Universitat Jaume I

Department d'Economia

Macro-prudential policy *via* network analysis

PhD Thesis

Giulia Provenzano

Supervisors: Dr. Simone Alfarano and Dr. Thomas Lux

Castelló de la Plana, July 2017

Contents

Pessimismo dell'intelligenza, ottimismo della volontá
Antonio Gramsci

Chapter 1

Introduction

The global financial crisis of 2007-2009 has demonstrated that the financial system should be regarded as a complex network whose nodes are financial institutions and links are financial dependencies. Since the financial system is becoming more complex and interconnected, network theory provides the ideal toolkit to study systemic crises and contagion events, proving to be crucial for the design of an appropriate macroprudential policy.

Indeed, while a shock might affect only a small number of banks, the interconnectedness of the financial system can trigger a default cascade, where contagion is transmitted to banks not directly exposed to the initial shock. So far, most of network literature focuses exclusively on a single kind of bank-to-bank connection: unsecured interbank lending. However, banks are directly linked to each other through a wide variety of financial obligations: CDS protections, payment transfers and secured (repo) lending. Recently, indirect layers of network connectivity have been also added. In particular, indirect connections form the backdrop for a new and fast growing literature that explores: (i) how the reduction in value of an asset spreads across banks investing in the same assets (ii) how the default of a real sector firm generates distress to its creditor banks.

Within this evolving literature, the aim of the paper **The topology of the bank-firm credit network in Spain, 1997-2007** is to provide an empirical analysis of the Spanish bank-firm network to study the structure of the connections between banking sector and real economy. We use data from the SABI database to obtain the firm-bank relationships for a sample consisted of over three hundred banks and over two thousand firms in the period 1997-2007. Particular emphasis is given to the network of co-financing banks, which is of primary importance to understand how the exposure to the same firms could create a potential source of systemic risk for the banking sector. We find three key results. First, credit extended to non-financial firms shares basic stylized

facts documented for other countries: the distribution of the number of lending relationships is much more fat-tailed for banks than for firms and the number of lending relationships increases with size for both bank and firms. Second, the network of co-financing banks is close to a core-periphery structure. Third, a community structure has been detected in the most significant banks' co-financing relationships.

While the previous paper focuses on how the topology of bank interconnections affects systemic risk, the paper **The controllability of the repo market in Mexico** investigates how the structure of interbank networks influences the control of central banks on money markets. By using Mexico's Central Bank data on interbank transactions in the repo market from 2005 to 2014, we study the control properties of the Mexican interbank network to detect which banks control the reallocation of central bank liquidity in the repo market. Our main finding is the identification of a set of controller banks, which does not correspond with the most connected banks or the largest lenders of the Mexican repo market. These results point out that a too big to fail policy might be inadequate to predict the systemically important banks. Moreover, we find that the repo market is more difficult to control over shorter time horizons, adding further evidence that central bank actions need to be adjusted to short-run and long-run policy goals.

In the last paper **Stress-testing the UK banking system: a network approach to cope with portfolio overlaps**, we go further beyond the traditional channel of contagion created via interbank credit and we analyze the contagion effects of a toxic asset in the balance sheets of banks. Indeed, the UK buy-to-let mortgage boom provides us with an excellent experimental setting to study this contagion channel. By using BankScope balance sheet data for over two hundred banks at the end of 2015, we assess the resilience of the UK banking system to a shock to real estate loans by estimating the extent to which portfolio overlaps amplify losses and defaults. The model that we consider for our stress-testing exercise is the Huang et al. (2013) algorithm, recently introduced in network literature. We find that the systemic impact of real estate loans has increased in the period 2007-2015. The stress-test results indicate that market conditions determine the size of the default cascade regardless of the intensity of the initial shock to real estate loans. This finding underlines the importance of the interconnectedness between banks' portfolios in stressed markets, that is also confirmed by the inadequacy of bank-specific features to predict defaults.

Introdución

La crisis financiera mundial de 2007-2009 ha demostrado que el sistema financiero debe considerarse como una red compleja cuyos nodos son instituciones financieras y que los vínculos son dependencias financieras. Dado que el sistema financiero se está volviendo más complejo y interconectado, la teoría de la red proporciona el conjunto de herramientas ideal para estudiar las crisis sistémicas y los eventos de contagio, demostrando ser crucial para el diseño de una política macroprudencial adecuada.

De hecho, si bien un choque puede afectar sólo a un pequeño número de bancos, la interconexión del sistema financiero puede desencadenar una *default cascade*, donde el contagio se transmite a los bancos no directamente expuestos al choque inicial. Hasta ahora, la mayor parte de la literatura de la red se centra exclusivamente en una único tipo de conexión de banco a banco: el crédito interbancario no garantizado. Sin embargo, las conexiones entre bancos son de varios tipos: protecciones CDS, transferencias de pago y crédito interbancario garantizado. Recientemente, *layers* indirectos de conectividad de red han estado también investigando. En particular, las conexiones indirectas forman el telón de fondo de una literatura nueva y de rápido crecimiento, que explora: (i) cómo la reducción en el valor de un activo se propaga entre los bancos que invierten en los mismos activos (ii) cómo el *default* de una empresa del sector real genera problemas financieros a sus bancos acreedores.

Dentro de esta literatura en evolución, el objetivo del artículo **The topology of the bank-firm credit network in Spain, 1997-2007** es proporcionar un análisis empírico de la red banco-empresa española para estudiar la estructura de las conexiones entre el sector bancario y la economía real. Utilizamos datos de la base de datos SABI para obtener las relaciones banco-empresa para una muestra de más de trescientos bancos y dos mil empresas en el período 1997-2007. Se hace especial atención en la red de bancos cofinanciadores, que es de primordial importancia para comprender cómo la exposición a las mismas empresas podría crear una fuente potencial de riesgo sistémico para el sector bancario. Encontramos tres resultados clave. En primer lugar, el crédito concedido a las empresas no financieras comparte *stylized facts* documentados para otros países: la distribución del

número de relaciones crediticias es mucho más *fat-tailed* para los bancos que para las empresas y el número de relaciones crediticias aumenta con el tamaño tanto por los bancos que por las empresas. En segundo lugar, la red de bancos cofinanciadores está cerca de una estructura *core-periphery*. En tercer lugar, se ha detectado una *comunity* en las relaciones de cofinanciación bancarias más significativas.

Mientras que los articulo anteriore se centra en cómo la topología de las diferentes interconexiones bancarias afecta el riesgo sistémico, el artículo **The controllability of the repo market in Mexico** investiga cómo la estructura de las redes interbancarias influye en el control de los bancos centrales en los mercados monetarios. Utilizando los datos del Banco de México sobre las transacciones interbancarias en el mercado de repos desde 2005 a 2014, estudiamos las propiedades de control de la red interbancaria mexicana para detectar qué bancos controlan la reasignación de liquidez del banco central en el mercado de repos. Nuestro principal resultado es la identificación de un grupo de bancos controladores, que no corresponde con los bancos más conectados o con los mayores prestamistas del mercado de repos mexicano. Estos resultados señalan que una política *too big to fail* podría ser inadecuada para predecir los bancos sistemicamente importantes. Por otra parte, encontramos que el mercado de repos es más difícil de controlar sobre horizontes de tiempo más cortos, añadiendo más pruebas de que las acciones de los bancos centrales deben ajustarse a los objetivos de política a corto y largo plazo.

En el ultimo artículo **Stress-testing the UK banking system: a network approach to cope with portfolio overlaps** vamos más allá del canal tradicional de contagio creado a través del crédito interbancario y analizamos los efectos de contagio de un activo tóxico en los balances de los bancos. De hecho, *the buy-to-let mortgage boom* en Reino Unido nos proporciona un excelente escenario experimental para estudiar este canal de contagio. Mediante el uso de datos de BankScope sobre las balanzas de más de doscientos bancos del Reino Unido a finales de 2015, evaluamos la resiliencia del sistema bancario del Reino Unido a un shock a los préstamos inmobiliarios al estimar la medida en que los *portfolio overlaps* amplifican las pérdidas y los *defaults*. Como modelo de riferimento, utilizamos el modelo de red de activos bancarios de Huang et al. (2013), recientemente introducido en la literatura de la red. Encontramos que el impacto sistémico de los préstamos inmobiliarios ha aumentado en el período 2007-2015. Los resultados indican que las condiciones del mercado determinan los *defaults* independientemente de la intensidad del choque inicial en los préstamos inmobiliarios. Este resultado señala la importancia de la interconexión entre las carteras de los bancos, lo que se confirma también por la insuficiencia de las características específicas de los bancos para predecir los *defaults*.

Chapter 2

The topology of the bank-firm credit network in Spain, 1997-2007

Co-Authored with: Mattia Montagna (European Central Bank), Thomas Lux (University of Kiel and University Jaume I) and Simone Alfarano (University Jaume I)[1]

[1]The views expressed in this paper do not reflect those of the European Central Bank (ECB). The authors thank Manuel Illueca for sharing his data set of bank-firm financial relations collected from the SABI (Sistema de Análisis de Balances Ibéricos) registry.

2.1 Introduction

The 2007-2008 financial crisis and the collapse of the interbank lending brought a surge of interest in the topology of the interbank market and formal modelling of interbank linkages via network analysis. This approach allows to take advantage of empirical concepts and theoretical models developed in network theory to study the structure of the banking system and its implications for systemic risk. Examples of empirical applications of network theory to the interbank market are: Bech et al. (2010) for the US Federal funds market, Boss et al. (2004) for the Austrian interbank market, Iori et al. (2008) for the Italian interbank market, Martinez-Jaramillo et al. (2014) for the Mexican interbank market and Soromäki et al. (2007) for the US Fedwire network.

However, direct interbank credit is only one particular source of systemic risk. The joint exposure via loans to non-financial firms is another important channel of distress in the banking system that so far has received much less attention than interbank credit. To the best of our knowledge, references are limited to: De Masi and Gallegati (2012) for the Italian bank-firm network De Masi et al. (2011), Marotta et al. (2015) for Japanese data and De Castro Miranda and Tabak (2013) for the Brazilian credit network. They have found that: (i) bank-firm networks are quite large (i.e. 39,194 firms and 508 banks for the Italian dataset in 2003) (ii) banks have more links and a more heterogeneous degree distribution than firms (iii) the number of links tends to increase with the size for both banks and firms, although there is no monotonic relationship between size and degree (iv) clustering coefficients are usually quite small and there is an inverse relationship between degree and clustering.

In these studies, the set of bank-firm credit relationships is represented as a *bipartite* network. Bipartite networks are composed of two disjoint sets of nodes, such that every link connects a node in the first set (banks) with a node of the second set (firms). Most of the literature on bank-firm networks focuses on, as it was previously mentioned, the bank-projection. The bank- projection is the network of co-financing banks, which is obtained by setting a link between banks when they share common firms in the original bipartite network, providing an insight into the structure of interconnectedness across the interbank system via its exposure to borrowers in the real sector.

However, while monopartite networks have been extensively studied, considerably less attention has been paid to bipartite networks (for review see Dormann et al. (2009)). Many prior studies actually tend to analyze the projections with the current null models available for monopartite networks. But such an approach suffers from several limitations because the information on the heterogeneity in degree associated with the two sets of nodes in the original bipartite network is partially lost in the one-mode projections, as demonstrated in the emerging literature on null models for bipartite networks. For example, Saracco et al. (2015) extend the null model for monopartite networks by Squartini and Garlaschelli (2011) to bipartite networks. They generate randomized ensembles of bipartite networks based on the maximum entropy principle, where each network realization of the

randomized ensemble is assigned a probability that maximizes the Shannon-Gibbs entropy under the constraint that the randomized ensemble has on average the same degree sequences of the two sets of nodes as the real network. By contrast, Tuminello et al. (2011) do not enforce the constraint on the degree sequences in the bipartite network on average. They rather use the hypergeometric distribution as null model for each link of the one-mode projections, which allows to take into account the degrees of the endpoints in the bipartite network in order to statistically validate the presence of a link in the projections.

In this paper, we provide an empirical analysis of the network structure of the Spanish credit market from 1997 to 2007 based on the SABI (Sistema de Análisis de Balances Ibéricos) database. We describe the yearly bank-firm relationships in terms of binary bipartite networks (loan exposures are not available). To anticipate the results, our basic data-analytical findings are in line with previous literature: banks exhibit a fat-tailed distribution of degrees, while firms tend to have a single or small number of lending relationships and the number of credit relationships is size dependent (non monotonically) for firms. We also investigate the properties of the bank-firm network in terms of degree correlation and clustering against a null-hypothesis preserving bank and firm degrees, but the results are statistically indistinguishable from the null-model. We also find that a significant persistence in bank-firm relationships from one year to the next exists.

We go beyond previous analyses by applying a core-periphery approach to the one-mode projection for banks. As it turns out, this framework provides a robust characterisation of the linkages created via joint exposures with a small error score that could not be spuriously explained by other popular network generating mechanisms such as purely random or scale-free networks. Our results also hold against random networks preserving the degree sequences of the bipartite network of banks and firms. To our knowledge this is the first time, that a core-periphery model has been applied to this particular layer of connectivity of financial institutions. Another important finding is that the structure of the bank-projection is strongly affected by the bank degrees in the original bipartite network. We follow the approach by Tuminello et al. (2011) and we provide evidence that around 90% of bank relationships are explained as a consequence of the bank lending activity and it would be therefore misleading to identify them as banks sharing preferentially the same borrowers.

This paper is organized as follows. Section 2 introduces the data set for the Spanish bank-firm network. Section 3 describes the network measures used to describe the bipartite bank-firm network. Section 4 presents the topological analysis for the bank-firm network and Section 5 for the bank network. Section 6 reports the results of the estimation of the core-periphery model to the bank network. Section 7 introduces the method used to track the time evolution of the preferential bank relationships and the results obtained. Section 8 concludes.

2.2 The data set and the structure of the Spanish banking sector

During the period from 2000 to 2007 the Spanish economy has experienced a period of high growth (the so-called "Spanish economic miracle"), followed by a deep crisis. The Spanish system of banks and firms has some peculiar aspects. First of all, we observe the presence of savings banks (the so-called cajas de ahorros). They are historically focused on retail banking, especially on mortgage lending to households, with little experience and expertise in commercial lending. The *cajas* have historically been non-profit organization, since their profits must be either retained or distributed in cultural and social community programs. The governance of savings banks is shared among representatives of several stakeholder groups: public authorities (from local and regional government), the founding entity, depositors and workers. Moreover, the banking system is characterized by commercial banks and credit cooperatives. The former are for-profit organizations under shareholders control, performing universal banking-services according to reasonable efficiency standards compared with banks from other European countries. Credit cooperatives are the smallest and most numerous of Spain's financial institutions. They are quite different from other banks, although they compete under bank legislation in loan, deposit and financial service markets. On one side, they are customer-based in comparison to product-based banks, as they are created specifically for small-size entrepreneurs based rurally in the least privileged areas of the country. On the other, they are owned by their members, since every depositor becomes an owner.

In 1989 a national law allowed saving banks to open branches nationwide, removing any entry restrinctions in geographical markets, hence their legal status became close to that of a commercial no-profit organization, with a growth strategy in commercial lending. Therefore, the Spanish banking system transformed from a strongly regulated oligopoly in the 1960s and 1970s into a highly liberalised system in the early 1990s, when commercial and savings banks can freely compete on prices and services.

As a consequence of the liberalisation, savings banks increased significantly the number of lending relationships with SMEs from 26 percent in 1996 to 35 percent in 2005 (Montoriol-Garriga (2008)), mainly in order to expand to new provinces.[2] Illueca et al. (2014) have documented how lending standards have deteriorated in the years before the crisis due to the attempts of the cajas to get market access to other provinces.

As the global financial crisis started with the consequent burst of the real estate in Spain, a considerable number of savings banks was converted into banks (cajas-banks) or merged with healthier savings banks after a rescue operation financed by the bank restructuring fund (FROB) of the Bank of Spain, amounting to about 16 billion euros (Etxezarreta et al. (2011)). This is a clear evidence in support of the crucial role of financial liberalization processes (and the corresponding

[2]Most notably, lending activity of savings banks was concentrated on the building and real estate sector and mortgages. Moreover, many of them had even their own real estate companies.

fast geographic expansion of savings banks) in the Spain's boom-bust.

Another point of interest of Spanish economy is the presence of few global player firms in strategic sectors, the so-called national champions, such as: Telefonica, Banco Santander, Banco Bilbao Vizcaya Argentaria (BBVA), Zara, Mango, Repsol, Ferrovial, Sacyr, Abertis and Aguas de Barcelona. Most of them were previously public companies, that were privatised during the 90s. During the boom period, the oligopoly power (or monopoly power in some cases) in the domestic market gave them the chance to expand overseas throughout the credit provided by the international markets. Clearly, it implied a growth of the external debt, due to an increase in both the deficit in current account (at least in the short run) and the debt of financial institutions.

It seems interesting under this historical perspective to study the salient features and the evolution of the bank-firm credit network as it has developed over this time of major structural changes. Here, we analyze a comprehensive data set extracted from the SABI (Sistema de Análisis de Balances Ibéricos) database, which is based on the public commercial registry in Spain. This database contains accounting information and additional information on firm characteristics on 268,485 Spanish firms for the period from 1997 to 2007, which is the final phase of savings bank geographic expansion. The available information we have on firms is: the size (total asset), the headquarter location (province), the sector and the identity of banks with whom they have financial relations (Bank of Spain Code). Most notably, in our analysis a firm is assumed to have a lending relationship with a bank when a firm reports the bank ID number, although the data does not strictly indicate a lending relationship, but just a provision of financial services (following Illueca et al. (2009) among others). Moreover, we point out that we do not know the name of the bank, though the variable "bank ID number" tracks the type of financial institution, i.e. commercial bank, savings bank and credit cooperative.[3] The data does not provide bank and loan information, i.e. the number of loans, their maturity (short term or long term) and their amount. In total, 263 banks (128 commercial banks, 52 savings banks and 83 credit cooperatives) and 268,485 firms were active in the lending market at least once during the period of analysis, while only 22,149 firms were active in all of the years in the sample period. The sample is comprised of 2,129,733 bank-firm links corresponding to 542,913 bank-firm relationships. The maximum number of active firms is reached in 2006 with 206,106 firms corresponding to 349,054 bank-firm relationships.[4] At the yearly level on average firms exhibit 1.86 bank relationships and firms borrow from no more than ten banks and 53% of firms have a single bank relationship, while 90% of firms maintain a relationship with a number between one and three of banks. The average number of bank relationships is, therefore, very close to the numbers reported for Italy and Japan using similarly comprehensive data sets (de Masi et al. (2011), de Masi and Gallegati (2013)).

[3]The first digit of the bank ID number identifies the type of financial institution, i.e. the first digit is equal to 0 or 1, 2 and 3 if the bank is a commercial bank, savings bank or credit cooperative respectively.

[4]Foreign banks are not present in the data set.

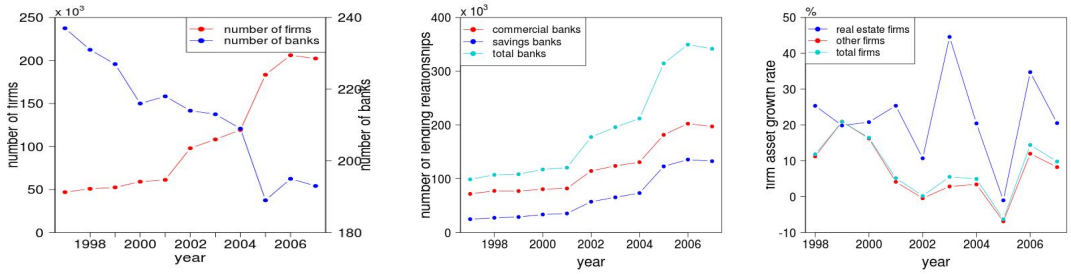


Figure 2.1: The number of banks and firms (left), the number of lending relationships for each type of bank (commercial bank and savings bank) and for the total banks (middle) and the asset growth rate of the real estate and construction firms, the other firms and the total firms (right) in each year from 1997 to 2007.

From the left panel of Figure 2.1, we see a clear upward trend in the number of firms over time, along with a slight and constant decrease in the number of banks. The overall numbers reflect the boom in the period 2005-2007: in 2005 the number of firms was roughly four times higher than in 1997, while the number of banks fell to their lowest level. However, we should note that the overall downward trend of the number of banks was due to the the decrease in the number of the commercial banks, while the number of savings banks remained almost unchanged. This suggests that the main reasons for the decrease of the number of commercial banks are the mergers and acquisitions (M&A) within the commercial banking.[5] It is important to mention that the ownership structure of savings banks implies that they cannot be acquired by other banks (i.e. commercial banks) or companies outside the savings bank sector, as they are owned by the respective local or regional public authority. Moving to the number of lending relationships in the central panel of Figure 2.1, we observe, as expected, the same general pattern as with the number of firms: the number of lending relationships increases susbstantially over time, especially between 2005 and 2007 it is roughly three times higher than in 1997. Interestingly, the lending activity of savings banks grew at a considerably faster rate than that of commercial banks (4.48% between 1997 and 2007, which was roughly twice the increase of lending by commercial banks).

If we consider the evolution of capital of real estate and construction firms in the right panel of Figure 2.1, we also observe that these growth rates soared during the years of higher increase of wholesale financing, peaking in 2003 with a growth rate of 44.6%. Compared to real estate and construction firms, the asset growth rates of the other firms are lower and much more stable, with an average annual growth rate of around 8%. We remark that we focus on total assets of non-financial firms[6] active in any of the years in the sample period, which are winsorized at the

[5]For instance, Montoriol-Garriga (2008) documents that in the period 1996-2005 40 mergers and acquisitions (M&A) took place: 24 mergers were between two commercial banks, 8 acquisitions of a commercial bank by a savings bank, 4 mergers between two savings banks and 4 mergers involved one official credit institution.

[6]In particular, we dismiss companies from the following sectors: Depository Institutions, Non-depository Credit In-

98th percentile of the distribution of the average annual asset growth rate. As a result, our final sample consists of 22,149 firms. This assumption is, however, quite restrictive, as the composition of banks changes over time.

2.3 Network analysis

We start with some basic definitions: a *network (or graph)* g is defined by a pair of sets *(N, E)*, which stands for the set $N = 1, ..., n$ of *nodes (or vertices)* and the set E of *edges* (or *links*) respectively. For an edge $e = uv \in g$, the vertices u and v are its ends.

The *order* of g is the number of nodes and the *size* of g is the number of edges established in the network.

A graph is *directed* if the edges are a set of ordered pairs. That is, the edges have a direction $uv \neq vu$. Otherwise, the graph is *undirected*. A further generalization is to assign a value $w(e)$ to each one of the edges e. In this case, we would speak of a *weighted* graph.

A network g is a *bipartite* (or *two-mode*) network if the set N has a partition to two subsets X and Y such that each edge $uv \in g$ connects a vertex of X and a vertex of Y. That is, all the edges are between nodes that are separated into two classes of nodes (*two-mode*). In this case, g is *(X, Y)-bipartite* and can be "projected" into the two *one-mode projections*. The X and Y-projections are two new graphs composed of nodes belonging to the same set, where there is a link between two nodes if they share common links in the original (X,Y)-bipartite graph.

In addition to the graphical visualization, a network is represented from a mathematical point of view by an *adjacency matrix* $\mathbf{A}$, where the element $a_{ij} = 1$ means that an edge between nodes i and j exists. Otherwise, $a_{ij} = 0$.[7] For a weighted graph, a w_{ij} element of the *weighted matrix* $\mathbf{W}$ is the weight (a real number) associated with the edge between nodes i and j.

In the following, we will introduce a list of metrics which are commonly used to describe the topology of financial networks.

2.3.1 Density

The *density* (or *connectivity*) of a network is defined as the proportion of actual links (E) relative to the total possible number of links. For an undirected network, the density is computed as:

stitutions, Security and Commodity Brokers, Dealers, Exchanges and Services, Insurance Carriers, Insurance Agents, Brokers and Services (NACE Rev.2 codes 64,65,66).

[7] In everything that follows, bold capital letters refer to matrices and bold lower-case letters to vectors.

$$\rho = \frac{2E}{N(N-1)} \tag{2.1}$$

This metric has to be divided by two for a directed graph. The index is closer to 1, as the network gets denser. In the extreme case where all the possible edges are established, the graph is *complete* and the index is equal to 1. Conversely, the graph with no edges has an index of 0.

2.3.2 Degree

The *degree* of a node i measures the number of nodes that a node is connected to. The degree k_i of the node i is defined as:

$$k_i = \sum_{j \in N(i)} a_{ij} \tag{2.2}$$

where $N(i)$ is the set of neighbors of vertex i.

In a directed network, the *out-degree* k_i^{out} is the number of links which originate from the node i and the *in-degree* k_i^{in} is the number of links which end at the node i.

2.3.3 Assortativity

Assortativity or degree correlation refers to the similarity between the degree of a node i and the degrees of its neighbours (Newman (2003)):

$$k_{nn}(i) = \frac{1}{k_i} \sum_{j \in N(i)} k_j \tag{2.3}$$

A network is *assortative* if nodes with high-degree are more likely to have links with nodes of similar degree, i.e. $k_{nn}(i)$ is increasing with k_i. Conversely, a network is *disassortative* when nodes with high degrees are more likely to be connected to nodes with low degrees (and vice versa), i.e. $k_{nn}(i)$ is decreasing with k_i. In a *uncorrelated* network, there is no dependence between the degree of a node and the degrees of its neighbors.

For bipartite networks, the *Pearson correlation* is widely used to compute the degree assortativity because in bipartite networks nodes are divided in two classes of nodes with two different sets of degree α and β:

$$r(\alpha,\beta) = \frac{E^{-1}\sum_i \left[(j_i^\alpha - \overline{j^\alpha})(h_i^\beta - \overline{h^\beta})\right]}{\sigma^\alpha \sigma^\beta} \tag{2.4}$$

where E the number of edges, j_i^α and h_i^β the α and β degree of the end vertices of edge i, $\overline{j^\alpha} = E^{-1}\sum_i j_i^\alpha$, $\sigma^\alpha = \sqrt{E^{-1}\sum_i (j_i^\alpha - \overline{j^\alpha})^2}$, $\overline{h^\beta}$ and σ^β similarly defined.

2.3.4 Clustering coefficient

The *clustering coefficient* is a measure of the density of the connections around a vertex. In a binary undirected network the clustering coefficient of a vertex i is defined as (Watts and Strogatz (1998)):

$$c_i = \frac{2}{k_i(k_i-1)} \sum_{j,h} a_{ij} a_{ih} a_{jh} \tag{2.5}$$

The clustering coefficient represents the proportion of nearest neighbors of a node i, which are linked to each other and it is calculated as the fraction of triangles present in the network out of all the possible triangles that could be formed. The *average clustering coefficient* C measures the density of the triangles in the system:

$$C = \frac{1}{N}\sum_i^N c_i \tag{2.6}$$

The traditional clustering coefficient cannot be used to quantify the clustering pattern of a bipartite network since it always gives a zero value. Lind and Herrmann (2006) propose a variant counting the rectangular relations instead of triadic clustering, which can be applied to bipartite networks. The *density of cycles of size 4* surronding a node i computes the number of squares over the total number of possible squares including the node i and its pairs of neighbors, say m and n with degree k_m and k_n respectively:

$$cc2(i) = \frac{\sum_{m=1}^{k_i} \sum_{n=m+1}^{k_i} q_{imn}}{\sum_{m=1}^{k_i} \sum_{n=m+1}^{k_i} \left[(k_m - \eta_{imn})(k_n - \eta_{imn}) + q_{imn}\right]} \tag{2.7}$$

where q_{imn} is the number of common neighbors between m and n (not counting i) and $\eta_{imn} = 1 + q_{imn} + \theta_{mn}$ with $\theta_{mn} = 1$ if neighbors m and n are connected with each other and 0 otherwise.

2.3.5 Components of a network

An undirected network is *connected* if there is a path connecting every pair of nodes, i.e. every pair of nodes in the network is reachable. Otherwise, the network is *disconnected*. A disconnected network can be partioned into a set of *connected components* (CC). All nodes within each component connect to each other via undirected paths, while no paths between nodes in different components exist. Let g be a graph and *C(g)* the set of components of g:

$$g = \bigcup_{g \in C(g)} g \tag{2.8}$$

Analogously, in a directed network a component can be *strongly connected* (SCC) if each pair of nodes is connected via directed path or *weakly connected* (WCC) if only undirected paths are considered.

2.3.6 Similarity measure

For binary data, it is standard to use the *Jaccard index* as the similarity measure between two networks. It can be calculated as:

$$JI = \frac{S_{11}}{S_{01} + S_{10} + S_{11}} \tag{2.9}$$

where S_{11} refers to the number of links which are present both in the two networks. S_{01} denotes the number of links which are not present in the first network, but they are present in the second one and vice versa for S_{10}.

2.4 The topology of the bank-firm network

Given the restrictions of our data set to binary observations, the structure of the set of firm-bank relationships can be studied in terms of the statistical and topological properties of an undirected bipartite network. We start by considering the degree distributions of banks and firms.

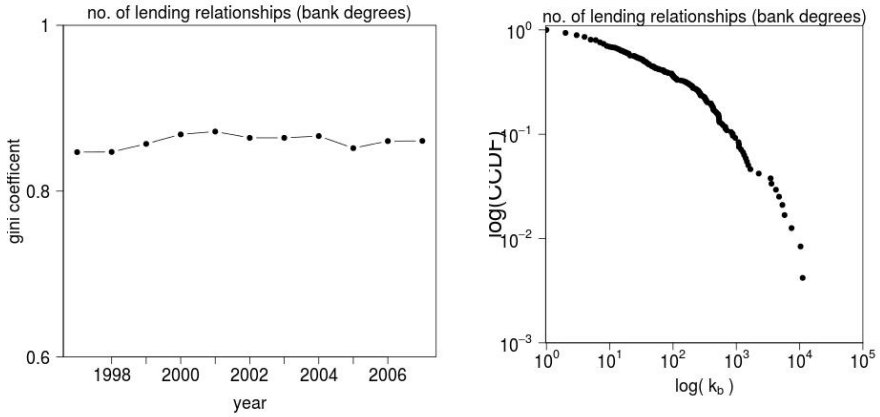


Figure 2.2: Gini coefficient of bank degree distribution between 1997 and 2007 for bank-firm network (left) and complementary cumulative distribution function (CCDF) of bank degrees k_b in logarithmic scale for bank-firm network in year 1997 (right).

degree distribution for bank-firm network		
statistics	bank	firm
median	34.50	2.00
mean	415.55	2.04
max	11,180	10.00
75th percentile	256.25	3.00

Table 2.1: Descriptive statistics of the distribution of bank degrees (left column) and firm degrees (right column) for the bank-firm network in the year 1997.

The complementary distribution function of the degree k_b and firm degree k_f for the year 1997 are plotted in the right panel of Figures 2.2 and 2.3 respectively.

The reason we will restrict ourselves to comment on the results concerning only the year 1997 is that they are comparable to those for the other years in the sample period (not reported), since the inequality of the distribution of bank degrees and firm degrees remains substantially unchanged over time, as indicated by the Gini coefficient in the left panel of Figures 2.2 and 2.3 respectively.

In 1997 our data set is quite similar in its order of magnitude to that of De Masi and Gallegati (2012) on the Italian bank-firm system: 46,938 versus 39,194 firms, 237 versus 508 banks. As

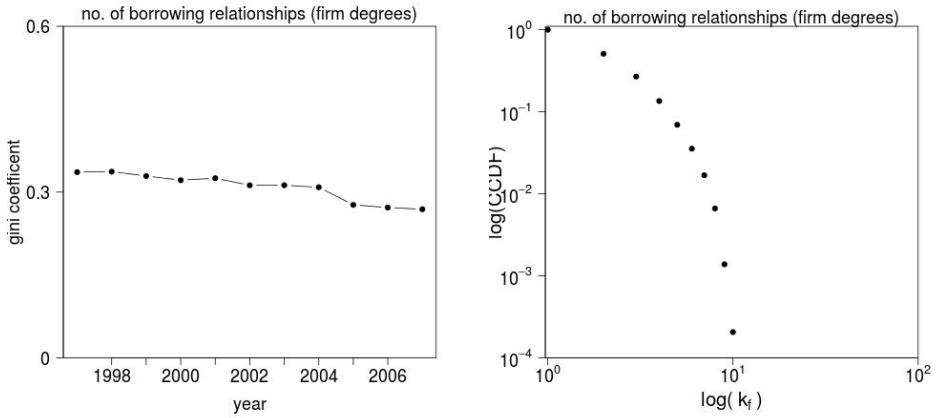


Figure 2.3: Gini coefficient of firm degree distribution between 1997 and 2007 for bank-firm network (left) and complementary cumulative distribution function (CCDF) of firm degrees k_f in logarithmic scale for bank-firm network in year 1997 (right).

shown by De Masi and Gallegati (2012), firms tend to have a single or a small number of lending relationships (the largest k_f is 10). For instance, Table 2.1 shows that 75% of firms have less than three lending relationships and the median number of relationships is 2. In contrast, we observe a much more heterogeneous behaviour among banks: the bank degree ranges over four orders of magnitude from 1 to 11,180. In addition, we see that the banks' degrees in the right panel of Figure 2.2 display a fat-tailed distribution. Most of the banks indeed finance a restricted number of firms, while some banks finance a substantially larger number of firms. These are the hubs of the network. The default of such a hub could spread to firms through network connection creating default cascades in both banking and real sector.

We then analyze the correlation between the firm degrees (number of credit relationships) versus the firm size. We observe that size is weakly correlated with the total degree of firms (correlation coefficient around 0.2). This indicates that there is only a weak relationship between the two variables. Indeed, single and multiple lending relationships are present for both small and large firms. In Figure 2.4 we consider the degree distributions for firms, dividing them into four equally populated groups on the basis of their balance sheet size.

According to the colours specified in the caption, we see an heterogeneous lending behaviour of firms related to the same group: both single ($k_f = 1$) and multiple ($k_f > 1$) bank relationships are present in all the four groups. Nevertheless, the tendency to be connected to multiple lenders is stronger for large firms, since the distribution of firm degree is shifted towards higher values of k_f as the size increases.

We now turn to the network metrics, introduced in the previous section, for bipartite networks

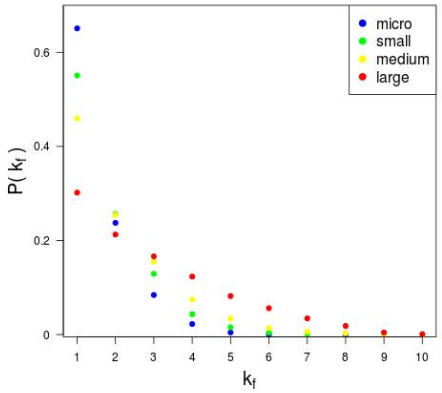


Figure 2.4: Frequency distribution of firm degrees k_f for the year 1997. We indicate firms in increasing order of the size with the colours: blue (less than 450), green (450-1,165), yellow (1,165-3,475) and red (over 3,475). The size is expressed in thousands.

to characterize the structure of the bank-firm lending market. Three measures of the network structure with direct economic interpretation are: the degree assortativity, the Jaccard index and the clustering coefficient. The degree assortativity, quantified by the Pearson correlation coefficient, captures the lending pattern and two bordeline cases are possible. The first is that highly (lowly) connected banks prefer to lend to highly (lowly) connected firms. The second is that highly (lowly) connected banks prefer to lend to lowly (highly) connected firms. The preference of the first lending pattern is captured by a positive Pearson correlation coefficient, while the second lending pattern is captured by a negative Pearson correlation coefficient. The Jaccard index of two networks for adjacent years measures the stability of the bank-firm relationships over time. The persistence of links implies that borrowers are not selected "randomly", but stable lending relationships between firms and their "house bank" do exist. Finally, the clustering coefficient, in its bipartite form proposed by Lind and Herrmann (2006), is the probability that a bank lends to two firms, which have a common bank and it thus measures the tendency of banks to finance a common subset of firms.

In order to be able to control for spurious effects, we compare the observed statistics of the bank-firm network with the properties expected from a null model. The chosen null model preserves the degree of each node in the real network (i.e. randomizing the network while preserving its row and column-wise sums) and it thereby enables us to test if the properties of the network are a mere consequence of the particular (bivariate) degree distribution or whether they convey additional information on the structural characteristics of the network. We generate the random

networks by adapting the switching algorithm by Milo et al. (2004) to bipartite networks. The switching algorithm uses a Markov chain and proceeds through a series of Montecarlo switching-steps to generate random networks starting from a one-mode network without changing its degree distribution. The same randomization algorithm can be easily adapted to a bipartite network g containing a set E of edges and two classes of nodes X and Y as follows:

1. randomly select two edges (a, b) and $(c, d) \in E$ with $a, c \in X$ and $b, d \in Y$
2. if $a \neq c$, $b \neq d$, $(a, d) \notin E$ and $(c, b) \notin E$ then:
   - add edges (a, d) and (c, b) to E
   - remove edges (a, b) and (c, d) from E

   otherwise do not modify the network
3. iterate steps 1 and 2 QN times where $Q = 100$ and N the number of edges. Note that switching steps which are not performed are still counted.

Milo et al. (2004) empirically show that the minimum number K of switching steps required is 100 times the number of edges ($K = 100N$) in order to minimize the initial bias of the Markov chain. Since our adjacency matrix is very large (i.e. for the year 2006 195 banks and 206,453 firms), the rewiring procedure is computationally expensive and we therefore generate only 100 random networks.[8] To assess the significance of empirical statistics, we calculate their one-sided p-value as the fraction of cases (i.e. the randomized networks with preserved degree sequences) that have a equal or higher value of the network metrics.

Figures 2.5, 2.6 and 2.7 summarize the results. As can be seen from the plots, we observe zero-correlation between bank and firm degrees that is virtually zero, high persistence of bank-firm relationships from one year to the next and relatively small clustering coefficient of about 0.1. However, all the results are significant at 1% statistic level, but only the results of the Jaccard index are distinctly different from the null model: the rewiring procedure completely destroys the stability of bank-firm relationship, which is observed empirically between consecutive years from 1997 to 2007. As shown in Figure 2.6, the Jaccard index at lag one between the original networks is 10 times higher than the Jaccard index at lag one between the rewired networks, indicating that the stability of bank-firm relationships from one year to the next is not imposed by the constraints of the degree distributions but results from active choices of counterparties. In particular, we see a high degree of stability of lending relationships with a Jaccard index around 0.7 between two successive years[9] and a progressive decline with the increase of the lag. Note that our random

[8] The random networks are generated by a greedy algorithm by using C++ Boost Graph Library.

[9] Note that the lending boom in 2005 clearly leads to a decline of the Jaccard index.

networks could not reproduce either the exact values of the degree correlation or the exact values of the clustering coefficient. It seems interesting that the real network displays systematically a slightly higher clustering tendency than the random networks. This is because more banks and firms are simultaneously active in the temporal vicinity rather than for years much earlier or later.

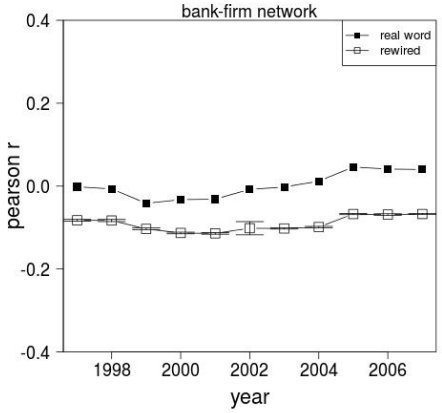


Figure 2.5: Degree correlation (Pearson correlation) of the bank-firm network and the average Pearson correlation and standard deviations (error bars) of 100 realizations of random bank-firm networks with preserved degree-sequences.

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
pearson r	-0.002*	-0.007*	-0.042*	-0.033*	-0.031*	-0.0079 *	-0.002*	0.012*	0.046*	0.041*	0.040*
simulation average	-0.082	-0.083	-0.103	-0.113	-0.114	-0.102	-0.102	-0.099	-0.067	-0.069	-0.067
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)

Table 2.2: Degree correlation (Pearson correlation) of the bank-firm network and average Pearson correlation of 100 realizations of random bank-firm networks with preserved degree-sequences with the associated standard errors in parenthesis. The asterix denotes statistical significance at 1% confidence level.

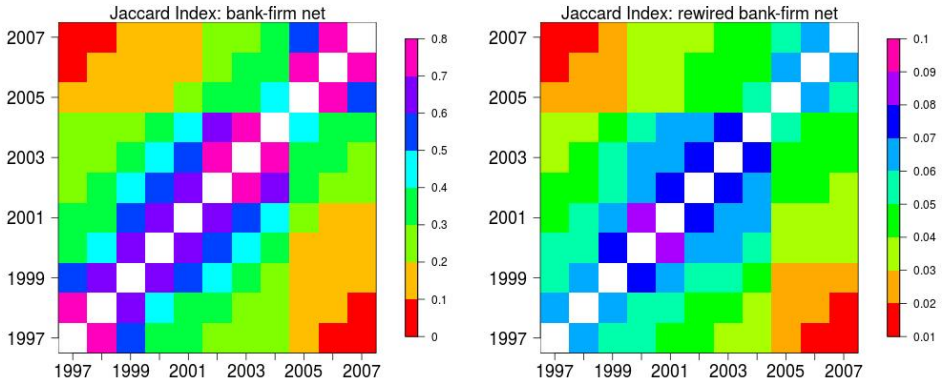


Figure 2.6: Contour plot of the Jaccard Index between all pairs of years in the sample period 1997-2007. Jaccard Index for the original bank-firm network (left) and the average Jaccard Index of 100 realizations of random bank-firm networks with preserved degree-sequences (right).

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
1997		0.742*	0.546*	0.398*	0.335*	0.295*	0.256*	0.215*	0.116*	0.091*	0.088*
1998			0.657*	0.475*	0.397*	0.336*	0.290*	0.244*	0.131*	0.112*	0.099*
1999				0.615*	0.514*	0.415*	0.356*	0.299*	0.154*	0.132*	0.118*
2000					0.693*	0.543*	0.461*	0.384*	0.188*	0.162*	0.144*
2001						0.624*	0.525*	0.435*	0.212*	0.181*	0.161*
2002							0.796*	0.638*	0.303*	0.259*	0.226*
2003								0.769*	0.357*	0.303*	0.263*
2004									0.429*	0.357*	0.306*
2005										0.732*	0.579*
2006											0.755*

(a) Real Jaccard Index

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
1997		0.061	0.058	0.053	0.049	0.041	0.038	0.034	0.020	0.019	0.017
		(0.0004)	(0.0005)	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
1998			0.063	0.059	0.054	0.045	0.041	0.037	0.023	0.021	0.019
			(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
1999				0.076	0.069	0.057	0.052	0.047	0.029	0.026	0.024
				(0.0005)	(0.0005)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
2000					0.084	0.069	0.062	0.057	0.035	0.032	0.030
					(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
2001						0.074	0.067	0.061	0.038	0.035	0.032
						(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0002)	(0.0001)
2002							0.076	0.069	0.044	0.041	0.038
							(0.0002)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
2003								0.075	0.048	0.045	0.041
								(0.0004)	(0.0003)	(0.0002)	(0.0002)
2004									0.053	0.049	0.045
									(0.0002)	(0.0002)	(0.0003)
2005										0.060	0.054
										(0.0002)	(0.0002)
2006											0.062
											(0.0002)

(b) Rewired Jaccard Index

Table 2.3: Jaccard Index between all pairs of years in the sample period 1997-2007 and the asterix denotes statistical significance at 1% confidence level (Panel A). The average Jaccard Index of 100 realizations of random bank-firm networks with preserved degree-sequences with the associated standard errors in parenthesis (Panel B).

2.5 The bank network

We now extract the bank-projection from the bipartite network of banks and firms. The bank-projection is the network of co-financing banks, which is only composed of one kind of node (banks) and a link between two banks exists whenever they are jointly exposed via loans to the same

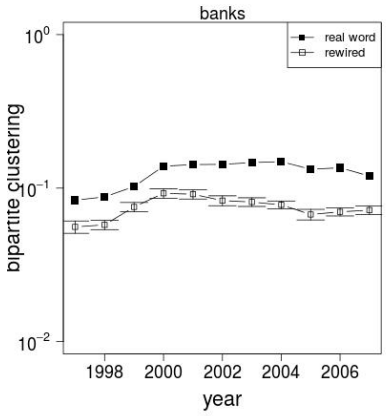


Figure 2.7: Bipartite clustering coefficient in logarithmic scale for banks and average bipartite clustering coefficient and standard deviations (error bars) of 100 realizations of random bank-firm networks with preserved degree-sequences.

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
clustering	0.083*	0.088*	0.10*	0.139*	0.142*	0.143*	0.147*	0.149*	0.133*	0.136*	0.119*
simulation average	0.056	0.058	0.075	0.092	0.091	0.083	0.081	0.078	0.067	0.070	0.072
	(0.005)	(0.004)	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)

Table 2.4: Bipartite clustering coefficient for banks and average bipartite clustering coefficient of 100 realizations of random bank-firm networks with preserved degree-sequences with the associated standard errors in parenthesis. The asterix denotes statistical significance at 1% confidence level.

firms. The bank network is undirected and unweighted and can be represented as a $(n \times n)$ square matrix $\mathbf{B}$, where n is the number of banks. An element b_{ij} of this matrix equal to 1 indicates that bank i and bank j have at least one common firm. The bank network is graphed in Figure 2.8. The Kamada-Kaway algorithm (Kamada and Kawai, 1989) has been employed, which positions more interconnected nodes spatially closer. At first glance, the bank network seems to have a core-periphery structure: a small set of core nodes highly connected and a periphery, where the nodes are not connected to each other but only to the core. Moreover, we see a noticeable color patterns from the core to the periphery. The core is dominated by the commercial banks (yellow nodes), which are highly interconnected (closer placement on the layout) and linked to the peripheral nodes. The rings of nodes standing just a notch outside the core are mainly the savings banks (green nodes), while the outer peripherical nodes are principally the credit cooperatives (red nodes). Moreover, the banks in the periphery are connected to the other banks in the core, but not to each other, suggesting a strongly connected network on the one hand and a very sparse adjacency matrix on the other hand. Indeed, the bank projection turns out to be: (i) a fully connected network in each

period, i.e. there are no isolated components and every bank is reachable from every other bank by 'travelling' along the edges[10]; (ii) a sparse network with an average density equal to 0.168 ± 0.012 (mean $\pm$ standard deviation over the sample period).

Next, we report the distribution of the number of the number of co-financing banks at annual time scale. Again, we will focus on the year 1997 because the Gini coefficient in the left panel of Figure 2.9 shows that dispersion of degrees remains almost constant over the sample period. In the right panel of Figure 2.9 we observe a fat-tailed degree distribution: a small number of banks finances at least one firm in common with a large number of banks.

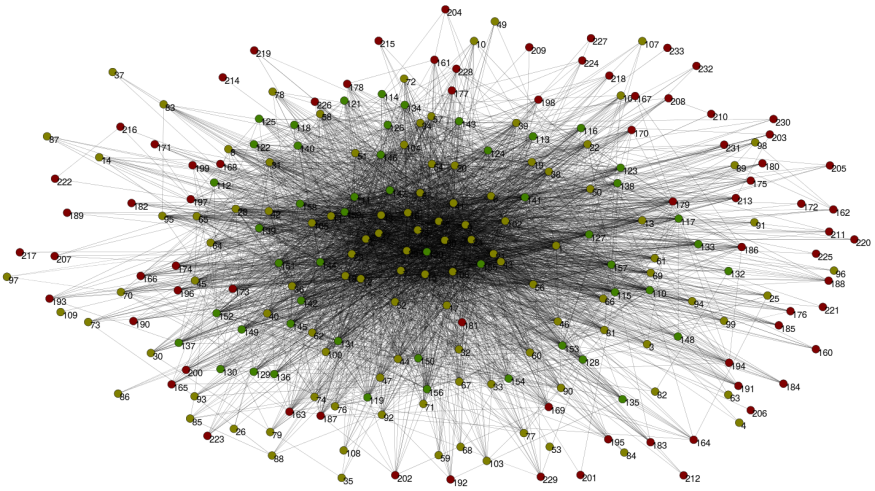


Figure 2.8: The bank network in the year 1997. 233 nodes and 4,312 undirected links. Nodes are colored according to the type of bank: yellow commercial banks (109), green savings banks (50) and red credit cooperatives (74). The network is represented using Kamada-Kaway algorithm.

Indeed, a very small number of banks has at least one firm in common with 83% of the total banking system (192 banks), while 75% of banks have firms in common with less than 50 banks (see Table 2.5). This discontinuity in degree is a further indication for a distinction between core and periphery in the bank network.

[10]Note that the same result holds for the firm projection, that forms a fully connected network in each year of the sample except for some disconnected pairs of firms.

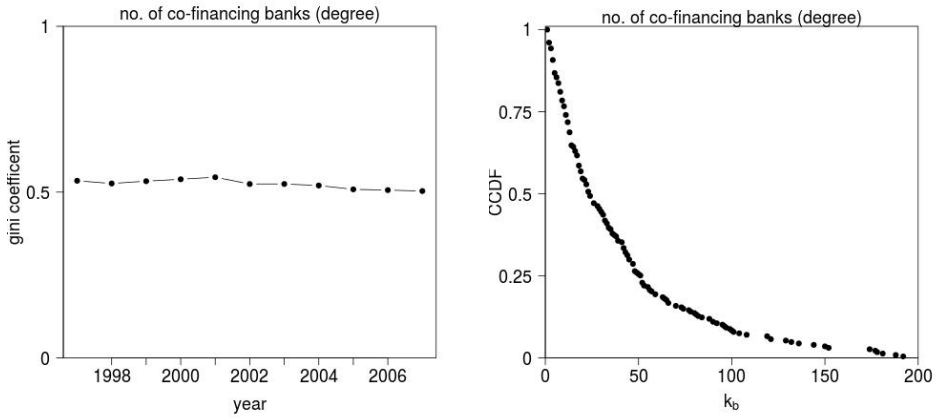


Figure 2.9: Gini coefficient of degree distribution between 1997 and 2007 for bank network (left) and complementary cumulative distribution function (CCDF) of degrees k_b for bank network in year 1997 (right).

bank network		
statistics	degrees	weights
median	23.00	152
mean	37.81	614.28
max	192.00	9,395
75th percentile	50.50	428

Table 2.5: Descriptive statistics of the degree (left column) and the weight (right column) distribution in the year 1997 for the bank network. The degrees are the number of co-financing banks and the weights are the number of common firms.

We now turn to the degree correlation. This is measured here by the correlation between the number of banks with which a bank shares the same firms versus the number of banks with which its co-financing banks share the same firms. We describe the degree correlation of the bank network in terms of the assortativity coefficient (Newman (2003)). Figure 2.10 reports the assortativity of the bank projection together with the average assortativity calculated on 100 simulations of the bank projection preserving the bank and firm degrees of the original bipartite network. As we can see, both the bank network and the simulated networks display disassortativity around values of -0.4 with negligible variance over the sample period. This implies that the tendency of high-degree banks (banks financing firms in common with several other banks) of sharing firms with low-degree banks (banks financing firms in common with a few banks) is a consequence of the degree distributions of the bank-firm network and it is not an authentic property of the bank projection. Finally, the fact that the bank network displays disassortativity is an additional indication of a core-periphery structure because it reflects the tendency of low-degree banks (the periphery) to be attached to high-degree banks (the core).

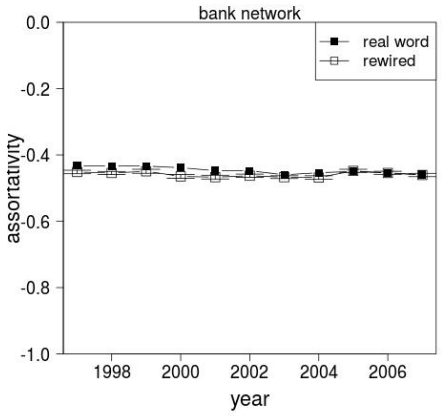


Figure 2.10: Assortativity of the bank projection and the average assortativity and standard deviations (error bars) of 100 simulations of the bank projection preserving the bank and firm degree-sequences of the original bipartite network.

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
assortativity	-0.433*	-0.433*	-0.433*	-0.438*	-0.447*	-0.448 *	-0.459*	-0.453*	-0.449*	-0.454*	-0.459*
simulation average	-0.451	-0.454	-0.449	0.465	-0.467	-0.463	-0.467	-0.468	-0.448	-0.454	-0.461
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table 2.6: Assortativity of the bank projection and average assortativity of 100 simulations of the bank projection preserving the bank and firm degree-sequences of the original bipartite network. Standard errors indicated in parentheses. The asterix denotes statistical significance at 1% confidence level.

2.6 Core-Periphery structure in the bank network

In the previous section we have argued that the disassortavity and the heterogeneity in degrees provide a first indication that the bank network displays a core-periphery (C-P) structure. Now we use the following framework and we estimate the structure of a C-P model and its statistical significance against standard random networks.

Core-periphery network models were introduced by Borgatti and Everett (2000). Craig von Peter (2014), Fricke and Lux (2015) and van Lelyveld and in't Veld (2012) have recently applied the C-P models to interbank networks. The basic idea of a C-P model is that a network can be divided into a core and a periphery: the core consists of a small set of N_c nodes highly connected to each other and the periphery is made of N_p nodes which are connected to the core without reciprocal connections.

In blockmodelling terms, an "ideal" C-P structure $\mathbf{M}$ contains a core/core region as a 1-block and a periphery/periphery region as a 0-block in the upper left-hand side and lower right-hand side

of the adjacency matrix, respectively. The 1-block denotes a $(N_c \times N_c)$ submatrix of ones and the 0-block denotes a $(N_p \times N_p)$ submatrix of zeros.

$$M = \begin{pmatrix} CC & CP \\ PC & PP \end{pmatrix} = \begin{pmatrix} 1 & CP \\ PC & 0 \end{pmatrix}$$

The CC-block is a fully connected subnetwork where each core-node is connected to all core-nodes and the PP-block is an empty subnetwork where there are no links between the periphery-nodes. The relations within the core and the periphery can be specified or alternatively treated as missing data without any constraints on the off-diagonals blocks CP and PC, as it is recommended by Borgatti and Everett (2000).

We use this approach and we use a C-P model without any constraints on the density of the CP and PC block. The objective is therefore to identify the optimal core-periphery bipartition C^*, which maximizes the density in the core-block and minimizes the density in the periphery- block, without regard for the density in the off-diagonal blocks. In this way, a data-driven identification of the 'core' banks is achieved without prespecifying their exact number. We employ the algorithm by Lipp (2013) to fit the C-P model to the bank network, where the subdivision of nodes into a core and a periphery depends exclusively on the degree distribution of the network, moving systematically nodes with a high degree into the core. The measure of distance between the bank network **B** and the 'ideal' C-P model **M** of the same dimension is the total error score. The total error score aggregates the inconsistencies between the bank network and the ideal C-P structure. A missing link in the CC-block represents an inconsistency with respect to the model, as a core bank would not be connected with another core bank. Similarly, the existence of a link in the PP-block represents an inconsistency, as a periphery bank would be connected with another periphery bank.

The aggregate inconsistencies across the CC and PP-block read:

$$E = \begin{pmatrix} E_{CC} & E_{CP} \\ E_{PC} & E_{PP} \end{pmatrix} = \begin{pmatrix} [c]N_c(N_c - 1) - \sum_{i,j \in C} a_{ij} & 0 \\ 0 & \sum_{i,j \notin C} a_{ij} \end{pmatrix}$$

The total error score $e(C)$ aggregates the inconsistencies across the main diagonal blocks CC and PP and we normalize it to the number of links N:

$$e(C) = \frac{E_{CC} + E_{PP}}{N} \tag{2.10}$$

The optimal core C^* is therefore the set of banks that produces the smallest error score:

$$C^* = \text{argmin } e(C) = \{C \in \Gamma \mid e(C) \leq e(c) \forall c \in \Gamma\} \tag{2.11}$$

where Γ denotes all strict and non-empty subset of population $\{1, 2, \ldots, n\}$.

Next, we report the results of fitting the discrete C-P model with arbitrary structure of the CP and PC blocks to the bank network. For the reference year 1997, we find an optimal core C^* of 37 banks, which corresponds to 15.54% of banks active in the lending market. The error score is 19.2% of network links, where the majority of errors occur because there are links within the periphery[11]. Most importantly, the core we estimated is highly persistent. In Figure 2.11 we show that the size of the core and the associated error score are stable over time with a core size of 32.8 ± 2.2 banks and an error score of $19.8 \pm 0.6\%$.

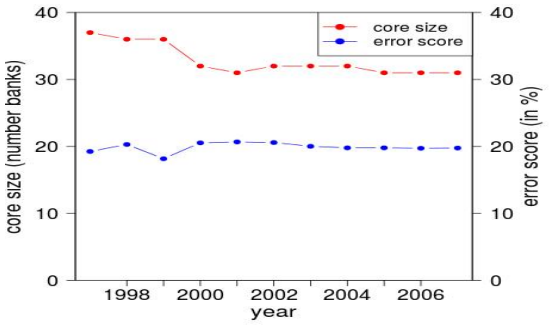


Figure 2.11: The size of the estimated core (left axis) and the associated error score expressed as a percentage of links (right axis) for the bank network.

Moreover, the composition of banks within the core is stable over the sample. On average 94.4% of the core banks remains in the core the next year. Likewise, the banks in the periphery also tend to remain in the periphery over time, as shown by the values close to unity on the main diagonal of the transition matrix in Table 2.7. In Figure 2.12 we also show the transition probabilities between core and periphery for each single year, which provides further evidence that the C-P structure is highly stable over time.[12]

[11] The total number of errors is 830, comprising 43 errors for the CC-block (missing links within the core) and 787 for the PP-block (links within the periphery).

[12] Table 2.7 and Figure 2.12 contain 'entry/exit' as an additional category since we have to take into account the changing composition of the population.

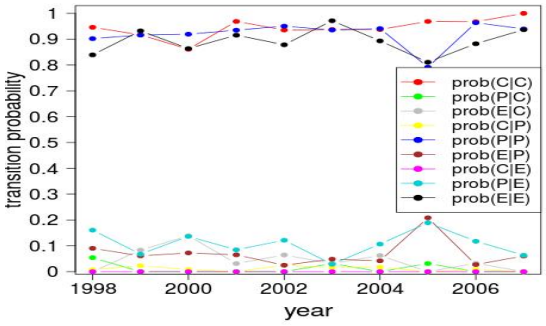


Figure 2.12: Transition probabilities for each year of the sample. For instance, prob(C|P) in yellow indicates the probability that a bank moves from the periphery to the core the next year.

	C_t	P_t	E_t
C_{t-1}	0.944	0.012	0.044
P_{t-1}	0.010	0.919	0.070
E_{t-1}	0.000	0.108	0.892

Table 2.7: Transition matrix. The first row indicates the probability of a core bank in $(t-1)$ (C_{t-1}) of remaining in the core in t (C_t), moving to the periphery (P_t) or exiting from the market (E_t). The column E takes account of the exit from the sample of banks having co-financing relationships and the reason could be (i) exit from the lending market (ii) M&A or failures (iii) not sharing common lending relationships anymore. These probabilities are aggregate values over the sample, for instance $P(C_t \mid C_{t-1})$ reads: on average 94.4% of the core banks remain in the core the next year.

As a last step, we test the significance of the core we estimated for the reference year 1997 against standard network structures as explanatory models. The first class of networks is the Erdös-Rényi (ER) random graphs. An ER random graph is obtained by connecting any two nodes with a fixed and independent probability p. Any realization of such a network has an expected density p. The degree distribution follows a Binomial, converging to Poisson distribution for large networks. The second class of standard networks is the scale-free (SF) networks whose degree distribution follows a power law. The most popular mechanism to generate SF networks is preferential attachement introduced by Barabasi and Albert (1999), where new nodes preferentially attach to existing nodes with probability proportional to the degrees of the target node.

We therefore generate 1,000 ER random graphs and SF networks with the same dimension and density as the bank network ($n = 233$, $d = 0.15$). For the SF networks we have used the algorithm by Goh et al. (2001) with scaling parameter $\alpha = 2.3$. We then fit the C-P model to each realization and we compute the error score distribution for both random networks. Figure 2.13 compares the

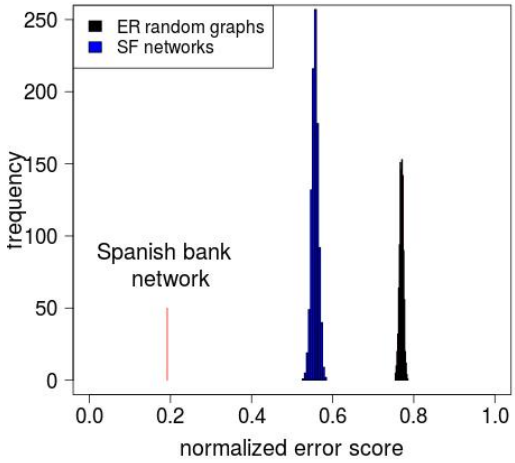


Figure 2.13: Error score from fitting the C-P model to the bank network for the reference year 1997 ($e(C^*) = 0.192$ shown as a red arrow) versus the empirical distribution function of error score for 1,000 Erdös-Rényi (ER) random graphs (black) and scale-free (SF) networks (blue) with the same dimension and density as the bank network. Error scores have been normalized to the number of links.

error score of the bank network against the bootstrap distributions obtained for the two random networks. The ER random graphs exhibit error scores concentrated around 0.77, while the SF networks scores around 0.56. As expected, the error score of SF networks is closer because SF networks are generated by preferential attachement mechanism producing a few highly connected hubs, which leads to network structures similar to C-P structures. Overall, the error score of the empirical bank network $e(C^*) = 0.192$ is lower than any possible percentile of the error distribution for both random networks and, hence, the identified C-P structure could not have been obtained spuriously from a random structure of links following one of these two generating mechanisms.

It is interesting to underline that the fit of the C-P model to the bank projection (in terms of the error score) seems to be even better than the fit of the C-P model to bank credit networks (e.g. Fricke and Lux, 2015). Overall, this result indicates that the C-P model might be an useful model for different layers of the complex universe of connections of the banking sector. Note that the closeness to a C-P structure, however, does not imply that the core would be the same for different network layers. Indeed, Langfield et al. (2013) demonstrate that different banks turn out to constitute the core of the UK banking system for different dimensions of network connections.

Finally, we also test the C-P structure against random bank networks preserving bank and firm degrees of the original bipartite network. The average error score over 100 rewired networks

is 0.056 ± 0.003, which represents an almost perfect fit. However, the average size of the core is 74.72 ± 0.87 banks (approximately the double of the core of the real network) and we, therefore, conclude that the fit of the bank projection to a C-P structure is not attributed to the degree distributions of the bank-firm network.

2.7 Quantifying preferential bank relationships

While we have used the binary projection for the estimation of the C-P model, we now consider the weighted projection of the bank network. We, therefore, define a $(n \times n)$ square matrix $\mathbf{W}$, where the entry w_{ij} indicates the number of firms that bank i and bank j finance in common.

The distribution of the number of common firms for the year 1997 is plotted in the right panel of Figure 2.14. Again we have restricted our analysis to just one year because the Gini coefficient in the left panel of Figure 2.14 shows that the dispersion of the weights (number of common firms) is almost constant over the sample period. We see that the distribution of the weights is fat-tailed: a small percentage of banks shares around 9,000 firms in common with other banks. In contrast, the median and the 75th percentile values are considerably smaller (152 and 428 respectively), pointing towards a pronounced heterogeneity in the distribution of the number of common firms.

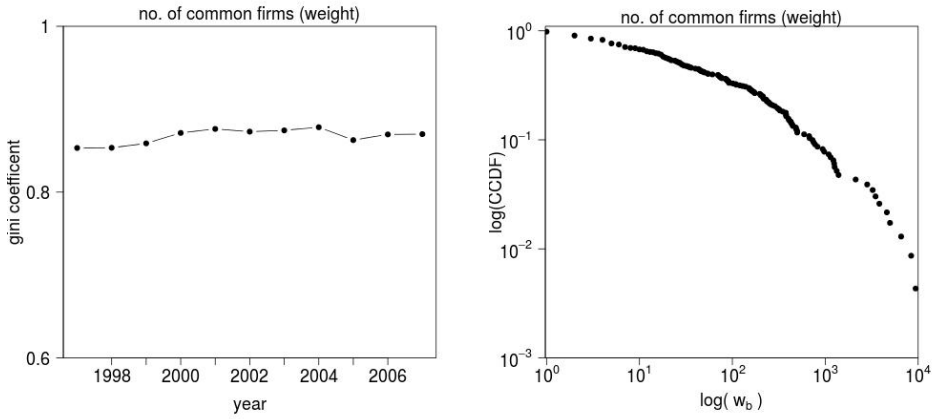


Figure 2.14: Gini coefficient of weight distribution between 1997 and 2007 for bank network (left) and complementary cumulative distribution function (CCDF) of weights (w_b) in logarithmic scale for bank network in the year 1997 (right). The weights are the number of common firms.

The fat-tailed distribution indicates clearly that a small set of banks shares a very large number of firms with other banks and this is a hint at the existence of preferential bank relationships, i.e.

pairs of banks preferentially sharing the same firms.

However, the distribution of the number of common firms could merely reflect the strong heterogeneity in the number of lending relationships of banks (see Figure 2.2). Again, the question is to isolate the effect of the heterogeneity in the degrees of banks in order to identify structural features that are not an automatic consequence of the construction of a network composed of two populations (banks and firms) with very heterogenous members. In order to asses the significance of the observed co-financing bank relationships, we thus compare the actual bank network against a random null hypothesis, which takes into account the heterogeneity of banks in terms of degrees in the bank-firm network. For each pair of banks, say i and j, we test the observed number of common firms η^{ij} against an hypergeometric null hypothesis by using the statistical test by Tumminello et al. (2011).

Under the null hypothesis the probability of observing η^{ij} common firms between bank i and j, given their number of lending relationships η^i and η^j respectively, reads:

$$H(\eta^{ij}|N,\eta^i,\eta^j) = \frac{\binom{\eta^i}{\eta^{ij}}\binom{N-\eta^i}{\eta^j-\eta^{ij}}}{\binom{N}{\eta^j}} \tag{2.12}$$

The basic idea is that $H(\eta^{ij}|N,\eta^i,\eta^j)$ provides the probability that η^{ij} are common firms with bank i in a random sample of η^j firms drawn from the N firms present in the lending market, where η^i firms belonging to bank i are present.

All in all, the hypergeometric distribution is an appropriate random benchmark because first it takes into account the heterogeneity of banks in terms of the number of lending relationships. Second, the hypergeometric distribution gives the same probability of bank j having η^{ij} common firms with bank i and bank i having η^{ij} common firms with bank j given their number of lending relationships, since we get the same probability exchanging η^i and η^j:

$$H(\eta^{ij}|N,\eta^i,\eta^j) = H(\eta^{ij}|N,\eta^j,\eta^i) \tag{2.13}$$

We therefore associate the right tail p-value with the observed number η^{ij} of common firms between bank i and bank j as:

$$p(\eta^{ij}) = \sum_{X=\eta^{ij}}^{Min[\eta^i,\eta^j]} H(X|N,\eta^i,\eta^j) \tag{2.14}$$

We calculate the right tail p-value because we are interested in testing whether the co-financing relationship between bank i and bank j is over-represented in the data respect to the null hypothesis of independence of the financing decisions of each pair of banks. Given a significance level s, the co-financing relationship between bank i and j is statististically significant if $p(\eta^{ij}) < s$. In this case, we say that bank i and bank j preferentially share the same firms. By calculating the p-value for all the E edges in the bank network, we therefore run E statistical tests and we use the False Discovery Rate (FDR) as a correction for multiple tests at a statistical significance level of 0.05. We thus refer to the network of the significant bank relationships as the FDR network.

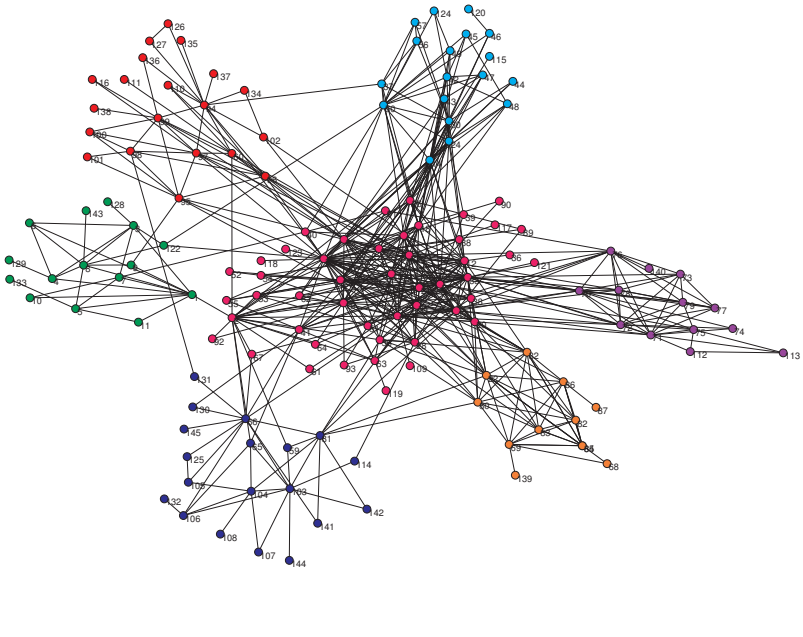


Figure 2.15: The FDR network for the year 1997. The colours of the nodes indicate the node membership to the 7 largest partitions detected by InfoMap (unweighted option).

In Figure 2.15 we show the FDR network for the year 1997. The nodes in the network are coloured according to the 7 largest partitions detected by using the InfoMap method by Rosvall and Bergstrom (2008). We see that reducing the weighted bank projection into a simpler binary network composed only of the significant co-financing bank relationships reveals a clustered structure. Figure

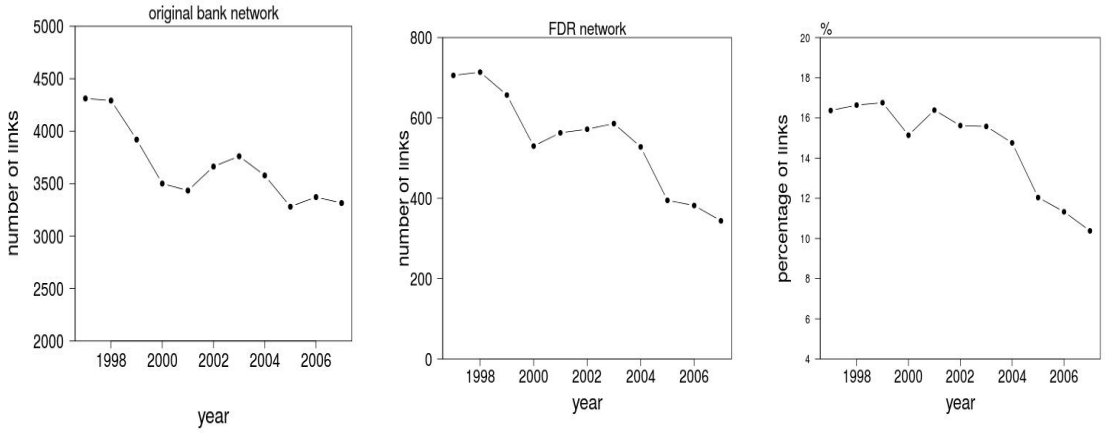


Figure 2.16: The number of links in the original bank network (left) and in the FDR network (middle) from 1997 to 2007. The ratio between the number of links in the FDR network and in the original bank network expressed as a percentage from 1997 to 2007 (right).

2.15 also shows that a core of the clusters emerges from the FDR network. We therefore compare this core with the core of the binary projection estimated in Section 2.6. We find that half of the banks in the core of the bank projection is also present in the core of the clusters, meaning that this small sample of banks (around 15 banks) identifies the authentic lobby group of the credit market, that acts as "influencer" of the lending decisions. The middle panel of Figure 2.16 reports the evolution over time of the significant bank relationships. The FDR network retains only $14.6 \pm 2.2\%$ of the links of the original bank network for the considered time period (right panel of Figure 2.16), suggesting that the heterogeneity of banks in the number of lending relationships plays a crucial role in determining the structure of the co-financing bank relationships. Moreover, the original network and the FDR network show a similar trend with an approximately constant decline of the number of links. As expected, when considering the ratio between the number of links of the FDR network and the original network we find a drop in ratio of significant links after the Spanish lending boom.

Note, however, that the structure of Figure 2.16 emerges only after discarding the majority of links which are not "conspicuous" or are not in full agreement with a random draw of the bank-firm links given the existing degree distributions.

2.8 Conclusions

The structure of the credit relationships in Spain has been scrutinized, applying a network approach to a yearly series of bank-firm data from December, 1997 to December, 2007.

The Spanish credit market has the following main topological features. First, banks have more lending relationships and a more heterogeneous distribution of the lending relationships than firms. Second, the number of credit relationships is size dependent (non monotonically) for firms. Third, bank-firm relationships display a high degree of persistence from one year to the next.

A stable and significant core-periphery structure against standard random networks and rewired networks with prescribed degree sequences has been found in the network of co-financing banks. Finally, around 15% of the co-financing bank relationships appear unusually strong under the null model of a hypergeometric distribution of the number of common firms preserving the degrees in the bank-firm network. By taking the hypergeometric distribution as null model, the members of the core acting as the lobby banks of the lending market are also validated. This findings provide evidence that the analysis of the bank projection requires special care because the weights of their edges incorporate information on the bipartite data instead of displaying truly structural features of the projection.

Further related research work may come in several forms. First, it is imperative to test the robustness of the results for the estimation of the core-periphery structure in the bank network. Second, it is advisable to employ a more complex null model to test the significance of the properties of the bank-firm network instead of a null hypothesis obtained by rewiring the original network, such as the maximum entropy model for bipartite networks with prescribed degree sequences proposed by Saracco et al. (2015). Third, it may be interesting to investigate the network of the co-financing bank relationships which are not explained by the heterogeneity of bank degrees in much more details.

Bibliography

Bech, M.L. and Atalay, E., 2010. The topology of the federal funds market. *Phys. A: Stat. Mech. Appl.* 389(22), 5223-5246

Barabasi, A.L. and Albert, R., 1999. Emergence of scaling in random networks. *Science* 286(5439), 509–512

Borgatti, S.P. and Everett, M.G., 2000. Models of core/periphery structures. *Social Networks* 21(4), 375-395

Boss, M., Elsinger H., Summer M. and Thurner, S., 2004. Network topology of the interbank market. *Quantitative Finance* 4, 677-684

Caldarelli, G., 2007. *Scale-free networks*. Oxford University Press, Oxford

Craig, B.R and von Peter G., 2014. Interbank tiering and money center banks. *Journal of Financial Intermediation*, in press

De Castro Miranda, C. and Tabak, B.M., 2013. Contagion risk within firm-bank bivariate networks. Working Paper Series 322, Central Bank of Brazil

Dell'Ariccia, G. and Marquez, R., 2006. Lending booms and lending standards. *Journal of Finance* 61, 2511-2546

De Masi, G., Fujiwara, Y., Gallegati, M., Greenwald, B. and Stiglitz JE., 2011. An analysis of the Japanese credit network. *Evolutionary and Institutional Economics Review* 2 (7), 209-232

De Masi, G. and Gallegati, M., 2012. Bank-firm topology in Italy. *Empirical Economics* 43, 851-866

Dormann C.F., Fründ J., Blüthgen N.and Gruber B., (2009). Indices, Graphs and Null Models: Analyzing Bipartite Ecological Networks. *The Open Ecology Journal* 2, 7-24

Etxezarreta, M., Navarro, F., Ribera, R. and Soldevila, V., 2011. Boom and (deep) crisis in the Spanish economy: the role of the EU in its evolution. Communication for 17th Workshop on Alternative Economic policy in Europe, Vienna, September 2011

Fricke, D. and Lux, T., 2015. Core-periphery structure in the overnight money market: Evidence from the e-MID trading platform. *Computational Economics* 45, 339-395

Goh, K.I., Kahng, B. and Kim D., 2001. Universal behavior of load distribution in scale-free networks. *Physical Review Letters* 87, 278701

Illueca, M., Norden, L. and Udell, G., 2009. Liberalization, corporate governance and saving banks. Mo.Fi.R. Working Papers 17, Money and Finance Research group (Mo.Fi.R.) - Univ. Politecnica Marche - Dept. Economic and Social Sciences

Iori, G., De Masi, G., Precup, G., Gabbi, G. and Caldarelli G., 2008. A network analysis of the Italian overnight money market. *Journal of Economic Dynamics and Control* 32, 259-278

Lip, S., 2011. A fast algorithm for the discrete core-periphery bipartitioning problem. Working Paper, University of Cambridge

Kamada, T. and Kawai, S., 1989 An algorithm for drawing general undirected graphs. *Information processing letters* 31, 7-15

Langfield, S., Liu, Z. and Oja, T., 2013, Mapping the UK banking system. Bank of England, Working Paper

Lux, T., 2016. A model of the topology of the bank–firm credit network and its role as channel of contagion. *Journal of Economic Dynamics and Control* 66, 36-53

Lind PG., González MC. and Herrmann HJ., 2005. Cycles and clustering in bipartite networks. *Physical review E* 72, 56127

Marotta, L., Micciche, S., Fujiwara, Y., Iyetomi, H., Aoyama, H., Gallegati, M. and R. Mantegna, 2015. Bank-firm credit network in Japan: An analysis of a bipartite network. *PLoS ONE* 10 (5)

Martinez-Jaramillo, S., Alexandrova-Kabadjova B., Bravo-Benitez B. and Solórzano-Margain, J.P, 2014. 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

Martin-Oliver, A., 2013. Financial integration and structural changes in Spanish banks during the pre-crisis period. *Banco de España, Estabilidad Financiera*, no. 24

Montoriol-Garriga, J., 2008. Bank mergers and lending relationships. Working Paper Series 0934, European Central Bank

Milo, R., Kashtan, N., Itzkovitz, S., Newman, M.E.J. and Alon U., 2004. On the uniform generation of random graphs with prescribed degree sequences. arXiv:cond-mat/0312028v2

Newman, M.E.J., 2003. Mixing patterns in networks. *Physical Review E* 67, 026126

Salas, V. and Saurina, J., 2002. Credit Risk in two instutional regimes: Spanish commercial and savings banks. *Journal of Financial Services Research* 22, 203-224

Rosvall, M. and Bergstrom, C. T., 2008. Maps of random walks on complex networks reveal community structure. *Proc. Natl. Acad. Sci. USA* 105, 1118-1123

Saracco, F., Di Clemente R., Gabrielli A. and Squartini T., 2015. Randomizing bipartite networks: the case of the World Trade Web. *Scientific Reports* 5(10595)

Soromäki, K., Bech, M.L., Arnold J., Glass R.J, Beyeler W.E., 2007. The topology of interbank payment flows. *Phys. A: Stat. Mech. Appl.* 379(1), 317-333

Squartini T. and Garlaschelli D., 2011. Analytical maximum-likelihood method to detect patterns in real networks. *New. J. Phys.* 13, 083001

Tumminello, M., Lillo, F., Piilo J. and Mantegna R.N., 2011. Statistically validated networks in bipartite complex systems. *PLoS ONE* 6, e17994

van Lelyveld, I. and in't Veld D., 2012. Finding the core: network structure in interbank markets. Working Paper Series 348, Central Bank of Netherlands

Watts, D.J. and Strogatz S.H., 1998. Collective dynamics of small-world networks *Nature* 393, 440-442

Chapter 3

The controllability of the repo network in Mexico

This work was carried out during my internship at Banco de México, Directorate of Payment Systems.[1]

[1]I am thankful to Javier Pérez Estrada, Serafin Martínez Jamarillo and Stefano Battiston for their comments and suggestions. The views and conclusions presented in this article are exclusively the responsibility of the author and do not necessarily reflect those of Banco de México.

3.1 Introduction and related literature

Control theory is a mathematical branch of engineering with applications to electric circuits, manufacturing processes, communication systems, aircraft and robots. The main objective of control theory is to control a system. The controllability of a system can be defined as the ability of an external controller to move the state of a system from any initial state to any other final state throughout the appropriate manipulation of a few input variables in a finite time interval. A toy example of controllability could be the control of the speed of a car, where the driver is the external controller who can change the initial speed to the target speed by manipulating the pedals.

Liu et al. (2011) have recently applied the concept of controllability to networks. The idea is that the whole network can be controlled by acting on a specific set of nodes. Put differently, the state of a system of N nodes can be steered from an initial state to a desired final state if an external signal is applied to a selected minum number of nodes (the so-called *drivers*). This method can be applied to the interbank market in order to study its role for monetary policy transmission. Imagine that a central bank wants to increase the amount of money in the banking system and uses open market operations (OMOs) as monetary policy instrument. Central banks aim at manipulating the amount of money in the banking system (or equivalently, the interbank interest rate) by injecting liquidity to the minimum number of banks to meet their monetary goal. Since the effectiveness of the OMOs crucially depends on the key players of the interbank market, the concept of driver can be readily employed to identify the main conduits of the monetary policy.

In this paper we apply such a methodology to the Mexican repo network. By using a detailed dataset available at the Mexico's Central Bank containing all the repo transactions between any two Mexican banks from January 2005 to December 2014, we explore the central bank's control over the interbank market from three different perspectives. First, the time horizon because a central bank *de facto* changes its strategy with respect to short-run and long-run policy goals. Second, the stability of the set of driver banks, which is a critical issue from a central bank's perspective. If a group of banks constantly play a pivotal role in providing credit, a central bank needs to monitor them carefully to ensure that the monetary policy is effective. Third, the size of the driver banks in order to test whether the too big to fail (TBTF) policy is adequate to predict the systemically important banks (SIBs) or whether complexity and interconnectedness go beyond the bank size.

Our main finding is the identification of a highly persistent set of drivers, which does not correspond with the most connected banks or the largest lenders of the Mexican repo market. These results shed light on the existence of a group of banks contributing constantly to the monetary transmission, which would have *not* detected by looking at TBTF banks. From a regulator point of view, network controllability appears to be a feasible method for identifying the SIBs. We systematically find a fraction of drivers below the 60%, which is a reasonable number of banks receiving liquidity during the OMOs. Moreover, we find that the fraction of drivers decreases with the time

horizon. This is in line with the empirical evidence of a lower volatility of the interbank funds over longer time horizons, which clearly leads to a lower number of banks needed to achieve full control over the interbank market.

Over the last few years, the research on interbank networks has been growing fast. Network theory has been used to analyze the connections between banks and to explain how they impact on financialy stability and on systemic risk. A well-established set of stylized facts has been found: disassortativity, power law degree distribution and core-periphery structure. Some of the most relevant studies on the structure of interbank networks are: Boss et al. (2004) for the Austrian interbank market, van Lelyveld and in't Veld (2014) for the Dutch interbank market, Craig and von Peter (2014) for the German interbank market, De Masi et al. (2006), Fricke and Lux (2015a), Fricke and Lux (2015b) and Iori et al. (2008) for the the Italian e-MID, Martinez-Jaramillo et al. (2014) and Molina-Borboa et al. (2015) for the Mexican interbank market and Langfield et al. (2014) for the UK interbank market.

Besides the empirical literature, theoretical models have been also developed to study the link between the structure of interbank networks and contagion. Since the pioneering work by Allen and Gale (2000), research has made considerable efforts to understand which topological properties make a financial system more resilient or fragile. Nowadays, the literature widely agrees on the *robust-yet-fragile* nature of financial systems and in particular on the role of connectivity in amplifying its stability or instability under certain connectivity thresholds (Nier et al. (2008), Gai et al. (2011) and Haldane and May (2011), among others).

However, despite the large literature on the impact of the network structure on systemic risk, researchers have only recently examined how the network structure affects the control of the interbank lending. So far, only three papers have applied control theory to an interbank network. Delpini et al. (2013) use the e-MID data to explore the controllability of the Italian interbank market across different time horizons. They characterize the drivers in terms of degrees (the number of lending relationships) and out-strength (the total amount lent). They estimate a fraction of drivers below the 70 % at any time scale and a power law decay of the fraction of drivers as function of the time horizon. Moreover, the set of drivers they find are neither the high degree nodes of the network nor the largest lenders. Galbiati et al. (2013) analyze the controllability properties of the TARGET2 (T2), the payment system for the Eurozone. Again, the drivers do not correspond with the standard centrality measures. The T2 drivers are not the most central nodes (according to centrality measures based on the centrality of the neighbors, the so-called feedback centralities) or the banks processing the largest value of payments. Finally, Galbiati and Stanciu-Vizeteu (2015) characterize the set of drivers for a simulated T2 network. They simulate a payment network with very simple rules, where the willingness of a bank to make a payment depends on its own stock of liquidity and on the bilateral netting against the counterparty. Like for the real network, they find that the drivers are not the most central nodes or the largest banks.

Although the Mexican repo network displays a completely different structure with respect to the e-MID network, namely homogeneous the first and scale-free the second (De Masi et al. (2006)), our results are similiar to those by Delpini et al. (2013). This may suggest that these properties of the drivers are constant across the markets regardless of the structure of the network, providing new stylized facts for interbank markets.

Finally, there is also another paper including the central bank in an interbank network, that is the one by Léon et al. (2014). They use data from the Colombian large-value payment system and they merge the interbank loans and the central bank's repos into a single network. They identify the drivers (or the central bank liquidity *super-spreaders* employing Léon et al. (2014)'s terminology) of the interbank funds and central bank's repo network by using a centrality measure based on the HITS (Hypertext Induced Topic Search) algorithm by Kleinberg (1998). This metrics is based on two distinct eigenvector centralities, namely the authority centrality and the hub centrality, which are associated with the in-degree and the out-degree of a node respectively. By using this approach, the SIBs are identified not only by their role as lenders in the interbank market, but also by their ability to access to central bank funding.

This paper is organized as follows. Section 2 describes the repo market and its regulatory system. Section 3 presents the dataset. Section 4 provides a short introduction to networks. Section 5 reports the main properties of the Mexican repo network. Section 6 briefly introduces control theory and presents the results of the controllability analysis. Section 7 concludes and drives the policy implications.

3.2 The repo market and the institutional background

Before describing the institutional structure of the Mexican repo market, we give a brief description of the main features of a repo contract.

A repurchase agreement, or repo, involves the simultaneous sale and forward agreement to repurchase the same, or a similar, security on a specified future date at a prearranged price. It is equivalent to a collateralized loan in which a cash-lender (repo-buyer) lends to a cash-borrower (repo-seller) and receives securities as collateral until the loan is repaid. At the beginning of the contract the repo-seller sells the security to the repo-buyer and takes cash. At maturity the repo-seller returns the capital lent plus interest (repo rate) to the repo-buyer and receives back the security. In order to protect the cash-lender from the risk of a reduction in the value of the collateral, repos usually involve overcollateralization and the difference between the value of the collateral and the value of the cash is the haircut, which is expressed as a percentage of the market value of the collateral. For example, a repo loan of MXN 95m with a collateral of MXN 100m is equivalent to an haircut of 5% .

According to the degree of intermediation between the cash-borrower and the cash-lender, repos can be distinguished in three types: OTC bilateral, tri-party repos and central clearing counterparty (CCP)-based. On the OTC market, the cash-lender and the cash-borrower transact directly with each other. Haircuts are part of the negotiation and both parties bear the risk of the transaction, in addition to settling the transaction themselves. Bilateral repos can allow for general collateral or impose strict standard securities. In tri-party repos, a clearing bank stands as an agent between the cash-borrower and the cash-lender. The clearing bank provides settlement and collateral management services to the parties, such as the selection and the valuation of the collateral and the transfer of cash and securities. However, the clearing bank does not take the risk that a counterparty defaults and the counterparty risk remains with the repo traders. Non-government bonds and equity are predominantly employed as collateral in tri-party repos. Finally, in CCP-based repos a central counterparty bears the counterparty risk and the lender is therefore protected from losses in case of default of the borrower. The transaction takes place anonymously via electronic platforms and the cash-lender and the cash-borrower do not have any direct exposure to each other. The CCP interposes itself between the cash-borrower and the cash-lender: it borrows the security (and lends cash against it) from the cash-borrower, and lends the security to the cash-lender (and borrows cash in exchange). The haircuts in this market are set centrally by the CCP, which is also in charge of the risk management duties. The operations are settled through a multilateral netting scheme and it means that the CCP nets the transactions between the participants, resulting in smaller net payment and collateral delivery obligations between the banks and the CCP. In contrast to tri-party repos, government bonds and other relatively safe securities are the most commonly-used type of collateral in CCP-based repos.

Repo clearing and settlement arrangements vary considerably across countries. In Europe CCP-based repos constitute the majority of the interbank repo market with a market share of 70 % (ECB, 2015). The assets financed are exclusively government bonds and the market is dominated by banks. By opposite, the tri-party market is the most important segment of the U.S. repo market, where they are currently two third parties: Bank of New York and JP Morgan Chase. Moreover, the cash-borrowers are not banks, but non-bank institutions without access to central bank's liquidity. Hence, the repo market represents for the shadow banking system an essential short-term funding source. Another difference between the two markets is that in U.S. the collateral is concentrated in private assets that are not backed by the U. S government or its agencies and that exhibit significant price volatility and illiquidity in crisis times. Indeed, several studies, such as Gorton and Metrick (2009) and Krishnamurthy et al. (2014), identify the uncertainty about the quality of the asset-backed securities (ABS) used as collateral as the main cause of the freezing of the U.S. repo market during the last financial crisis.

In the Mexican case repos are traded over the counter (OTC) and government securities are the most used collateral. Most of the lending is at very short maturities, typically from one to

three days. Credit institutions (commercial banks and development banks) and brokerage houses are allowed to lend and to borrow in the Mexican repo market, while the other participants, such as physical and legal persons (companies and persons), mutual funds, insurance companies, investment funds and the government's treasury, can act only as lenders. Repos are also an important monetary policy instrument for the Mexico's Central Bank. Central bank's repos correspond to the liquidity provided by the Mexico's central bank to the banking system throughout open market operations (OMOs), where the eligible collateral is mainly local sovereign securities.

3.3 Data set

Our dataset is provided by the Mexico's Central Bank and it consists of the repo transactions in Mexico during the period 2005-2014. For each repo trade we have access to: the institutions' ID number (the identity is anonymised), the type of institution, the repo position (repo-buyer or repo-seller), the maturity, the amount (expressed in Mexican Pesos) and the date of the transaction.

As can be seen from Figure 3.1, our data confirms that repos are almost exclusively overnight loans in Mexico and the collateral is concentrated in government securities. Moreover, the right panel of Figure 3.1 decomposes the repo market by type of participant. Participants are grouped into: banks (commercial banks and development banks), authorities (the Mexico's Central Bank), other financial institutions (brokerage houses, insurance companies and investment funds), physical and legal persons and public sector (the government's treasury and state-owned companies). Physical and legal persons and banks are the two largest participants and they account together for nearly 80 % of the value of the transactions. Overall, it is also striking that the amount of repos traded has a positive trend with a significant increase after the end of the financial crisis: it peaks in 2012 and then remains substantially at the same level.

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
banks	35	37	42	45	43	41	43	44	45	45
volume (MXN billion)	78.04	79.55	80.10	82.11	84.18	93.78	117.36	137.79	133.95	127.87
trades	59,825	59,898	60,095	62,868	65,213	69,332	72,196	72,624	71,750	70,759
links	493	529	548	547	516	508	573	566	554	578
trades per link	121.35	113.23	109.66k	114.93	126.38	136.48	125.99	128.31	129.51	122.42

Table 3.1: Evolution of the repo market from 2005 to 2014 in terms of: number of active banks, volume (MXN billion), number of trades, number of links and number of trades per link.

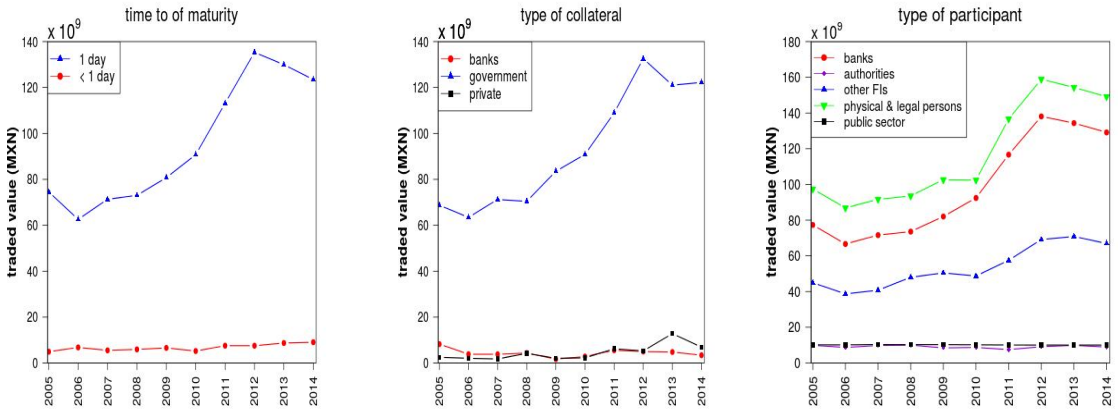


Figure 3.1: Evolution of the total value of repo loans in billion Mexican pesos (MXN) in the period 2005-2014. In the left panel by time to maturity: 1 day and greater than 1 day. In the middle panel by type of collateral: bank securities, private securities and government securities. In the right panel by type of participant: banks (commercial banks and development banks), authorities (the Mexico's Central Bank), other financial institutions (brokerage houses, insurance companies and investment funds), physical and legal persons and public sector (the government's treasury and state-owned companies).

For our study we include only overnight transactions, that involve commercial banks and development banks because by Mexican law they are the only credit institutions which can get funding through the repo market. Moreover, exclusively repos with sovereign securities are considered. The number of transactions covered in our dataset is 5,412, which corresponds to an amount approximately equal to MXN 999.73 billion. Table 3.1 presents some basic descriptive statistics for our sample. The number of banks is remarkably stable, ranging from 35 to 45, while the trading volume tends to increase over time, even though the number transactions remains constant, which implies an increase in the size of the repos. Finally, Figure 3.2 reports information on the total amount and on the number of loans made and received by type of participant. As expected, commercial banks account for 75 % of the repo market in terms of volume and number of transactions, reflecting a minor role for development banks.

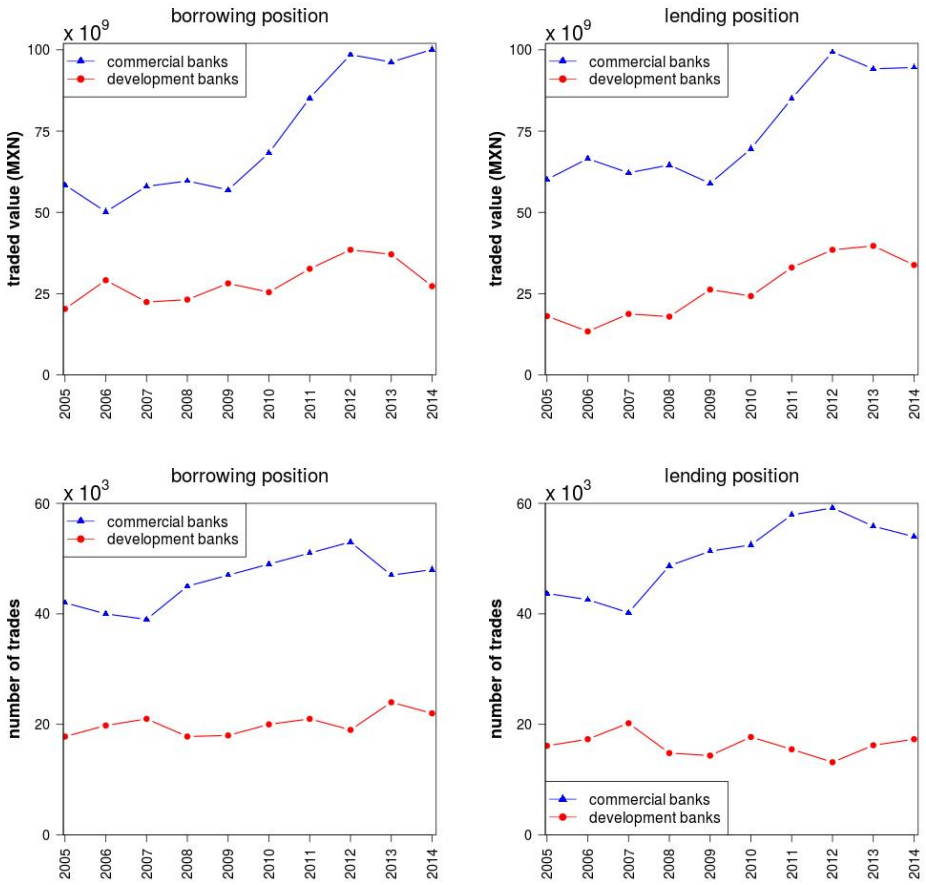


Figure 3.2: The total amount (top) and the number of transactions (bottom) from the borrower and lender-side for each type of bank (commercial bank and development bank).

3.4 Network analysis

All these overnight repos form a network $G = (V, E)$, where the set V of vertices are the banks and the set E of links represents the repo contracts among them. The network is directed and weighted, which means that the links have a direction (from one node to another) and an associated value according to their importance. Namely, a link goes from the repo-buyer to the repo-seller and a weight is attached to each link in terms of the total amount of money lent by the repo-buyer to the repo-seller. From a mathematical point of view, a repo network of N banks is represented by

a square matrix **A** of size N.[2] Since the network is directed, the adjacency matrix is asymmetric and a nonzero element a_{ij} in **A** represents the amount of repo loans that bank i extends to bank j.

Degree. In network theory the number of links of a node is known as its degree and in the case of a directed network there are two kinds of degree: in-degree and out-degree. The in-degree of a node is the number of its incoming links and the out-degree of a node is the number of its outcoming links. Formally, the in-degree k_i^{in} and the out-degree k_i^{out} of a node i in a network of size N can be defined as follows:

$$k_i^{in} = \sum_{j=1}^{N} a_{ji} \tag{3.1}$$

$$k_i^{out} = \sum_{j=1}^{N} a_{ij} \tag{3.2}$$

In our study $a_{ij} = 1$ if bank i provides a repo loan to bank j and $a_{ij} = 0$ otherwise. The in-degree of a bank is the number of banks from whom receives a repo loan, namely the number of its repo-buyers. Similarly, the out-degree of a bank is the number of banks to whom extends a repo loan, namely the number of its repo-sellers.

Strength. In a weighted network the strength of a node is computed as the sum of the weights of all the links attached to it. Similarly to the degree, two kinds of strength are defined in the case of a directed network: in-strength and out-strength. The in-strength of a node is the sum of the weights of its incoming links and the out-strength of a node is the sum of the weights of its outcoming links. Formally, the in-strength s_i^{in} and the out-strength s_i^{out} of a node i in a network of size N can be defined as follows:

$$s_i^{in} = \sum_{j=1}^{N} w_{ji} \tag{3.3}$$

$$s_i^{out} = \sum_{j=1}^{N} w_{ij} \tag{3.4}$$

In our study w_{ij} is the amount of repo that bank i extends to bank j. The in-strength s_i^{in} of bank i corresponds to the total amount of money borrowed by bank i in the repo market and the out-strength s_i^{out} of bank i corresponds to the total amount of money lent by bank i in the repo market.

Reciprocity. In a directed network reciprocity refers to the fraction of links present in both

[2]In everything that follows, bold capital letters refer to matrices and bold lower-case letters to vectors.

directions out of the total number of links. Formally, reciprocity for a network with $E^{\leftrightarrow}$ bidirectional links is defined as:

$$r = \frac{E^{\leftrightarrow}}{E} \tag{3.5}$$

For a repo network reciprocity can be seen as the number of repos where banks enter both as repo-buyer and repo-seller.

Density. The density of a network is defined as the proportion of actual links E relative to the total possible number of links. For a directed network of size N, the density is computed as:

$$\rho = \frac{E}{N(N-1)} \tag{3.6}$$

The index is closer to 1, as the network gets denser. In the extreme case where all the possible edges are established, the graph is complete and the index is equal to 1. For a repo network the density can be computed as the ratio between the number of realized repo contracts and the number of all the possible repo contracts.

3.5 The Mexican repo network

Based on the concepts and metrics described before, in this section we present the main properties of the Mexican repo network. We study the repo network for: daily, weekly, monthly and quarterly time scale. We obtain the weekly, monthly and quarterly networks by aggregating the daily loans over a 5, 21 and 63-day period respectively. The four networks at the different time aggregations are displayed in Figure 3.3. Each vertex corresponds to a financial institution and each arrow represents the existence of a repo transaction between two financial institutions. The direction of the arrow corresponds to the direction of the original transaction (i.e. from the repo-buyer to the repo-seller).

From the plots it is clear that the four networks are particularly sparse and they become denser as the time aggregation increases. This is expected because the networks are obtained by summing the edges across time and they therefore become denser with the time scale by construction. Table 3.2 provides further evidence of the decay of the density with the time aggregation. An average density of around 10% for the daily and weekly networks has been found, whereas an average density of around 20% for the monthly and quarterly networks.

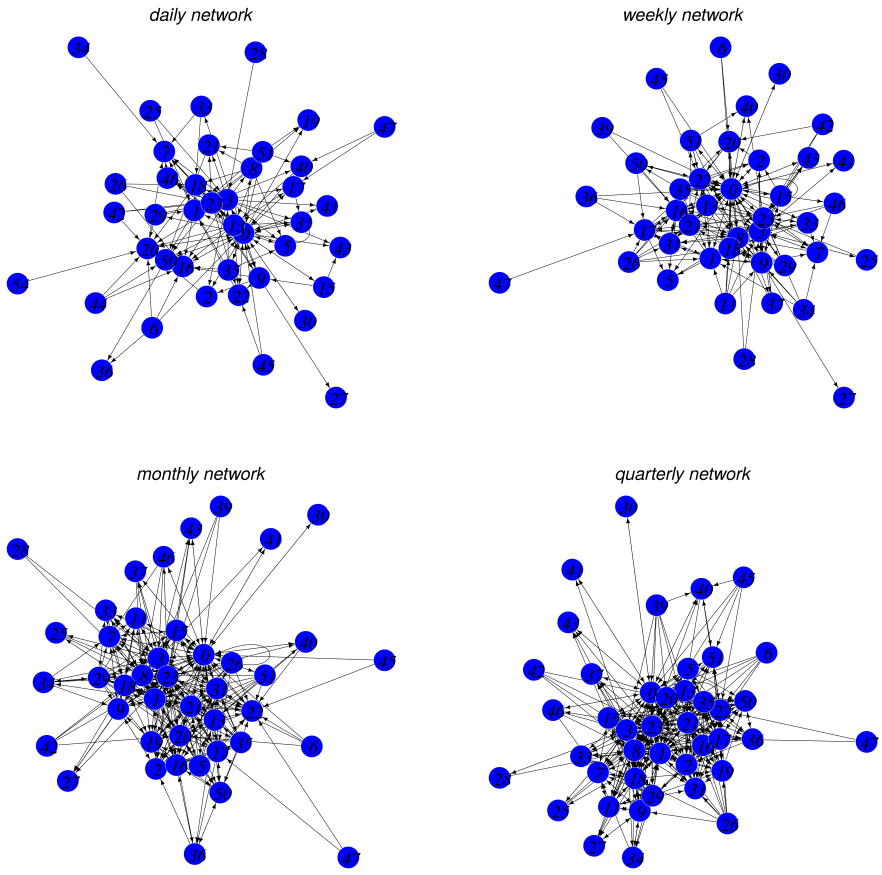


Figure 3.3: Repo network for the different time scales. Top left: the daily network for April 15$^{\text{th}}$, 2013. Top right: the weekly network for the second week of April, 2013. Bottom left: the monthly network for April, 2013. Bottom right: the quarterly network for 2013 second quarter. The network are represented by using Kamada-Kaway algorithm.

density	day	week	month	quarter
mean	0.08	0.13	0.19	0.24
standard deviation	0.006	0.01	0.015	0.018

Table 3.2: Density for the daily, weekly, monthly and quarterly network. Averages and standard deviations based on all the networks at the different time scales.

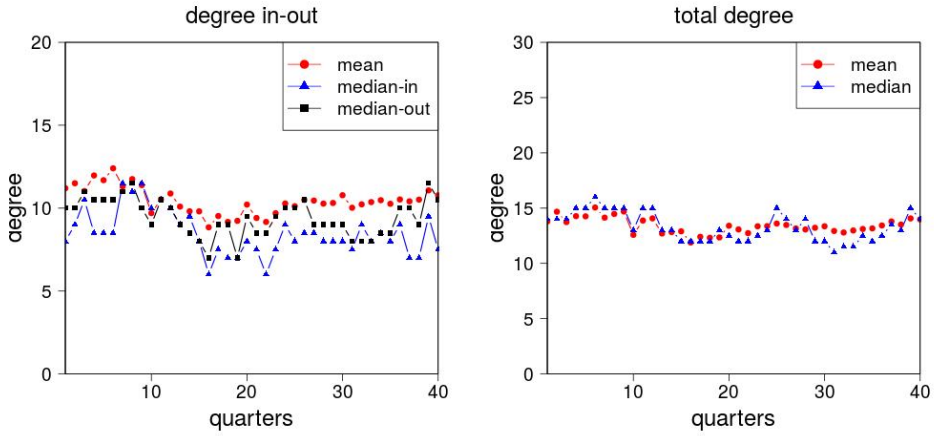


Figure 3.4: Left: mean and median for in- and out-degrees (directed network). Right: mean and median for total degrees (undirected network).

Before analyzing the distributions of the number of lending relationships (degrees) across banks, we briefly illustrate their evolution over time by using data on a quarterly basis. Left panel of Figure 3.4 shows the in- and out-degrees from the directed networks. The median value for in-degree and out-degree is approximately the same, suggesting that in- and out-degree follow the same distribution. In addition, the Kolmogorov-Smirnoff test is also performed for all the sample periods and the null hypothesis of the equality of the in- and out-degree distributions is not rejected at a signicance level of 1 %. Such results seem to indicate that the directed version of the network does not contain additional information on the dymamics of the degrees. Further evidence is provided by the correlation between in- and out-degrees and the link reciprocity, which are plotted in the left panel and in the right panel of Figure 3.5 respectively. The correlation between in and out-degrees is remarkably high 0.82 ± 0.04 and the reciprocity of the network is also high 0.73 ± 0.02. This means that: (i) a high level of correlation between the number of repos in which a bank is involved as a borrower and those in which is involved as a lender (ii) banks predominantly borrow and lend money from the same banks.

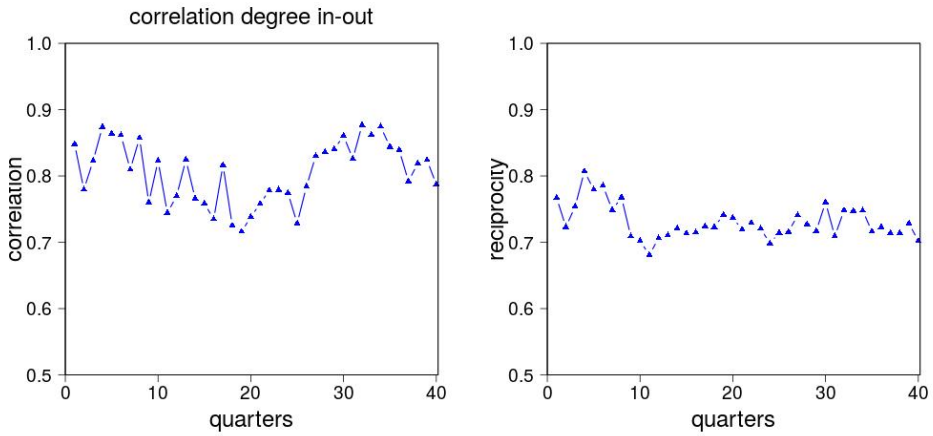


Figure 3.5: Left: correlation in- and out-degrees. Right: reciprocity.

Therefore, in the following we will work with the undirected version of the repo networks without a significant loss in accuracy.

The right panel of Figure 3.4 shows the evolution of the mean and the median for the degrees from the undirect networks. We do not observe any breaks in the mean and median values and we therefore analyze the degree distribution without splitting the data into different time-periods. Noticeably, the mean and the median are approximately the same, indicating that the distribution is symmetric. Figure 3.6 reports the degree distributions for daily, weekly, monthly and quarterly aggregation periods. The plots correspond to a single reference time period because, as formerly pointed out, we do not observe any breaks in the time series of the mean and the median. The plot read as follows. For a given point of the degree distribution function, the y-axis coordinate gives the number of banks holding the number of links equal to the value of the x-axis coordinate. It is evident that distribution of the degrees is quite homogeneous at any time scale, although the repo market is more hetereogeneous at shorter time scales. Moreover, the daily network displays a smooth right-skewed distribution with a skewness around 3, while the monthly and quarterly networks are perfectly symmetric with a skewness near zero.

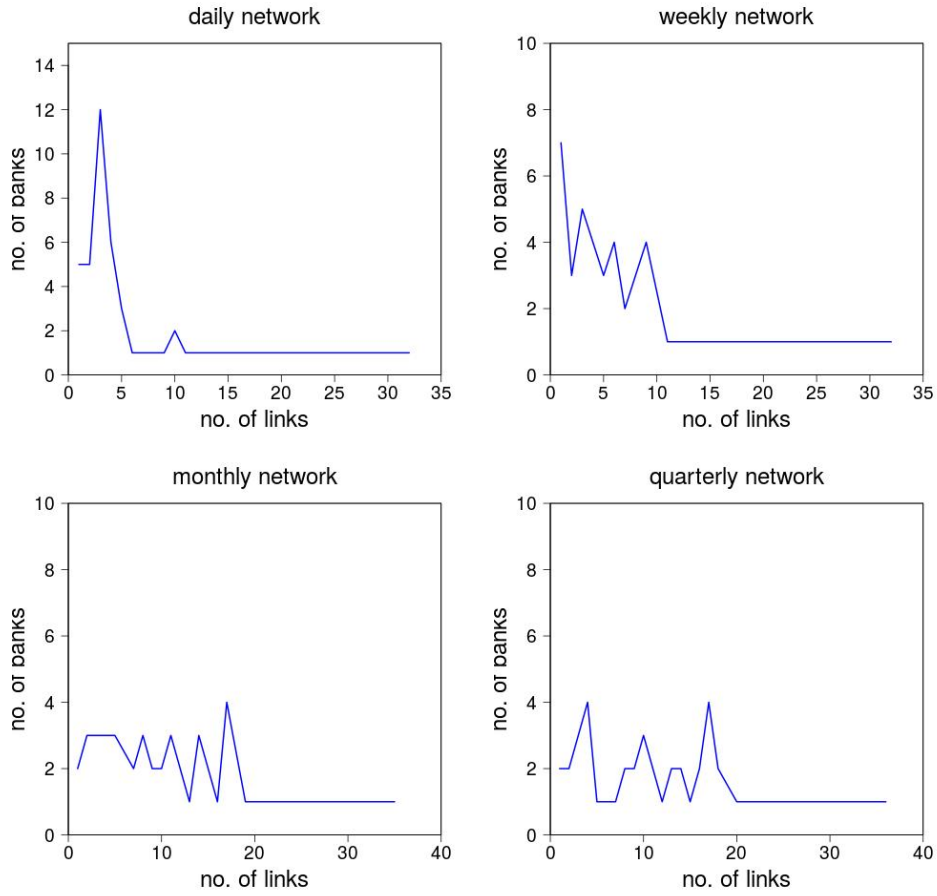


Figure 3.6: Degree distributions for the different time scales. Top left: the daily network for April 15$^{\text{th}}$, 2013. Top right: the weekly network for the second week of April, 2013. Bottom left: the monthly network for April, 2013. Bottom right: the quarterly network for 2013 second quarter.

3.6 Controllability analysis

3.6.1 Method

Given a weighted and directed graph $G = (V, E)$ of N nodes, we consider a linear time-invariant (LTI) continuous dynamics for the state $\mathbf{x}$ of the network

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \tag{3.7}$$

The vector $\mathbf{x}(t)$ of size N represents the state of the network at time t. $\mathbf{A}$ is a square matrix of size N and represents the transpose of the adjacency matrix of the network. $\mathbf{B}$ is a $N \times M$ matrix and provides the weights of the links between the nodes of the graph and ($M \leq N$) external signals, indicated by the vector $\mathbf{u}(t)$ of size M.

Controllability is equivalent to find the M signals $\mathbf{u}$ such that $\mathbf{x}$ is equal to the desired final state for a given network represented by the adjacency matrix $\mathbf{A}$. Note that the assumption of linearity in Eq. (7) refers exclusively to the control of the state of the network and do *not* specify any network formation process.

A sufficient and necessary condition for a network to be controllable is given by the Kalman rank condition. For a given realization of $\mathbf{A}$ and $\mathbf{B}$, the state of a network is controllable if and only if the $N \times NM$ controllability matrix $\mathbf{C}$

$$\mathbf{C} = (\mathbf{B}, \mathbf{A} \cdot \mathbf{B}, \mathbf{A^2} \cdot \mathbf{B}, \ldots, \mathbf{A^{N-1}} \cdot \mathbf{B}) \tag{3.8}$$

has full rank.

$$rank(\mathbf{C}) = N \tag{3.9}$$

However, the rank condition has no practical application for the control of a network. First, it does not specify how to identify the signals $\mathbf{u}$. In other words, which are the nodes to which the signals must be applied in order to achieve the full control of the network, the so-called *drivers*. Second, the computation of the rank is computationally demanding for large networks, since the rank of $\mathbf{C}$ must be calculated for $2^N - 1$ different combinations.

In order to overcome these limitations, Liu et al. (2011) have recently proposed a feasible method to identify the minimum number of drivers for a directed network of arbitrary size. They adopt the framework of *structural* control of Lin (1974), that is based on a more practical definition of control, which does not require the exact value of the elements of $\mathbf{A}$ in order to deal with incomplete knowledge of most real-world systems. The entries of $\mathbf{A}$ and $\mathbf{B}$ in Eq. (7) are therefore assumed to be fixed zeros and free parameters, which means that a network is controllable in the usual sense ($rank(\mathbf{C}) = N$) for any realization of $\mathbf{A}$ and $\mathbf{B}$ with the same non-zero elements. Note that structural control depends exclusively on the structure of the network (the presence or absence of links) and it is set apart from the weights of the links.

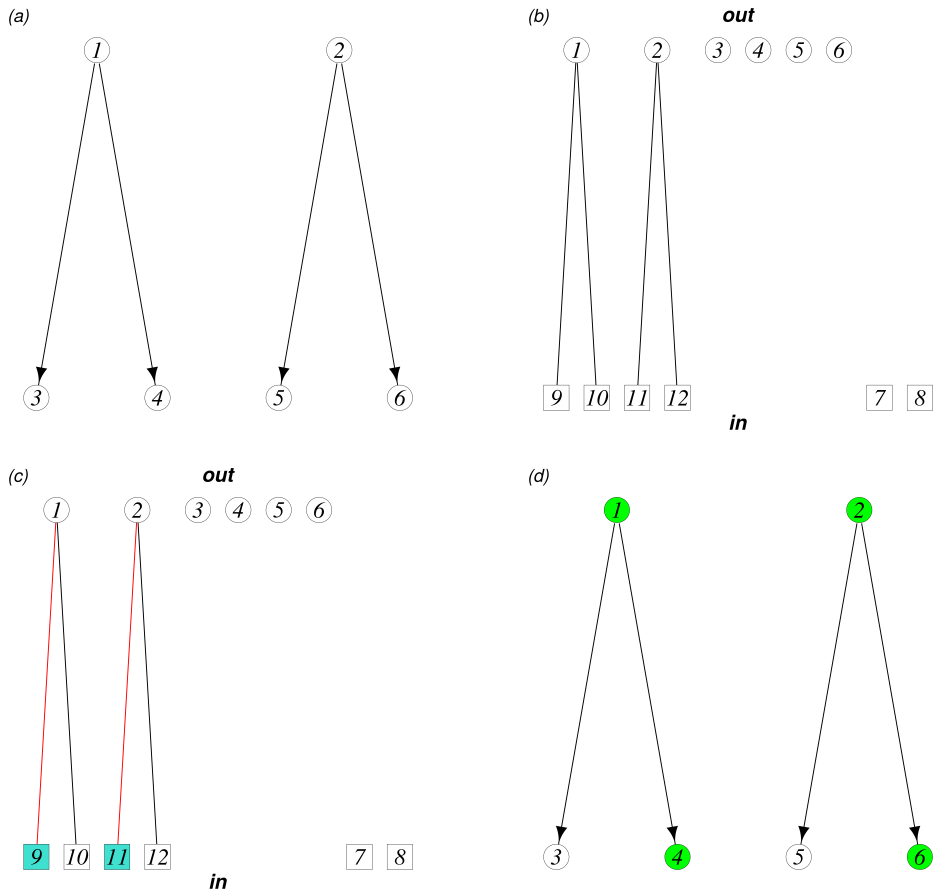


Figure 3.7: (a): a directed network with 6 nodes. (b): the bipartite representation of the directed network in (a). Nodes (1,2,3,4,5,6) of the *out* set correspond to the nodes (7,8,9,10,11,12) of the *in* set respectively. For instance, a directed link from node 1 to node 3 in (a) corresponds to a link between the node 1 in the *out* set and the node 9 in the *in* set. (c): a maximum matching of the bipartite network is shown in red. Matched nodes are shown in turquoise, while unmatched nodes in white. Note that (1-9, 2-11) is not the unique maximum matching. Other maximum matchings are: (1-9, 2-12), (1-10, 2-11), (1-10, 2-12). (d): the minimum number of drivers of (a) based on the maximum matching (1-9, 2-11). The drivers (marked in green) are the unmatched nodes.

A schematic representation of the Liu et al. (2011)'s methodology is reported in Figure 3.7. The method encompasses two steps. First, a directed network (Figure 7*a*) is transformed into an equivalent bipartite network (Figure 7*b*), which is obtained in the following way. Nodes are divided into two distinct sets *out* and *in*. A link between a node i in the *out* set and a node j in the *in* set exists if a link goes from node i to node j in the original network. Second, a maximum matching is computed in the bipartite network (Figure 7*c*). For a bipartite graph (or, more generally, for an undirected graph), a matching M is a subset of edges without common vertices. A matching that contains the largest possible number of edges is a maximum matching.[3] A vertex is matched if it is a node in the *in* set, which is incident to an edge in the matching. Otherwise, the vertex is unmatched.

The authors proved that the minimum number of drivers need to fully control a network is equal to the number of unmatched nodes in any maximum matchings, providing an efficient methodology for network controllability. As it is evident from the toy example in Figure 3.7, the theorem (the so-called *Minimum Input Theorem*) has an intuitive explanation. For instance, consider the maximum matching (1-9, 2-11) in Figure 7*c*. The matched nodes 3 and 5 are internally controlled[4] (Figure 7*d*) because node 1 points to 3 and node 2 to 5. By contrast, the unmatched nodes 1,2,4 and 6 are out of control. Nodes 1 and 2 have no nodes pointing to them, whereas nodes 4 and 6 share the unique node, which points to them with a matched node: node 4 shares 1 with 3 and node 6 shares 2 with 5. Therefore, the unmatched nodes identify exactly which nodes need to be externally controlled. That is, where the signals must be injected into a network in order to ensure that each node has its own controller, so that the full control is achieved.

As a last step of their analysis, Liu et al. (2011) explore the network characteristics that affect controllability. It turns out the degree distribution mainly determines the controllability, measured by the density of the drivers. The sequence of random networks preserving the in-degree and the out-degree sequences predicts extremely well the density of the drivers of several real networks. The simulation results are also validated analytically by using the cavity method (Mézard and Parisi (2001)), revealing that hetereogeneous and sparse networks are the hardest to control.

In this paper we apply the method developed by Liu et al. (2011) to detect the SIBs of the Mexican interbank market, focusing on the repo segment. In this context, controllability is equivalent to select the minimum number of driver banks that, if refinanced by the central bank, steer the liquidity in the interbank market into the defined target for the overnight rate. Figure 3.8 provides a stylised illustration of how central bank liquidity is redistributed in the interbank market throughout the driver banks. We identify the state $\mathbf{x}$ of the repo network with the aggregate lending, namely the total amount of money lent in the repo market. $\mathbf{A}$ is the transpose of the adjacency matrix of the repo network, where an element a_{ij} corresponds to the amount of repo

[3] Note that, a maximum matching is not necessarily unique. See, for example, Figure 7*c*.

[4] The nodes 3 and 5 are the univariate representation of the nodes 9 and 11 respectively.

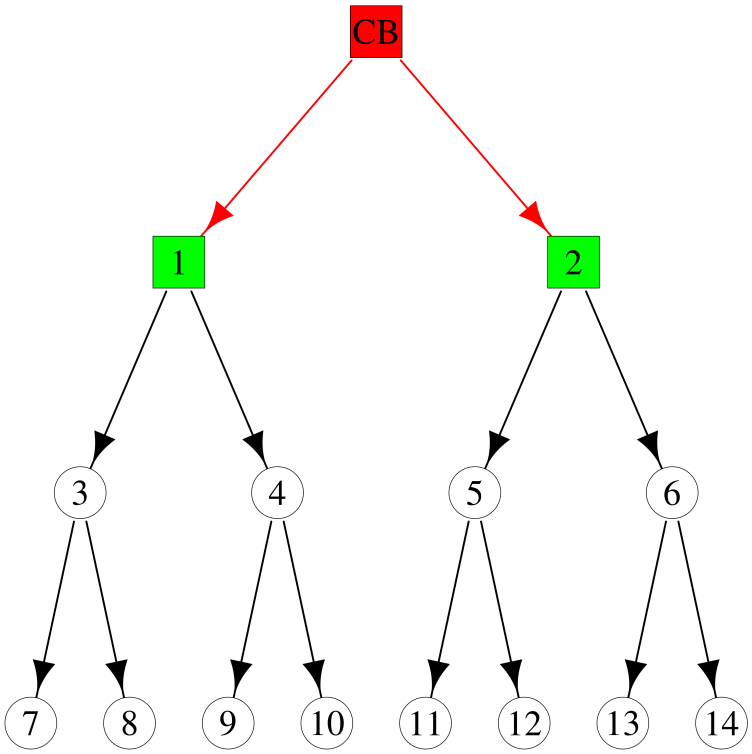


Figure 3.8: Example of the control of an interbank market. The bank funding can be completely controlled whether the central bank refinances the driver banks. The control of the interbank funds works *hierarchically*. The central bank provides liquidity directly to the driver banks (green nodes). Then, the drivers lend to other banks (white nodes, first layer), which, in turn, provide funding to the rest of the interbank market (white nodes, second layer).

borrowed by bank i from bank j. The external controller is the central bank. The entries of the vector $\mathbf{u}$ identify the central bank refinancing operations with $\mathbf{B}$ providing the amounts. We emphasize that the refinancing operations must be strictly at the initiative of the central banks, such as open market operations, because in our framework the signals are imposed by the external controller.[5]

It is worthwhile to highlight that the dynamics for the control of the interbank funds in Eq. (7)

[5] As a matter of fact, standing facilities are excluded.

rely on three strong assumptions. First, the lending volume of a bank is proportional to the lending volume of its lenders. While it is reasonable to hypothesize that the funding of a bank depends on the funding of its lenders, it is certainly too simplistic to assume a linear relation. Second, the lending relationships are assumed constant, since the zero and non-zero entries of $\mathbf{A}$ are fixed. Third, the SIBs are selected merely on the basis of the structure of the lending relationships (who lends to whom/who borrows from whom), without taking into account neither the amount of the interbank loans nor the amount of the central bank refinancing operations. In practice, this means that for any interbank market with a given set of lending relationships, the banks offering full control over the interbank funds are the same regardless how much they lend to the other banks and how much they take from the central bank.

In the following, we provide some technical details how we estimated the driver banks. We use the well-known Hopcroft-Karp algorithm to find a maximum matching in the bipartite version of the repo network. The edges in a matching are recursively increased by finding the so-called *augmenting* paths. The augmenting paths are sequences of edges, which start at a matched node and end at an unmatched node, by alternating edges inside and outside the matching. The algorithm iteratively swaps which edges are inside and which edges are outside the matching through a Breadth-First Search (BFS) in order to produce the matching with the largest possible number of edges. Finally, from the different maximum matchings we select the maximum matching with the maximum sum of the weights of the edges to incorporate somehow the information on the amount of the loans. Intuitevely, including the largest bilateral exposures we are maximizing the impact of the SIBs on the interbank market.

3.6.2 Results

In this section we report the results of our study on the controllability of the Mexican repo network in the period 2005-2014 for the overnight interbank market.

We estimate the network controllability for the daily, weekly, monthly and quarterly scale. We are interested in estimating the controllability of the repo network for different time horizons because different time horizons reflect different supervision purposes and different network structures. We then proceed to compare the set of driver banks with the standard network centrality measures, such as degrees (the number of counterparties) and out-strength (the total amount of repos lent). In the last step we analyze the stability of the drivers over time. We use the fraction of the driver banks n_D as the metrics of the control efficiency of the network. A large value of n_D indicates a low degree of the controllability of the network because a high number of driver nodes is required to control the state of the whole network.

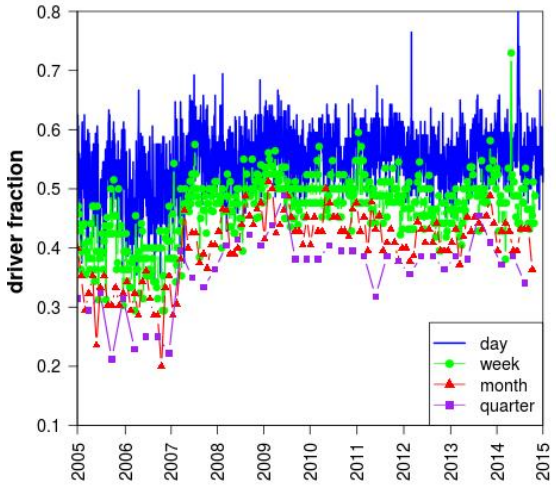


Figure 3.9: Evolution of the fraction of drivers for daily, weekly, monthly and quarterly networks.

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
day	0.523	0.488	0.548	0.543	0.567	0.563	0.572	0.543	0.556	0.567
	(0.052)	(0.054)	(0.052)	(0.044)	(0.039)	(0.034)	(0.034)	(0.037)	(0.039)	(0.047)
week	0.416	0.373	0.463	0.481	0.501	0.495	0.498	0.468	0.478	0.490
	(0.045)	(0.040)	(0.049)	(0.034)	(0.033)	(0.027)	(0.031)	(0.034)	(0.034)	(0.051)
month	0.324	0.308	0.381	0.436	0.453	0.444	0.436	0.409	0.424	0.423
	(0.039)	(0.041)	(0.048)	(0.032)	(0.034)	(0.025)	(0.028)	(0.023)	(0.025)	(0.030)
quarter	0.286	0.261	0.325	0.401	0.428	0.391	0.372	0.377	0.396	0.377
	(0.044)	(0.032)	(0.064)	(0.023)	(0.040)	(0.010)	(0.031)	(0.012)	(0.035)	(0.025)

Table 3.3: Average fraction of drivers for daily, weekly, monthly and quarterly networks. Standard deviations in parenthesis.

Figure 3.9 plots the evolution of the fraction n_D at the different time scales. The most evident observation is that the controllability of the repo network increases monotonically with the time scale. The reason is that the denser and homogeneous a network is, fewer drivers are needed to control it. Liu et al. (2011) show analytically that the minimum number of drivers required to control a network is mainly determined by the degree distribution, finding that heterogeneous and sparse networks report a higher number of drivers. As it turns out, the density and the homogeneity in degrees increase with the aggregation period (see Section 5), which makes it easier for central

banks to control interbank funds over longer time horizons. However, an intuitive explanation is that interbank funds are less volatile in the long run, and for this reason they are easier to control.

However overall, the repo network exhibits a high degree of controllability at any time scale with variations below 10% (see Table 3.3). Even the daily network requires only about the 50% of driver banks to achieve the full control over the network. From a monetary policy perspective, a fraction of drivers below 60% at any time scale reflects the possibility for central banks of ad-hoc liquidity injections to a feasible number of banks. Our results are similar to those by Delpini et al. (2013) for the Italian interbank market, who find a high level of controllability of the system at any time scale and a power-law decay of the fraction of drivers with the time horizon. In the following, we characterize the driver banks with standard network centrality measures, namely degrees and out-strength. For the sake of completeness, we consider the total degrees (the sum of in-degrees and out-degrees) from the direct networks, which provide information on the number of relationships in which a bank acts as repo-buyer and as repo-seller. In the left panel of Figure 3.10 we plot the average degree of the drivers against the average degree of the network. It is clear that the drivers are the low-degree nodes of the network at every time scale because the average degree of the drivers is constantly lower than the average degree of the network.

Next, we move to the volumes and we assess the extent to which the drivers correspond with the largest lenders of the repo market. For simplicity, we identify the top-lenders as the first m_{top} lenders whose volume accounts for approximately 80% of the total lending. In the right panel of Figure 3.10 we plot the fraction of top-lenders which are also the drivers of the network. Surprisingly, not all the top-lenders are drivers. On average around 10 % of the top-lenders are drivers at quarterly scale and the percentage gradually increases at shorter time scales. Again,

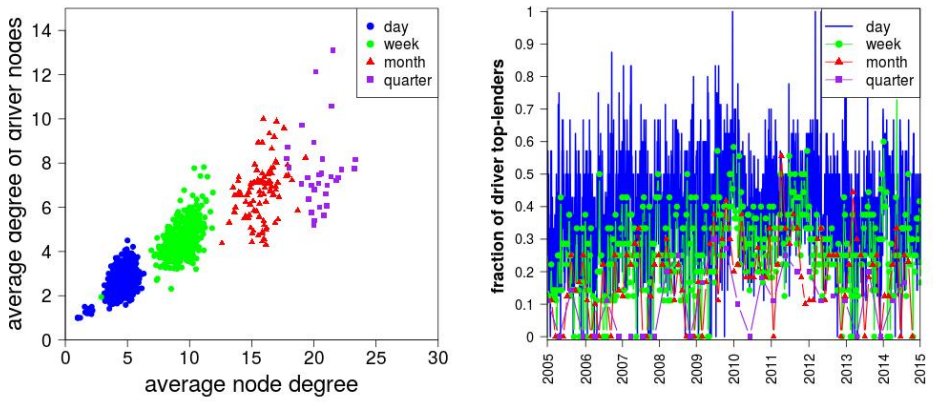


Figure 3.10: Left: scatterplot of the average degree of the driver nodes *versus* the average network degree. Right: fraction of the top-lenders being also the drivers of the network.

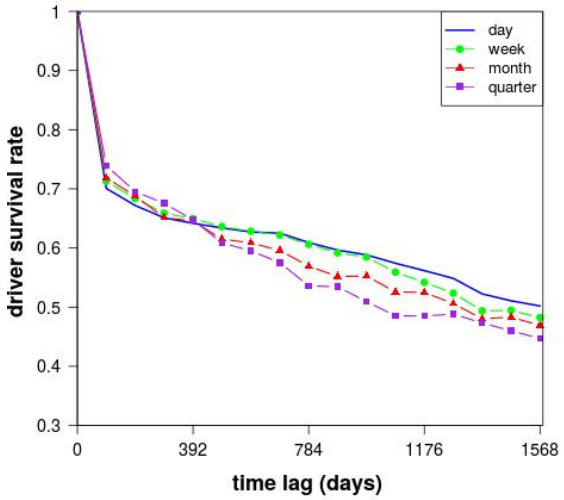


Figure 3.11: Empirical survival function for the drivers. The plot can be read as follows. For a given point of the curve, the y-axis coordinate gives the average proportion of drivers which survives after a number of days equal to the value of the x-axis coordinate.

similar results are reported by Delpini et al. (2013), who also find that the drivers are typically neither the largest lenders nor the high degree nodes of the network. Overall, our analysis supports the view that the traditional focus on TBTF institutions could be misleading and the interconnections among banks should be also considered for the design of a robust policy guide. Standard centrality metrics, such as the degrees and the size, do not capture the topology of the network and control theory could be therefore of interest to central banks for a too-interconnected-to-fail (TITF) assessment. Moreover, we would like to point out that even though the Mexican network is homogeneous and the Italian network is scale-free (De Masi et al. (2006)), our results are similiar to those by Delpini et al. (2013). This may suggest that these properties of the drivers are a common feature of interbank markets, which exist independently of the network-specific characteristics. Finally, in Figure 3.11 we test whether the driver banks remain stable over time. No significant changes in the composition of the drivers are observed when we track the survival function: around 65 % of the drivers are still the drivers of the network after 392 days (around 18 months) at any time scales. Thus, the flow from and to the driver set is only a small percentage, which supports the evidence that we have identified a truly structural feature of the interbank market. Moreover, the fact that the persistence of the drivers decreases with the time scale makes sense and could be due to the more structural changes of the interbank relations over longer time horizons.

3.7 Conclusions and policy implications

The main findings of our paper are the following. We find a highly persistent set of drivers in the Mexican interbank network for the repo market in the period 2005-2014. The identified drivers do not coincide with the most connected banks or with the largest lenders. Our results are similar to those by Delpini et al. (2013) for the Italian interbank market, albeit the different types of network topology. Scale-free the Italian network and homogeneous the Mexican network. Evidence advocates that the finding of a stable set of drivers with the above-mentioned features is unlikely to be a coincidence. Therefore, we expect that other interbank markets display similar properties, which might be classified as new "stylized facts" of the controllability of interbank networks.

Although network controllability is not ready yet for a formal policy design and concrete policy realizations, it certainly provides interesting insights into the financial institutions, which should be monitored. In our view, there are mainly three reasons why control theory might offer a convenient way to identify the SIBs of the Mexican repo market. First, it does not require the exact knowledge of the size of the loans, which is a particularly nice feature for regulators because a complete picture of the interbank interactions is often missing. Second, it identifies a fraction of drivers below the 60%, which makes it possible for the central bank to implement ad hoc liquidity injections to a feasible number of banks. Third, it takes into account the time dimension of the monetary policy actions. The fraction of drivers changes with the time scale and this is an appealing result for regulators because the provision of liquidity is reasonably adjusted to the time horizon of interest.

Finally, the following extensions are possible. An useful test is to compare the drivers with the banks, which receive money from the Mexico's Central Bank throughout OMOs. Such a comparative exercise may provide an understanding of the cause and effect relationship between the identification of the SIBs and the effectiveness of Mexico's Central Bank monetary policy. In this respect, it is also advisable to match our results with the dates of OMOs and with the interbank interest rates. Furthermore, it would be interesting to compare the drivers with centrality measures, which also include the dimension of interconnectedness in order to control that driverness is not a redundant property, but provides additional explanatory power. Possible candidates could be the so-called feedback centralities, such as Katz centrality and PageRank, which give a measure of the direct and indirect links in the network by ranking the centrality of a node on the basis of the centrality of its neighbors.

Bibliography

Adrian, T., Begalle, B., Copeland, A. and Martin, A., 2012. Repo and securities lending. National Bureau of Economic Research

Allen, F. and Gale, D., 2000. Financial contagion. *Journal of political economy* 108(1), 1-33

Copeland, A., Martin, A. and Walker, M., 2014. Repo Runs: Evidence from the Tri-Party Repo Market. *The Journal of Finance* 69 (6), 2343-2380

Bech, M.L. and Atalay, E., 2010. The topology of the federal funds market. *Phys. A: Stat. Mech. Appl.* 389(22), 5223-5246

Craig, B.R and von Peter, G., 2014. Interbank tiering and money center banks. *Journal of Financial Intermediation* 23(3), 322-347

Delpini, D., Battiston, S., Riccaboni, M., Gabbi, G., Pammolli, F. and Caldarelli, G., 2013. Evolution of controllability in interbank networks. *Scientific reports*, 3

De Masi, G., Iori, G. and Caldarelli, G., 2006. Fitness model for the Italian interbank money market. *Physical Review E* 74 (6)

ECB, 2015. European Money Market Survey

Fricke, D. and Lux, T., 2015*a*. Core–periphery structure in the overnight money market: evidence from the e-MID trading platform. *Computational Economics* 45(3), 359-395

Fricke, D. and Lux, T., 2015*b*. On the distribution of links in the interbank network: evidence from the e-MID overnight money market. *Empirical Economics* 45(4), 1463-1495

Gai, P., Haldane, A. and Kapadia, S., 2011. Complexity, concentration and contagion. *Journal of Monetary Economics* 58(5), 453-470

Galbiati, M., Delpini, D. and Battiston, S., 2013. The power to control. *Nature Physics*

Galbiati, M. and Stanciu-Vizeteu, L., 2015. Communities and driver nodes in the TARGET2 payment system. *The Journal of Financial Market Infrastructures* 3 (4)

Gorton, G. and Metrick A., 2012. Securitized banking and the run on repo. *Journal of Financial economics* 104 (3), 425-451

Haldane, A.G. and May, R.M., 2011. Systemic risk in banking ecosystems. *Nature* 469(7330), 351-355

Iori, G., De Masi, G., Precup, G., Gabbi, G. and Caldarelli, G., 2008. A network analysis of the Italian overnight money market. *Journal of Economic Dynamics and Control* 32, 259-278

Hopcroft, J.E. and Karp R.M., 1973. An $5^{5/2}$ algorithm for maximum matchings in bipartite graphs.*SIAM Journal on computing* 2, 225-231

Kamada, T. and Kawai, S., 1989. An algorithm for drawing general undirected graphs. *Information processing letters* 31, 7-15

Kleinberg, J.M, 1998. Authoritative sources in a hyperlinked environment. *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*

Krishnamurthy, A., Nagel, S. and Orlov, D., 2014. Sizing up repo. *The Journal of Finance* 69 (6), 2381-2417

Langfield, S., Liu Z. and Ota, T., 2014. Mapping the UK interbank system. *Journal of Banking and Finance* 45, 288-303

Léon, C, Machado, C.L. and Sarmiento, M., 2014. Identifying central bank liquidity superspreaders in interbank funds networks. *Borradores de Economía, No. 816, Banco de la República*

Lin, C.T, 1974. Structural controllability. *IEEE Transactions on Automatic Control* 473 (7646), 19(3), 201-208

Liu, Y-Y., Slotine J-J and Barabasi A-L., 2011. Controllability of complex networks. *Nature* 473 (7646), 167-173

Martínez -Jaramillo, S., Alexandrova-Kabadjova B., Bravo-Benitez B. and Solórzano-Margain, J.P, 2014. 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

Mézard, M. and Parisi, G., 2001. The Bethe lattice spin glass revisited. *The European Physical Journal B-Condensed Matter and Complex Systems* 20(2), 217-233

Molina-Borboa, J. L., Martínez -Jaramillo, S., López-Gallo, F. and van der Leij, M., 2015. A multiplex network analysis of the Mexican banking system: link persistence, overlap and waiting times. *Journal of Network Theory in Finance* 1(1), 99-138

Nier, E., Yang, J., Yorulmazer, T. and Alentorn, A. (2007). Network models and financial stability. *Journal of Economic Dynamics and Control* 31(6), 2033-2060

Soromäki, K., Bech, M.L., Arnold J., Glass R.J, Beyeler W.E., 2007. The topology of interbank payment flows. *Phys. A: Stat. Mech. Appl.* 379(1), 317-333

van Lelyveld, I. and in't Veld D., 2014. Finding the core: network structure in interbank markets. *The Journal of Banking and Finance*, 49, 27-40

Chapter 4

Stress-testing the UK banking system: a network approach to cope with portfolio overlaps

4.1 Introduction and related literature

The emerging risks in the housing market are the central focus of the Bank of England's 2017 stress test.[1] Indeed, low interest rates and an easing of mortgage constraints have led to a substantial increase in UK household debt in recent years. The greater mortgage debt raises concerns about the vulnerabilities of the household sector to changes in income and interest rates with strong implications for the stability of the financial system.

In this respect, the Prudential Regulation Authority has introduced a loan-to-income limit for mortgages exceeding £100 million per annum, that requires that no more than 15 % of mortgages issued should exceed a loan-to-income ratio of 4.5.

Another important policy issue is the correlation between credit (mortgage) cycle and house prices. Mortgage market expansions often coincide with housing price booms, creating a procyclical mechanism of financial fragility, where more credit available allows more potential buyers to bid up house prices.[2]. Housing prices in United Kingdom, are soaring rapidly[3] and it is the question whether or not this is supported by economic fundamentals. Of particular concern are tier 1 cities, such as London and Oxford, where house price boom goes hand-in-hand with an increasing mortgage lending to buy-to-rent (known as buy-to-let) purchasers, who are driving *de facto* the housing boom into a bubble (Bracke 2015; Rafa Batista et al. 2016).

This paper proposes a stress test to assess the resilience of the UK financial system to an hypothetical burst of the housing bubble. We hypothesize that a sharp drop in house prices leads to a devaluation of residential mortgages (Mayer et al. 2009; Cho 1995 among ohers), which may in turn trigger a cascade of defaults through mark-to-market losses. Suppose a bank defaults and its assets are entirely liquidated. The price of these assets will then be driven down, which generates losses to the other banks exposed to the same assets, triggering further defaults and price falls in a vicious downward spiral (Shleifer and Vishny 2011).

In spirit of the state-of-the-art stress tests carried out at central banks (see Constâncio 2015), our framework emphasizes the role of portfolio overlaps as a source of contagion by taking full account of the second round effects that arise from common exposure to the same assets. Such contagion dynamics does not depend on the direct connections between banks (i.e. interbank credit (Allen and Gale 2000; Gai et al. 2011; Battiston et al. 2012; Co-Pierre 2013), CDS protections (Markose et al. 2012; Peltonen et al. 2014; Puliga et al. 2014) or payment obligations (Galbiati and Soramäki 2012; Diehl 2013; Pröpper 2013; Léon and Berndsen 2014) but it depends on the indirect connections, that arise when defaulted banks are forced to liquidate their portfolios.

[1] See: http://www.bankofengland.co.uk/financialstability/Documents/stresstesting/2017/keyelements.pdf

[2] US, Spain, Ireland and other Eurozone economies experienced credit and house price booms prior to the 2007-2009 financial crisis.

[3] For more details, see: the Land Registry: http://http://landregistry.data.gov.uk/app/ukhpi

In order to capture this layer of contagion, we embed our stress testing exercise into a bipartite network model composed of banks and assets. In such type of networks, nodes represent banks and assets, while links connect banks to the assets in their portfolios, with the caveat that the links are not allowed within banks and assets. As our workhorse model, we use the cascading default algorithm by Huang et al. 2013 that is simple enough to allow us to study the importance of portfolio overlaps as a contagion channel under different scenarios by varying the intensity of the initial shock and market liquidity. It is also in line with top-down stress testing practice (see for a review Henry and Kok 2013) to disentangle first-round and second-round effects in order quantify separately the contribution of the initial shock and contagion to systemic risk. Moreover, modelling contagion through portfolio overlaps has the advantage of using balance sheet data, providing an useful alternative to overcome the usual shortcomings in the availability of bank-by-bank bilateral exposures (Anand et al. 2015; Hałaj and Kok 2013; Mistrulli 2011; Upper 2011).

The literature on contagion via portfolio overlaps is sizeable and we therefore mention only some important recent contributions. Levy-Carciente et al. 2015 present a dynamic version of the Huang et al. 2013 model to assess the impact of the legal reforms of the banking sector on the asset allocation decisions in Venezuela for the period 1998-2013, detecting changes in the stability of the financial system. Caccioli et al. 2014 develop a network model to study contagion via portfolio overlaps, providing an analytical estimation of the stability region of the bank-asset system where no cascade failures occur. Caccioli et al. 2015 present a stress-testing framework for the Austrian banking system where the contagion mechanism is a two-way interaction between portfolio overlaps and direct interbank credit. Greenwood et al. 2015 propose a model in which fire sales propagate shocks across banks' balance sheets, providing macroprudential metrics based on the vulnerability and the contribution to systemic risk of each single financial institution. An empirical application of the model by Greenwood et al. 2015 to US banks has been done by Duarte and Eisenbach 2015.

The rest of the paper is organized as follows. In Section 1 we present the Huang et al. 2013 model and we show how it can be applied to a stress-test framework for the UK banking system. In Section 2 we present the dataset and we discuss the results of our stres-test exercise. In Section 4 we conclude and we drive the policy implications.

4.2 Model

We propose a simple model of financial contagion based on Huang et al. 2013 to design a consistent stress scenario for the UK banking system. Since the burst of the housing bubble represents a severe yet plausible adverse scenario, we are interested in assessing the resilience of UK banks to a shock to real estate loans by estimating the extent to which portfolio overlaps amplify losses and defaults. This approach allows us to go beyond the first-round effects of an initial shock to mortgage market

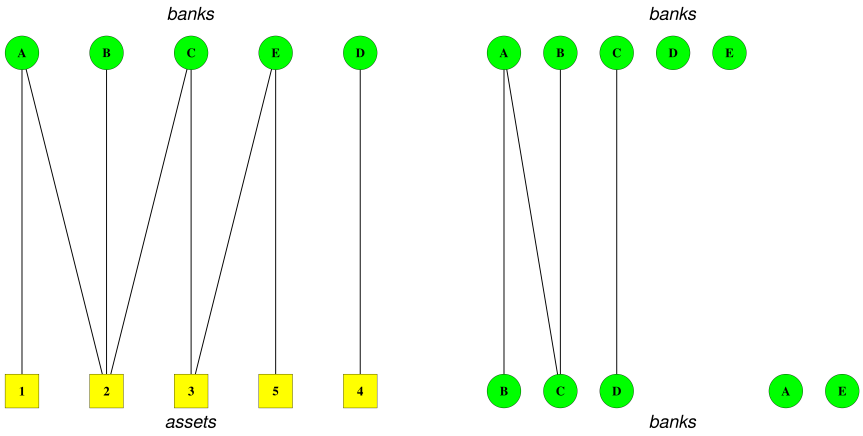


Figure 4.1: Left panel: a bipartite network of 5 banks and 5 assets. Banks (A,B,C,D,E) are denoted by green cycles and assets (1,2,3,4,5) by yellow squares. Edges indicate ownership of assets by banks. Right panel: the common exposure network extracted from the network of exposures in the left panel. Edges indicate overlapping exposures to the assets.

and to include the second-round effects throughout interbank contagion.

The stress test methodology can be summarized as follows. We consider a bipartite network of N banks (indexed by $i = 1, \ldots, N$) and M assets (indexed by $j = 1, \ldots, M$). Whenever a bank i holds an asset j in its portfolio, a link connects bank i to asset j. Links are only allowed between banks and assets, meaning that banks are never directly linked to other banks and assets are never directly linked to other assets. A graphical representation of a bipartite network of banks and assets is reported in Figure 4.1.

At time t, we populate the bank-asset system by banks' balance sheet data and each bank i is initially endowed with assets A_i^t, cash C_i and liabilities L_i. C_i and L_i are simply constant book-keeping entities, while the time-dependent variable A_i^t is the key variable in our contagion process via portfolio overlaps. The assets A_i^t of each bank i are allocated over M asset classes, whose value at time t is equal to:

$$A_i^t = \sum_{j=1}^{M} a_{i,j} p_j^t \tag{4.1}$$

where the constant $a_{i,j}$ is the number of shares of asset j held by bank i and p_j^t the price of asset j at time t. In our dynamics, the number of shares is assumed constant and banks sell assets only when they go bankrupt and they are forced to liquidate their portfolios. This allows us to isolate the impact of fire-sales on asset prices, by excluding asset sales for regular portfolio adjustments.

At time $t+1$, the stress scenario starts and an exogeneous shock $s \in [0,1]$ is applied to real estate loans. Let the asset class j be the real estate loans. If bank i holds $a_{i,j}$ amount of real estate loans, then the value of its assets evolves from time t to time $t+1$ as follows:

$$A_i^{t+1} = A_i^t - s \cdot a_{i,j} p_j^t \tag{4.2}$$

If the equity E_i^{t+1} becomes negative, then bank i defaults. The default condition is therefore formally expressed as:

$$\begin{aligned} &E_i^{t+1} < 0 \\ &E_i^{t+1} = A_i^{t+1} + C_i - L_i \end{aligned} \tag{4.3}$$

When a bank defaults, we assume that its assets are immediately liquidated. Fire-sales of defaulted banks induce a decrease in the value of their assets and thus a reduction of equity of the banks holding the same assets in their portfolios. Let us formalize the mechanism. If bank i defaults at time $t+1$, all its assets A_i^{t+1} are entirely liquidated. Let $a_{i,j'}$ be the amount of asset j' that has been liquidated ($j \neq j'$). At time $t+2$, its value $V_{j'}^{t+2}$ is reduced to:

$$\begin{aligned} &V_{j'}^{t+2} = V_{j'}^{t+1} - \alpha \cdot a_{i,j'} p_{j'}^{t+1} \qquad\qquad \forall j' : E_i^{t+1} < 0 \\ &V_{j'}^{t+2} = \sum_{i=1}^{N} a_{i,j'} p_{j'}^{t+2} \end{aligned} \tag{4.4}$$

where the parameter $\alpha \in [0,1]$ is the market illiquidity, namely the difficulty of the market to sell an asset without causing a drastic change in the price.[4]

Eq. (4.4) should be interpreted as follows: fire-sale of an asset j' at time $t+1$ generates a devaluation at time $t+2$, that is proportional to the amount sold $a_{i,j'}$ with a positive multiplier α, that amplifies the asset devaluation as markets become less liquid. In other words, we assume that assets are devalued according to a simple linear price impact function with market liquidity being a constant parameter. In our fire-sale scenario, therefore, only the amount sold determines how much the price will fall and the heterogeneity in the liquidity of asset classes is not taken into account.[5]

[4] If N banks default, Eq. (4.4) can be written as follows: $V_{j'}^{t+2} = V_{j'}^{t+1} - \sum_{i=1}^{N} \alpha \cdot a_{i,j'} p_{j'}^{t+1}$

[5] It is certainly too simplistic to assume that all assets have identical liquidity. In future work, we plan to relax this strong assumption and to introduce heterogeneity in the liquidity of asset classes.

At time $t+2$, contagion *via* the devaluation of asset j' propagates to, say, bank k according to:

$$A_k^{t+2} = A_k^{t+1} - \left(1 - \frac{V_{j'}^{t+2}}{V_{j'}^{t+1}}\right) a_{k,j'} \tag{4.5}$$

where $a_{k,j'}$ is the number of shares of asset j' hold by bank k. Eq. (4.4) describes how contagion spreads across banks throughout the exposure to the same assets under the market-price channel. From the point of view of contagion dynamics, this is a mechanistic model of loss-redistribution, where behavioral rules are missing and banks record their losses without any adjustment strategies.

The contagion process stops once all the remaining banks are solvent ($E > 0$) after the last default or when all banks have defaulted. A schematic representation of the contagion process is reported in Figure 4.2.

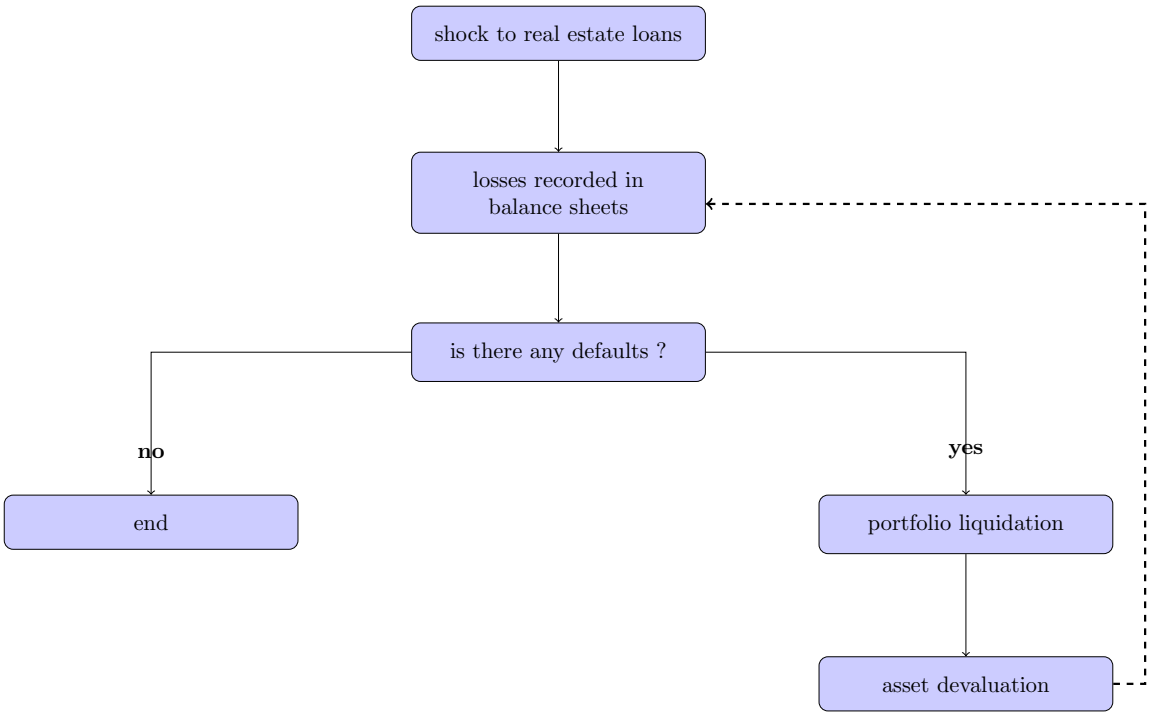


Figure 4.2: The dynamics of the stress-test model in stilized form.

4.3 Stress test results

For our stress test exercise, we use BankScope balance sheet data for 214 UK financial institutions[6] (depositories, broker-dealers, insurance companies, non-depositary institutions and real estate companies) at the end of 2015, which contains information on: total assets, total liabilities and total equity. The assets are divided into 10 classes: residential mortgage loans, corporate and commercial loans, other loans, consumer and retail loans, loans to public institutions, held-to-maturity securities, available for sale securities, other securities, fixed assets and total debt securities.

Table 4.1 and Table 4.2 report over the period 2007-2015 the size (in terms of market share) and the Herfindhal index respectively, providing information on the systemic importance of an asset class and its critical function for the stability of the financial system.[7] We see from Table 4.1 that mortgage loans represent the largest asset class over the whole period 2007-2015. After a stagnation during the 2007–2009 crisis, mortgage loans steadily increase and they account for around 30% of the market value of the UK financial system after 2013. Figure 4.3 plots the evolution of house prices, that is coherent with the upward trend of mortgage loans after 2013, underlining the crucial importance in assessing the impact of a devaluation of mortgage loans in times of house price boom.

With regard to asset concentration, we observe an Herfindhal index in range of 0.1 to 0.5 for the 10 asset classes over the whole period 2007-2015 (see Table 4.2), meaning that the 10 asset classes are evenly distributed across all the banks in the system. This clearly shows that balance sheets are highly interconnected and banks have strong overlapping exposures, demonstrating that portfolio overlaps are relevant channels of contagion when a bank defaults and liquidates its assets in portfolio. Therefore, our data seems to confirm that mortgage loans are systemically important and overlapping portfolios play a key role in contagion, adding further evidence that our stress-test methodology is able to capture a realistic scenario for the UK financial sysytem.

In the following, we report the results of our stress-test exercise for the UK banking system at the end of 2015. We begin by comparing the results of the 2015 stress test with the 2009 stress test in order to emphasize that the vulnerability of the system to the real estate sector has increased after the mid-2013's housing boom. We then change the initial shock of the 2015 stress test by imposing a devaluation to other asset classes to test whether a deterioration of real estate loans significantly contributes to systemic risk. Figure 4.4 and Figure 4.5 plot the fraction of defaults for the year 2009 and 2015 respectively by varying the level of s and α for the four largest asset classes: residential mortgage loans, other loans, other securities and debt securities.

[6] Hereafter, referred to as banks.

[7] The Herfindahl index (Hi) is defined as the sum of the squares of the market shares of all the banks in the system, where the market shares are expressed as fractions. As a general rule, an Hi below 0.2 signals a lowly concentrated asset, while an index between 0.2 and 0.5 signals a moderately concentrated asset. For an Hi higher than 0.5, an asset is considered an highly concentrated asset.

size of the asset class

asset type	index	balance sheet item	2007	2008	2009	2010	2011	2012	2013	2014	2015
real estate loans	0	residential mortgage loans	0.194	0.167	0.159	0.227	0.263	0.275	0.293	0.325	0.325
other loans	1	corporate and commercial loans	0.105	0.110	0.139	0.131	0.125	0.131	0.155	0.165	0.105
	2	other loans	0.154	0.222	0.128	0.196	0.111	0.134	0.084	0.086	0.151
	3	consumer and retail loans	0.092	0.038	0.035	0.036	0.036	0.031	0.030	0.030	0.037
	4	loans to public institutions	0.002	0.002	0.003	0.002	0.085	0.007	0.003	0.002	0.002
other assets	5	held-to-maturity securities	0.084	0.085	0.098	0.035	0.037	0.038	0.037	0.031	0.049
	6	available for sale securities	0.074	0.074	0.104	0.087	0.078	0.067	0.068	0.069	0.071
	7	other securities	0.140	0.106	0.154	0.138	0.119	0.197	0.158	0.122	0.157
	8	fixed assets	0.005	0.004	0.004	0.005	0.005	0.004	0.004	0.004	0.004
	9	total debt securities	0.124	0.164	0.143	0.123	0.122	0.115	0.123	0.113	0.084

Table 4.1: Size (in terms of market share) by asset class, period 2007-2015. Source: BankScope. Note that only residential mortagage loans are recorded in BankScope for the category real estate loans, i.e. no loans for constructions or commercial mortgage loans are available.

Herfindhal index of the asset class

asset type	index	balance sheet item	2007	2008	2009	2010	2011	2012	2013	2014	2015
real estate loans	0	residential mortgage loans	0.145	0.142	0.131	0.132	0.147	0.151	0.130	0.108	0.106
other loans	1	corporate and commercial loans	0.184	0.244	0.215	0.198	0.164	0.165	0.153	0.148	0.142
	2	other loans	0.237	0.245	0.145	0.195	0.182	0.200	0.131	0.099	0.098
	3	consumer and retail loans	0.297	0.200	0.190	0.163	0.152	0.141	0.150	0.145	0.148
	4	loans to public institutions	0.392	0.385	0.426	0.381	0.415	0.490	0.315	0.308	0.229
other assets	5	held-to-maturity securities	0.500	0.501	0.551	0.478	0.523	0.529	0.518	0.468	0.456
	6	available for sale securities	0.124	0.128	0.128	0.148	0.169	0.179	0.205	0.192	0.191
	7	other securities	0.116	0.116	0.116	0.116	0.116	0.116	0.116	0.116	0.118
	8	fixed assets	0.196	0.197	0.185	0.160	0.146	0.126	0.117	0.101	0.102
	9	total debt securities	0.235	0.255	0.320	0.237	0.258	0.272	0.261	0.234	0.195

Table 4.2: Herfindhal index by asset class, period 2007-2015. Source: BankScope. Note that only residential mortagage loans are recorded in BankScope for the category real estate loans, i.e. no loans for constructions or commercial mortgage loans are available.

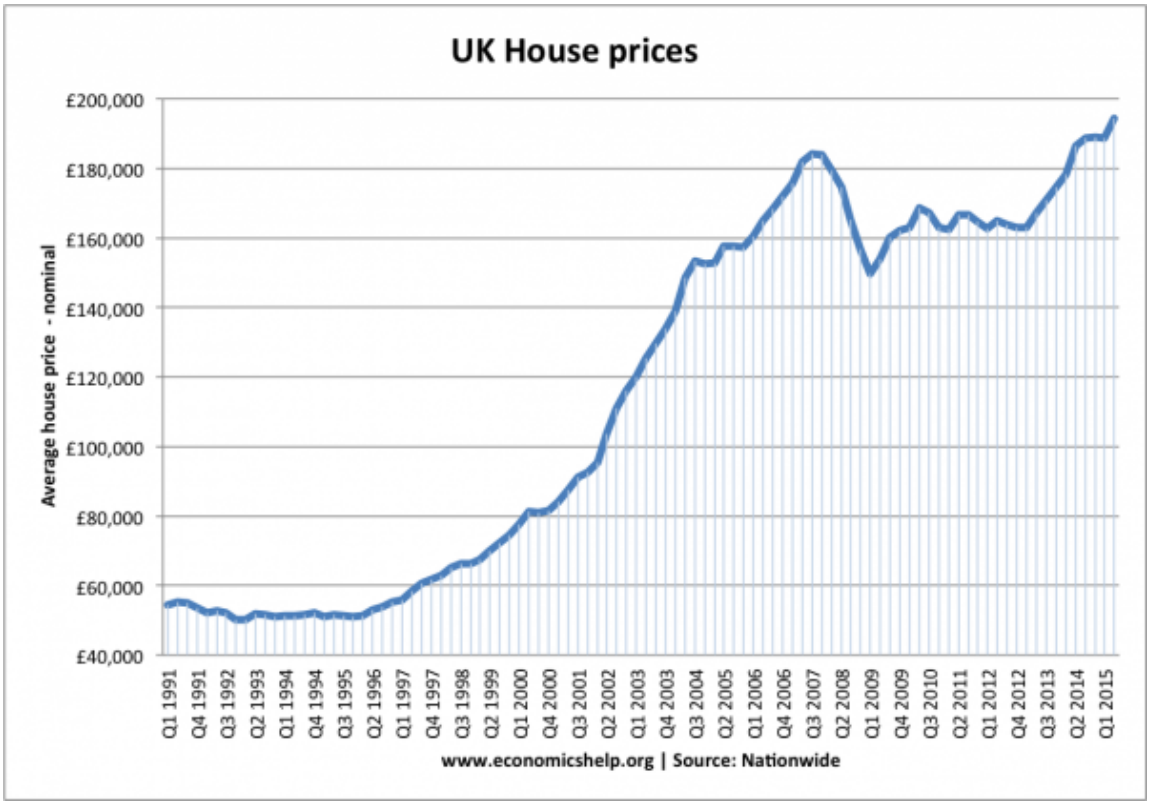
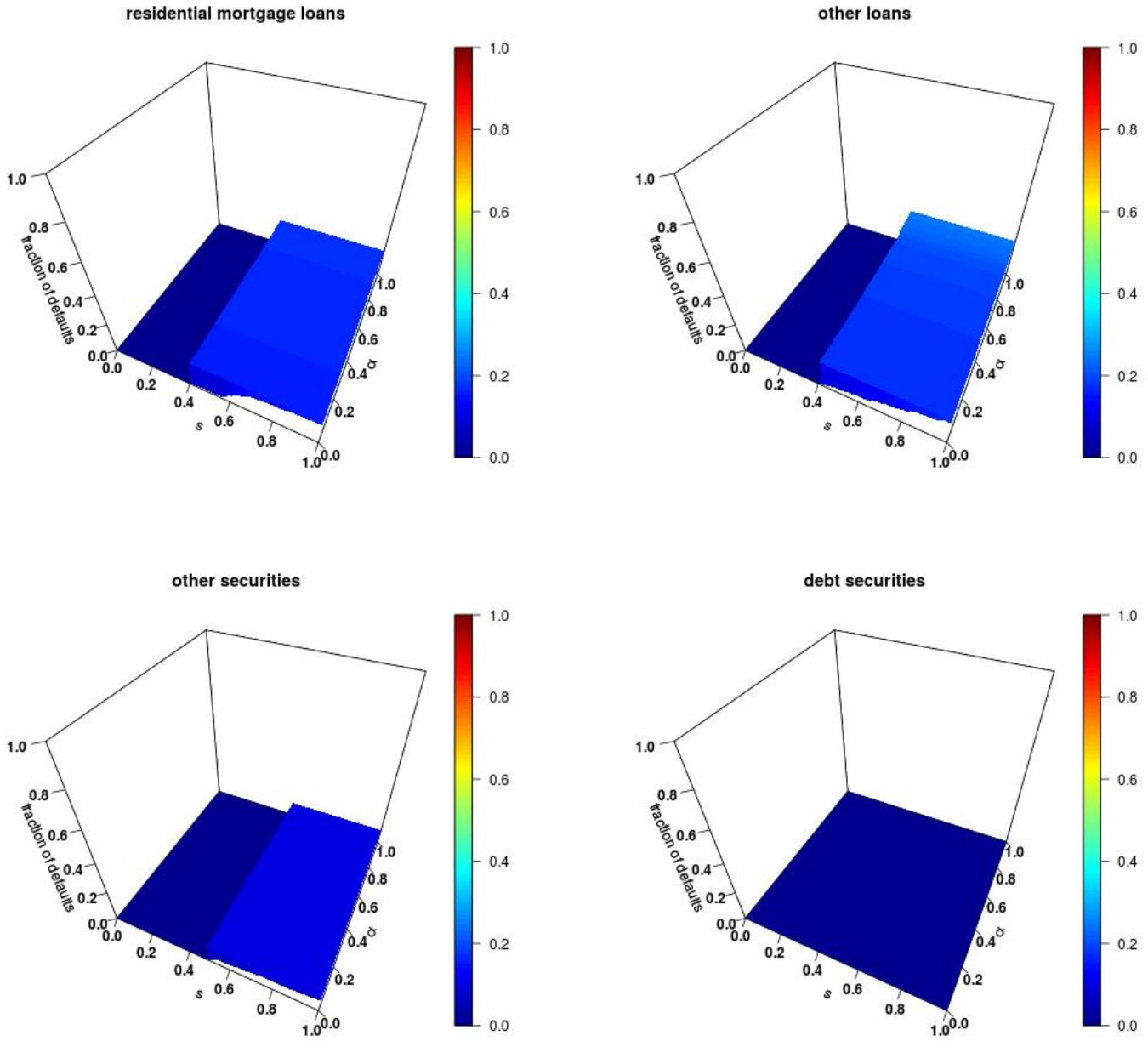


Figure 4.3: Average house price in United Kingdom, period 1991-2015. Source: Natiowide.

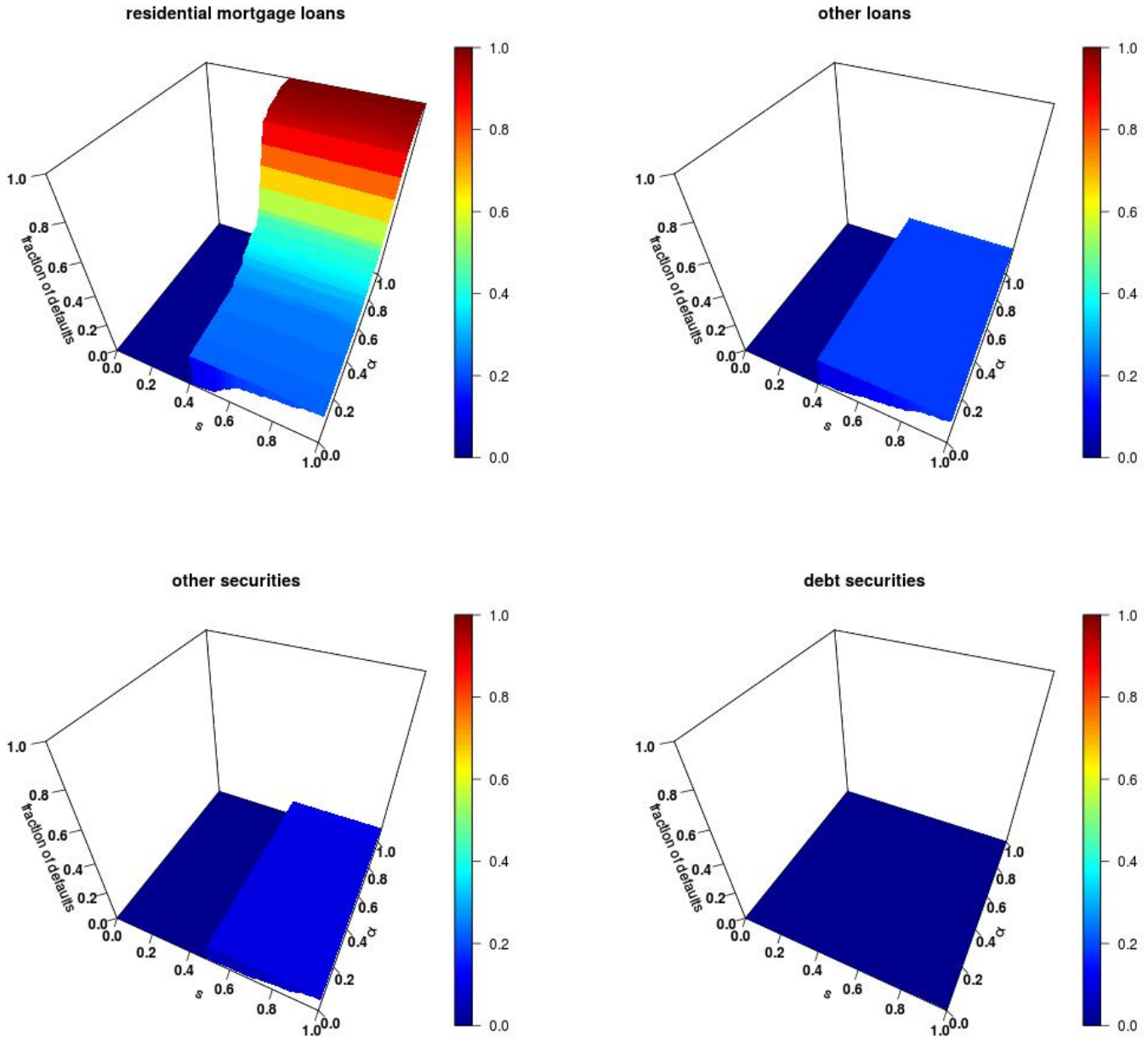
The most evident observation is that when residential mortgage loans are shocked in 2009 (before the mid-2013's housing boom), the number of defaults is below 20% for any level of s and α (upper-left panel of Figure 4.4). By contrast, when we shock residential mortgage loans in 2015 (after the mid-2013's housing boom), the number of defaults is highly sensitive to s and α. As we can see from the upper-left panel of Figure 4.5, no defaults occur for a shock to residential mortgage loans lower than 0.42. The fraction of defaults increases with α implying a strong sensitivity of the fragility of the system to the increase of the illiquidity. Indeed, defaults are not correlated with the initial shock s and market conditions are therefore the driver of the cascade size as long as $s \geq 0.42$. This is because a shock to residential mortgage loans causes a small number of defaults at every level $s \geq 0.42$, that triggers a large number of defaults (i.e. higher than 80%) in the second round only when markets are stressed ($\alpha > 0.6$).

This is confirmed in Figure 4.6, where we observe that the number of defaults in the first round represents only a small number for increasing levels of the initial shock s and the second round defaults account for a significant percentage of defaults when $\alpha > 0.6$. Specifically, Figure 4.6



2009

Figure 4.4: Before the mid-2013's housing boom. Fraction of defaults as a function of the market illiquidity α and the initial shock s to the following asset classes: residential mortgage loans (upper-left), other loans (upper-right), other securities (lower-left) and debt securities (lower-right).



2015

Figure 4.5: After the mid-2013's housing boom. Fraction of defaults as a function of the market illiquidity α and the initial shock s to the following asset classes: residential mortgage loans (upper-left), other loans (upper-right), other securities (lower-left) and debt securities (lower-right).

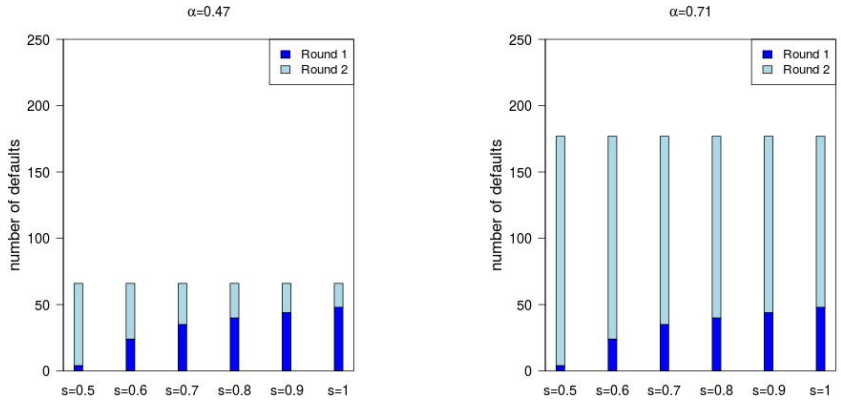


Figure 4.6: Decomposition of the first and the second round defaults in 2015 with $\alpha = 0.47$ (left) and $\alpha = 0.71$ (right). The initial shock s to residential mortgage loans ranges from 0.5 to 1 (no defaults occur for $s < 0.42$).

displays the decomposition of the first round and the second round defaults in 2015 for $\alpha = 0.47$ (left) and $\alpha = 0.71$ (right), which correspond to the minimum α^* such that the rate of defaults is greater than or equal to 30% and 80% respectively. It is clear from Figure 4.6 that the 2015 stress testing results can be misleading if we do not take into account portfolio overlaps as a source of contagion when markets are illiquid ($\alpha > 0.6$). When $\alpha = 0.71$ (right panel), only 4 banks default in the first round, while 173 banks default in the second round for an initial shock $s = 0.5$ to residential mortgage loans.

With respect to the other asset classes (i.e. other loans, other securities and debt securities), in both 2009 and 2015 we observe a number of defaults below 20% regardless of s and α (see Figure 4.4 and Figure 4.5 respectively), confirming that the systemic importance of mortgage loans changes significantly between 2009 and 2015 in comparison to the other asset classes.

Finally, we provide further evidence that the 2015 stress testing results can be inaccurate if we do not consider the interconnectedness between banks' portfolios by showing that bank-level data are not able to assess the systemically important financial institutions in stressed liquidity conditions. In Figure 4.7 we examine whether bank-specific features, such as the weight of real estate loans in portfolio and the equity-to-asset ratio, are informative on which banks default when $\alpha = 0.71$. The top panel of Figure 4.7 shows the complementary cumulative density function (CCDF) of the weight of real estate loans in portfolio for banks defaulted in the first round (top-left) and in the second round (top-right). Banks defaulted in the first round tend to have the portfolio concentrated in real estate loans, while banks defaulted in the second round do not report large exposures to real estate loans. As we can see in the top-right panel, around 50% of banks defaulted in the second round

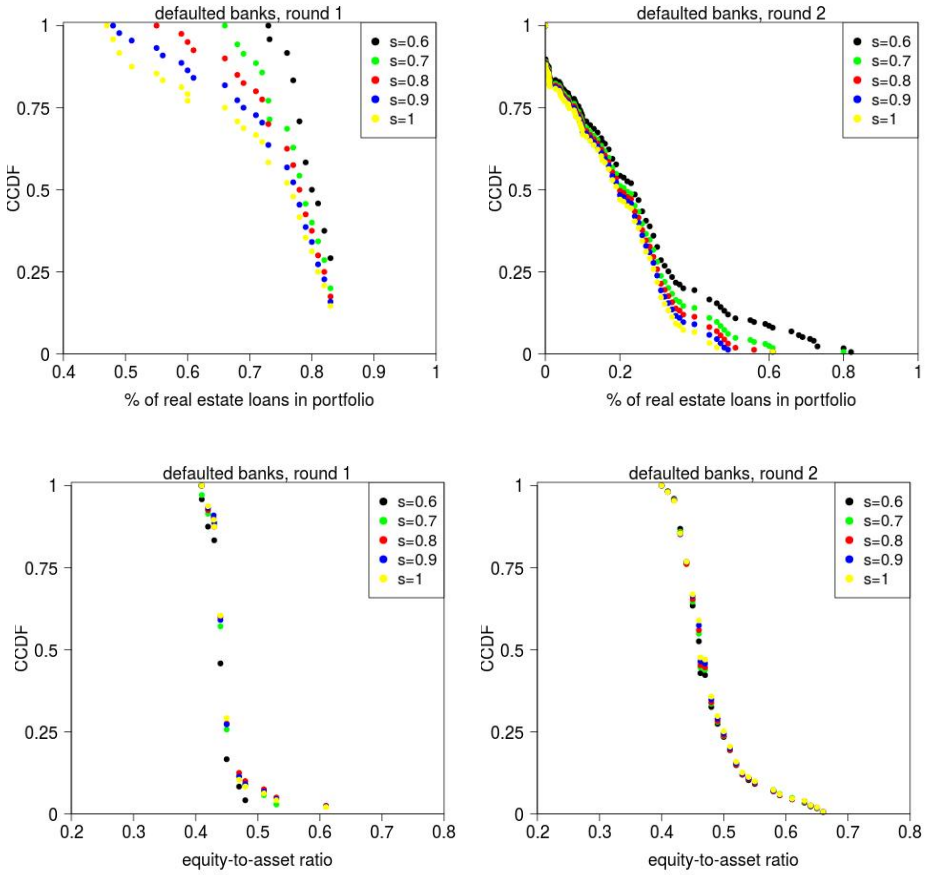


Figure 4.7: Bank-specific features of the defaults in round 1 and in round 2 in 2015 for $\alpha = 0.71$ and s ranges from 0.6 to 1 (no defaults occur for $s < 0.42$). Complementary cumulative density function (CCDF) of portfolio's weight of real estate loans (top) and equity-to-asset ratio (bottom).

hold less than 20% of real estate loans in their portfolios for any level s of a shock to real estate loans. Therefore, it is evident that portfolio concentration in real estate loans is not an adequate indicator of systemic importance and bank interconnectedness needs also to be taken into account when markets are illiquid.

This is confirmed in the bottom panel of Figure 4.7, where the complementary cumulative density function (CCDF) of the equity-to-asset ratio for banks defaulted in the first round (bottom-left) and in the second round (bottom-right) is plotted. In both round 1 and round 2 almost the totality of defaulted banks have a equity-to-asset ratio higher than 40%, meaning that: controlling for the level of equity-to-asset ratio, portfolio concentration in real estate loans determines the defaults in

the first round, while portfolio interconnections determine the defaults in the second round.

4.4 Conclusions and policy implications

In this paper, we present a stress test framework for the UK banking system. In the context of the recent housing bubble, this work contributes to the ongoing debate on the risks of a drop in house prices for the UK economy.

We focus on the role of portfolio overlaps in amplifying an initial shock to real estate loans under different stress scenarios. We find that the intensity of the initial shock does not matter in contrast with what one could think ex ante and market conditions are the main drivers of systemic cascades. When markets are illiquid, the number of defaults in the second round are more than three times higher than in the first round, pointing out the importance of interconnections among banks investing in the same assets. The inadequacy of *standing-alone* balance sheet indicators in predicting systemic defaults provides further support that stress test methodologies should be adapted to incorporate the complexity of the financial system.

The present work aims to provide a basic stress-testing framework and can be extended in several directions. First, one could introduce heterogeneity in the liquidity of assets classes. Indeed, the liquidity actually varies strongly across assets (i.e. available for sale securities are considered to be a more liquid asset class than mortgage loans) and our framework does not account for this heterogeneity. Second, one could model the asset devaluation endogeneously by defining a different price impact function for each asset class. Finally, another open question in our stress-testing exercise concerns the optimal asset allocation which minimizes contagion due to portfolio overlaps for the design of ad hoc macroprudential requirements.

Bibliography

Allen, F. and Gale, D., 2000. Financial contagion. *Journal of political economy* 108(1), 1-33

Anand, K., Craig, B. and Von Peter, G., 2015. Filling in the blanks: Network structure and interbank contagion. *Quantitative Finance*, 15(4), 625-636

Baptista, R., Farmer J.D, Hinterschweiger, M., Low, K., Tang D. and Uluc A., 2016. Macroprudential policy in an agent-based model of the UK housing market. Bank of England Working Paper

Battiston S., Delli Gatti D., Gallegati M., Greenwald B.C. and Stiglitz J.E., 2012. Default cascades: When does risk diversification increase stability? *Journal of Financial Stability* 8(3), 138-149

Bracke, P., 2015. How much do investors pay for houses?. Bank of England Working Paper

Caccioli, F., Shrestha, M., Moore, C. and Farmer, J. D., 2014. Stability analysis of financial contagion due to overlapping portfolios. *Journal of Banking & Finance* 46, 233-245

Caccioli, F., Farmer, J. D., Foti, N. and Rockmore, D., 2015. Overlapping portfolios, contagion, and financial stability. *Journal of Economic Dynamics and Control* 51, 5063

Cho, M., 1996. House price dynamics: A survey of theoretical and empirical issues. *Journal of Housing Research*, 145-172

Constâncio, V., 2015. The role of stress testing in supervision and macroprudential policy. Keynote speech at the London School of Economics Conference on "Stress Testing and Macroprudential Regulation: a Trans-Atlantic Assessment". London, 29 October 2015.

Diehl, M., 2013. Measuring free riding in large-value payment systems: the case of TARGET2. *The Journal of Financial Market Infrastructures*, 1(3), 31-53

Duarte, F. and Eisenbach, T. M., 2015. Fire-sale spillovers and systemic risk. Federal Reserve Bank of New York Working Paper

Gai, P., Haldane, A. and Kapadia, S., 2011. Complexity, concentration and contagion. *Journal of Monetary Economics* 58(5), 453-470

Galbiati, M. and Soramäki K., 2011. An agent-based model of payment systems. *Journal of Economic Dynamics and Control* 35(6), 859-875

Georg, C. P., 2013. The effect of the interbank network structure on contagion and common shocks.*Journal of Banking & Finance* 37(7), 2216-2228

Greenwood, R., Landier, A. and Thesmar, D., 2015. Vulnerable banks. *Journal of Financial Economics* 115(3), 471-485.

Hałaj, G. and Kok, C., 2013. Assessing interbank contagion using simulated networks. *Computational Management Science* 10(2-3), 157-186.

Henry, J. and Kok, C., 2013. A macro stress testing framework for assessing systemic risk in the banking sector. ECB Occasional Paper No. 152

Huang, X., Vodenska, I., Havlin, S. and Stanley, H. E., 2013. Cascading failures in bi-partite graphs: model for systemic risk propagation. *Scientific Report* 3, 1219-1229

Léon, C. and Berndsen, R. J., 2014. Rethinking financial stability: challenges arising from financial networks' modular scale-free architecture.*Journal of Financial Stability* 15, 241-256

Levy-Carciente, S., Kenett, D. Y., Avakian, A., Stanley, H. E. and Havlin, S., 2015. Dynamical macroprudential stress testing using network theory. *Journal of Banking & Finance* 59, 164-181

Mayer, C., Pence, K. and Sherlund M., 2009. The rise in mortgage default. *The Journal of Economic Perspectives* 23.1, 27-50

Markose, S., Giansante, S. and Shaghaghi, A. R., 2012. Too interconnected to fail financial network of US CDS market: topological fragility and systemic risk. *Journal of Economic Behavior & Organization* 83(3), 627-646

Mistrulli, P. E., 2011. Assessing financial contagion in the interbank market: Maximum entropy versus observed interbank lending patterns. *Journal of Banking & Finance*, 35(5), 1114-1127

Peltonen, T. A., Scheicher, M. and Vuillemey, G., 2014. The network structure of the CDS market and its determinants.*Journal of Financial Stability*, 13, 118-133

Puliga, M., Caldarelli, G.and Battiston, S., 2014. Credit default swaps networks and systemic risk. *Scientific reports*, 4

Shleifer, A. and Vishny, R., 2011. Fire sales in finance and macroeconomics. *The Journal of Economic Perspectives* 25(1), 29-48.

Upper, C., 2011. Simulation methods to assess the danger of contagion in interbank markets. *Journal of Financial Stability* 7(3), 111-125

Chapter 5

Conclusions

Viewing the financial system as a complex network may seem, at least initially, a fuzzy approach to provide guidance to regulators and policymakers in defining a macroprudential policy agenda. In this thesis, we have attempted to show that this perspective can indeed offer an intuitive understanding of how the network structure affects systemic risk and monetary policy decisions. Looking at the financial system from this vantage point, we actually emphasize the distinction between microprudential and macroprudential policy. While the first targets the stability and soundness of individual financial institution, the second safeguards the stability of the financial system as a whole by looking at the interactions and feedback in the financial sector and with the rest of the economy. By going beyond a micro approach, we assess the health of a financial institution in connection with the other parts of the financial system and with the real economy by addressing a number of important questions.

In the first paper **The topology of the bank-firm credit network in Spain, 1997-2007**, we consider the indirect channel of contagion that arises when banks provide credit to the same firms. We investigate the relationship between the structure of bank-firm networks and the vulnerability of the financial system to a shock in the real sector. As a case study, we use the Spanish bank-firm system in the period 1997-2007. With this approach we have been able to show the existence of a stable core of the credit market, that involves an over-expression of a group of firms that is jointly financed. The robustness of our result has been checked against standard random networks and the bank-firm degree distributions. Although a significant and persistent core has emerged, the economic reason why the banking sector organizes itself around a lobby group is not clear. Do large banks influence small banks in selecting firms ? Or do large banks influence each other ? Do banks have any economic incentives that encourage to finance the same firms ? In order to answer to

these questions, an economic model is therefore required in order to bridge the emprical evidence with economic theory.

In the second paper **The controllability of the repo market in Mexico**, we study the relationship between the structure of interbank networks and the effectiveness of monetary policies. Our investigation is performed with tools of control theory, where we detect the banks that control the provision of the intraday liquidity by using an algorithm, that takes into account recursively of all the paths in the network. As an application, we analyze a comprehensive dataset on repo transactions in the Mexican banking system during 2005-2014. We show that a group of banks constantly controls the interbank liquidity, which does not coincide with the most connected banks or the largest lenders. To go beyond these suggestive results, advances in the research on the control of interbank networks are actually needed. For instance, a promising agenda is to study which network architecture can make the interbank market more controllable. Once one understands the impact of the network structure on the control, a more ambitious challenge is certainly to take into account the amount of the loans, without stopping at the paths of the lending relationships.

In the last paper **Stress-testing the UK banking system: a network approach to cope with portfolio overlaps** we carry out a stress test on UK financial institutions. Given the negative outlook for 2017 real estate market, we check the resilience of the UK banking system to a decrease in the value of residential mortgage loans by exploring how contagion spreads via portfolio overlaps. Indeed, recent studies as well as empirical evidences, such as the dot-com bubble or the subprime mortgage crisis, suggest that the main contribution to systemic risk comes actually from the exposure to common assets. Overall, our results confirm an increase in the vulnerability to the real estate sector in the period 2007-2015, providing evidence that portfolio overalaps is an important channel of contagion. The mechanics of distress propogation that we introduce is very simple: when a bank defaults and liquidates its assets, distress propagates to the banks holding the same assets, which in turn suffer losses and so on. Therefore, establishing a contagion mechanism for distressed-but-non-defaulting institutions opens several possible directions for future research. Moreover, here we have employed an elementary representation of the balance sheets (i.e. assets are divided into only 10 classes) and we obviously improove the accuracy of our results by using more granular balance sheet data. Hence, another future extension of the framework could be based on extracting a higher number of asset classes with a close calibration of the asset liquidity.