MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Company induced systemic risk in the Austrian economy"

verfasst von / submitted by

Abraham Hinteregger, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2017 / Vienna, 2017

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 876
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Physik
Betreut von / Supervisor:	Univ.-Prof. Mag. DDr. Stefan Thurner
Mitbetreut von / Co-Supervisor:	Univ.-Prof. Dr. Christoph Dellago

Erklärung zur Verfassung der Arbeit

Abraham Hinteregger, BSc
Harmoniegasse 1/4, 1090 Wien

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 1. Oktober 2017

Abraham Hinteregger

Danksagung

Als erstes möchte ich Stefan Thurner und Sebastian Poledna für ihre Hilfe und zahlreichen Ratschläge bei der Erstellung dieser Arbeit danken. Vielen Dank außerdem auch den anderen Mitgliedern des Instituts für Wissenschaft Komplexer Systeme der Medizinischen Universität Wien sowie des Complexity Science Hubs in Wien für ihre vielen Hilfeleistungen, Beratungen und interessanten Vorschläge und Vorträge. Weiters bin ich Christian Feuersänger dankbar für TikZ. Besonderer Dank gilt auch meiner Mutter, meinen Schwestern und meinen anderen Verwandten für ihre Unterstützung und Geduld. Vielen Dank auch an Nina, für die Versorgung mit gesundem Essen, die Gesellschaft beim Burgerbro und für viele andere Dinge.

Acknowledgements

I would first like to thank Stefan Thurner and Sebastian Poledna for their help, guidance and remarks during the time I worked on this thesis. Additionally, I would also like to thank the other members of the Institute for the Science of Complex Systems of the Medical University of Vienna and the Complexity Science Hub Vienna for their helpful and often interesting input. I would also like to thank Christian Feuersänger for making TikZ as great as it is. Furthermore, I want to express my sincere gratitude for the support and patience of my mother, sisters, Nina and my extended family. Apart from that, I would like to apologize to Jakob and Daniel for not showing up more often for a sweet bike ride in the last few months.

Contents

Part I

Econophysics and generalized eigenvector analysis of networks

CHAPTER 1

Introduction

Econophysics is an interdisciplinary area of research that applies methodical approaches from statistical mechanics to economic problems. One of these problems is the assessment of systemic risk in financial networks, which, in the wake of the financial crisis of 2007/08, has been extensively studied in recent years.

As the risk factors that lead to a market collapse are not constrained to only affect financial institutions, the assumption of this work is that private companies may also have a significant contribution to the systemic risk in the economy.

The goal is therefore to identify companies that would have the biggest impact on the economy in the case of default. For this purpose, we construct a liability network of the Austrian economy for the year 2008 by extracting the relevant information from empirical data provided to the commercial register. This network is then analyzed with the DebtRank method proposed by Battiston et al. [BPK+12] in 2012 that is frequently used to quantify systemic risk in financial networks.

1.1 Outline

In the first part we will introduce the concepts needed to understand DebtRank, a method to assess systemic risk based on node centrality for graphs. For this purpose, we will provide a brief overview of elementary graph theory as well as centrality measures that are conceptually similar to the DebtRank method in chapter 2. In chapter 3, applications of the graph theoretic concepts for econophysics will be discussed. Furthermore, we will present an in-depth explanation of the DebtRank method. In the second part of this thesis we will report how we employed the methods covered in the first part to construct a liability network for the Austrian economy using empirical data. Subsequently, we will quantify the contribution of private companies to the systemic risk of the economy using the DebtRank method and present the results in chapter 6. In the final chapter we will summarize and discuss the results.

CHAPTER 2

Graph Theory

This chapter starts with a brief introduction to basic graph theory. Subsequently, various measures of node centrality are introduced. These measures provide the foundation of the DebtRank method—the main measure of interest for this work—which will be explained in section 3.3.

2.1 Basic definitions

A graph is a structure that consists of a set of objects with a set of connections between them. Unless otherwise specified, all graphs in this thesis have directed edges.

Figure 2.1: Undirected graph on the left, directed graph on the right. Note that each vertex may be connected to zero, one or many other vertices.

Definition 2.1 (Undirected graph) *A set of vertices V connected by a set of edges E (also called arcs) form an undirected graph $G = (V, E)$. Each edge is a set that contains the two vertices it connects.*

$$V = \{v_1, v_2, \ldots, v_n\} \qquad \textit{short:} \qquad V = \{1, 2, \ldots, n\}$$
$$E = \{\{v_{i_1}, v_{j_1}\}, \ldots, \{v_{i_m}, v_{j_m}\}\} \qquad \textit{short:} \qquad E = \{\{i_1, j_1\}, \ldots, \{i_m, j_m\}\}$$

Definition 2.2 (Directed graph) *In a directed graph an edge is not a set that contains the two vertices it connects but an ordered pair. The vertex set is defined the same way as for an undirected graph.*

$$E = \{(v_{i_1}, v_{j_1}), \ldots, (v_{i_m}, v_{j_m})\} \qquad \textit{short:} \qquad E = \{(i_1, j_1), \ldots, (i_m, j_m)\}$$

For a given edge $e = (v_i, v_j)$ in a directed graph we refer to the first (v_i) and second component (v_j) of the edge with e_1 and e_2. The first component is the start vertex of the edge, the second component is the end vertex of the edge.

Definition 2.3 (Multigraph) *In a multigraph the set of edges E is a multiset. Two vertices a and b may then be connected with zero, one or multiple edges.*

Definition 2.4 (Loop) *An edge that connects a vertex with itself is called a loop.*

Definition 2.5 (Simple graph) *A simple graph is a graph without loops and without multiple edges.*

Remark 2.1 *A loop in an undirected graph is not a set but a multiset as it contains the same vertex twice.*

Definition 2.6 (Weighted graph) *If each edge has an associated value, the resulting graph is called a weighted graph. A weighted graph comprises a graph*

$G = (V, E)$ as well as a weight function $w : E \to \mathbb{R}$ that maps each edge (v_i, v_j) to its associated weight W_{ij}.

$$w(v_i, v_j) = \begin{cases} W_{ij} & \text{if } v_i \text{ and } v_j \text{ are connected} \\ 0 & \text{if } v_i \text{ and } v_j \text{ are not connected} \end{cases}$$

Remark 2.2 *For a multi-graph the weight function must be adapted to enable edges that connect the same two vertices with different weights.*

Definition 2.7 (Bipartite graph) *A graph $G = (V, E)$ is a* bipartite graph *if and only if the set V can be partitioned into two disjunct sets V_1 and V_2 such that all edges $e = (v_1, v_2)$ connect a vertex from one set with a vertex from the other set.*

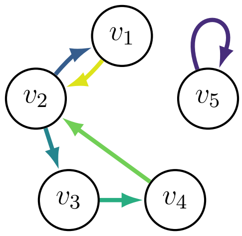
2.2 Graph representation

Until now, a graph was represented by a tuple of sets containing its elements. Another way to store this information is the adjacency matrix.

Definition 2.8 (Adjacency matrix) *A graph $G = (V, E)$ can be represented as a $N \times N$ matrix where $N = |V|$ is the number of vertices. This matrix is called the adjacency matrix. In a graph without multiple edges the adjacency matrix A has entries that are either 1 ($A_{ij} = 1 \Leftrightarrow (v_i, v_j) \in E$) or 0 ($A_{ij} = 0 \Leftrightarrow (v_i, v_j) \notin E$).*

Remark 2.3 *For undirected graphs the adjacency matrix is symmetric as for every pair of vertices v_i and v_j there is an edge from v_i if and only if there is an edge from v_j to v_j and thus $A_{ij} = A_{ji}$.*

Below is an example of a graph and its adjacency matrix representation. The entries in the adjacency matrix have the same color as the edge associated with the entry.



$$
A_{5\times 5} = \begin{array}{cc} & \begin{array}{ccccc} v_1 & v_2 & v_3 & v_4 & v_5 \end{array} \\ & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{array} \end{array}
$$

For weighted graphs the weighted adjacency matrix W results from multiplying each entry A_{ij} of the adjacency matrix with the weight of the corresponding edge, given by $w(v_i, v_j)$.

2.3 Paths on graphs

Paths on graphs were of interest since the very beginning of graph theory in the 18th century when it was introduced by Leonhard Euler in a paper about the "Seven Bridges of Königsberg"-problem [Eul35] which asks for the existence of a path that uses all edges exactly once (thus subsequently known as *Eulerian path*).

Definition 2.9 (Walk) *A walk on a graph is a sequence of edges where every two consecutive edges have a common vertex—the target of the first edge is the source of the second et cetera.*

Remark 2.4 (Number of walks of certain length) *The adjacency matrix A of a graph $G = (V, E)$ has entries $A_{ij} = 1$ for all pairs $(i, j) \in E$, i.e. for all pairs of vertices that are reachable with a path using only one edge. Taking the n-th power of the adjacency matrix yields a matrix A^n where the entries $(A^n)_{ij}$ correspond to the number of walks from v_i to v_j.*

Definition 2.10 (Path) A path *is a walk where no vertex is visited more than once, i.e. any vertex v_i occurs in at most two edges (one edge leading to v_i and one* *edge leading away from v_i)*[1].

Definition 2.11 (Connected vertices) *Two vertices are* connected *iff there exists a path from the first to the second vertex. This relationship is only necessarily symmetric for undirected graphs.*

Definition 2.12 (Connected component) *For an undirected graph $G = (V, E)$ a set of vertices $C \subseteq V$ is a connected component iff for every pair of nodes $x, y \in C$ there exists a path from x to y using edges from E.*

Definition 2.13 (Strongly connected component) *For a directed graph $G = (V, E)$ a strongly connected component is a set of vertices $C \subseteq V$ iff for every pair of vertices $x, y \in C$ there exists a path from x to y using edges from E.*

Definition 2.14 (Weakly connected component) *For a directed graph $G = (V, E)$ a set of vertices $C \subseteq V$ is a weakly connected component iff for every pair of vertices $x, y \in C$ there exists a path from x to y using edges from $E' = \{\{v_i, v_j\} | (v_i, v_j) \in E\}$ (E' is a set of undirected edges).*

Definition 2.15 (Distance) *For two vertices v_i and v_j in a graph $G = (V, E)$ the distance $d(v_i, v_j)$ is defined as the number of edges on the shortest path with v_i as the first and v_j as the last vertex.*

Remark 2.5 (Distance for directed graphs) *For a directed graph the distance usually takes the edge directions into account, therefore $d(v_i, v_j)$ is not necessarily equal to $d(v_j, v_i)$.*

Remark 2.6 (Distance for unconnected vertices) *If a graph consists of more than one connected component there are vertices that are not connected by a path. If v_i and v_j are such vertices, the distance is usually defined as $d(v_i, v_j) = \infty$.*

[1]This definition allows paths that start and end at the same vertex. These are called *cycles*.

2.4 Centrality measures

In a graph, some vertex or some vertices may be more important than others. For this purpose several centrality measures are used to highlight different properties of interest.

The most basic centrality measure is the degree centrality (section 2.4.1) which only considers local information. Closeness and betweenness centrality (sections 2.4.2 and 2.4.3) are path based measures. The last class of centrality measures outlined here are variants of the eigenvector centrality (sections 2.4.4, 2.4.5 and 3.3) where the centrality of a vertex depends on the centrality of the other vertices.

2.4.1 Degree centrality

Degree centrality is a very basic centrality measure that only counts the number of edges that are incident upon a vertex. In an undirected graph the degree centrality C_D for a vertex v_i is thus defined as

$$C_D(v_i) = |\{e \in E | v_i \in e\}| \ .$$

The degree centrality of a vertex is often called the degree of a vertex or $deg(v_i)$. In fig. 2.2 the degree centrality of a graph is illustrated.

Source: [Jac17, Lecture 16, 2.4]

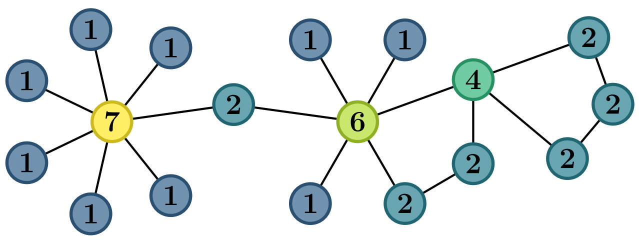


Figure 2.2: Each vertex labelled and colored according to its degree centrality.

For a directed graph, there exist two additional measures of degree centrality, one for the outgoing and one for the ingoing arcs. The sum of the in-degree and out-degree centrality is the degree centrality as defined before:

$$
\begin{aligned}
C_{Di}(v_i) &= |\{e \in E | v_i = e_2\}| \\
C_{Do}(v_i) &= |\{e \in E | v_i = e_1\}| \\
C_{D\ }(v_i) &= |\{e \in E | v_i = e_1 \lor v_i = e_2\}| \quad .
\end{aligned}
$$

In fig. 2.3 the degree centrality is illustrated for an undirected and for a directed graph.

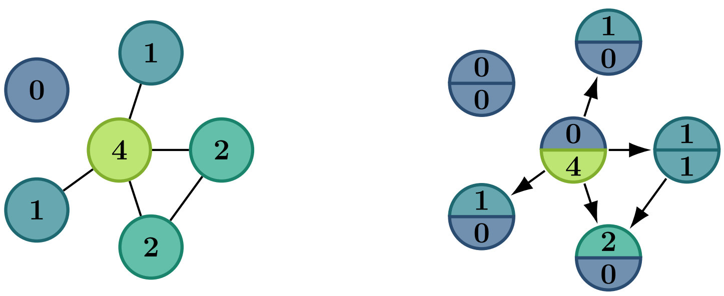


Figure 2.3: On the left is an undirected graph, on the right is the same graph with directed edges. Vertices of the undirected graph are labelled with degree centrality C_D; vertices of the directed graph are labelled with in-degree C_{Di} (top) and out-degree C_{Do} (bottom) centrality. Note that the sum $C_D = C_{Di} + C_{Do}$ for each vertex of the right graph is the value of the degree centrality of the corresponding vertex on the left.

Applications

Source: [Eul35]

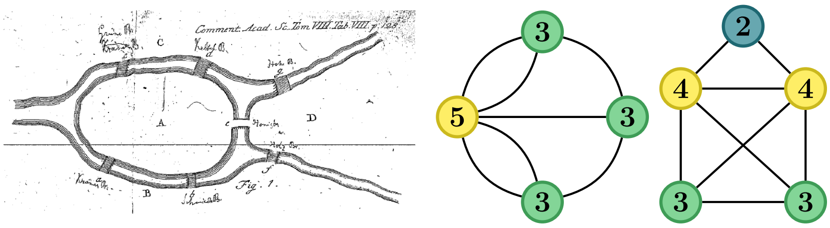


Figure 2.4: Left: Eulers illustration of the “Seven Bridges of Königsberg” problem. Center: a graph of the “Seven Bridges…” problem where each vertex is labelled with its degree centrality. Right: The “Haus des Nikolaus” riddle where only the two vertices with odd degree centrality can be used as start and end vertices to draw the figure in a single stroke.

In Leonhard Eulers paper on the “Seven Bridges of Königsberg” problem [Eul35], the fact that the graph had more than two vertices with odd degree centrality was used to prove that there is no Eulerian path that uses all edges in the graph. In fig. 2.4 the sketch from the original publication as well as a simplified version where vertices are labelled with their degree centralities are depicted. Since there are more than two vertices with odd degree there cannot exist a path that uses each bridge/edge as only the start and end vertex of the path may have an odd degree. A better-known version of this is the “Haus des Nikolaus” riddle where the two bottom vertices have an odd degree centrality and the other corners have even degree centrality (see fig. 2.4). It follows that each successful attempt has to start at one of the two bottom vertices and end at the other one.

2.4.2 Closeness centrality

The closeness centrality was introduced in 1950 by Alex Bavelas [Bav50] who analyzed different communication patterns in groups. He defined the “relative centrality" of a vertex v_i as

$$C_C(v_i) = \frac{\sum_{v_k \in V} \sum_{v_l \in V} d(v_k, v_l)}{2 \sum_{v_k} d(v_i, v_k)} \quad ,$$

where the denominator is the sum of the distances from v_i to all other vertices and the numerator is the sum of the distances for all vertices summed up. The factor two in the denominator is due to the fact that the double sum counts every distance twice, once as $d(v_k, v_l)$ and once as $d(v_l, v_k)$.

In an undirected network a high closeness centrality of a vertex corresponds to a short average path length to other vertices. In a communication network of politicians, this translates to a capacity for more efficient communication, which, according to Hafner-Burton et al. [HBM10], can be leveraged for more power and influence.

For directed graphs the closeness centrality may be either based on distance from or on distance to other vertices.

$$C_{Cf}(v_i) = \frac{\sum_{v_k \in V} \sum_{v_l \in V} d(v_k, v_l)}{2 \sum_{v_k} d(v_i, v_k)}$$

$$C_{Ct}(v_i) = \frac{\sum_{v_k \in V} \sum_{v_l \in V} d(v_k, v_l)}{2 \sum_{v_k} d(v_k, v_i)} \quad ,$$

An example for this would be a politician that can use public broadcasting to efficiently communicate his ideas to a population that can only respond by casting a vote once every few years (and, thanks to the wonders of our time, by ranting on social media).

Each of these measures must be adapted in some way for graphs that consist of more than one (strongly) connected component, as the distance of two vertices that are not connected is usually defined to be ∞ (see remark 2.6).

2.4.3 Betweenness centrality

Betweenness centrality is measures the importance of a vertex for connecting other vertices introduced by Freeman [Fre77]. For a vertex v_i of the graph $G = (V, E)$ it is defined as the sum of the number of shortest paths connecting all pairs of vertices in the set $V' = V \setminus \{v_i\}$ that pass through v_i (p_i) divided by the number of paths that do not pass through v_i (p):

$$C_B(v_i) = \sum_{v_k \in V'} \sum_{v_l \in V'} \frac{p_i(v_k, v_l)}{p(v_k, v_l)} \quad .$$

According to Hafner-Burton et al. [HBM10] betweenness centrality can be used for brokerage power.

2.4.4 Eigenvector centrality

The eigenvector x of a matrix A is defined by the formula

$$\lambda x = Ax, \tag{2.1}$$

where λ is the eigenvalue corresponding to the eigenvector x.

When the matrix A is the adjacency matrix (see definition 2.8) of the graph $G = (V, E)$ and λ is the largest real eigenvalue[2] of A, the i-th entry x_i of the eigenvector x corresponds to the eigenvector centrality $C_E(v_i)$ of the of the vertex v_i.

Reformulating eq. (2.1) yields

$$C_E(v_i) = \frac{1}{\lambda} \sum_j A_{ij} C_E(v_i) = \frac{1}{\lambda} \sum_{v_j \in N(v_i)} C_E(v_i) \ ,$$

where $N(v_i)$ is the set of neighbors of the vertex v_i. From this equation, the centrality notion that eigenvector centrality [Bon72] captures is more obvious:

[2]Using the largest real eigenvalue guarantees real and strictly positive eigenvector entries as asserted by the Perron–Frobenius theorem.

Being connected to vertices with high centrality leads to a higher centrality. This is illustrated in fig. 2.5.

Source: [Jac17, Lecture 16, 2.4]

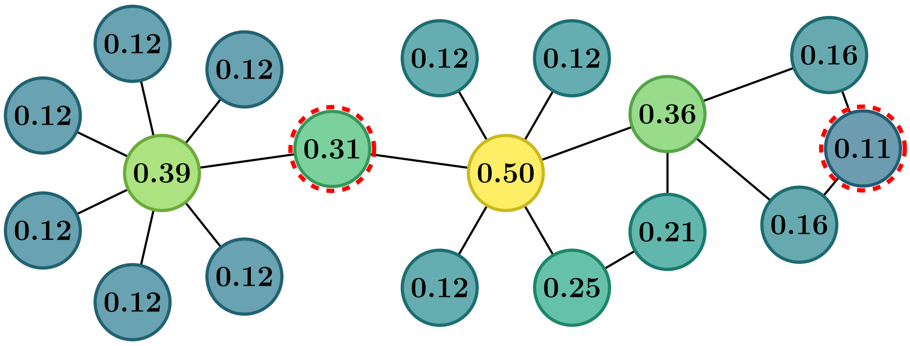


Figure 2.5: Vertices labelled with their eigenvector centrality. Note that the two vertices highlighted with a dashed red border both have a degree centrality of 2 but the left vertex has a considerably higher eigenvector centrality due to being connected to more central vertices.

2.4.5 Katz centrality

Whereas eigenvector centrality only takes the centrality of neighboring vertices into account to calculate the centrality of a vertex v_i, Katz centrality [Kat53] extends this by also including the centralities of vertices that are reachable from v_i with walks of arbitrary length.

Using the property from remark 2.4 Katz centrality is defined as

$$C_K(v_i) = \sum_{k=1}^{\infty} \sum_{v_j \in V} \alpha^k (A^k)_{ij} \tag{2.2}$$

where α is an attenuation factor that weights paths according to their length. This factor must be chosen to be smaller than the inverse of the absolute value of the largest eigenvalue of A to ensure that C_K converges. Using a smaller value for α reduces the contribution of walks with longer distance. This is a desired property as for example in a social context having a a more direct connection to an influential person increases the possible benefit from this connection.

Katz centrality can be calculated in closed form by using the geometric series

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots = \sum_{i=0}^{\infty} x^i \quad .$$

Substituting αA for x and reducing the result by the multiplicative identity yields the operator

$$(I - \alpha A)^{-1} - I$$

which can be applied to $\vec{I}$ (a vector with each entry equal to 1) to get a vector $\vec{K}$ with entries $K_i = C_K(v_i)$

$$\vec{K} = ((I - \alpha A)^{-1} - I)\vec{I} \quad . \tag{2.3}$$

CHAPTER 3

Econophysics

Econophysics is an interdisciplinary research field where methods with roots in statistical physics are applied to economic systems to gain a better understanding of various phenomena, including but not limited to scaling laws [SAB+96] and failure modes of economic systems, such as the market crashes that provide the motivation for the analysis in this work.

In the first section of this chapter we will outline the idea of the science of complex systems and its relation to physics. The second section is a brief overview of how the graph theoretic concepts from chapter 2 are applied to the study of economic systems. Furthermore, we will explain the model of the financial network used for the analysis in this work and its properties.

The third section is an introduction to the DebtRank method and is followed in the final section of this chapter by a selection of notable applications of the DebtRank method.

3.1 Complex Systems

In statistical mechanics, macroscopic properties of a material emerge from the microscopic interactions of its molecules. For this purpose, statistical methods are applied to derive the macroscopic properties from interactions based on the fundamental forces.

The science of complex systems is an extension of physics where instead of the fundamental interactions a multitude of possibly non-linear interactions are studied. These interactions often do not work on a homogeneous space with a distance metric but on a network of components with arbitrarily complex neighborhoods.

Spin glass models (e.g. the Ising model, a model for ferromagnetism) in physics are similar insofar as they consist of interacting components (the spins), and the interactions of these components are limited to a discrete neighborhood. When first devised by Ernst Ising [Isi25], the spins were laid out on a regular lattice with interactions between a spin and its nearest neighbors. More recently, Ising models on general graphs were also studied [Bre15].

3.2 Economic networks

The economy is a complex system that consists of a multitude of entities (factory owners, factories, employees, products and consumers) interacting in various ways, for example:

- *investing, owning*: Factory owners invest in factories
- *being employed by*: employees are employed by factory owners
- *working at*: employees work at factories

- *producing*: employees produce products
- *buying*: consumers buy products

This can be modelled by a graph that represents the entities by sets of vertices (one set for each entity type) and the interactions by edges (again, a set for each possible interaction type). The resulting graph is then:

$$G = (V, E, V_1, V_2, \ldots, E_1, E_2, \ldots) ,$$

with $V = \bigcup_i V_i$ and $E = \bigcup_i E_i$. A vertex v could possibly be in two different sets V_i and V_j at the same time, for example a person could own a factory and be a consumer.

3.2.1 Financial networks

A financial network is a special case of an economic network where the links between nodes only represent financial dependencies, for example liabilities or investments.

Liability and asset network

The financial network of interest for this work is only between companies and banks which interact through lending money. The companies and banks are referred to as nodes, each of which is represented by a vertex in the graph that represents the financial network. The nodes are subdivided into two disjunct sets, the banks B and companies C and their union is the set of vertices $V = B \cup C$.

The money-lending and borrowing connections are modelled as edges connecting the vertices representing the nodes involved in the transaction. Therefore edges represent outstanding debts, either in the form of loans from banks or as deposits from companies at banks.

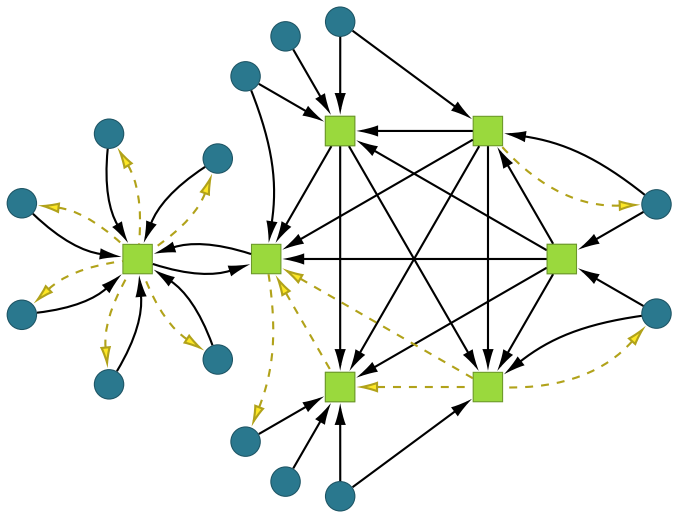


Figure 3.1: Illustration of a network of banks and companies . Banks have connections to companies and to each other, whereas companies do not interact with each other but only with banks. Connections are either loans , or deposits .

Each outstanding debt can be either viewed as an asset or as a liability: From the point of view of the lending party the debt is an asset, if viewed from the borrowing party the debt must be paid back and is therefore a liability.

If the graph $G = (V, E)$ represents the assets network, the transposed graph $G' = (V, E')$ (which is obtained by reversing the direction of all edges) is the liability or exposure network. The weighted adjacency matrices of these two graphs are therefore called assets matrix (A) and liability (exposure) matrix (L) where $A = L^\top$ and $L_{ij} = A_{ji}$.

Figure 3.1 is an illustration of a bank-company network.

3.3 DebtRank

A financial crisis from the viewpoint of econophysics results from the spread of an initial market-shock in an ecosystem. Depending on the resilience of the ecosystem, the spread of the shock can either be locally constrained or extend over large parts of the system.

The method used to analyze these dynamics, called DebtRank, was introduced by Battiston et al. [BPK+12] in 2012. It is a form of generalized eigenvector analysis that identifies systemically important nodes by assessing the fraction of the economic value of the system that is affected by a shock diffusing from a given node or a set nodes.

3.3.1 Idea of the DebtRank method

In a financial liability network a node d may default and therefore not be able to satisfy its liabilities to other nodes. If j is one of the nodes that was exposed to the default-risk of node d it incurs a loss on its loan (or deposit/investment) to c. Node j then tries to mitigate its own default by compensating the loss with its capital. If the capital of node j is not sufficient to offset the incurred loss, node j also defaults. From here the process continues recursively until the set of defaulting nodes remains stable.

3.3.2 Formal definition

The financial dependencies of the nodes in the network are given in a liability matrix L with entries L_{ij} denoting that node j has given node i a loan (or investment/deposit) of size L_{ij}. Additionally, there is a capital (or equity) vector C with entries C_i denoting the capital of node i.

The relative economic value of a node i is given by

$$v_i = \frac{L_i}{\sum_j L_j} \tag{3.1}$$

where $L_i = \sum_j L_{ji}$ is the sum of the outstanding liabilities of node i.

The default of node i then affects all nodes j where $L_{ij} > 0$. The impact of the default of i on j is defined as

$$W_{ij} = \min\left(\frac{L_{ij}}{C_j}, 1\right) \ . \tag{3.2}$$

The impact of a shock is thus measured as the fraction of capital loss due to the credit default. It is therefore a value in the range $[0, 1]$ with the semantics $W_{ij} = 0$ if the default of node i does not affect node j, and $W_{ij} = 1$ if the default of node i results in a loss that matches or exceeds the capital of node j.

The economic value of the impact is obtained by multiplying the impact with the relative economic value from eq. (3.1). The economic value of the impact of i on its neighbors is therefore given by

$$I_i = \sum_j W_{ij} v_j \tag{3.3}$$

Though if the neighbors of i do not have enough capital to compensate for the default of i they default themselves, leading to reverberations in the network. It is possible to account for this by modifying eq. (3.3) in a way resembling the adaption of eigenvector centrality (see section 2.4.4) for Katz centrality (see section 2.4.5):

$$I_i = \sum_j W_{ij} v_j + \alpha \sum_j W_{ij} I_j \ . \tag{3.4}$$

Equation (3.4) can be simplified similar to eqs. (2.2) and (2.3). Also, the same restrictions apply for the dampening factor α.

The problem with formula eq. (3.4) though is that the impact from a default spreads along walks of arbitrary length on the graph defined by the matrix W. If W contains a cycle the impact from the default of node i spreads along the edges until it reaches i again (due to the cycle), possibly resulting in another default of i.

To fix this, Battiston et al. proposed to only take walks without repeating edges[1].

[1] Walks without repeating edges are sometimes referred to as *trails*

This is done by introducing two additional state variables for each node, s_i and h_i (both dependent on the time step t). The variable s_i takes one of three values:

$s_i(t)$	Interpretation
U	Node i is undistressed at time t
D	Node i is in distress at time t
I	Node i is inactive at time t

The variable h_i has a value in the range $[0, 1]$ and is the level of distress, with 0 meaning *undistressed* and $h_i(t) = 1$ in the case of default. The value of $h_i(t)$ is defined as

$$h_i(t) = \min\left(1, h_i(t-1) + \sum_{j|s_j(t-1)=D} W_{ji} h_j(t_1)\right) , \tag{3.5}$$

whereas $s_i(t)$ is given by:

$$s_i(t) = \begin{cases} D & \text{if } h_i(t) > 0 \wedge s_i(t-1) \neq I, \\ I & \text{if } s_i(t-1) = D, \\ s_i(t-1) & \text{otherwise} \end{cases} \tag{3.6}$$

To calculate the DebtRank of a node d (d for *defaulting*), the distress h_i and status s_i at time step $t = 1$ are initialized as follows:

$$h_i(1) = \begin{cases} 1 & \text{if } i = d \\ 0 & \text{otherwise} \end{cases} \qquad s_i(1) = \begin{cases} D & \text{if } i = d \\ U & \text{otherwise} \end{cases} . \tag{3.7}$$

Then the values of s_i and h_i are calculated for every node i and time step t, according to eqs. (3.5) and (3.6) until all nodes are either inactive or undistressed (see fig. 3.2 for an example of the resulting dynamics).

When the distress levels of the nodes are converged at time step $t = T$, the DebtRank of node d can be calculated as the sum of the distress level multiplied

Source: [BPK^{+}12, Supplementary information, Figure 5]

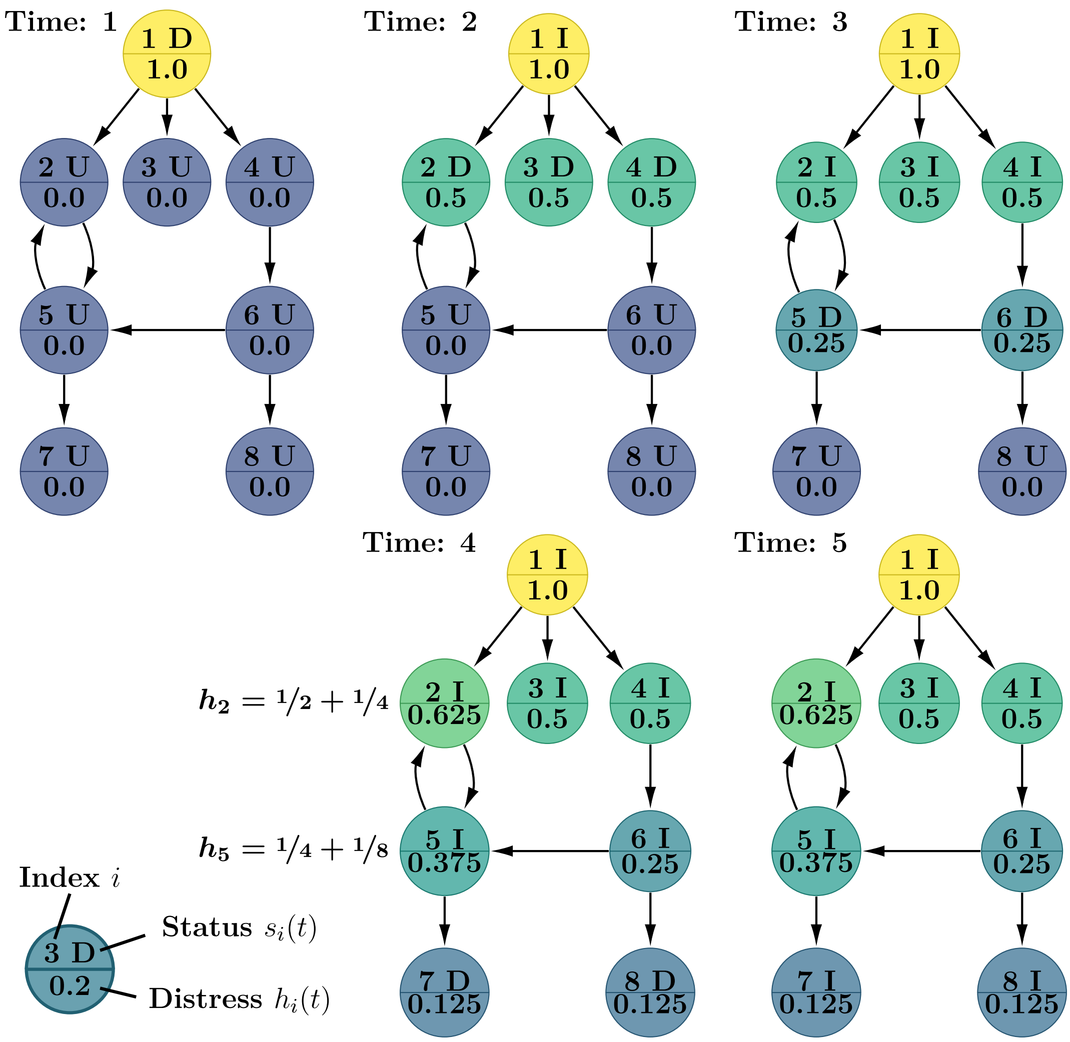


Figure 3.2: Example of DebtRank computation. Each node is labelled with its index i, $s_i(t)$ and current distress level $h_i(t)$ (see legend on bottom left). Graph is based on the impact matrix W where all entries W_{ij} are either 0 (no edge from i to j) or 0.5. In each step the neighbors of nodes that were distressed at the previous time step get a contribution to their level of distress. The first time h_i is updated from 0.0 to a value > 0, the node goes in distress. If the node was in distress at the previous time step, its status is updated to inactive and does not further contribute to the distress level of adjacent nodes (at $t = 4$ at node i the contribution is $1/2 \times h_5(3) = 1/8$ and does not increase further at $t = 5$ because v_5 is already inactive when its distress rises from $1/4$ to $3/8$).

by the relative economic value of each node except node d:

$$R_d = \sum_i h_i(T)v_i - h_d(1)v_d \tag{3.8}$$

Simultaneous default of multiple nodes

It would be possible to calculate the combined DebtRank of a set S of nodes by replacing the $i = d$ conditions in the initialization (eq. (3.7)) by $i \in S$ and changing eq. (3.8) to one of the following equations:

$$R_S = \sum_i h_i(T)v_i - \sum_{d \in S} h_d(1)v_d \tag{3.9}$$

$$R_S = \sum_i h_i(T)v_i \tag{3.10}$$

Equation (3.9) excludes the impact of the initial shock, whereas eq. (3.10) does not.

In this work, DebtRank is calculated on two different networks, the interbank network B and the full network F. We use R^F and R^B to discern between the two measures.

3.4 Applications of the DebtRank method

Since its introduction, the DebtRank has been frequently used to measure either systemic risk itself or effectiveness of policies that try to rein in systemic risk. In this section we will present a selection of notable applications of the DebtRank method.

3.4.1 DebtRank transparency [TP13]

Individual institutions cannot decide whether another institution is systemically risky and therefore they cannot decide whether borrowing from or lending to them would be advisable. Thurner and Poledna propose to alleviate this by making the systemic risk of financial institutions accessible to enable financial institutions to make better informed decisions.

3.4.2 Systemic risk transaction tax [PT16]

Poledna and Thurner developed a method to quantify the marginal contribution of systemic risk by single transactions and proposed to introduce a tax on systemically risky transaction, thus providing an incentive for banks and companies to mitigate not only transactions with high credit default risk but also transactions with high systemic risk.

They run simulations with an agent based model to compare their proposed policy with a financial transaction (Tobin-like) tax. Their results suggest that their method has a smaller impact on the transaction volume of the interbank market and is simultaneously more effective at reducing systemic risk in the network.

3.4.3 Systemic risk in multi-layer networks [PMBMJ+15]

In contrast to the analysis of a single layer financial network in this thesis, Poledna et al. analyzed the systemic risk in the Mexican interbank network with four layers, reflecting the different types of exposure between banks (derivatives, securities, foreign exchange and liabilities). They show that considering only one layer (as done in this thesis) vastly underestimates the systemic risk.

Part II

DebtRank analysis of the Austrian economy using empirical data

CHAPTER 4

Introduction

Lending and borrowing between financial institutions introduces exposures from one institution to another and enables propagation of market shocks in the financial network. Due to this, the default of a credit may not only affect the directly involved parties but also institutions with exposures to them, potentially rendering large portions of the system non-functioning.

The crisis of 2007–2008 has demonstrated the societal costs of neglecting this systemic risk in financial markets. This sparked the developments of methods for quantifying systemic risk [BPK+12, SC13, AB16, APPR17, PL16]. These methods were used to assess systemic risk in a multitude of financial networks [TSGo13, PMBMJ+15, PCB14] and for research on the effectiveness of crisis resolution mechanisms[KPFT15], regulations [PBT17] as well as policies and taxing schemes that counteract the build-up of systemic risk [TP13, PT16, AB16, APPR17, LPT16].

As market shocks are propagated via exposures between financial institutions, interconnectedness is a necessary condition (either directly through financial dependencies [BPK+12] that default or via mutually held assets that deprecate due to synchronized behavior, e.g. fire sales [PBT17, PPT17, AB16]) for systemic collapses and thus systemic risk. Therefore the analysis of systemic risk is until now primarily focussed on the interbank market due to its role as the "disease vector" for the contagion of market disturbances. This neglects the possible influence of companies as origin and propagator of systemic collapses. My goal is therefore to expand upon the existing research by analyzing a financial network that not only includes interbank liabilities but also liabilities and deposits between banks and companies. We do this by using empirical data from the commercial register to build a bank-company liability network which will then be connected with the interbank liability network, provided by the Austrian national bank (see chapter 5). We then use the DebtRank method proposed by Battiston et al. [BPK+12] (see section 3.3) to identify systemically important companies. Furthermore, we will investigate how adding the bipartite bank-company to the interbank network affects the ranking of banks. In chapter 6 the results of the analysis will be presented, followed by a discussion in chapter 7.

CHAPTER 5

Statistics and reconstruction from empirical data

The data used for this analysis consists of two main parts, annual financial statements of companies and balance sheets of banks. The company data is from the commercial register of Austria and extracted and collected by the Bureau van Dijk. The bank data is from the Austrian Central Bank (OeNB). The public financial statements were collected from the official homepage[1], the interbank network was also provided by the Austrian Central Bank as an anonymized and linearly transformed data set.

In the first section we will provide some statistics describing the different data sets. Then we will outline how the financial network as described in section 3.2.1 was constructed from the raw data.

[1]https://www.oenb.at/jahresabschlusski/jahresabschlusski

5.1 Statistics

As the data sets for the interbank liability network—and the commercial register data—only overlapped for 2008, only data for this year was considered.

Figure 5.1 shows a stacked bar chart of the the different liability types of the 106 919 companies in the commercial register grouped by the number of bank connections for the calendar year 2008 (individual accounts only, most recent account if multiple balance sheets submitted in same year).

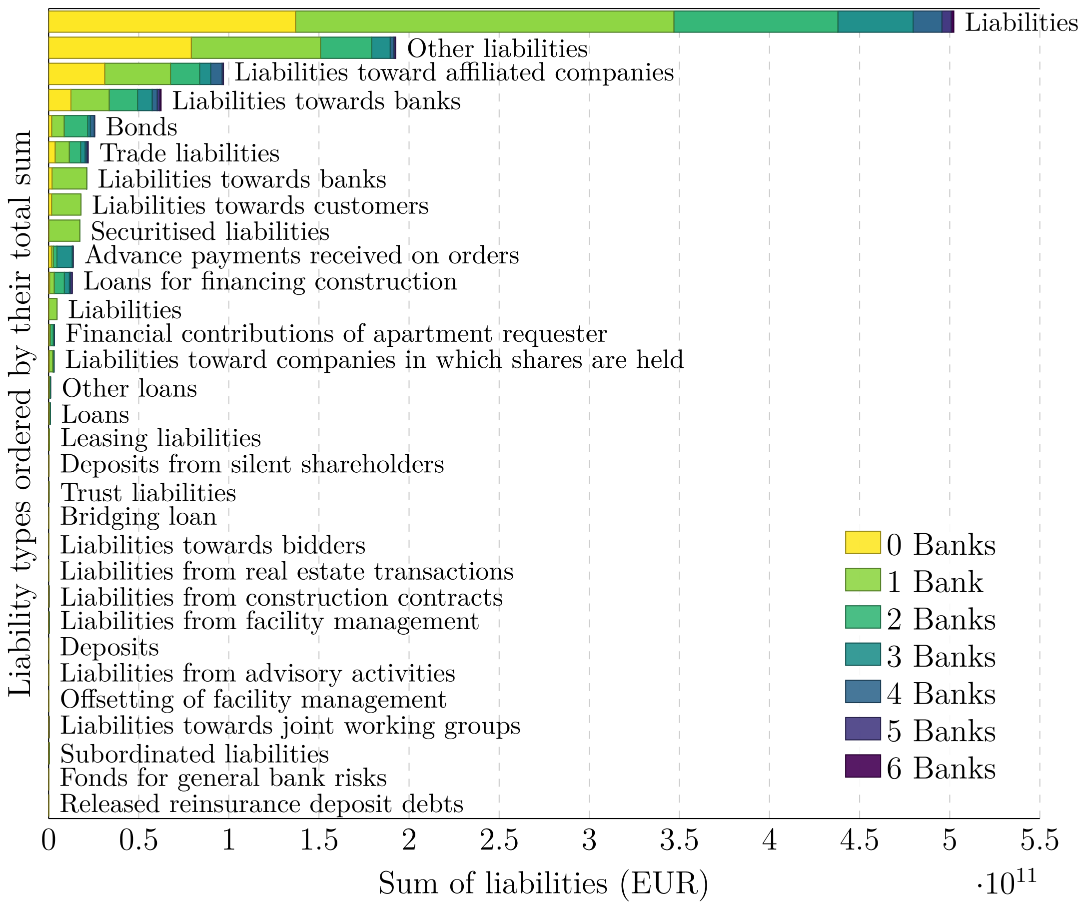


Figure 5.1: Different liability types in the commercial register aggregated over all companies with a given number of bank connections. Liability types are sorted according to the total sum per type. Companies without bank connections are discarded, which affects approximately 27.3% of the liabilities.

In fig. 5.2 the distribution of the number of unique bank connections per company is shown. Approximately 51.4% of the companies provide no bank connection and therefore cannot be used for the DebtRank analysis. The impact of this missing-data problem is alleviated by the fact that primarily smaller companies are affected by it. As fig. 5.1 shows only 27.3% of the total liabilities are associated with companies dropped due to not providing information about their bank connections.

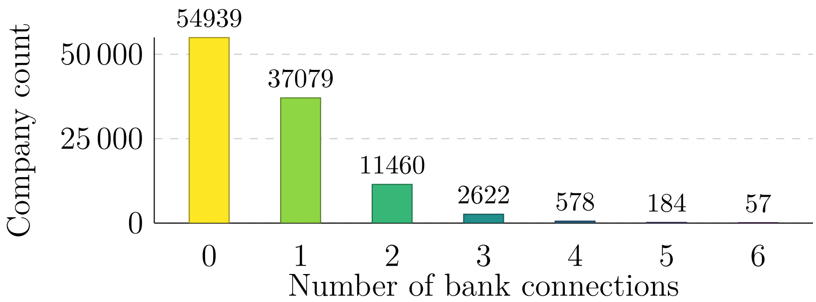


Figure 5.2: Number of companies that provide a certain number of bank connections in the commercial register. Up to six banks are possible.

As most companies are not legally obliged to provide an exact breakdown of their different liability types, the "Liabilities towards banks", which was needed for our further analysis, was estimated by average ratio of companies working in the same line of business, as indicated by their OeNACE code.

5.2 From raw data to the liability network

The liability network was reconstructed from the unweighted network of companies, banks and their connections by assigning a weight to each edge. This company-bank liability network was then connected with the interbank liability network provided by the Austrian Central Bank.

5.2.1 Data cleaning

Four banks had a leverage (loan-equity ratio) of 10.15, 11.34, 61.84 and 293.03, above the allowed threshold of 10. We assume that this is due to an error with the reporting of their equity (Tier 1 capital). This was fixed by increasing the equity of the affected banks to ensure that the leverage of all banks was ≤ 10.

Apart from that we encountered one company that had liabilities, which increased by a factor of approximately 10 from 2007 to 2008 and decreased again to the level from 2007 a year later. We assumed that this was due to a misplaced decimal point and removed the company from our data set.

5.2.2 Company-bank adjacency matrix

The unweighted network connecting companies with banks was created using the bank connections provided by the companies in the commercial register. The result was a bipartite graph $G = (V, E, B, C)$ with B (the set of banks) and C (the set of companies) being two disjunct sets of nodes. The sets have the following cardinalities:

$$\begin{aligned} |V| &= 52\,776 \\ |E| &= 71\,439 \\ |C| &= 51\,980 \\ |B| &= 796 \end{aligned}$$

The adjacency matrix of the graph G is called L^0 subsequently.

5.2.3 Weighted company-bank network

To get from the adjacency matrix to the weighted liability network used for the DebtRank-analysis, the edges were assigned weights by distributing the aggregated liability/asset data from each node to its incident[2] connections as follows:

- For every company c, take the liabilities L_c from its balance sheet submitted to the commercial register. If the company provides detailed data take the "Liabilities towards banks", else use the average ratio of companies in the same line of business to estimate the share of liabilities to banks from total "Liabilities".

- Take the set of aggregated loans (referred to as assets, or A_i where i is the index of the bank) of all banks from their balance sheets.

- For every company c calculate the vector ℓ^c in the following way:

$$\ell_b^c = \begin{cases} 0 & \text{if } L_{cb} = 0 \\ A_b & \text{else} \end{cases}$$

  For every company c the vector ℓ^c has an entry for every bank b, where the entry ℓ_b^c equals the aggregated assets of bank b if the company is connected to the bank and zero else.

- Normalize the resulting vector with the L1 norm.

$$\hat{\ell}^c = \frac{\ell^c}{\sum_i |\ell_i^c|}$$

- Partition the aggregated liabilities L_c of each company c with the distribution $\hat{\ell}^c$ to get the entries for the company-bank liability network:

$$L_{:c} = L_c \odot \hat{\ell}^c$$

  where $\odot$ is the operator for element wise multiplication.

[2]In a graph the edges starting or ending at a vertex are incident to this vertex.

This essentially partitions the liabilities of each company to the banks it is connected to according to the relative size of the banks.

This rather naïve approach was used for the network reconstruction because the row and column sums of the data did not match. This is due to the fact that this analysis only considers loans from banks to companies in Austria whereas the banks only provide their aggregated data which additionally include e.g. loans to private citizens or to companies abroad. The applied method assumes that this inflation is the same for all banks. If more precise data were available this should be incorporated into the reconstruction and more advanced methods should be applied [MSFG14, BEST04].

These steps were repeated for the bank deposits of the companies to get the liabilities of the banks to the companies.

5.2.4 The full liability network

To get the complete bank-company liability network, the interbank network provided by the Austrian National Bank and the weighted bipartite asset and liability networks of the companies were combined and padded with zeros to get the full liability network.

The resulting liability matrix L has the following form:

$$L_{n\times n} = \begin{pmatrix} BB_{b\times b} & BC_{b\times c} \\ \\ CB_{c\times b} & CC_{c\times c} = 0 \end{pmatrix}, \qquad n = c + b \ . \tag{5.1}$$

Here c and b denote the number of companies and banks. The matrix consists of the following four parts:

- interbank network BB connecting banks with banks

- bank-company network BC containing information about deposits companies have at financial institutions
- company-bank network CB containing information about liabilities companies have to financial institutions
- company-company network CC with inter-company liabilities, omitted in this work thus $CC_{xy} = 0 \quad \forall x, y \in [1, c]$.

Each entry L_{ij} indicates the liability that node i (which is a bank if $i \leq b$ and a company if $i > b$) has to node j (the same applies here).

5.3 Graph visualizations

We used the graph layout algorithms from Hu Yifan [Hu06] and Shawn et al. [SBB11], implemented in Gephi [BHJo09] to create the graph illustrations in this section and in chapter 6.

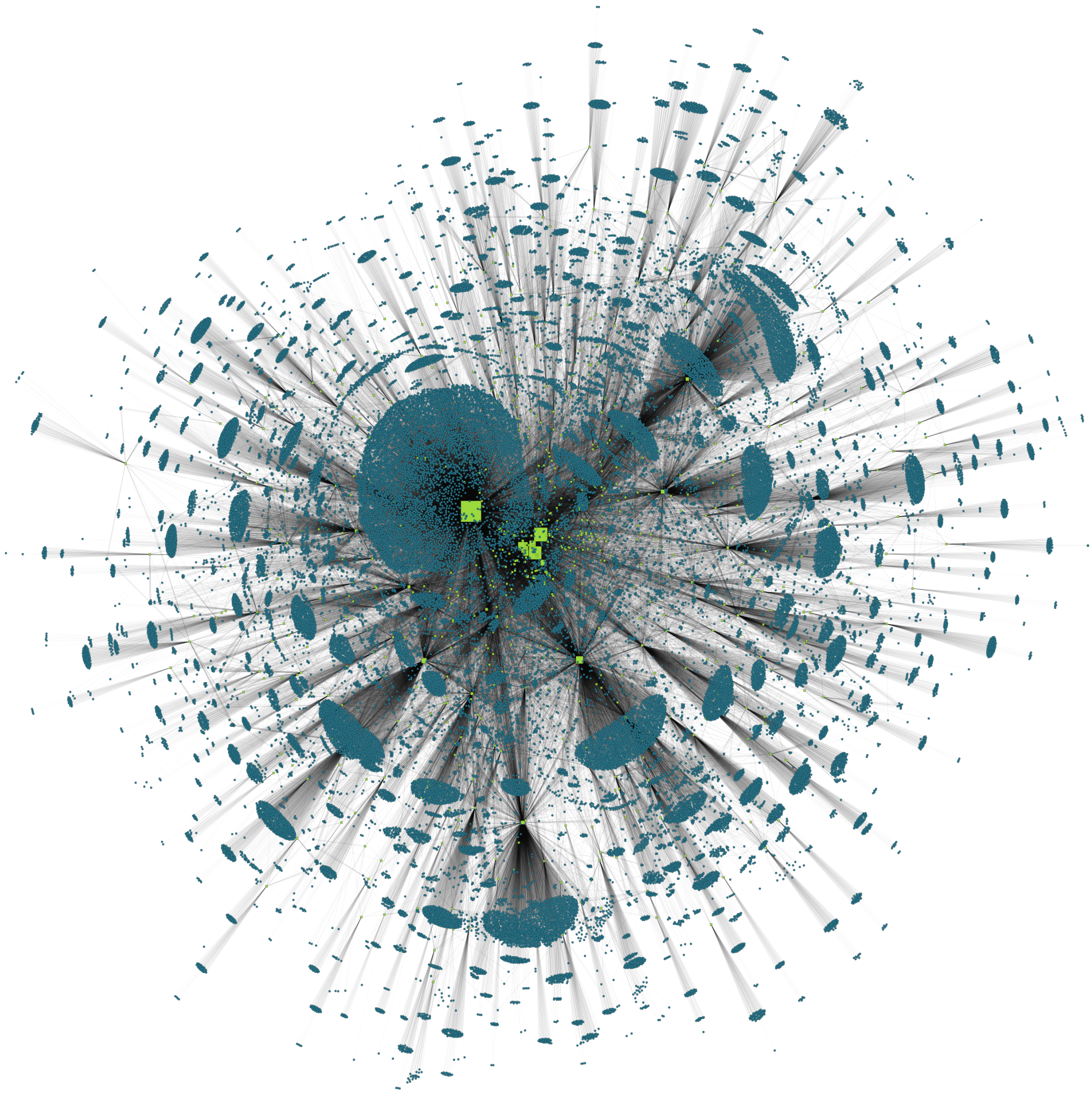


Figure 5.3: Plot of liability network containing all 796 banks and 51 980 companies . Nodes have sizes corresponding to their total assets. In the center mostly banks that operate nationally and companies with ties to them are located. The clusters further out are companies operating in certain regions being connected to the regional branch of a bank in close vicinity. Hierarchies stem from the federal structure of Austria.

Figure 5.4: Plot of liability network containing all 796 banks and the 5 000 companies with the biggest liabilities to financial institutions. Node sizes correspond to their total assets. This network was used for the DebtRank analysis in chapter 6. As the total assets of the companies are small compared to the biggest banks in the data set, most of the nodes have the minimum node size in the visualization. Figure 6.5 shows the distribution of asset sizes of the analyzed network.

5.4 Network statistics

The degree distributions of the banks in the networks analyzed in this work are illustrated in figs. 5.5 to 5.7. The in- and outdegree distributions are depicted in fig. 5.5 for the full network F and in fig. 5.6 for the interbank network B. Figure 5.7 shows the undirected degree in both, the full network and the interbank network for banks. In figs. 5.5 to 5.7 the outer plots show the whole range, whereas the insets provide additional details about the ranges with higher density.

Figure 5.8 shows the degree distribution of companies in the full network (similar to fig. 5.2, but restricted to companies with degree > 0 and only containing the 5000 companies with the biggest liabilities). Note, that the in- and outdegree for companies are identical as the bank connections provided to the commercial register were used for both types of connections between companies and banks (deposits & liabilities).

The undirected and unweighted global clustering coefficients $\langle C_i \rangle$ (calculated according to [WS98]) of the full network and of the interbank network are significantly higher than the the clustering coefficient of a random graph with identical number of nodes and vertices $\langle C_i \rangle^{\mathrm{rand}}$, as shown in table 5.1.

Network	Nodes	Edges	$\langle C_i \rangle$	$\langle C_i \rangle^{\mathrm{rand}}$
Full network F	5796	28127	0.77	0.005
Interbank network B	796	12783	0.89	0.043

Table 5.1: Undirected and unweighted global clustering coefficient $\langle C_i \rangle$ of networks analyzed in this work and global clustering coefficient of random graph with same number of nodes and edges. Random graph has considerable lower $\langle C_i \rangle$. as both networks.

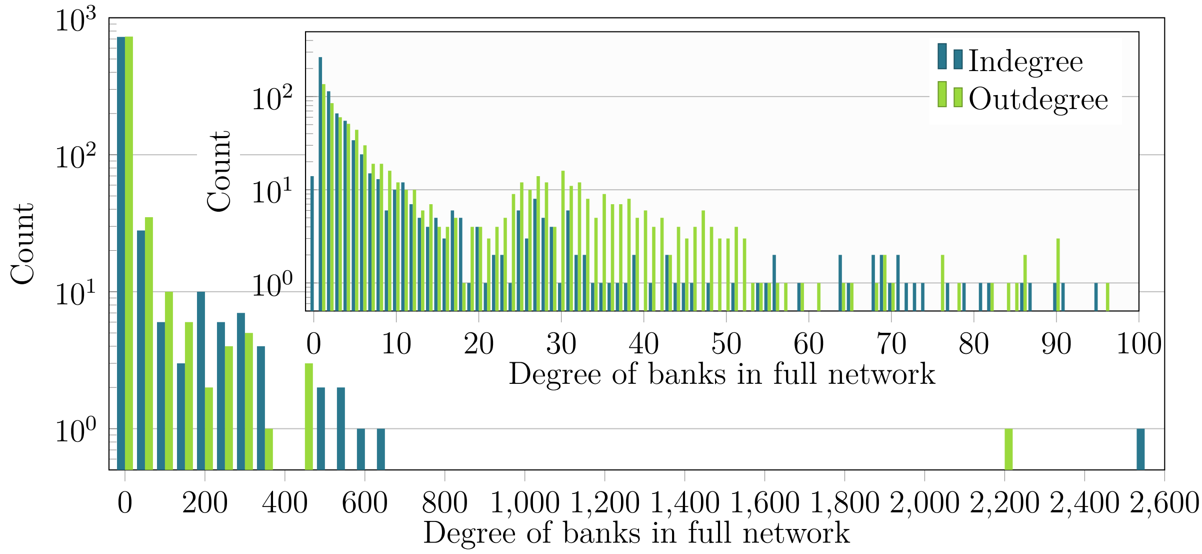


Figure 5.5: In- and outdegree distribution of banks in the full company-bank network F. Outer plot has 60 uniform bins on the interval $[0, 3000]$, the inset shows the distribution of degrees in the range $[0, 100]$.

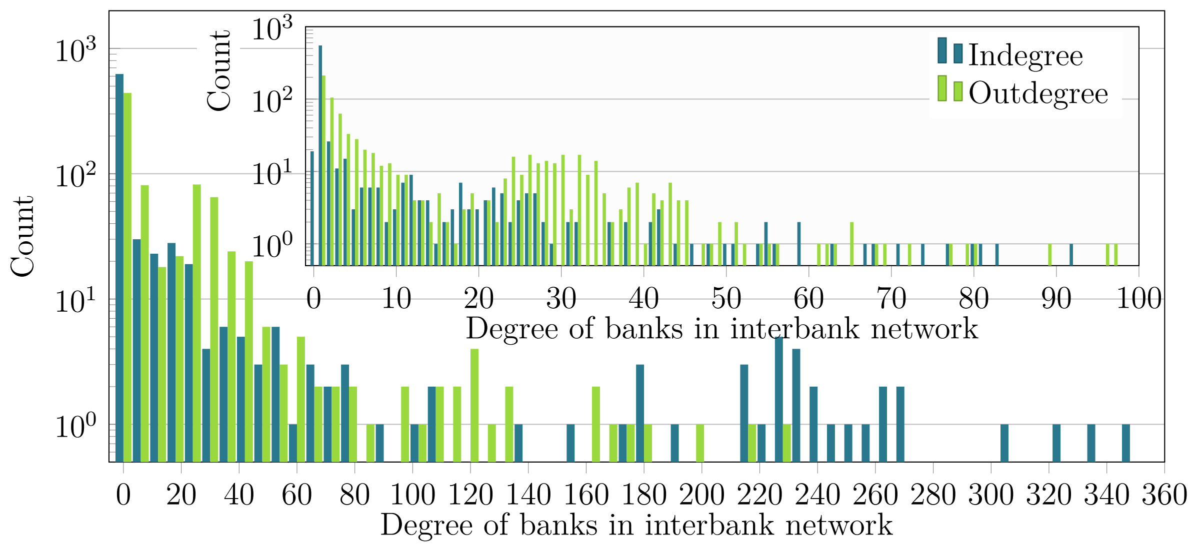


Figure 5.6: In- and outdegree distribution of banks in the interbank network B. Outer plot has 60 uniform bins on the interval $[0, 360]$, the inset shows the distribution of degrees in the range $[0, 100]$.

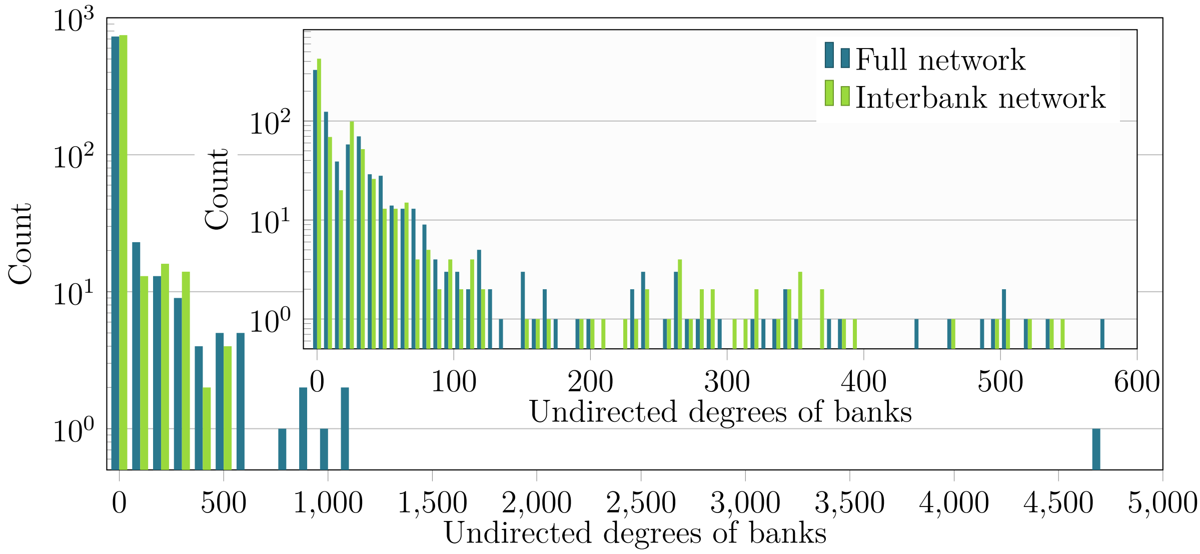


Figure 5.7: Undirected degree distribution of banks in the full network F [ ] and in the interbank network B [ ]. Outer plot has 50 uniform bins on the interval $[0, 5000]$, the inset shows the distribution of degrees in the range $[0, 600]$ (75 bins).

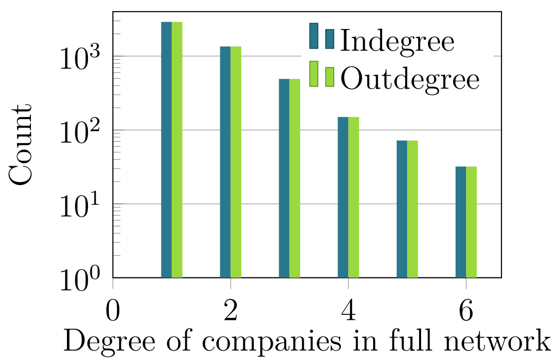


Figure 5.8: In- and outdegree distribution of companies in full network. Values are identical for all companies because loans and deposits (BC and CB network, see section 5.2) use the same adjacency matrix, based on the bank connections in the commercial register.

CHAPTER 6

Results

The results from the analysis of the liability network with the DebtRank method will be presented in this chapter and then discussed and summarized in the final chapter. For these results the liability graph L was reduced to the subgraph induced by the union of the set of all 796 banks and the 5 000 biggest companies ordered by their total liabilities. This graph is referred to as the "full graph"—the DebtRank calculated on this graph is therefore abbreviated with R^F (F for full). For the second part of the analysis (see section 6.2) additionally the interbank network consisting of the 796 banks and their connections to each other was considered. The DebtRank calculated on this bank-only network is abbreviated with R^B (B for bank).

6.1 DebtRank of companies and banks

Figure 6.1 visualizes the DebtRanks R^F of the full graph, though only companies and banks with a DebtRank $R^F \geq 0.01$ are depicted to improve legibility. Nodes are either banks (squares) or companies (circles), with sizes corresponding to their total assets and colored according to their DebtRank.

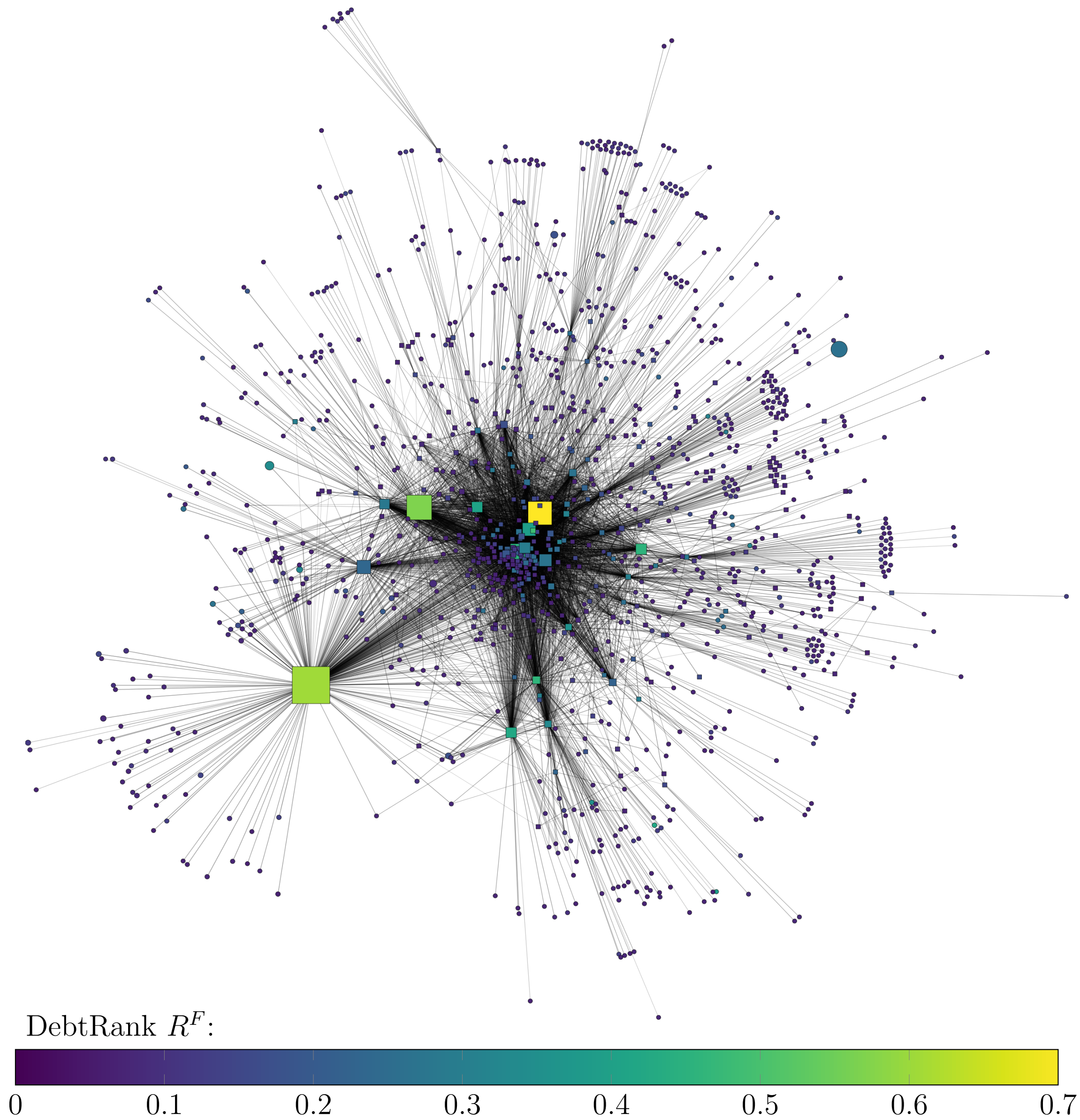


Figure 6.1: Subgraph of the analyzed network (fig. 5.4) with nodes restricted to those having a DebtRank $R^F \geq 0.01$. Banks are depicted as squares, companies as circle. Both, banks and companies are colored according to their DebtRank R^F and have sizes corresponding to their total assets. Most companies have a similar size as their total assets are small compared to the total assets of the biggest banks. Figure 6.5 shows the distribution of asset sizes of the analyzed network.

Figure 6.2 shows the distribution of DebtRanks of banks and companies with a histogram (range = $[0, 0.7]$, 70 bins). The DebtRanks of companies tend to be lower but there is considerable overlap in the distributions.

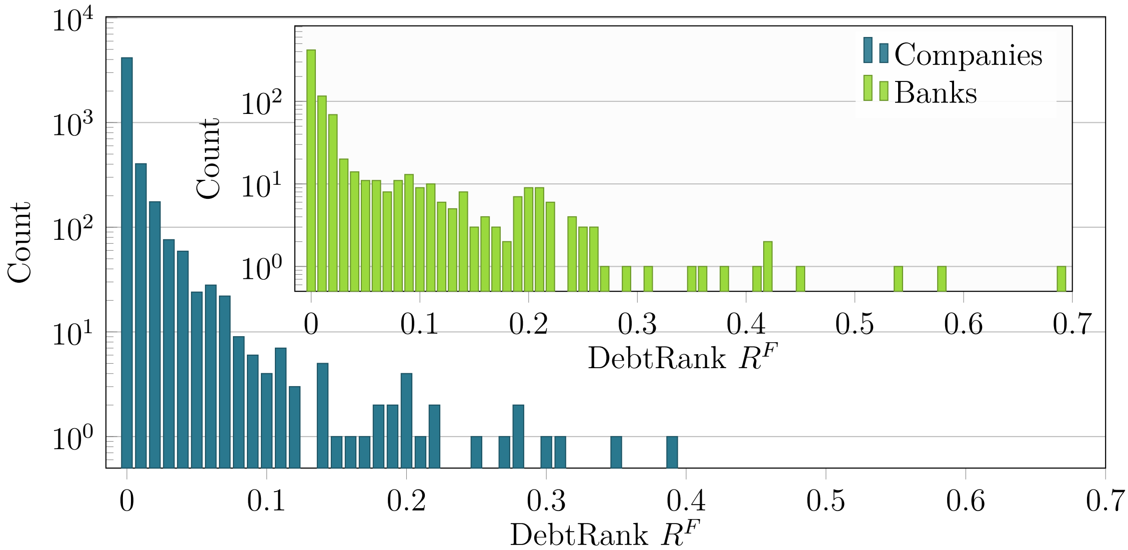


Figure 6.2: Histogram of DebtRanks R^F in the full network of banks [ ] and companies [ ]. Banks and companies have a qualitatively similar DebtRank-distribution. The highest DebtRank of a company is 0.39

Figure 6.3 is a bar plot of the first 45 banks and companies (left) and companies (right) ranked according to their DebtRank. The right plot only depicts companies and is labelled with the first character of the ONACE code (a system used in Austria to classify companies according to the sector they operate in, see [Wir08]). Banks are not labelled due to confidentiality. Comparing the profiles of both distributions seems to confirm the hypothesis that the impact from the default of a company may also affect a significant portion of the economy.

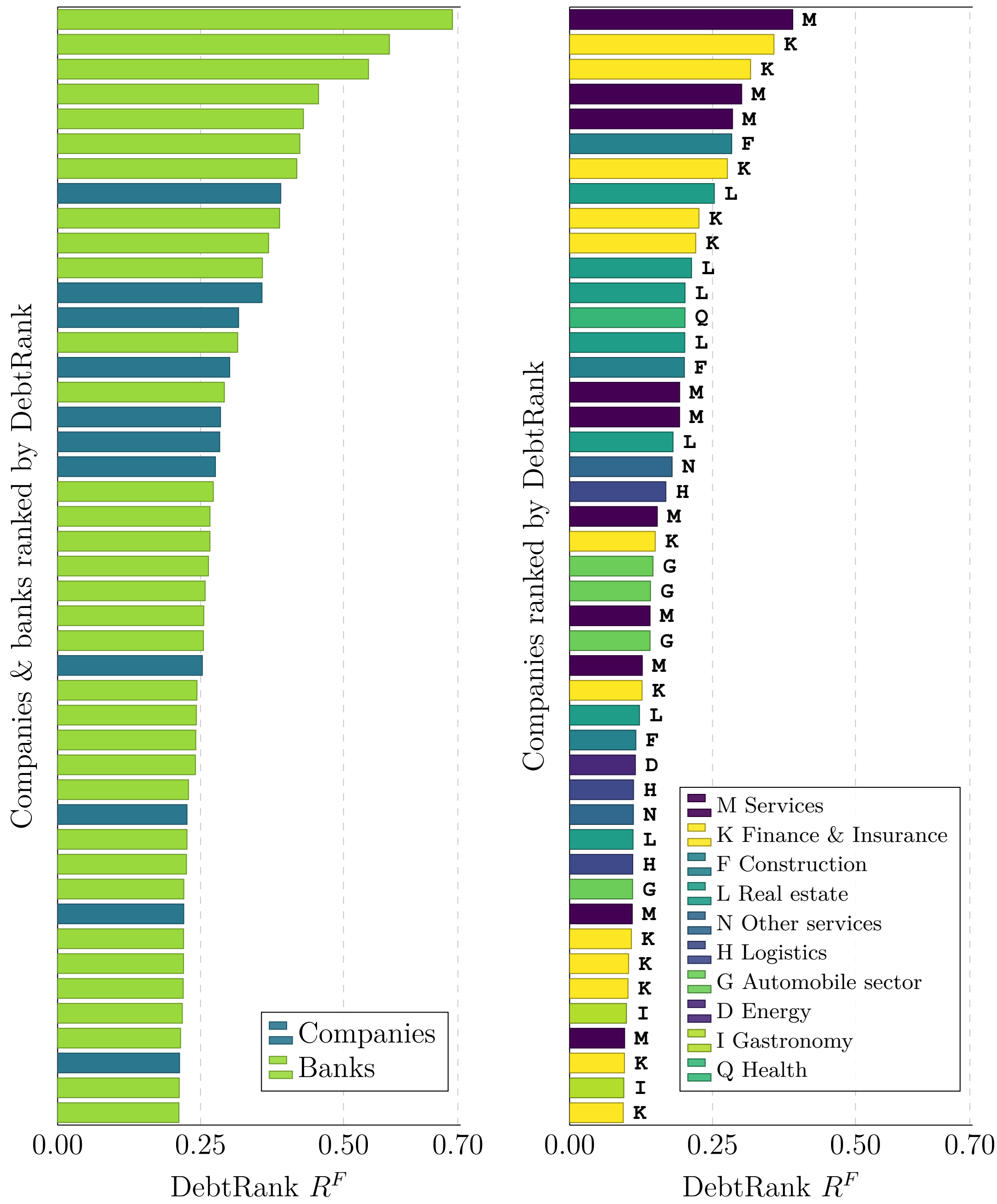


Figure 6.3: DebtRank of companies and banks sorted by their DebtRank from top to bottom in decreasing order. Distribution of banks and companies (left) similar to distribution of companies only (right). In the right plot the bars are labelled and colored according to the economic sector in which the companies operate, as per the first character of their OeNACE code (a system used in Austria to classify companies by sector [Wir08]). The banks are unlabelled due to confidentiality.

DebtRanks of companies and banks plotted against their total liabilities are depicted in depicted in fig. 6.4. Figure 6.5 is the DebtRank of companies and banks plotted against their total assets. The area with dashed box in the left figure is magnified in the right figure for both plots. The relation between liabilities and DebtRank (fig. 6.4) is more pronounced as between total balance and DebtRank (fig. 6.5).

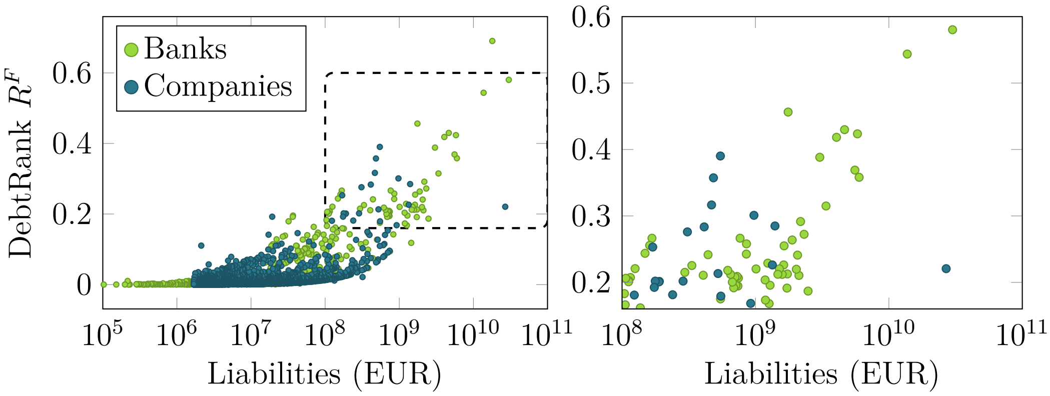


Figure 6.4: DebtRank of companies and banks depending on their total liabilities. Companies have a cutoff at 2×10^6 EUR as only the first 5 000 companies ranked by their total liabilities were used.

6.2 Impact of companies on the DebtRanks of banks

Figure 6.6 and fig. 6.7 illustrate how adding the bipartite company bank network influences the DebtRank of the 796 banks in two different ways.

Each bank b_i has a DebtRank $R^F(b_i)$ in the full bank-company network, as well as a DebtRank $R^B(b_i)$ in the bank-only network. Then $r^F(b_i)$ and $r^B(b_i)$ are defined as the ranks of bank i when the banks are sorted in ascending order according to R^F or R^B, respectively.

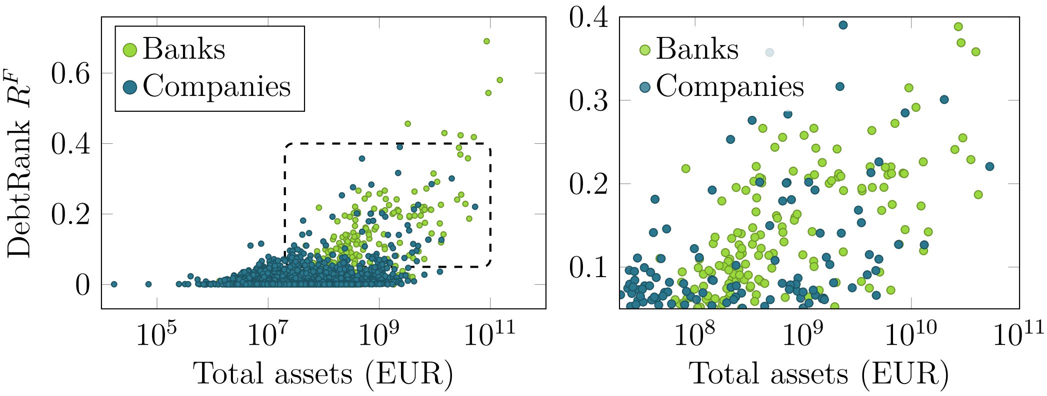


Figure 6.5: DebtRank of companies and banks plotted against their total assets in Euro. Distribution of banks and companies do not seem to be qualitatively different. Companies as well as banks with more assets tend to have a higher DebtRank. Nevertheless, companies with similar DebtRank have asset sizes that differ by multiple orders of magnitude.

The line plot of fig. 6.6 [—] uses $r^F(b_i)$ as its x-axis value whereas the scatter plot [●] uses $r^B(b_i)$. The y-value in both cases is $R^F(b_i)$. The resulting figure illustrates the impact of the bank-company network on the DebtRank of the banks: The horizontal distance between a point of the scatter plot and the line is the change of rank whereas the vertical distance is related to the change of DebtRank value. Although it should be noted that the DebtRank values cannot be compared directly as DebtRank is a relative measure and the denominator (the economic value in the network) differs.

Figure 6.7 shows the two different rankings $r^F(b_i)$ and $r^B(b_i)$ plotted against each other. If the ranking of the banks would not change by adding the companies to the network the points of the scatter plot [●] would be on the $y = x$ line [—].

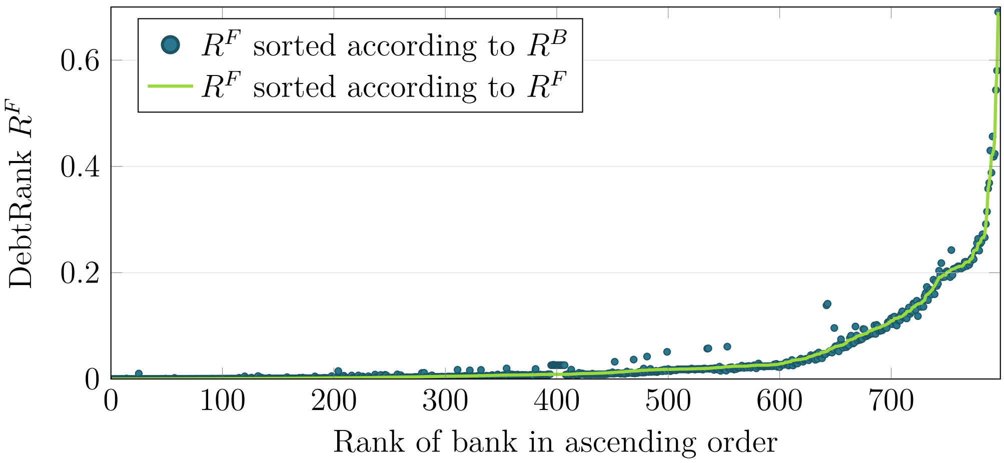


Figure 6.6: Sorting the DebtRank values of the banks in the network of companies and banks yields the curve [——]. Using the DebtRank of the bank in the subgraph induced by the banks only (i.e. the interbank network, IB) results in [●]. For the points above the curve, using the DebtRank in the interbank network R^B underestimates the impact from a possible default.

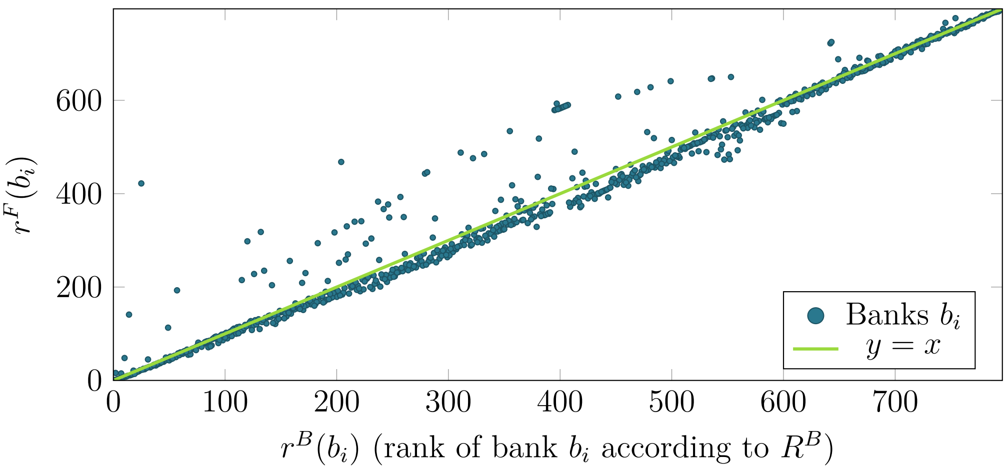


Figure 6.7: Rank of banks with and without bank-company network plotted against each other. The rank of some banks is vastly underestimated if only the interbank network is considered. The marks of the banks [●] are closer to the line [——] when the DebtRank is not influenced by taking only the bank-induced subgraph into account.

6.3 Quantifying the total contribution from companies and banks

As usually only the interbank network is used for the analysis of systemic risk we try to quantify how severely this underestimates systemic risk. For this purpose we define two ratios, one for assessing the split of the systemic risk in the full network and one for comparing the total systemic risk of the interbank network and of the full network.

Q_1 is therefore the ratio of the sum of the DebtRanks of all companies divided by the sum of all DebtRanks in the full network:

$$Q_1 = \frac{\sum_{i \in C} R_i^F}{\sum_{i \in F} R_i^F} \ . \tag{6.1}$$

To compare the systemic risk of the interbank network with the systemic risk of the full network we define a similar ratio Q_2, though it is necessary to take the different economic value of the two networks into account as DebtRank is a relative measure.

$$Q_2 = \frac{V^B \sum_{i \in B} R_i^B}{V^F \sum_{i \in F} R_i^F} \tag{6.2}$$

where V^B and V^F refer to the total economic values of the interbank network and the full network, respectively.

We get $Q_1 = 0.55$, which means that companies contribute 55% of the systemic risk in the full network. For Q_2 we get $Q_2 = 0.29$, i.e. the total systemic risk from the interbank network amounts to 29% of the total systemic risk of the full network. The remaining systemic risk stems from the liabilities between companies and banks and from the increased economic value V^F of the full network.

CHAPTER 7

Discussion

Using methods partly inspired from statistical mechanics, we analyzed a network of companies and banks reconstructed from empirical data. By extending the analysis from the interbank network to (large parts of) the whole economy, we were able to identify companies that have significant contributions to the overall systemic risk. In the full network, the systemic risk contribution of companies exceeds that of banks, with 55% and 45% respectively. Furthermore, we find that only taking the interbank network into account considerably underestimates the total systemic risk. According to our analysis, only 29% of the total systemic risk is due to the interbank network.

For both—companies and banks—it is possible to have a comparable DebtRank while having total assets that differ by multiple orders of magnitude. A few companies contribute a significant amount of systemic risk to the system, despite being rather small in terms of total assets. As systemically risky companies induce systemic risk in a similar way as banks, it seems that the concept of "too connected to fail", as coined by Battiston et al. [BPK$^+$12], does also apply to companies and the notion of systemically important banks (G-SIBs [Ban13]) could be generalized to companies. As some companies inherit a significant part of the systemic risk through their financial connections (loans from risky banks) to banks, it may be advisable to evaluate the extension of regulations and tax policies that try to rein in systemic risk in financial networks from banks to the real economy. The companies that would be subject to those macro-prudential regulations should be chosen depending on their systemic impact, for which company size is an insufficient proxy.

Furthermore, we compared the DebtRank of banks in two networks: the interbank network, containing only banks, as well as the full network comprised of banks and companies. We found that for some banks the DebtRank of the interbank network R^B vastly underestimates the systemic importance compared to the DebtRank in the full network R^F. In most cases the DebtRanks in both networks are similar, nevertheless it may be beneficial to take the whole economy into account for the analysis of systemic risk or its possible countermeasures.

A possible avenue for further research would be the extension of the analysed network with additional layers that have not been considered in this work, such as financial dependencies between companies (shadow banking [AS09]), exposures via mutually held assets or the shareholder network.

Bibliography

[AB16] Tobias Adrian and Markus K. Brunnermeier. CoVaR. *The American Economic Review*, 106(7):1705–1741, 2016.

[APPR17] Viral V. Acharya, Lasse H. Pedersen, Thomas Philippon, and Matthew Richardson. Measuring systemic risk. *The Review of Financial Studies*, 30(1):2–47, 2017. URL: `https://academic.oup.com/rfs/article-abstract/30/1/2/2682977`.

[AS09] Tobias Adrian and Hyun Song Shin. The Shadow Banking System: Implications for Financial Regulation. SSRN Scholarly Paper ID 1441324, Social Science Research Network, Rochester, NY, July 2009. URL: `https://papers.ssrn.com/abstract=1441324`.

[Ban13] Bank for International Settlements. Global systemically important banks: updated assessment methodology and the higher loss absorbency requirement, July 2013. URL: `https://www.bis.org/publ/bcbs255.htm`.

[Bav50] Alex Bavelas. Communication Patterns in Task-Oriented Groups. *The Journal of the Acoustical Society of America*, 22(6):725–730, November 1950. URL: `http://asa.scitation.org/doi/10.1121/1.1906679`, `doi:10.1121/1.1906679`.

[BEST04] Michael Boss, Helmut Elsinger, Martin Summer, and Stefan Thurner. Network topology of the interbank market. *Quantitative Finance*, 4(6):677–684, December 2004. URL: http://www.tandfonline.com/doi/abs/10.1080/14697680400020325, doi:10.1080/14697680400020325.

[BHJo09] Mathieu Bastian, Sebastien Heymann, Mathieu Jacomy, and others. Gephi: an open source software for exploring and manipulating networks. 2009. URL: http://www.aaai.org/ocs/index.php/ICWSM/09/paper/download/154/1009.

[Bon72] Phillip Bonacich. Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2(1):113–120, 1972. URL: http://www.tandfonline.com/doi/pdf/10.1080/0022250X.1972.9989806.

[BPK+12] Stefano Battiston, Michelangelo Puliga, Rahul Kaushik, Paolo Tasca, and Guido Caldarelli. DebtRank: Too Central to Fail? Financial Networks, the FED and Systemic Risk. *Scientific Reports*, 2:541, August 2012. URL: http://www.nature.com/srep/2012/120802/srep00541/full/srep00541.html, doi:10.1038/srep00541.

[Bre15] Guy Bresler. Efficiently learning Ising models on arbitrary graphs. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, pages 771–782. ACM, 2015. URL: http://dl.acm.org/citation.cfm?id=2746631.

[Eul35] Leonhard Euler. "The Seven Bridges of Konigsberg" IN NewmanJames-1957v01. *UNZ.org*, 1735. URL: `http://www.unz.org/Pub/NewmanJames-1957v01-00573`.

[Fre77] Linton C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977. URL: `http://www.jstor.org/stable/3033543`.

[HBM10] Emilie M. Hafner-Burton and Alexander H. Montgomery. Centrality in Politics: How Networks Confer Power. 2010. URL: `http://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1007&context=pnconfs_2010`.

[Hu06] Yifan Hu. Efficient, high-quality force-directed graph drawing. 2006. URL: `https://pdfs.semanticscholar.org/be33/ebd01f336c04a1db20830576612ab45b1b9b.pdf`.

[Isi25] Ernst Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 31(1):253–258, February 1925. URL: `https://link.springer.com/article/10.1007/BF02980577`, `doi:10.1007/BF02980577`.

[Jac17] Matthew O. Jackson. Social and Economic Networks: Models and Analysis, 2017. URL: `https://www.coursera.org/learn/social-economic-networks`.

[Kat53] Leo Katz. A new status index derived from sociometric analysis. *Psychometrika*, 18(1):39–43, 1953. URL: `http://www.springerlink.com/index/71033wl12lw13744.pdf`.

[KPFT15] Peter Klimek, Sebastian Poledna, J. Doyne Farmer, and Stefan Thurner. To bail-out or to bail-in? Answers from an agent-based model. *Journal of Economic Dynamics and Control*, 50:144–154, 2015. URL: http://www.sciencedirect.com/science/article/pii/S0165188914002097.

[LPT16] Matt V. Leduc, Sebastian Poledna, and Stefan Thurner. Systemic risk management in financial networks with credit default swaps. 2016. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2713200.

[MSFG14] Rossana Mastrandrea, Tiziano Squartini, Giorgio Fagiolo, and Diego Garlaschelli. Enhanced reconstruction of weighted networks from strengths and degrees. *New Journal of Physics*, 16(4):043022, 2014. URL: http://stacks.iop.org/1367-2630/16/i=4/a=043022, doi:10.1088/1367-2630/16/4/043022.

[PBT17] Sebastian Poledna, Olaf Bochmann, and Stefan Thurner. Basel III capital surcharges for G-SIBs are far less effective in managing systemic risk in comparison to network-based, systemic risk-dependent financial transaction taxes. *Journal of Economic Dynamics and Control*, 77:230–246, April 2017. URL: http://linkinghub.elsevier.com/retrieve/pii/S0165188917300398, doi:10.1016/j.jedc.2017.02.004.

[PCB14] Michelangelo Puliga, Guido Caldarelli, and Stefano Battiston. Credit Default Swaps networks and systemic risk. *Scientific Reports*, 4, November 2014. URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4219172/, doi:10.1038/srep06822.

[PL16] Daniele Petrone and Vito Latora. A hybrid approach to assess systemic risk in financial networks. *arXiv preprint arXiv:1610.00795*, 2016. URL: `https://arxiv.org/abs/1610.00795`.

[PMBMJ+15] Sebastian Poledna, José Luis Molina-Borboa, Serafín Martínez-Jaramillo, Marco Van Der Leij, and Stefan Thurner. The multi-layer network nature of systemic risk and its implications for the costs of financial crises. *Journal of Financial Stability*, 20:70–81, 2015. URL: `http://www.sciencedirect.com/science/article/pii/S1572308915000856`.

[PPT17] Anton Pichler, Sebastian Poledna, and Stefan Thurner. Minimization of Systemic Risk as an Optimal Network Reorganization Problem - The Case of Overlapping Portfolio Networks in the European Government Bond Market. 2017.

[PT16] Sebastian Poledna and Stefan Thurner. Elimination of systemic risk in financial networks by means of a systemic risk transaction tax. *Quantitative Finance*, 16(10):1599–1613, October 2016. URL: `http://dx.doi.org/10.1080/14697688.2016.1156146`, `doi:10.1080/14697688.2016.1156146`.

[SAB+96] Michael HR Stanley, Luis AN Amaral, Sergey V. Buldyrev, Shlomo Havlin, and others. Scaling behaviour in the growth of companies. *Nature*, 379(6568):804, 1996. URL: `http://search.proquest.com/openview/aed4602d3978e18589cbca21ecad4be6/1?pq-origsite=gscholar&cbl=40569`.

[SBB11] Martin Shawn, W. Micheal Brown, and Kevin W. Boyack. OpenOrd: An open-source toolbox for large graph layout. In *Proc. SPIE*,

2011. URL: http://dx.doi.org/10.1117/12.871402, doi:10.1117/12.871402.

[SC13] Kimmo Soramäki and Samantha Cook. SinkRank: An algorithm for identifying systemically important banks in payment systems. *Economics: The Open-Access, Open-Assessment E-Journal*, 7(2013-28):1–27, 2013. URL: https://www.econstor.eu/handle/10419/77858, doi:10.5018/economics-ejournal.ja.2013-28.

[TP13] Stefan Thurner and Sebastian Poledna. DebtRank-transparency: Controlling systemic risk in financial networks. *Scientific Reports*, 3:1888, May 2013. URL: http://www.nature.com/srep/2013/130528/srep01888/full/srep01888.html, doi:10.1038/srep01888.

[TSGo13] Benjamin M. Tabak, Sergio RS Souza, Solange M. Guerra, and others. Assessing systemic risk in the Brazilian interbank market. Technical report, 2013. URL: http://www.bcb.gov.br/secre/apres/Paper%20benjamin%20Tabak.pdf.

[Wir08] Wirtschaftskammer Österreich. ÖNACE - Klassifikation der Wirtschaftstätigkeiten, 2008. URL: https://www.wko.at/service/zahlen-daten-fakten/oenace.html.

[WS98] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, June 1998. URL: http://www.nature.com/nature/journal/v393/n6684/full/393440a0.html?foxtrotcallback=true, doi:10.1038/30918.

List of Figures

Kurzfassung

In dieser Arbeit haben wir Methoden verwendet, die von Diffusion aus der statistischen Physik inspiriert sind, jedoch auf diskrete Netzwerke angewandt werden, um das systemische Risiko in Österreich zu untersuchen. Man spricht von systemischem Risiko, wenn durch einen Marktschock nicht nur direkt betroffene, sondern durch kaskadierende Effekte (möglicherweise große) Teile des Systems außer Kraft gesetzt werden. In einem finanziellen Netzwerk führen Kredite von Banken untereinander zu systemischem Risiko, da Verleihen von Geld dazu führt, dass man dem Marktrisiko anderer Institute ausgesetzt ist. Methoden wie DebtRank wurden in den vergangenen Jahren verwendet, um die möglichen Auswirkungen von systemischen Risiko in einem finanziellen Netzwerk zu analysieren.

In der vorliegenden Arbeit erweiterten wir das untersuchte System von Banken und ihren Krediten untereinander um Firmen und ihre Verbindlichkeiten und Einlagen bei Banken. Basierend auf diesem derart erweiterten Netzwerk verwendeten wir die DebtRank Methode um das systemische Risiko von Firmen zu ermitteln. Außerdem ermittelten wir, wie sich die Bewertung des systemischen Risikos von Banken durch die Erweiterung des untersuchten Netzwerks verändert.

Der Beitrag von Firmen zum systemischen Risiko im erweiterten Netzwerk beträgt 55%. Ein Beschränken des untersuchten Netzwerks auf Banken führt dazu, dass man das gesamte systemische Risiko sogar um etwa 70% unterschätzt. Des Weiteren

können selbst verhältnismäßig kleine Firmen überproportional systemisch riskant sein, wenn ihre Verbindlichkeiten zu einer ungünstiger Auswahl an Banken bestehen. Es könnte daher ratsam sein, Maßnahmen, die versuchen der Entwicklung von systemischem Risiko entgegenzuwirken, auch bei Firmen anzuwenden.

Abstract

We used methods partly inspired from statistical physics to analyze systemic risk in the Austrian economy. Systemic risk is the notion that a shock may not only render directly affected components but (possibly large) parts of the system non-functioning due to cascading effects. In a financial network this systemic risk stems from the exposures of lending institutions to one another. Network measures such as DebtRank were used to estimate the possible impact on the interbank network resulting from a market shock. In this work, we extended the analysis of systemic risk from financial institutions to the whole economy and found that the DebtRank distribution of companies and of financial institutions is qualitatively similar. Companies may have exposures to a disadvantageous set of banks, thus enabling the propagation of an impact through large parts of the network. Even though larger companies tend to have a higher DebtRank, there are firms sharing a very similar DebtRank with total assets that span multiple orders of magnitude. We found that the systemic risk of the companies amounts to 55% of the total systemic risk and that only using the interbank network underestimates the total systemic risk by approximately 70%. The results suggest that the default of a systemically important company can be similarly harmful as the default of a financial institution. It may therefore be beneficial to take companies into account when designing and assessing policies that try to alleviate systemic risk. Possibly even the extension of regulations to systemically risky companies may be advisable.