after one round of default.

By multiplying with $\mu(j,k)$ and summing over all j and k, the expression $\sum_{j,k} jk\,\mu(j,k)\, p(j,k,1)$ gives the total expected number of shocks to the system after one round of default.

Note that the expected number of shocks to the system is not necessarily bounded above by λ, which is the total number of links in the system. This is because for a node of

type $p(j,k,1)$, it is possible that more than one outgoing link can lead to its default (even though only one counterparty default is required to make the node itself default as well). This results in 'double-counting' the node's incoming links when looking at the number of shocks the system receives.

However, this is not a drawback of the network resilience measure. It is necessary to keep this information because of the fact that the order in which a node's counterparties default is random. A node with more than one counterparty that could cause it to default should be deemed more risky than another node with only a single counterparty that could cause it to default.

Finally, by dividing the expected number of shocked links in the system by the total number of links in the system, it is possible to account for the relative size of the system. For example, if a system with 10 links expected 5 shocks after one round, it would be less resilient than a system with 20 links expecting the same number of shocks.

The term $\sum_{j,k} \frac{jk}{\lambda} \mu(j,k)\, p(j,k,1)$ is a measure of the susceptibility of the network rather than its resilience, because the term increases for networks that are deemed more susceptible to contagion.

In order to measure resilience, the term $-\sum_{j,k} \frac{jk}{\lambda} \mu(j,k)\, p(j,k,1)$ should to be considered instead. It is added to one in the definition of the network resilience measure, though this is just a matter of shifting the range of the resilience measure to the right.

A possible reason for incorporating this shift could be that if the expected number of shocked links is more than the total links, then we would have that $\sum_{j,k} \frac{jk}{\lambda} \mu(j,k)\, p(j,k,1) > 1$ and the system could be seen as being 'unsafe'. Alternatively if the expected number of shocked links is less than the total links in the system, then $\sum_{j,k} \frac{jk}{\lambda} \mu(j,k)\, p(j,k,1) < 1$ and the system could be deemed as being relatively 'safe'. This explanation is supported further by theorems C.1 and C.2 below [8, Theorems 3.10 and 3.11].

Theorem C.1. *Let* $(\mathbf{e}_n, \vec{\gamma}_n)_{n\geq 1}$ *be a sequence of financial networks. Then if*

$$1 - \sum_{j,k} \frac{jk}{\lambda} \mu(j,k)\, p(j,k,1) > 0, \tag{C.2}$$

the final fraction of defaults becomes negligible with high probability if the initial fraction of defaults tends to zero.

This shows that if the contagion does not spread too widely after the first round of default, then the system will be able to withstand the shock.

Theorem C.2. *Let* $(\mathbf{e}_n, \vec{\gamma}_n)_{n\geq 1}$ *be a sequence of financial networks. If*

$$1 - \sum_{j,k} \frac{jk}{\lambda} \mu(j,k)\, p(j,k,1) < 0, \tag{C.3}$$

then with high probability there exists a (non-empty) set of nodes such that any node belonging to this set can trigger the default of all nodes in the set. Furthermore there is a directed path of contagious links[1] *from any node in this set to another node in the set.*

Note that this set of nodes will not necessarily trigger the default of all the nodes in the entire system. Only nodes belonging to the set will default as a result of a single defaulting node in the set. In contrast to the hypothesis of theorem C.1, the final defaulting fraction of nodes will not be negligible, given that one of the nodes in the aforementioned set defaults.

Branching process motivation

The results given by theorems C.1 and C.2 can be motivated by considering the connected clusters in the network and approximating them by a branching process.

Definition C.2 (Connected cluster)**.** The *connected cluster* of a node ν is the set of all nodes for which there exists a path of edges to ν. For our purposes, only incoming edges are considered when determining connected clusters. In other words, any node in this set can be reached from a path of incoming edges, starting at ν's incoming edges.

For illustration purposes, fix an arbitrary n and suppose that G is a sample of the random network $(\mathbf{E}_n, \vec{\gamma}_n)$, where all nodes are still in their non-defaulted state. Each connected cluster in G can be found by means of a branching process approximation. The branching process is not constructed directly, but a random walk which is representative of a branching process is constructed.

For example, suppose we want to find the connected cluster of the node ν. The root of the branching process is the node ν, and the children in the n^{th} generation are all the nodes that can be reached via n incoming links[2] starting at ν.

As mentioned above, a random walk representation of a branching process is constructed. For this stochastic process, we consider each node as being either active, inactive,

[1] A link is said to be contagious if it represents an exposure larger than the capital of the associated creditor node.

[2] Note that if a node η can be reached via either τ_1 or τ_2 incoming edges, then η will be in generation $\min\{\tau_1, \tau_2\}$ only, and in no other generation. This will become clear from the construction of the branching process.

or neutral, and we are interested in the number of active nodes at each time step. A node is said to be active if it has been added to the branching process, but the nodes connected to its incoming links still need to be added to the process. A node is said to be inactive if it has already been added to the branching process, and the nodes connected to its incoming links have also been added. Neutral nodes have not been added to the branching process yet.

Let S_t denote the number of active nodes at time t. Then the process is constructed as follows:

1. The node, say v, whose connected cluster we want to determine represents the root of the branching process. At time 0, v is the only active node in the system, while all the other nodes are neutral. Therefore $S_0 = 1$.

2. Suppose now that v has k_0 incoming edges. At time 1 these k_0 nodes become active, and v becomes inactive. Therefore

$$S_1 = k_0 = S_0 + k_0 - 1. \tag{C.4}$$

3. Now randomly choose any one of the currently active nodes. Suppose that this node has k_1 incoming edges. The chosen node now becomes inactive, while the k_1 nodes connected to its incoming edges become active. Now

$$S_2 = S_1 + k_1 - 1, \tag{C.5}$$

and for the n^{th} randomly chosen active node we have that

$$S_n = S_{n-1} + k_{n-1} - 1. \tag{C.6}$$

The random walk constructed above becomes zero when all the nodes belonging to the connected cluster of v have become inactive.

The set of inactive nodes at any time t represents the part of the connected cluster that has been determined up until time t. As only one node can be added to the connected cluster at each time step, then at time t there will be t nodes in the connected cluster of v (including v itself).

If S_t becomes zero at some point, then the population of active nodes dies out, and the connected cluster of v is a strict subset of the whole system. As the size of the system grows to infinity, the fraction that the connected cluster represents becomes negligible.

This corresponds to the case dealt with by theorem C.1, as losses cannot spread outside of the connected cluster.

On the other hand, if S_t does not become zero, the population of active nodes does not die out. In branching processes, populations usually either die out, or become very big. Therefore if the population does not die out with probability one, then the connected cluster has the potential of becoming very large, so that losses could potentially spread through the whole system. This corresponds to the case dealt with by theorem C.2.

However, theorems C.1 and C.2 consider nodes with default threshold one. Therefore the definition of a connected cluster needs to be modified slightly. Here, we need the connected cluster of a node ν to be the set of all nodes with default threshold one, for which there exists a path of incoming edges starting at ν.

In order to evaluate the extinction probability of a branching process, we need to evaluate the probability generating function of the offspring distribution (the distribution of the number of offspring produced by each member of the active population). In the context of this setting, a node's offspring is the nodes connected to it via its incoming edges. Therefore the in-degree distribution of a randomly chosen member of the active population is required.

The node ν has k_0 incoming edges with probability $\sum_j \mu(j,k)$. However, the degree distribution of the rest of the nodes are determined differently. Each subsequent addition to the active node population is randomly chosen from the incoming edges of the current active population. The chosen incoming edge will be an outgoing edge for the next node to be added to the active population.

Therefore we need the probability that a randomly chosen outgoing edge will belong to a node of degree (j,k). Recall from the explanation below equation (3.19) that $\frac{\mu(j,k)k}{\lambda}$ represents the fraction of incoming edges belonging to nodes of degree (j,k). Using a similar argument, it can be deduced that $\frac{\mu(j,k)j}{\lambda}$ represents the fraction of outgoing edges belonging to nodes of degree (j,k). In other words if one randomly chooses an outgoing edge, then the node whose outgoing edge was chosen will have degree (j,k) with probability $\frac{\mu(j,k)j}{\lambda}$.

As only nodes with a default threshold of one are included in the connected cluster, this needs to be allowed for as well. The probability that a randomly chosen outgoing edge will belong to a node with degree (j,k) and default threshold one is given by $\frac{\mu(j,k)j}{\lambda}p(j,k,1)$. The corresponding probability that a randomly chosen outgoing edge will belong to a node with k in-coming links and a default threshold of one is $\sum_j \frac{\mu(j,k)j}{\lambda}p(j,k,1)$. This gives the required offspring distribution, as it gives the distribution of the number of new nodes added to the active population at each time step.

Now that the in-degree distribution of the active population is determined, the extinc-

tion probability can be determined. From [61, chapter 1, theorem 6.1], it is known that the extinction probability q of the branching process is given by the smallest non-negative root of the equation

$$q = \sum_k p_k \, q^k. \tag{C.7}$$

Note that $q = 1$ will always be a solution to this equation, and therefore the smallest non-negative solution will always be contained in the interval $[0, 1]$. For the purpose of this example, the above equation can be rewritten as follows:

$$q = \sum_{j,k} \frac{\mu(j,k)\, j}{\lambda} p(j,k,1) \; q^k. \tag{C.8}$$

If the average number of offspring (or in-degrees in the cluster) is less than one, i.e.

$$\sum_{j,k} \frac{\mu(j,k)\, jk}{\lambda} p(j,k,1) < 1, \tag{C.9}$$

then the smallest non-negative solution is $q = 1$. In this case the population dies out with probability one.

In this context, it means that the connected cluster of a node v consisting of nodes with default threshold one constitutes a negligible fraction of the system. As the above inequality is independent of v, it means that all the connected clusters consisting of nodes with default threshold one will each constitute a negligible fraction of the system. Therefore if the initial fraction of defaults is small, then the spread of contagion should be limited to a small number of these clusters. As the nodes connected to the defaulting clusters have higher default thresholds, they should become less likely to default as the initial fraction of defaults decrease. This is therefore a heuristic argument to support the statement of theorem C.1.

If the average number of in-degrees in a cluster is greater than one, i.e.

$$\sum_{j,k} \frac{\mu(j,k)\, jk}{\lambda} p(j,k,1) > 1, \tag{C.10}$$

then the smallest non-negative solution to (C.8) is $q < 1$. The population does not necessarily die out here, so that there is a positive probability that a cluster can keep growing indefinitely. Such a cluster should in general constitute a positive fraction of the system.

Each node in the system can, with probability q, be the root of a cluster which constitutes a positive fraction of the system. As the number of nodes tends to infinity, the

probability of there being such a cluster tends to one. As the default of one node in the set will trigger the default of all the nodes after it in the branching process, this is serves as support for the statement of theorem C.2.

Approximating the final fraction of defaults by a function of susceptibility and connectedness of the initial defaulting nodes. The purpose of this part is to identify factors that determine how a small initial shock to the system will affect the final fraction of defaults. To do this, an approximation is found for the final fraction of defaults based on the fraction of initially defaulting nodes, $p(j,k,0) = \epsilon$. Let π^*_ϵ be the least fixed point of I based on ϵ.

Assume that the condition

$$1 - \sum_{j,k} \frac{jk}{\lambda} \mu(j,k)\, p(j,k,1) > 0, \tag{C.11}$$

from theorem C.1 is satisfied. Then since the final fraction of defaults becomes negligible for small enough ϵ, and the possibilities for π^*_ϵ covered by theorem C.1 are exhaustive, it implies that $\pi^*_\epsilon < 1$ and that

$$\alpha_n(\mathbf{E}_n, \vec{\gamma}_n) \to \sum_{j,k} \mu(j,k) \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j, \pi^*_\epsilon, \theta) \tag{C.12}$$

weakly as $n \to \infty$. The expression on the right can be regarded as a function of π^*_ϵ, say $f(\pi^*_\epsilon)$. If we let $f(\pi^*_\epsilon) = \sum_{j,k} \mu(j,k) \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j, \pi^*_\epsilon, \theta)$, then $f(\pi^*_\epsilon)$ can be approximated by

$$f(\pi^*_\epsilon) \approx f(0) + h'(0)\, \pi^*_\epsilon + o(\pi^*_\epsilon)\,, \tag{C.13}$$

where

$$\begin{aligned} f(0) &= \sum_{j,k} \mu(j,k) \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j,0,\theta) \\ &= \sum_{j,k} \mu(j,k) \sum_{\theta=0}^{j} p(j,k,\theta)\, \mathbb{P}(X \geq \theta)\,, \qquad \text{where } X \sim \text{Bin}\,(j,0) \\ &= \sum_{j,k} \mu(j,k)\, p(j,k,0) \\ &= \epsilon \sum_{j,k} \mu(j,k) = \epsilon. \end{aligned} \tag{C.14}$$

For $h'(0)$, note that

$$\begin{aligned}
h'(\pi_\epsilon^*) &= \frac{d}{d\pi_\epsilon^*}\left(\sum_{j,k}\mu(j,k)\sum_{\theta=0}^{j}p(j,k,\theta)\,\bar{B}(j,\pi_\epsilon^*,\theta)\right)\\
&= \sum_{j,k}\mu(j,k)\sum_{\theta=0}^{j}p(j,k,\theta)\,\frac{d}{d\pi_\epsilon^*}\bar{B}(j,\pi_\epsilon^*,\theta)\\
&= \sum_{j,k}\mu(j,k)\sum_{\theta=0}^{j}p(j,k,\theta)\,\frac{d}{d\pi_\epsilon^*}\left(\sum_{l\geq\theta}^{j}\binom{j}{l}(\pi_\epsilon^*)^l\,(1-\pi_\epsilon^*)^{j-l}\right)\\
&= \sum_{j,k}\mu(j,k)\sum_{\theta=0}^{j}p(j,k,\theta)\ldots\\
&\qquad\sum_{l\geq\theta}^{j}\binom{j}{l}\left[l\,(\pi_\epsilon^*)^{l-1}\,(1-\pi_\epsilon^*)^{j-l}-(j-l)\,(\pi_\epsilon^*)^l\,(1-\pi_\epsilon^*)^{j-l-1}\right].
\end{aligned}\tag{C.15}$$

If $\pi_\epsilon^* = 0$, then the terms inside the third summation above is only non-zero for $l = 0$ and $l = 1$, which in turn can only occur when $\theta = 0$ or $\theta = 1$. Therefore

$$\begin{aligned}
h'(0) &= \sum_{j,k}\mu(j,k)\sum_{\theta=0}^{1}p(j,k,\theta)\sum_{l=\theta}^{1}\binom{j}{l}\left[l\,(\pi_\epsilon^*)^{l-1}\,(1-\pi_\epsilon^*)^{j-l}-(j-l)\,(\pi_\epsilon^*)^l\,(1-\pi_\epsilon^*)^{j-l-1}\right]_{\pi_\epsilon^*=0}\\
&= \sum_{j,k}\mu(j,k)\left[p(j,k,0)\left[-j+j\right]+p(j,k,1)\,j\right]\\
&= \sum_{j,k}j\,\mu(j,k)\,p(j,k,1)\,.
\end{aligned}\tag{C.16}$$

Substituting the relevant expressions into (C.13), we find that

$$\sum_{j,k}\mu(j,k)\sum_{\theta=0}^{j}p(j,k,\theta)\,\bar{B}(j,\pi_\epsilon^*,\theta)=\epsilon+\pi_\epsilon^*\sum_{j,k}j\,\mu(j,k)\,p(j,k,1)+o(\pi_\epsilon^*)\,.\tag{C.17}$$

To expand this further, one can find an approximation for π_ϵ^* by finding a similar first-order approximation for $I(\pi_\epsilon^*)$:

$$I(\pi_\epsilon^*)=I(0)+\pi_\epsilon^*I'(0)+o(\pi_\epsilon^*)\,.\tag{C.18}$$

Here,

$$\begin{aligned} I(0) &= \sum_{j,k} \frac{\mu(j,k)\,k}{\lambda} \sum_{\theta=0}^{j} p(j,k,\theta)\,\bar{B}(j,0,\theta) \\ &= \sum_{j,k} \frac{\mu(j,k)\,k}{\lambda} p(j,k,0) \\ &= \epsilon \frac{1}{\lambda} \sum_{j,k} \mu(j,k)\,k = \epsilon \end{aligned} \tag{C.19}$$

and by using a similar procedure as for $h'(\pi^*_\epsilon)$,

$$\begin{aligned} I'(\pi^*_\epsilon) &= \frac{d}{d\pi^*_\epsilon} \left(\sum_{j,k} \frac{\mu(j,k)\,k}{\lambda} \sum_{\theta=0}^{j} p(j,k,\theta)\,\bar{B}(j,\pi^*_\epsilon,\theta) \right) \\ &= \sum_{j,k} \frac{\mu(j,k)\,k}{\lambda} \sum_{\theta=0}^{j} p(j,k,\theta) \ldots \\ &\qquad \sum_{l\geq\theta}^{j} \binom{j}{l} \left[l\,(\pi^*_\epsilon)^{l-1}\,(1-\pi^*_\epsilon)^{j-l} - (j-l)\,(\pi^*_\epsilon)^{l}\,(1-\pi^*_\epsilon)^{j-l-1} \right]. \end{aligned} \tag{C.20}$$

Therefore by the same reasoning as above,

$$\begin{aligned} I'(0) &= \sum_{j,k} \frac{\mu(j,k)\,k}{\lambda} \sum_{\theta=0}^{j} p(j,k,\theta) \sum_{l\geq\theta}^{j} \binom{j}{l} \left[l\,(\pi^*_\epsilon)^{l-1}\,(1-\pi^*_\epsilon)^{j-l} - (j-l)\,(\pi^*_\epsilon)^{l}\,(1-\pi^*_\epsilon)^{j-l-1} \right]_{\pi=0} \\ &= \sum_{j,k} \frac{\mu(j,k)\,k}{\lambda} \left[p(j,k,0)\,[-j+j] + p(j,k,1)\,j \right] \\ &= \sum_{j,k} \frac{\mu(j,k)\,jk}{\lambda} p(j,k,1)\,. \end{aligned} \tag{C.21}$$

Substituting the expressions for $I(0)$ and $I'(0)$ into equation (C.18),

$$I(\pi^*_\epsilon) = \epsilon + \pi^*_\epsilon \sum_{j,k} \frac{\mu(j,k)\,jk}{\lambda} p(j,k,1) + o(\pi^*_\epsilon)\,. \tag{C.22}$$

However as π^*_ϵ is a fixed point of I, then $I(\pi^*_\epsilon) = \pi^*_\epsilon$. By substituting this into the above approximation and solving for π^*_ϵ, we obtain the following:

$$\pi^*_\epsilon = \frac{\epsilon}{1 - \sum_{j,k} \frac{\mu(j,k)jk}{\lambda} p(j,k,1)} + \frac{o(\pi^*_\epsilon)}{1 - \sum_{j,k} \frac{\mu(j,k)jk}{\lambda} p(j,k,1)}$$

$$= \frac{\epsilon}{1 - \sum_{j,k} \frac{\mu(j,k)jk}{\lambda} p(j,k,1)} + o(\pi_\epsilon^*)\,. \tag{C.23}$$

This can now in turn be substituted into equation (C.17) to obtain

$$\begin{aligned}
&\sum_{j,k} \mu(j,k) \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j, \pi_\epsilon^*, \theta) \\
=&\epsilon + \pi_\epsilon^* \sum_{j,k} j\,\mu(j,k)\, p(j,k,1) + o(\pi_\epsilon^*) \\
=&\epsilon + \left(\frac{\epsilon}{1 - \sum_{j,k} \frac{\mu(j,k)jk}{\lambda} p(j,k,1)} + o(\pi_\epsilon^*) \right) \sum_{j,k} j\,\mu(j,k)\, p(j,k,1) + o(\pi_\epsilon^*) \\
=&\epsilon \left(1 + \frac{\sum_{j,k} j\,\mu(j,k)\, p(j,k,1)}{1 - \sum_{j,k} \frac{\mu(j,k)jk}{\lambda} p(j,k,1)} \right).
\end{aligned} \tag{C.24}$$

Therefore from theorem 3.2.3

$$\alpha_n(\mathbf{E}_n, \vec{\gamma}_n) \to \epsilon \left(1 + \frac{\sum_{j,k} j\,\mu(j,k)\, p(j,k,1)}{1 - \sum_{j,k} \frac{\mu(j,k)jk}{\lambda} p(j,k,1)} \right) \tag{C.25}$$

weakly as $n \to \infty$ and $\pi_\epsilon^* \to 0$.

Now consider the case where the initially insolvent nodes are only of degree (d^+, d^-), so that $p(d^+, d^-, 0) = \epsilon$ and $p(j,k,0) = 0$ for all $j \neq d^+$ and $k \neq d^-$.

The expression for $f(0)$ from equation (C.13) now becomes

$$\begin{aligned}
f(0) &= \sum_{j,k} \mu(j,k) \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j,0,\theta) \\
&= \sum_{j,k} \mu(j,k)\, p(j,k,0) \\
&= \mu(d^+, d^-)\, \epsilon
\end{aligned} \tag{C.26}$$

whereas $h'(0)$ remains the same. Therefore equation (C.17) becomes

$$\begin{aligned}
&\sum_{j,k} \mu(j,k) \sum_{\theta=0}^{j} p(j,k,\theta)\, \bar{B}(j, \pi_\epsilon^*, \theta) \\
=&\epsilon\, \mu(d^+, d^-) + \pi_\epsilon^* \sum_{j,k} j\,\mu(j,k)\, p(j,k,1) + o(\pi_\epsilon^*)\,.
\end{aligned} \tag{C.27}$$

The expression for $I(0)$ from equation (C.18) changes to

$$\begin{aligned} I(0) &= \sum_{j,k} \frac{\mu(j,k)\,k}{\lambda} p(j,k,0) \\ &= \frac{\mu(d^+,d^-)\,d^-}{\lambda}\epsilon \end{aligned} \tag{C.28}$$

and $I'(0)$ remains the same. Therefore equation (C.18) is now

$$I(\pi^*_\epsilon) = \epsilon\frac{\mu(d^+,d^-)\,d^-}{\lambda} + \pi^*_\epsilon \sum_{j,k} \frac{\mu(j,k)\,jk}{\lambda} p(j,k,1) + o(\pi^*_\epsilon)\,. \tag{C.29}$$

As before, the identity $I(\pi^*_\epsilon) = \pi^*_\epsilon$ is used to solve for π^*_ϵ in the above equation:

$$\begin{aligned} \pi^*_\epsilon &= \left(\epsilon\frac{\mu(d^+,d^-)\,d^-}{\lambda} + o(\pi^*_\epsilon)\right)\frac{1}{1-\sum_{j,k}\frac{\mu(j,k)jk}{\lambda}p(j,k,1)} \\ &= \frac{d^-}{\lambda}\frac{\epsilon\,\mu(d^+,d^-)}{1-\sum_{j,k}\frac{\mu(j,k)jk}{\lambda}p(j,k,1)} + o(\pi^*_\epsilon)\,. \end{aligned} \tag{C.30}$$

Substituting this into equation (C.27) results in

$$\begin{aligned} &\sum_{j,k}\mu(j,k)\sum_{\theta=0}^{j} p(j,k,\theta)\,\bar{B}(j,\pi^*_\epsilon,\theta) \\ =&\epsilon\,\mu(d^+,d^-) + \left(\frac{d^-}{\lambda}\frac{\epsilon\,\mu(d^+,d^-)}{1-\sum_{j,k}\frac{\mu(j,k)jk}{\lambda}p(j,k,1)} + o(\pi^*_\epsilon)\right)\sum_{j,k} j\,\mu(j,k)\,p(j,k,1) + o(\pi^*_\epsilon) \\ =&\epsilon\,\mu(d^+,d^-)\left(1 + \frac{d^-}{\lambda}\frac{\sum_{j,k} j\,\mu(j,k)\,p(j,k,1)}{1-\sum_{j,k}\frac{\mu(j,k)jk}{\lambda}p(j,k,1)}\right), \end{aligned} \tag{C.31}$$

so that

$$\alpha_n(\mathbf{E}_n,\vec{\gamma}_n) \to \epsilon\,\mu(d^+,d^-)\left(1 + \frac{d^-}{\lambda}\frac{\sum_{j,k} j\,\mu(j,k)\,p(j,k,1)}{1-\sum_{j,k}\frac{\mu(j,k)jk}{\lambda}p(j,k,1)}\right) \tag{C.32}$$

weakly as $n \to \infty$ and $\pi^*_\epsilon \to 0$.

The right hand side of the above expression is increasing in d^- and the network's susceptibility as measured by $\sum_{j,k}\frac{\mu(j,k)jk}{\lambda}p(j,k,1)$ (all else being held fixed). This shows that the final fraction of defaults is increased by the interconnectedness of the initial default and the network susceptibility. Since small values of ϵ are considered, it shows that these factors result increased damage to the system due to small initial shocks.

Bibliography

[1] D. Acemoglu, A. Ozdaglar, and A. Tahbaz-Salehi. Networks, shocks, and systemic risk. Working paper 20931, National Bureau of Economic Research, 2015.

[2] D. Aikman, P. Alessandri, B. Eklund, P. Gai, S. Kapadia, E. Martin, N. Mora, G. Sterne, and M. Willison. Funding liquidity risk in a quantitative model of systemic stability. Working paper 555, Central Bank of Chile, 2009.

[3] R. Albert. Scale-free networks in cell biology. *Journal of Cell Science*, 118(21):4947–4957, 2005.

[4] R. Albert and A. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47, 2002.

[5] P. Alessandri, P. Gai, S. Kapadia, N. Mora, and C. Puhr. Towards a framework for quantifying systemic stability. *International Journal of Central Banking*, 5(3):47–81, 2009.

[6] F. Allen and A. Babus. Networks in finance. Working paper 08–07, Wharton Financial Institutions Center, 2008.

[7] Y. Altunbas, D. Marqués-Ibáñez, and S. Manganelli. Bank risk during the financial crisis: do business models matter? Working paper 1394, ECB, 2011.

[8] H. Amini, R. Cont, and A. Minca. Resilience to contagion in financial networks. *Mathematical Finance*, 26(2):329–365, 2016.

[9] H. Amini and A. Minca. Inhomogeneous financial networks and contagious links. Working papers, HAL, 2014. `https://hal.inria.fr/hal-01081559/document`.

[10] H. Amini, A. Minca, and A. Sulem. Optimal equity infusions in interbank networks. *Journal of Financial stability*, 31:1–17, 2017.

[11] K. Anand, B. Craig, and G. Von Peter. Filling in the blanks: Network structure and interbank contagion. *Quantitative Finance*, 15(4):625–636, 2015.

[12] K. Anand, P. Gai, and M. Marsili. Rollover risk, network structure and systemic financial crises. *Journal of Economic Dynamics and Control*, 36(8):1088–1100, 2012.

[13] N. Arinaminpathy, S. Kapadia, and R. M. May. Size and complexity in model financial systems. *Proceedings of the National Academy of Sciences*, 109(45):18338–18343, 2012.

[14] R. Ayadi, W. P. De Groen, I. Sassi, W. Mathlouthi, H. Rey, and O. Aubrey. Banking business models of 2015 — Europe. Technical report, Alphonse and Dorimène Desjardins International Institute for Cooperatives and International Research Centre on Cooperative Finance, 2016.

[15] J. Balthrop, S. Forrest, M. E. Newman, and M. M. Williamson. Technological networks and the spread of computer viruses. *Science*, 304(5670):527–529, 2004.

[16] A. Baranga. The contraction principle as a particular case of Kleene's fixed point theorem. *Discrete Mathematics*, 98(1):75–79, 1991.

[17] L. Bargigli, G. Di Iasio, L. Infante, F. Lillo, and F. Pierobon. The multiplex structure of interbank networks. *Quantitative Finance*, 15(4):673–691, 2015.

[18] Basel Committee on Banking Supervision. Basel III. Monitoring report, BIS, February 2017.

[19] S. Battiston, D. D. Gatti, M. Gallegati, B. Greenwald, and J. E. Stiglitz. Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk. *Journal of Economic Dynamics Control*, 36(8):1121–1141, 2012.

[20] S. Battiston, M. Puliga, R. Kaushik, P. Tasca, and G. Caldarelli. Debtrank: Too central to fail? Financial networks, the FED and systemic risk. *Scientific Reports*, 2(541):1–6, 2012.

[21] P. Behr and R. H. Schmidt. The German banking system: Characteristics and challenges. White Paper Series No. 32, Sustainable Architecture for Finance in Europe, 2015.

[22] U. W. Birchler and M. Facchinetti. Self-destroying prophecies? The endogeneity pitfall in using market signals as triggers for prompt corrective action. Research task force paper, BIS, 2007.

[23] D. Bisias, M. Flood, A. W. Lo, and S. Valavanis. A survey of systemic risk analytics. *Annual Review of Finance*, 4, 2012.

[24] C. Borio, M. Drehmann, and K. Tsatsaronis. Stress-testing macro stress testing: Does it live up to expectations? *Journal of Financial Stability*, 12:3–15, 2014.

[25] M. Boss, H. Elsinger, M. Summer, and S. Thurner. Network topology of the interbank market. *Quantitative Finance*, 4(6):677–684, 2004.

[26] N. Brink and C.-P. Georg. Systemic risk in the South African interbank system. Discussion paper DP/11/01, South African Reserve Bank, August 2011.

[27] S. P. Brooks and B. J. T. Morgan. Optimization using simulated annealing. *Journal of the Royal Statistical Society. Series D*, 44(2):241–257, 1995.

[28] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin. Catastrophic cascade of failures in interdependent networks. *Nature*, 464(7291):1025–1028, 2010.

[29] F. Caccioli, T. A. Catanach, and J. D. Farmer. Heterogeneity, correlations and financial contagion. *Advances in Complex Systems*, 15(supp02):1250058, 2012.

[30] D. O. Cajueiro and B. M. Tabak. The role of banks in the brazilian interbank market: Does bank type matter? *Physica A: Statistical Mechanics and its Applications*, 387(27):6825–6836, 2008.

[31] G. Caldarelli and M. Catanzaro. *Networks A very short introduction*. Oxford University Press, 2012.

[32] J. M. D. Canedo and S. M. Jaramillo. A network model of systemic risk: Stress testing the banking system. *Intelligent Systems in Accounting, Finance and Management*, 16(1-2):87–110, 2009.

[33] J. Caruana. Systemic risk: How to deal with it? Other publications, BIS, February 2010. `http://www.bis.org/publ/othp08.htm`.

[34] V. Cerra and S. C. Saxena. Booms, crises, and recoveries: A new paradigm of the business cycle and its policy implications. Working Paper WP/17/250, IMF, 2017.

[35] N. Chen, X. Liu, and D. D. Yao. An optimization view of financial systemic risk modeling: Network effect and market liquidity effect. *Operations Research*, 64(5):1089–1108, 2016.

[36] M. Chinazzi and G. Fagiolo. Systemic risk, contagion, and financial networks: A survey. Working paper series No. 2013/08, LEM, 2015.

[37] R. Cifuentes, G. Ferrucci, and H. S. Shin. Liquidity risk and contagion. *Journal of the European Economic Association*, 3(2–3):556–566, 2005.

[38] L. Clerc, A. Giovannini, S. Langfield, T. Peltonen, R. Portes, and M. Scheicher. Indirect contagion: the policy problem. Occasional Paper Series 9, ESRB, 2016.

[39] R. Cont, A. Moussa, and E. B. Santos. Network structure and systemic risk in banking systems. In *Handbook of Systemic Risk*. Cambridge University Press, 2012.

[40] L. Corrado and T. Schuler. Interbank market failure and macro-prudential policies. *Journal of Financial Stability*, 33:133–149, 2017.

[41] B. Craig and G. von Peter. Interbank tiering and money center banks. *Journal of Financial Intermediation*, 23(3):322–347, 2014.

[42] O. De Bandt and P. Hartmann. Systemic risk: A survey. Working paper 35, European Central Bank, November 2000.

[43] H. Elsinger, A. Lehar, and M. Summer. Using market information for banking system risk assessment. *International Journal of Central Banking*, 2(1):137–165, 2006.

[44] P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae*, 6:290–297, 1959.

[45] Federal Deposit Insurance Corporation. Regulatory capital rules. *Federal Register*, 78(175), 2013.

[46] K. Finger, D. Fricke, and T. Lux. Network analysis of the e-MID overnight money market: the informational value of different aggregation levels for intrinsic dynamic processes. *Computational Management Science*, 10(2–3):187–211, 2013.

[47] R. W. Floyd. Algorithm 97: Shortest path. *Communications of the ACM*, 5(6):345, 1962.

[48] D. Fricke and T. Lux. Core-periphery structure in the overnight money market: Evidence from the e-MID trading platform. *Computational Economics*, 45(3):359–395, 2015.

[49] D. Furceri and A. Mourougane. Financial crises: past lessons and policy implications. Working Paper 668, OECD Economic Department, 2009.

[50] S. Gabrieli. The microstructure of the money market before and after the financial crises: a network perspective. Research paper series 181, CEIS Tor Vergata, January 2011.

[51] P. Gai, A. Haldane, and S. Kapadia. Complexity, concentration and contagion. *Journal of Monetary Economics*, 58(5):453–470, 2011.

[52] P. Gai and S. Kapadia. Contagion in financial networks. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 466(2120):2401–2423, 2010.

[53] R. J. Garratt, L. Mahadeva, and K. Svirydzenka. Mapping systemic risk in the international banking network. Working paper 413, Bank of England, March 2011.

[54] C.-P. Georg. The effect of the interbank network structure on contagion and common shocks. Discussion paper Series 2: Banking and financial studies 12/2011, Deutsche Bundesbank, 2011.

[55] C.-P. Georg and N. Brink. Note on interlinkages in the South African interbank system. Financial stability review, South African Reserve Bank, 2011.

[56] C.-P. Georg and J. Poschmann. Systemic risk in a network model of interbank markets with central bank activity. Jena Economic Research Papers 2010-033, Friedrich Schiller University, 2010.

[57] P. Glasserman and H. P. Young. How likely is contagion in financial networks? *Journal of Banking & Finance*, 50:383–399, 2015.

[58] P. Glasserman and H. P. Young. Contagion in financial networks. *Journal of Economic Literature*, 54(3):779–831, 2016.

[59] M. Gofman. Efficiency and stability of a financial architecture with too-interconnected-to-fail institutions. *Journal of Financial Economics*, 124(1):113–146, 2017.

[60] A. G. Haldane and R. M. May. Systemic risk in banking ecosystems. *Nature*, 469:351 – 355, January 2011.

[61] T. E. Harris. *The theory of branching processes*. Courier Corporation, 2002.

[62] N. Hautsch, J. Schaumburg, and M. Schienle. Financial network systemic risk contributions. *Review of Finance*, 19(2), 2015.

[63] Z. He and A. Krishnamurthy. A macroeconomic framework for quantifying systemic risk. Working paper 19885, National Bureau of Economic Research, February 2014.

[64] X. Huang, I. Vodenska, S. Havlin, and H. E. Stanley. Cascading failures in bi-partite graphs: Model for systemic risk propagation. *Scientific Reports*, 3(1219):1–9, 2013.

[65] X. Huang, H. Zhou, and H. Zhu. Assessing the systemic risk of a heterogeneous portfolio of banks during the recent financial crisis. *Journal of Financial Stability*, 8(3):193–205, 2012.

[66] A.-C. Hüser. Too interconnected to fail: A survey of the interbank networks literature. Working Paper Series No. 91, Research Center SAFE, 2015.

[67] G. Iori, G. De Masi, O. V. Precup, G. Gabbi, and G. Caldarelli. A network analysis of the Italian overnight money market. *Journal of Economic Dynamics and Control*, 32(1):259–278, 2008.

[68] H.-J. Kim, Y. Lee, B. Kahng, and I.-M. Kim. Weighted scale-free network in financial correlations. *Journal of the Physical Society of Japan*, 71(9):2133–2136, 2002.

[69] J. Kim and T. Wilhelm. What is a complex graph? *Physica A: Statistical Mechanics and its Applications*, 387(11):2637–2652, 2008.

[70] A. Krause and S. Giansante. Interbank lending and the spread of bank failures: A network model of systemic risk. *Journal of Economic Behavior & Organization*, 83(3):583–608, 2012.

[71] S. Lenzu and G. Tedeschi. Systemic risk on different interbank network topologies. *Physica A: Statistical Mechanics and its Applications*, 391(18):4331–4341, 2012.

[72] L. Li, D. Alderson, J. C. Doyle, and W. Willinger. Towards a theory of scale-free graphs: Definition, properties, and implications. *Internet Mathematics*, 2(4):431–523, 2005.

[73] S. Markose, S. Giansante, M. Gatkowski, and A. R. Shaghaghi. Too interconnected to fail: Financial contagion and systemic risk in network model of CDS and other credit enhancement obligations of US banks. Discussion paper 683, University of Essex Department of Economics, February 2010.

[74] S. Martínez-Jaramillo, B. Alexandrova-Kabadjova, B. Bravo-Benítez, and J. P. Solórzano-Margain. An empirical study of the Mexican banking system's network and its implications for systemic risk. *Journal of Economic Dynamics and Control*, 40:242–265, 2014.

[75] R. M. May and N. Arinaminpathy. Systemic risk: the dynamics of model banking systems. *Journal of the Royal Society*, 7(46):823–838, 2010.

[76] R. M. May, S. A. Levin, and G. Sugihara. Ecology for bankers. *Nature*, 451(21), 2008.

[77] P. E. Mistrulli. Assessing financial contagion in the interbank market: Maximum entropy versus observed interbank lending patterns. *Journal of Banking and Finance*, 35(5):1114–1127, 2011.

[78] E. Nier, J. Yang, T. Yorulmazer, and A. Alentorn. Network models and financial stability. Working paper 346, Bank of England, April 2008.

[79] T. Opsahl and P. Panzarasa. Clustering in weighted networks. *Social Networks*, 31(2):155–163, 2009.

[80] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200, 2001.

[81] J. Persky. Retrospectives: Pareto's law. *The Journal of Economic Perspective*, 6(2):181–192, 1992.

[82] R. Roengpitya, N. Tarashev, and K. Tsatsaronis. Bank business models. Quarterly review, BIS, December 2014.

[83] T. Roukny, H. Bersini, H. Pirotte, G. Caldarelli, and S. Battiston. Default cascades in complex networks: Topology and systemic risk. *Scientific Reports*, 3(2759):1–8, 2013.

[84] E. B. Santos and R. Cont. The Brazilian interbank network structure and systemic risk. Working Paper Series 2019, Banco Central do Brasil, 2010.

[85] W. Silva, H. Kimura, and V. A. Sobreiro. An analysis of the literature on systemic financial risk: A survey. *Journal of Financial Stability*, 28:91–114, 2017.

[86] T. Snijders. The statistical evaluation of social network dynamics. *Sociological Methodology*, 31:361395, 2001.

[87] W. Souma, Y. Fujiwara, and H. Aoyama. Complex networks and economics. *Physica A: Statistical Mechanics and its Applications*, 324(1-2):396–401, 2003.

[88] V. Stoltenberg-Hansen, I. Lindström, and E. R. Griffor. *Mathematical theory of domains*, volume 22 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1994.

[89] M. Summer. Financial contagion and network analysis. *Annual Review of Financial Economics*, 5:277–97, 2013.

[90] T. Tao. *An introduction to measure theory*, volume 126 of *Graduate Studies in Mathematics*. American Mathematical Society, 2011.

[91] A. Temizsoy, G. Iori, and G. Montes-Rojas. Network centrality and funding rates in the e-MID interbank market. *Journal of Financial Stability*, 33:346–365, 2017.

[92] M. Teteryatnikova. Systemic risk in banking networks: Advantages of "tiered" banking systems. *Journal of Economic Dynamics and Control*, 47:186–210, 2014.

[93] C. Upper. Simulation methods to assess the danger of contagion in interbank markets. *Journal of Financial Stability*, 7(3):111–125, 2011.

[94] C. Upper and A. Worms. Estimating bilateral exposures in the German interbank market: Is there a danger of contagion? *European Economic Review*, 48(4):827–849, 2004.

[95] F. Vallascas and K. Keasey. Bank resilience to systemic shocks and the stability of banking systems: Small is beautiful. *Journal of International Money and Finance*, 31(6):1745–1776, 2012.

[96] D. Watts and S. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.

[97] S. Wells. Financial interlinkages in the United Kingdoms interbank market and the risk of contagion. Working paper 230, Bank of England, 2004.