Shock Propagation in Interacting Economic Networks

András Borsos

July 16, 2021

Supervisor: János Kertész
Co-supervisor: Ádám Zawadowski

A Dissertation Submitted in Partial Fulfillment of
the Requirements for the Degree of Doctor
of Philosophy in Network Science

Central European University
Budapest, Hungary

CEU eTD Collection

CEU eTD Collection

András Borsos: *Shock Propagation in Interacting Economic Networks*, © 2021
All rights reserved.

Researcher declaration

I, András Borsos, certify that I am the author of the work *Shock Propagation in Interacting Economic Networks*. I certify that this is solely my own original work, except where I have clearly indicated, in this declaration and in the thesis, the contributions of others. The thesis contains no materials accepted for any other degrees in any other institutions. The copyright of this work rests with its author. Quotation from it is permitted, provided that full acknowledgement is made. This work might not be reproduced without my prior written consent.

Statements of inclusion of joint work

I confirm that Chapter 2 and 3 are based on a paper that was written in collaboration with Martin Stancsics. My contributions to the paper were the following: I conceived the main conceptual ideas and devised the project. I obtained the data sets used for the research project. I performed the data cleaning and the analysis of the ownership network of Hungary. The cleaning and the preparation of the firm-level supply transaction data, as well as the analysis of the production network was performed together with Martin Stancsics. I wrote the manuscript to which Martin Stancsics provided critical feedback. The tables and figures in Chapter 2 and 3 are reproduced from the paper.

Signature of the PhD Candidate:

April, 2021

Signature of Martin Stancsics, endorsing statement of joint work:

April, 2021

CEU eTD Collection

COMPLEXITY
SCIENCE
HUB
VIENNA

Agreement on authorship contributions for the usage of joint research contributions in the dissertation of András Borsos

By signing this authorship agreement and contribution form, authors agree on the following points.

The core part of the algorithm for calculating systemic risk in the Hungarian buyer-supplier network (production network) was developed at the Complexity Science Hub Vienna (CSH) with specific contributions of Christian Diem and Stefan Thurner. The algorithm refers to recursive update equations, which model how the production level of the entire buyer-supplier network changes in response to an initial shock faced by firms. This includes taking into account downstream contagion from suppliers to buyers – inducing a drop of the production level due to a lack of inputs – and upstream contagion from buyers to suppliers – inducing a drop in the production level due to a lack of demand. Specifically this includes modelling a heterogeneous production process by accounting for different types of production functions for companies and how they affect the shock propagation. This also includes the separation between crucial and non-crucial inputs for the production functions and how they lead do different levels of shock propagation.

In a joint research project the authors, in alphabetical order, András Borsos, Christian Diem, János Kertész, Tobias Reisch, Stefan Thurner use this algorithm for analyzing the systemic risk in the Hungarian buyer-supplier network (production network).

The core algorithm was extended by András Borsos, Christian Diem and Tobias Reisch during the research stay of András Borsos at CSH. This extension specifically included the introduction of a supplier replaceability factor, specifying how easily a supplier can be substituted based on the suppliers' dynamic market share in the network for the goods it produces.

The authors agree that András Borsos can present the results of the joint research project in a chapter in his dissertation thesis. For this András Borsos must cite the joint research project either, as a technical report, preprint or research article (whichever is applicable) as the underlying work of this dissertation chapter and can highlight the supplier replaceability factor – including the empirical sensitivity analysis with respect to the introduction of this factor – as his original contribution in his dissertation thesis.

Christian Diem	Tobias Reisch	András Borsos
János Kertész	Stefan Thurner	23.10.20 Date

Enclosed: Technical report of joint research project

CEU eTD Collection

I confirm that Chapter 5 is based on a paper that was written in collaboration with Bence Mérő. My contributions to the paper were the following: I conceived the research idea and obtained and processed the supplier network data. I conceptualized and implemented the real economy feedback block of the model. I performed the econometric estimation of the parameters for the feedback mechanisms. The development and the programming of the banking system block of the model, furthermore, the application of the model for liquidity stress testing are the results of close collaboration with Bence Mérő. I conceptualized and implemented the applications for SIFI identification and real economy shock assessment. I wrote the manuscript to which Bence Mérő provided critical feedback. The tables and figures in Chapter 5 are reproduced from the paper.

Signature of the PhD Candidate:

April, 2021

Signature of Bence Mérő, endorsing statement of joint work:

April, 2021

CEU eTD Collection

Abstract

The economy can be considered a complex system, in which phenomena such as prosperity and crisis are the results of feedback-based interactions of many heterogeneous components. One of the most important factors in the emergence of economic outcomes is very often an underlying spreading process among these constituent parts, which facilitates the propagation of information, sentiments, risks, resources or losses. To be able to analyse these mechanisms in a formalized way, we need to use a wide range of modeling techniques. Among these, network science is one of the most promising approach, which can equip economists with the necessary tools to build models capable of capturing the intricacies of complex systems. This thesis contributes to these endeavours in a threefold way: (i) by providing insights into the hidden structure and the unique traits of micro-level firm network data; (ii) by proposing a model of shock spreading in firm-level production networks; (iii) and lastly, by offering a novel way of modeling feedback channels between the financial sector and the real economy in the context of interacting economic networks.

In order to be able to investigate firm networks, I obtained access to sensitive datasets about the ownership links and the supplier connections among Hungarian firms. This way, it has become possible to construct the multi-layer representation of the Hungarian firm network, which enabled us to gain insight into its previously unobserved structure. Network analysis provided suitable techniques to explore several topological traits on micro-, meso-, and macro-scale as well, which can be conducive to contagious mechanisms via supplier links. Furthermore, it was also possible to assess the significance of economic entities regarding the extent to which they can influence and control the economy via their ownership relations.

These pieces of information also enabled the simulation of shock propagation in the production network. The granularity of the data made it possible to rectify several shortcomings of industry-level supply chain analyses. The proposed model features heterogeneous production functions at the firm-level, differentiation in the importance of input types and replaceability of defaulting suppliers. With these advancements, the model is capable of quantifying short-term damages after supply chain disruptions, assessing the systemic risk of individual firms, and testing countermeasures, which has relevance for policy making.

Lastly, I propose a computational model of contagious mechanisms in the banking system complemented with feedback channels towards the real economy. The framework incorporates the interactions between the network of banks and the network of firms which systems are linked together via loan-contracts. The model has been embedded into the liquidity stress test of the Central Bank of Hungary, and the results proved the importance of the real economy feedback channel, without which systemic risks could potentially be severely underestimated. To illustrate the versatility of this modeling framework, two further applications have been elaborated. The model can be used to identify systemically important financial institutions (SIFIs), furthermore, it is also suitable to assess the financial stability impact of shocks originated in the real economy.

CEU eTD Collection

Acknowledgements

I would like to acknowledge and thank several people who contributed to this work in various ways. First and foremost I am extremely grateful to my supervisor, János Kertész. I could rely on his continuous support and guidance at every stage of my Ph.D. study. I am also thankful to my co-supervisor, Ádám Zawadowski for his insightful comments and suggestions.

I would like to offer my special thanks to my co-authors: Christian Diem, János Kertész, Bence Mérő, Tobias Reisch, Martin Stancsics and Stefan Thurner, who contributed immensely to the research presented in this thesis. I am also grateful to Sanjay Kumar and Eszter Borsos for their help to improve the manuscript.

I would like to thank my fellow Ph.D. students and the whole community of the Department of Network and Data Science for providing a friendly and inspiring environment. I am also grateful to my colleagues at the Applied Research and Stress Testing Department of the Central Bank of Hungary. I had the opportunity to work here with excellent experts and researchers, whose help was an invaluable contribution to my work. Furthermore, I am thankful to the people at the Complexity Science Hub (Vienna) and at the Institute for New Economic Thinking (Oxford) for hosting me. My time as a visiting student was not only fruitful professionally, but it was also an unforgettable personal experience.

Lastly, I am grateful to my family, my friends and my partner Erika for believing me and encouraging me during the whole time.

CEU eTD Collection

Contents

CEU eTD Collection

CEU eTD Collection

CEU eTD Collection

1 Introduction

The 2008 economic crisis shed light on a distinctive feature of the financial intermediary system: banks and other financial institutions are constituents of a multi-layer network, in which their interactions and feedbacks create non-linear processes. Since the recognition of this *complexity*[1] as an intrinsic and influential characteristic which requires special attention has become widely accepted, a vast amount of research (e.g. Gai and Kapadia (2010), Haldane and May (2011), Caballero and Simsek (2013), Elliott et al. (2014), Acemoglu et al. (2015), etc.) was conducted on network-based contagious mechanisms in the financial system. Beside the success of this research direction, a more general paradigm shift has also been ignited (Farmer & Foley, 2009), and the role of networks started to increase in other areas in economic research as well. Some of the most notable examples are game theory (e.g. Jackson and Zenou (2015), Galeotti et al. (2010)), trading in networks (e.g. Kranton and Minehart (2001), Choi et al. (2017)), labor market (e.g. Calvo-Armengol and Jackson (2004), Beaman (2012)), peer effects (e.g. Ballester et al. (2006), A. Banerjee et al. (2013)), but maybe the most rapidly growing area is nowadays the analysis of production networks (e..g. Acemoglu et al. (2012), Oberfield (2012)). In my research I attempted to further enrich the literature of network-based economic analysis by contributing to two topics which are currently in the center of attention in economics: (i) shock propagation in supplier networks and (ii) interactions between the real economy and the financial sector.

Regarding the first topic, the current COVID-19 crisis drew increased interest to shock propagation among firms as the instability of international and national supply chains has made their vulnerability obvious. The epidemic situation highlighted the strong dependence of firms on their suppliers and the fact that the non-availability of inputs inevitably leads to halts in productions, which can spread to other suppliers and customers. Similarly, the default of customers puts firms at risk of losing revenues and thus reducing their own demand for inputs. Inherently, the amplification of the initial economic shock in the supplier network can lead to cascading failures of firms along supply chains, which means that far more firms can be affected indirectly than one would expect without network effects. Although network analysis seems to be a suitable tool to tackle this phenomenon, there is a severe hindrance researchers often face in this research topic, namely, that in most countries these networks are not observable, and only coarse-grained, industry-level connections can be seen. To overcome this challenge, I obtained access to extremely rarely available supplier transaction information among almost all Hungarian firms, on which it was possible to superimpose their ownership links as well. As both the supplier and the owner-

[1] In this sense a complex system is not merely a synonym for a complicated, large, sophisticated structure. Complexity is a scientific theory which asserts that some systems display emergent phenomena that are completely inexplicable by any conventional analysis of the systems' constituent parts. The source of complexity is usually assumed to be the non-linear, feedback-based interaction of many heterogeneous components (Thurner et al., 2018).

CEU eTD Collection

ship layers are considered to be among the most influential shock-transmitting media, this data is ideal to gain insight into previously unobserved topological drivers of spreading processes. These pieces of information also enabled the building of a model of shock spreading on firm-level production networks, which makes it possible to rectify several shortcomings of industry-level models.

The second aspect of this thesis's contribution is concerned with embedding the above described firm network into a model which exhibits the interconnectedness of the financial sector and the real economy. Most of the network-based economic models focus only on one isolated economic system, and there is only very limited attention on the interactions between different spheres of the economy. Most importantly, the connections between the real economy and the financial sector – which topic is of central importance in mainstream economics at least since 2008 – is only discussed in very few network-based analyses, e.g. Gatti et al. (2010), Riccetti et al. (2013), Vitali et al. (2016), Silva et al. (2018), Gurgone et al. (2018) and Popoyan et al. (2020). A further problem is that most of these models have only a small number of simulated agents and they do not use empirically observed networks. This means that they are suitable to demonstrate and analyze a given mechanism of interest, but their validity is very limited in the case of actual policy analysis or in simulating realistic shock scenarios. To contribute to the development of the field from this point of view, I propose a novel way of modeling feedback channels between the financial sector and the real economy by using a computational microsimulation framework. This model incorporates the interactions between the actual network of banks in Hungary (exhibiting contagion mechanisms among them) and the almost complete network of firms (transmitting shocks to each other along the supply chain) which systems are linked together via actually observed loan-contracts. Additionally, the last part of the thesis offers policy relevant illustrations of how the feedback mechanisms in these coupled networks could amplify the losses in the economy way beyond the shortfalls expected when we consider the subsystems in isolation.

In the following two subsections I provide a more detailed description of the background of these research directions, while the last part of the introduction will give an overview about the structure of the dissertation.

1.1 Supplier networks

In the past decade we experienced a vast surge in interest towards modeling and analyzing interdependencies among companies. The structure of these networks is a key element in understanding the governing forces behind any kind of spreading phenomena among firms. The first part of this dissertation offers a general, descriptive exploration of the topological structure of multi-layer firm networks using Hungarian data. Although the topology of the underlying graph might play a different role for the various types of shocks, this work is relevant for a wide range of applications, such as productivity spillovers (Liu et

CEU eTD Collection

al. (2000), Gorg and Strobl (2001)), spreading of financial shocks (Demir et al. (2018), Costello (2020)), or upstream and downstream supply chain disruptions (Barrot and Sauvagnat (2016), Carvalho et al. (2016)).

1.1.1 Recent advancements in supplier network analysis

Examining the system of interfirm connections has been present in the economic literature at least since Leontief's seminal work on the structure of the American economy (Leontief, 1951). The roots of the recent increased enthusiasm in using more fine-grained, firm-level disaggregation are twofold: (i) developments in data availability and (ii) new conceptual innovations.

As a part of the universal pattern of increased accessibility to micro-level data, in some countries it became possible to obtain comprehensive datasets about firm-level connections. Previously, researchers who wanted to consider firm connections in their analyses could use either a sample of the given network or a higher aggregation level (e.g. industry or country). Both approaches turned out to suffer from serious limitations. When measured on a sample, even the most elementary characteristics of networks (e.g. the density or the average degree) require non-trivial corrections, which can be very different depending on the sampling method (Granovetter, 1976), while more sophisticated analyses on samples are hindered severely by potential distortions (Frank, 1971). The other option is to use a completely observed, but aggregated system, however, this approach has other caveats. One seemingly obvious drawback is that during the process of aggregation we lose information not only about the heterogeneity of the actors, but also about their connections among each other. However, it was not evident at all for a long time in economics (at least from the point of view of macroeconomics) whether disregarding the observation of firm-level events and characteristics is relevant or not.

As this debate flared up and gained a lot of attention recently, it has led to the second, more theoretical branch of factors giving popularity to granular firm network analysis in economics. An important milestone in the development of this field was the rejection of the traditional argument of Lucas Jr (1977) about the diversification of shocks in the economy. The former consensus was that firm-level idiosyncratic events do not have any influence on the macroeconomic scale as they cancel each other out based on the law of large numbers. However, Gabaix (2011) and Acemoglu et al. (2012) showed that due to the heterogeneity of the firms and the topology of their connections, stress events of the largest companies cannot be offset by smaller firms even if the shocks are uncorrelated[2].

A second reasoning supporting the irrelevance of network effects was pro-

[2] They showed that the distribution of company size (and also the direct and indirect demand towards a given company's products) could be described well using power-low distribution, in which there is a relatively high probability of extremely large observations. Depending on the exponent of the distribution, the assumptions of the law of large numbers could be violated.

CEU eTD Collection

posed by Hulten (1978). This argument claimed that the Domar weights (sales as a share of GDP) of firms (or industries) are sufficient statistics to assess the aggregate total factor productivity (TFP) impact of micro-level TFP shocks. Although this statement seems to be intuitively questionable[3], it has been serving as a justification for a long time to ignore granular connections in the production network. In the past decade, however, new results questioned this argument. Most notably, Baqaee and Farhi (2019) pointed out that even the Domar weights themselves can be influenced by the TFP shocks, hence, second-order effects should also be taken into account in the shock propagation process. Furthermore, Baqaee and Farhi (2020) proved that even if one considers only first-order impacts, the network structure can have an important role if there are frictions in the economy.

In parallel to these theoretical developments, the empirical literature on supply chain contagion gained popularity as well. In these papers researchers try to measure the extent of spreading of some exogenous event (e.g. natural catastrophes or policy shocks) on the supplier network. For example Bimpikis et al. (2018) and Bimpikis et al. (2019) have showed, that disruptions in the supplier network can result in suboptimal network formation which can amplify systemic risks. Baqaee (2018) showed in a general equilibrium model that shock propagation can be further amplified by the interconnectedness between industries. Luo (2019) establishes linkages between firms using both the production network and financial links due to delays in input payments, and shows that this multiplex network leads to the propagation of financial shocks in both upstream and downstream directions. Carvalho et al. (2016) provided further support for this mechanism using empirical data about the Great East Japan Earthquake. Barrot and Sauvagnat (2016) also uses natural disasters for the identification of firm-level shocks, and they found that suppliers can trigger considerable output losses for their customers. Further examples of supply chain disruption analyses are Demir et al. (2018) and Boehm et al. (2019), but Carvalho and Tahbaz-Salehi (2018) and Bernard and Moxnes (2018) offer reviews of the broader literature on production networks.

These results about the unexpected directions and rate of shock spreading proved that the interconnectedness of different economic actors is a vastly influential aspect of many processes of the economy. However, this observation cannot be simply interpreted as more connections mean higher potential for any kind of spreading phenomena. E.g the seminal work of Elliott et al. (2014) showed, that the contagion potential in a financial system depends on the network structure in a non-monotonic way: Diversification (having more economic partners) increases the size of the connected component in the network initially, but after a while this process actually leads to more diversified, more resilient systems. Similar logic can be observed in the case of dual source strategies in supply chain management, which means that firms often establish more than

[3]Consider TFP shocks to a large retail company and to electricity production. Both can have similar sales share in the GDP, but one would correctly expect that a shock to the electricity company would result in much more severe system-wide economic damage.

CEU eTD Collection

one link for a given input to decrease their sensitivity to disruptions in the supply chain.

1.1.2 Firm-level network data analysis

The developments described so far triggered a new wave for economic network analysis as network theory offered a novel way of thinking about the structure of the economy by representing it as a complex system. By now the network-based approach has become part of the mainstream in several areas. Nevertheless, the first step before one could integrate these - now sometimes almost fully observable - networks into economic models should be the thorough exploration of the data. Most importantly we have to examine the topological structure in high resolution, which is the distinguishing feature compared to the previous aggregate observations. However, one can find mainly theoretical works aiming to describe these systems (e.g. Benhabib et al. (2010), Dutta and Jackson (2003), Goyal (2012), Jackson (2010)), and just a very limited number of empirical papers.

In one of these works, Watanabe et al. (2015) offers a detailed analysis of trade connections of 400,000 Japanese firms. Although it is still about a sample of firms in the country, they were the first to analyze a supplier network of this extent. Dhyne et al. (2015) describes the production network of Belgium from the point of view of its integration into the world trade network. This work was one of the first in a series of papers about the Belgian production network (Kikkawa et al. (2019), Magerman et al. (2016), Tintelnot et al. (2018)). Additionally, Demir et al. (2018) used the Turkish, while Kumar et al. (2020) considered the Indian supplier network in their work, however, all of these research projects focused on economic questions with a higher abstraction level and not on the network itself.

Another branch of empirical research considers ownership relations among companies, which is also a recently often examined layer of firm networks. These papers usually use the Orbis database[4] to analyze the global ownership network of companies: Using the same or highly overlapping data sources Glattfelder (2013) offered a methodology to extract the backbone of the global ownership network, Vitali et al. (2011) showed that there is a very high concentration in this network with a group of core companies, Vitali and Battiston (2011) examined the ownership structures' embedding in the geographical space, Vitali and Battiston (2014) explored the community structure of the global ownership network, Heemskerk and Takes (2016) described the multipolar nature of the global political economy, and Garcia-Bernardo et al. (2017) tried to identify offshore financial centers. Some other layers of corporate connections were studied as well, however, often only on smaller samples of firms. Zajac and Westphal (1996), Battiston et al. (2003) and G. F. Davis et al. (2003) looked

[4]Orbis is a company database provided by Bureau van Dijk (which is a Moody's Analytics business information publisher).

CEU eTD Collection

at the network of interlocking board members and decision makers. Innovation dynamics have also been considered on networks of R&D partnerships e.g. by Tomasello et al. (2017), while there are numerous studies focusing on stock price correlation-based network of listed companies (e.g. Tumminello et al. (2010)).

Most of the papers until now dealt with only a single layer of corporate networks. A very recent exception is de Jeude et al. (2019) whose study features four layers: ownership links and board member overlaps among a very large sample of firms; furthermore R&D collaborations and stock correlation between listed companies. In this research I would like to contribute to this direction of the literature by attempting to unfold the non-trivial characteristics of the multi-layered network of firms using Hungarian data. At the Central Bank of Hungary, I was able to create a uniquely rich dataset by having access to supplier transaction information among firms, on which it was possible to superimpose their ownership links as well in the period between 2014-2017. The supplier information is coming from firms' VAT reports collected by the National Tax and Customs Administration of Hungary. In this data one can observe trade links among Hungarian firms where the tax content of the transactions between two firms exceeds EUR 3000 in the given year. Considering the ownership data I have used the OPTEN[5] dataset of more than 400 000 Hungarian firms. To further enrich the scope of our analysis, other micro-level datasets of the Central Bank of Hungary has been merged to the network data. These made it possible to use the detailed characteristics of firms (coming from their balance sheets and profit and loss statements) as additional attributes of the nodes.

Both the supplier and the ownership layers are among the most significant shock-transmitting media; thus, this dataset is ideal to explore the topological origins of the above described spreading phenomena. As these data are usually not collected for research purposes, and economists are often unfamiliar with the specific and unique characteristics of network data, this work has an additional contribution by describing the significant amount of preprocessing which is necessary in order to use these pieces of information in line with the economic expectations and interpretations. This way, the methodological approach of the analysis consists of elements coming not only from economics, but also from the network science literature (M. Newman (2018), Barabási et al. (2016)), which can provide us with suitable tools to explore the underlying structure of the firm network on micro-, meso- and macro-scales as well.

As a first step in the analysis, it was necessary to consider the ownership structure of the Hungarian economy in order to distinguish transactions within and between ownership groups. As this system is much sparser than the supplier layer, and it consists of many small components, it is not possible to analyze it on the level of a giant component. However, one can still measure its most important characteristics and gain insight into the typical motifs in the ownership structure of firms.

[5]OPTEN is a Hungarian firm-level data provider company.

CEU eTD Collection

Furthermore, the ownership data enables the exploration of the network of economic actors from another angle as it conveys invaluable information about the direct and indirect influence of the observed entities. Based on a measure of control proposed in Chapter 2, one can also analyze the distribution of control in the economy. Based on this analysis, more than 40% of the control is associated with the top 100 owners in the Hungarian economy. This investigation was also carried out at more aggregated levels revealing the role of different groups formed along numerous dimensions, such as the nationality, the legal category or the HQ location of the owners.

Regarding the supplier layer, the analysis of this system identified several topological patterns of the production network which can be responsible for facilitating contagious processes: (i) despite the low density of the network we can identify a giant component which encompasses more than 94% of the nodes. (ii) The average shortest path length among the firms in this component is around five[6], which indicates small-worldness in the network. (iii) The long-tailed degree distribution ensures the presence of hubs, that can be key actors in spreading shocks. (iv) Contagions can be further promoted by micro-level motifs: there is an unexpectedly high probability of reciprocal dyads and closed triangles, which can amplify shocks via local feedback loops.

One could gain further valuable insights about the system by exploring its meso-level configuration. This type of examination identified a well-defined and occasionally overlapping community structure, which reflects closely the production chains of different segments in the Hungarian economy. This grouping allows us to assess firms' capacity to connect communities, which measure can be used as a proxy for shock transmitting ability between the otherwise separated chains of production. The results showed that firms in the transportation and infrastructure sectors, and firms with high productivity and high export rate have the most important role in connecting different blocks of the economy. In addition to these topological traits, the network also demonstrated strong homophily[7] based on several firm attributes, most notably in the case of productivity, profitability and geographical location. However, these traits are much weaker in terms of separating the network than the supply chains identified by the community detection procedure.

1.1.3 Modeling shock spreading in firm-level supplier networks

The current advent of granular firm-level data can uncover the supply chains of entire nations. This allows researchers for the first time to study the propagation of economic shocks on the level of firms, instead of on the sector level as usually done in input–output analyses. One of the first attempts into this direction was done by Magerman et al. (2016), who study micro-level shocks and

[6] Here we did not consider the directions of the links as shocks can spread in both directions depending on the process.

[7] The tendency of the formation of links between similar nodes.

CEU eTD Collection

their effects on aggregate fluctuations. They found that 90% of the volatility is driven by the largest 100 firms in the Belgium supply chain network. Inoue and Todo (2019b) uses a more recent version of the Japanese supply chain with 1.1 million nodes and 5 million buyer–supplier relationships and introduce a weighting scheme based on sector-level input–output tables and firm revenue data. They adapt the model of Hallegatte (2008) and introduce heterogeneous inventory levels and a different rationing mechanism for supply shortages. Inoue and Todo (2019b) finds that shocks decreasing the production of only 10,000 nodes to 50% of the original level propagate through the network and amplify the initial shock by a factor of 16 after 30 days and a factor of 100 after 200 days. Furthermore, they find considerably larger shocks on the firm-level supply chain network compared to the input–output table shocks on sector level. Inoue and Todo (2019a) used the same method on the 2011 earthquake in Japan as an initial economic shock and showed that their model predicts indirect shocks in a magnitude consistent with the actual empirically measured shocks in value added. Fujiwara et al. (2016) applies the DebtRank algorithm onto the supply chain network and finds a nonlinear relationship between firm size and systemic losses measured by DebtRank, i.e. that large firms cause disproportionately large cascades.

Models for real economy shock propagation have also been used to analyse the current COVID-19 crisis. Inoue and Todo (2020) apply their model to simulate the effect of a lock-down of the Tokyo metropolitan area due to COVID-19. They find that for a 30 day lock-down the indirect shock on the other provinces is twice the size of the initial shock. Pichler et al. (2020), Pichler and Farmer (2021) and Pichler et al. (2021) study the effect of different sector lock-downs in the UK economy with a sector-level input–output model in spirit of Hallegatte (2008) and Inoue and Todo (2020). They managed to predict the GDP decreases of Q2 2020 more accurately than traditional GDP forecasting models, which underlines the importance of taking real economy shock propagation into account.

Overall the literature finds ample evidence that economic shocks spread in supply chain networks; indirect shocks outweigh the initial direct shocks severely in size; and that sector-level estimates of second round effects highly underestimate indirect shocks in comparison to firm-level supply chain networks. Chapter 4 of this dissertation presents a supply chain contagion model in this spirit. I contribute to the literature by introducing several novel features which grant a higher level of plausibility in the quantification of short term damages after shocks in the supplier network. The most important characteristic is that the propagation of exogenous initial economic shocks are based on the Hungarian firm level production network first described by Borsos et al. (2020). With the described data sources each firm can be prescribed with a production function that specifies the amount of outputs a firm can produce with its suppliers' inputs. This way, the model can quantify the systemic risk of firms by simulating the effects of distinct upstream and downstream spreading mechanisms on the production network in response to the firm's failure. The network dynamics

CEU eTD Collection

are based on an edge update scheme where a loss of inputs of a firm causes a reduction of its own production and in turn causes a loss of inputs for other firms.

This methodology addresses three usual shortcomings of sector level analyses. First, the analysis of the data showed that even within fine grained industry classifications (NACE 4) firms tend to have very heterogeneous inputs. This can lead to inaccurate results when using them for assessing shock propagation in production networks. Second, this is especially true if the crisis scenario does not affect all firms within a sector to the same extent (for example in the current COVID-19 crisis). The proposed model takes this into account and yields different cascades for shocks which would appear to be the same at industry level, but are distributed differently among firms within an industry. Third, in contrast with sector level models, each firm in the data has a specific production function based on its input vector and industry classification. The model uses a combination of linear and Leontief production functions, and the varying criticality and replaceability of inputs has also been taken into account. Overall, these features improve the accuracy of analysing how shocks spread in production networks, and consequently foster a more plausible estimation of the effects of crisis scenarios.

In the first application of this model we showed how large the systemic risk is that single firms pose to the entirety of the firm-level production network. The simulations show that only less than 100 firms have the potential to destroy more than five percent of the national production network, and hence, pose a considerable threat to the overall economy. However, even the default of a single one of these companies can affect 21% of the production in the system. (Using different production functions it was also possible to estimate bounds to the damage.) This indicates, that the knowledge of the systemic riskiness of single firms is crucial for understanding and preventing potentially large failure cascades in these networks.

1.2 Interactions between the real economy and the financial sector

The unexpected cascading spillovers in the global economy after the 2008 crisis fostered the emergence of network-based simulations as a popular modeling framework in financial economics (Elliott et al. (2014), Acemoglu et al. (2015)), but this recognition so far resulted mainly in numerous analyses about contagion channels only within the financial system. However, besides the complexity of this sector there is also another, more conventional line of reasoning which justifies the special role of the financial intermediation industry: it is connected to all other industries, which puts banks in a special position again from the point of view of shock propagation in the economy. This consideration (among

CEU eTD Collection

others[8]) led in the first place to the regulatory frameworks, which have been much stricter than one can experience in almost any other industry even before the crisis.

There are only a few papers in the literature pursuing the synthesis of these two regulatory considerations about banks (e.g. Gatti et al. (2010), Riccetti et al. (2013), Vitali et al. (2016), Gurgone et al. (2018) and Popoyan et al. (2020)), and these models usually use a small number of artificial agents to analyze a given mechanism in interest. Although there are also numerous papers using empirical networks both about the financial systems and about supply chains, the topic of interconnected empirical economic networks (although it was even listed as an objectives of the FuturICT project (Farmer et al., 2012)) remained so far largely unexplored. Results obtained in theoretical models suggest that the interconnected nature of networks causes qualitatively different behavior and alters the robustness of a complex system compared to the mere aggregation of its subsystems (Buldyrev et al. (2010), Leicht and D'Souza (2009)). Similarly, the feedback mechanisms between economic networks could amplify the losses beyond the shortfalls expected when we consider the interacting systems in isolation. Consequently, one can assume that to accurately assess financial systemic risks we need to consider the feedback channels between the interacting economic networks as well.

1.2.1 Shock propagation in the banking system with real economy feedback

The fifth chapter of the thesis demonstrates some of the consequences of the above mentioned intricacies on the financial stability of an economy by proposing a microsimulation based framework which is suitable to capture contagious mechanisms in an interconnected system of economic networks. More specifically, the model focuses on the interactions between the network of banks (exhibiting contagious mechanisms among them) and the network of firms (transmitting shocks to each other along the supply chain) which systems are linked together primarily via loan-contracts. This high resolution representation of the economy grants higher validity of the simulation results, which makes this tool potentially suitable for versatile policy purposes. According to my knowledge, this is the first model which integrates all the above mentioned mechanisms by using microsimulation jointly on empirical firm network data and the banking system.

This modeling framework consists of four blocks: (i) contagions in the banking sector, (ii) modeling credit supply shocks for firms, (iii) assessing the amplification of these shocks in the production network and (iv) estimating banks' losses on their corporate loan portfolios. The first block is basically a banking

[8]There are other characteristics of the financial sector which can justify its unique regulation as it operates in a highly leveraged way compared to other sectors and information asymmetries are present on multiple levels.

CEU eTD Collection

system contagion model with channels for interbank losses, liquidity hoarding and fire sales effects, however, it also incorporates several balance sheet adjustment mechanisms to take into account the realistic behavior of banks in a stress scenario. This feature makes it possible to expand the propagation of distress towards the real sectors by acknowledging the procyclicality of the banking sector. Furthermore, additionally to the capital adequacy ratio (CAR) default condition, the liquidity coverage ratio (LCR) is also included to account for defaults due to liquidity insufficiency. The other three blocks of the model are treated together in a spatial econometric model which gives estimates for the probability of default on corporate loan contracts. To carry out this estimation I borrowed tools from another stream of economic literature, which deals with shock propagation along the supply chain in production networks. As this research project does not aim to build a general economic model, only those channels between the banking system and the real economy has been elaborated, which seem to be the most influential from the point of view of financial stability[9].

1.2.2 Related literature on modeling interacting economic networks

The proposed model is most closely related to papers which connect the banking system and firms using loan contracts but do not consider the production network. One of the first attempts at this was done by Lux (2016), which study considered shock propagation via firms with multiple bank connections (similarly to the concept of contagion through overlapping portfolios). If a bank defaulted, the resulting credit crunch could force firms dependant on the banks' loan into bankruptcy. These firms then caused losses to their other bank connections. In their simulations they found that the joint exposures to counterparty risk in corporate lending is actually more important in the contagious spread of defaults than the interbank lending channel. A model in similar spirit was done by Silva et al. (2018), but in this case the simulations were run using empirical data as well. This paper extended a variant of the DebtRank model (Bardoscia et al., 2015) to incorporate lending connections between banks and firms to create additional channels of shock propagation (but without including links among in the firm network). They showed that without taking the links between the financial and the real sectors into consideration one can severely underestimate systemic risks. Recent developments in the European Central Bank also include real economy feedbacks within their stress testing framework (Budnik et al., 2019). In their work, they used a DSGE model to investigate how deleveraging the banking system affects the real economy, which effect feeds back into the aggregated macroeconomic variables. Additionally, they also consider cross-sectoral spillovers due to losses on claims of distressed banks, and then due to the equity holdings between sectors in the real economy (Dees & Henry, 2017). However, the DSGE approach entails some disadvantages: it

[9] Broer et al. (2010) offers a comprehensive summary of several other potential interactions.

CEU eTD Collection

produces only macro-level outcomes without revealing the heterogeneity of the economic actors and the role of the distinct components in the contagion along production chains. In a further related project Gross and Siklos (2020) consider spillovers of financial shocks in the real economy without articulating a feedback component. They are using network-based econometric tools to estimate the transmission of bank and sovereign risks to the non-financial corporate sector based on CDS spreads. Furthermore, some papers depict connections between the financial and the real sector in the form of indirect interconnectedness among banks via exposures to common asset holdings (Caccioli et al. (2014), Duarte and Eisenbach (2018), Cont and Schaanning (2019), Roncoroni et al. (2019)).

Additionally, there are also theoretical models of interconnected networks, which can give relevant insights into the behaviour of interacting economic systems. Buldyrev et al. (2010) found that a broader degree distribution can amplify the vulnerability of coupled systems to random failures, which is opposite to how a single network behaves. Furthermore, Leicht and D'Souza (2009) showed that the percolation threshold in an isolated subnetwork can be significantly lower when edges to other networks are also present. Although these results were obtained in theoretical models with a very high abstraction level, they suggest that accounting for the interconnected nature of economic networks can be crucial in systemic risk assessment.

Papers focusing solely on financial networks are also relevant to my work. The banking system block of the proposed model is most similar in its spirit to Georgescu (2015), Idier and Piquard (2017), Covi et al. (2019) and Coen et al. (2019), however, there is a vast amount of related papers concerning interbank contagions. E.g. Rogers and Veraart (2013) and Dietrich and Hauck (2020) focused on shock propagation in interbank networks, Gai and Kapadia (2010) and Gai et al. (2011) dealt with contagion through funding risk, and Bargigli et al. (2015), Caccioli et al. (2015), Poledna et al. (2015) and Montagna and Kok (2016) conducted research on contagion on multi-layer networks of banks. Upper (2011) and Jackson and Pernoud (2020) offer exhaustive summaries about further potential contagion channels. There are several other influential papers, which served as a starting point for these research projects: Furfine (2003) offered one of the first algorithmic solutions to the contagion mechanisms on a bank network, Eisenberg and Noe (2001)[10] managed to deal with the simultaneity problem of accounting for defaults and losses in a network, Battiston et al. (2012) offered a widely-used centrality measure to identify systemically important institutions and Barucca et al. (2016) improved on handling ex-ante valuation of claims among constituents of financial networks.

In order to create the model's microsimulation environment, several detailed datasets at the National Bank of Hungary and at the National Tax and Customs

[10] Csóka and Herings (2018) shows a decentralized approach for the clearing in Eisenberg and Noe (2001), generalized to the discrete setup, while Csóka and Herings (2020) offers an axiomatization for the clearing process.

CEU eTD Collection

Administration have been used. Most notably balance sheet data of the Hungarian banks and firms, bilateral exposures at the interbank market, information about the investment portfolios of banks, details of loan contracts between banks and firms, and transaction level data about the supply chain connections among firms[11]. This kind of data availability is not typical in the literature. Links between economic entities are very often confidential information and they are rarely accessible for academic institutions. There are only very few countries where fine-grained production network data is available, and the situation is not much better in the case of financial networks. There is a vast literature dealing with the reconstruction of the topological structure of financial institutions using only aggregate observations. An often used procedure for reconstruction is the Maximum Entropy (ME) approach (Upper and Worms (2004), Elsinger et al. (2013)). Distributing each bank's total interbank lending as evenly as possible also means that ME results in an unrealistic, almost complete network. Drehmann and Tarashev (2013) enhanced ME by adding random perturbations to the maximum entropy output matrix to generate results with higher concentration mimicking more closely the sparse structure of empirical networks. Another variant of ME is the Minimum Density (MD) approach developed by Anand et al. (2015), which method creates an interbank lending network using as few links as possible by imposing a cost on link formation. Mastrandrea et al. (2014) take the degrees of the nodes into account as well during the ME allocation. An alternative technique was applied by Baral and Fique (2012), whose paper used copulas to construct the interbank lending network. ME was also applied to recover input-output matrices by Golan et al. (1994), however, reconstruction or more generally even the use of granular entity-level linkages is much less prevalent in the case of firms than in the banking system. The accessibility of all the above listed data sources makes it possible to avoid the drawbacks of these methods and gain a more plausible picture of these networks.

1.2.3 Policy applications

As a first application, the model has been embedded into the Hungarian Central Bank's liquidity stress test (which is calibrated to the 2008 crisis). The results of the simulation indicate that in the Hungarian banking system the magnitude of feedback-based losses on the non-performing loan portfolio coming from the firm network is similar or in some cases even more severe than the losses caused by the usual firesales and interbank contagion channels. Additionally, the introduction of real economy feedbacks changed fundamentally the distribution of the losses among banks. The new contagion channels also made the interaction between solvency and liquidity problems more emphasized: some banks became unable to comply with the solvency criterion even in the case when only liquidity shocks were present in the stress scenario. By using firm-level granularity, it is also possible to assess some of the real economy consequences as well. In this

[11] Due to the sensitivity of these datasets, we could merge them together using anonymized identifiers.

CEU eTD Collection

particular application 0.5% of the firms in the model became non-performing on their loans.

A further important application of the model is to use it as a tool for identifying systemically important institutions (SIFIs). To construct a SIFI measure I embedded the model into a modified version of the Shapley value concept. This indicator can be decomposed into three elements: i) system-wide losses caused by the default of a given bank, ii) losses suffered by the given bank due to external shocks, and iii) the part of other banks' losses which were caused by the shock amplifier effect of the given bank. The importance of these three factors can greatly vary among banks. In some cases the systemic importance is rooted mainly in the vulnerability of a bank, while others can be resilient from this perspective, but their default can represent more serious systemic risk. The third factor is usually less pronounced, which indicates that the complexity of the Hungarian banking system might not be as high as that of some larger countries. However, in some cases the ability to amplify shocks can also have significant influence on the systemic importance of Hungarian banks.

The modeling framework makes it also possible to simulate the effects of shocks originated not necessarily in the banking system, but also those coming from the real economy. By assuming that firms in a given industry become non-performing on their loans, one could assess the significance of different economic sectors for financial stability. Following this logic, the preliminary assessment of the economic impacts of the COVID-19 pandemic was used to illustrate how shocks originated in the real economy can be analysed using this model. Although the necessary statistics are not yet available to make confident assumptions about some crucial parameters, the current results can still indicate a plausible range for the expected consequences.

1.3 Thesis outline

The remainder of the thesis is structured as follows.

- Chapter 2 starts with the topological analysis of the ownership network of Hungarian firms. The next step after this is the description of three corrections to the raw ownership information which are necessary to the proper assessment of economic entities from the point of view of their influence and control in the economy via ownership relations. The final section of this chapter describes these influence and control measures at a more aggregated level based on the different attributes of the owners.

- Chapter 3 is concerned with the supplier network layer of Hungarian firms. After a detailed topological description, an analysis of the connection between the supplier links and firm characteristics is discussed. This is followed by the explanation of the methodology and the results of the community structure analysis of the network. The last phase of the analysis

CEU eTD Collection

describes the identification of the bridge firms which connect the otherwise separated communities.

- Chapter 4 introduces a model of shock propagation based on the previously analyzed firm-level networks. The chapter gives detailed justification for firm-level input-output analysis and then a formal description of the proposed model. The last part of this chapter presents and discusses the results of different simulations and policy applications.

- Chapter 5 describes a model of shock propagation in the banking system with real economy feedback. Firstly, an intuitive description and justification of the model is provided. This is followed by the detailed formulation of the simulation steps. There will be special emphasis dedicated to the calibration of the key parameters of the model. Lastly, the results of the implemented applications are discussed.

- Chapter 6 summarizes and discusses the key results and contributions of Chapter 2-5, proposes directions for future research, and concludes the thesis.

CEU eTD Collection

2 The ownership network of Hungarian firms

In order to be able to investigate firm networks, we obtained access to sensitive datasets about the ownership links and the supplier connections of Hungarian firms. Using these sources we built the multi-layer representation of the Hungarian firm network which enabled us to gain insight into its previously unobserved structure. In the case of the ownership layer we used the OPTEN ownership database containing more than 400 000 Hungarian firms for the period between 2015-2019. We could also merge this data with other micro-level datasets in order to use additional characteristics of the companies. (Further description of the quality and the cleaning of the OPTEN data can be found in Appendix A.)

This chapter is based on the paper titled *Unfolding the hidden structure of the Hungarian multi-layer firm network (forthcoming)* by Andras Borsos and Martin Stancsics.

2.1 Network terminology and definitions

To accomplish a formal analysis of the ownership network we have to introduce some basic concepts to represent a network in a mathematically interpretable way:

- In graph theory, the number of links connected to a given node (i) is called the *degree* (k_i). If the links have directions, we distinguish between the *indegree* (k_i^{in}) and the *outdegree* (k_i^{out}), showing the number of links coming in and going out in the case of a particular node.
- The links can also have weights which correspond to the ownership share in our data. In the case of a weighted network, we can calculate the *strength* (s_i) of a node instead of its degree by summing up the weights of the links associated with the given node.
- If we want to refer to the whole network, the simplest – although computationally often very inefficient – way is to represent it as an *adjacency matrix* (A or in the case of weighted networks W), where $A_{i,j}$ (or $W_{i,j}$) corresponds to the ownership share of actor i in actor j. The size of this matrix is $(m+n) \times (m+n)$ where m and n are the number of firms and the number of individuals in the network respectively.
- The *density* of a (sub)graph is defined as the ratio of the number of edges to the number of possible edges in the network[12].
- A *(connected) component* is a subgraph of the network where at least one path exists between every pair of nodes. We can distinguish between

[12]This definition is valid in the case of simple graphs, where there are no self-links or multi-edges between the nodes.

CEU eTD Collection

strong and *weak* forms of connectedness. The former requires that only directed paths can be considered, while the latter ignores the direction of the links.

- The *local clustering coefficient* shows the probability that two neighbors of a node are connected to each other forming closed triads.

- The *average shortest path length* shows the average number of steps it takes to get from one member of the network to another. It is calculated by finding the shortest path between all pairs of nodes, and taking the average over them.

2.2 Topological analysis of the ownership network

Although we analyzed the data for every year which we could observe, the topological structure showed very similar results for all the observed periods; therefore, we present here only the description of the year 2017, which is the last year for which we have access to every datasets we are using during the analysis. (The basic description of the network in other years can be found in Appendix A.)

The network consists of more than 1 million nodes (firms and individuals as well) and almost the same number of edges (ownership relation between nodes), which implies that the average indegree (or outdegree) is somewhat less than one (Table 1). The first important observation is that the network is not connected, i.e. it consists of many (259 138) components (using the weak form of connectedness), among which even the largest one contains only around 11% of the nodes, while all the others have maximum a few 100 members (Figure 1). One can also consider strongly connected components, but they would capture only partial information about ownership structures. For instance, if two individuals are owners of a firm, we would observe only one of them within one component (as there is no directed path between the two owners). This way, the largest strongly connected component contains only 19 nodes.

Table 1: Basic description of the ownership network

	Total network	**Largest component**
Number of nodes	1 029 487	115 218
Number of edges	963 744	253 840
Density	9e-7	1.9e-5
Average degree	1.86	4.4
Shortest path lengths (avg.)	-	13.78
Shortest path lengths (st.dev.)	-	5.91
Local clustering (avg.)	-	0.18
Local clustering (st.dev.)	-	0.2

CEU eTD Collection

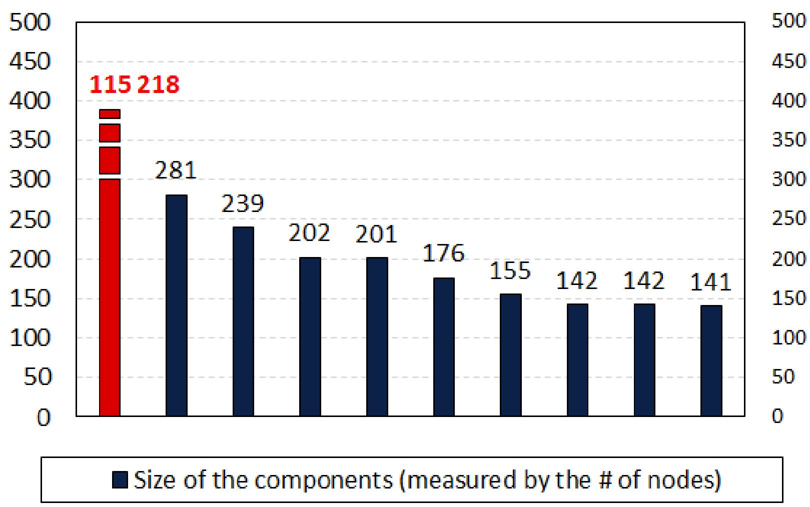


Figure 1: Size distribution of the 10 largest components (based on 2017 data)

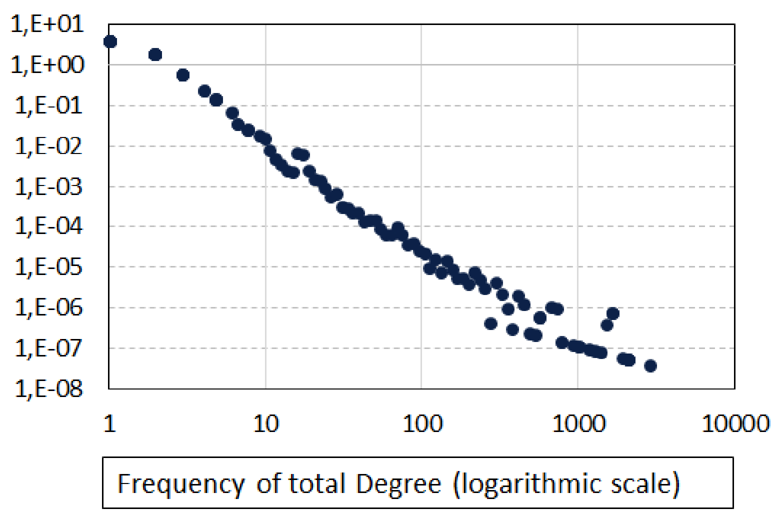


Figure 2: Degree distribution of the 2017 ownership network's giant component

CEU eTD Collection

Due to the low edge density of the network the size of the giant component remains rather limited as it encompasses only 11% of the nodes.[13] This largest component with an average degree of 4.4 is not as sparse as the network in general, and it also features some interesting characteristics:

- Its degree distribution is fat-tailed indicating the presence of actors with outstanding influence (Figure 2).
- The average shortest path length is 13.78. This value might seem low, however, it is not as low as it is typical in the case of many observed small-world networks. One additional reason for this high number can be the fact that a large portion of nodes represent individuals, which are restricted to have outgoing links only.
- We also calculated the local clustering coefficient, which is more than 0.18. It is a way higher probability than having a link between two randomly chosen nodes. This result reveals an important structural pattern in the network which is worth examining in more details by calculating motif statistics.

As it can be seen on Figure 3, there are 16 types of motifs consisting of three nodes (J. A. Davis & Leinhardt, 1967). In the case of a sparse network, the vast majority of the cases fall into the first, disconnected category, however, the distribution of the remaining motifs is very uneven. Although motif c) depicts only dyadic connections, it is interesting to observe that there are many reciprocal ownership relationships in the network. However, this number is actually less than the corresponding statistics of a directed configuration model which was generated using the empirical in- and outdegree sequences of this network. (The statistics for this null model are shown in parenthesis.) A possible explanation for this might be that in the ownership network these dyads can be formed only between firms (as individuals cannot be owned). Based on motifs d) and e) it occurs more often that a firm has more than one owner than having more than one firm in an actor's ownership, which implies difference in the *indegree* and *outdegree* distributions. While there are only 2896 observations of the simplest chain structure shown by motif f), we could find many ownership connections intertwined in more convoluted ways, e.g. following the patterns of motifs i), l) and m). Furthermore, it is surprising to notice the high number of instances in the case of motif p) which illustrates a fully connected triad with all the possible links among the nodes.

[13] In network theory the term *giant component* can sometimes be defined in a rigorous way, however, very often it is a rather loosely used concept. Here we simply mean the largest component which includes a significant portion of the nodes.

CEU eTD Collection

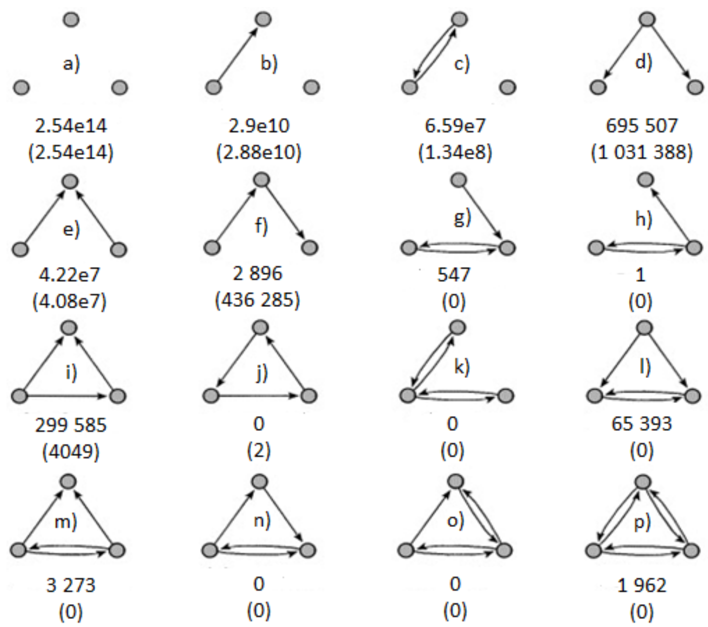


Figure 3: Motif statistics of the 2017 ownership network's giant component. (The numbers in parenthesis show the corresponding statistics for a directed configuration model generated using the empirical degree sequences.)

2.3 Measuring *influence* and *control* of owners

Our data makes it possible to assess the significance of economic entities from the point of view of the extent to which they can influence and control the economy via their ownership relations. In order to able to properly analyze this aspect of the ownership structure, we had to define measures for the manifestation of the economic actors' power. Although our methods sometimes differ from the analysis made by Glattfelder (2013), we often follow the approach and terminology of that paper in this section.

In order to carry out this analysis we have to apply a few corrections to the raw ownership information. Firstly, it is not obvious at all, how much actual power is entailed to a given ownership share. Secondly, we want to consider not only direct, but also indirect ownership links to gain accurate assessment about the influence of a given actor in the economy. Thirdly, we also need to take into account some measure of the sizes of owned firms. In the following subsections we describe our approach to deal with these points.

CEU eTD Collection

2.3.1 Distribution of ownership shares

To assess the influencing ability entailed to the observed ownership in our data we have to consider at least two potential distortions. The first one considers the assumption that ownership shares correspond to voting share. Although there are several common practices in corporate governance to deviate from the one-share-one-vote principle, it is credible to assume (for instance based on de Silanes et al. (1999)) that in the vast majority of cases we can use ownership as a proxy for influence manifested in the voting rights. The second bias, however, may require more effort to correct for. Owners or shareholders of a company can be considering not necessarily just as individuals but rather as rivaling voting blocks. In this mindset it is obviously incorrect to assume perfect proportionality between ownership and effective influence. The most common example to illustrate the difference between the two is the following distribution of ownership shares: 49% – 49% and 2%. In this case all three owners have practically the same influence as any two of them can form a block to gain majority.

One can find numerous similar examples, but it is far from being obvious how to create a correction which covers as many of these situations as possible, but it is still tractable computationally. Several so called power indices were proposed regarding this problem (see e.g.Leech (2002)), but there is no consensus in the literature on best practices. Because of its simplicity and efficacy we decided to apply the method proposed in Glattfelder (2013). The underlying idea of this measure is that the actual influence of an owner depends not only on its own ownership share, but also on the distribution of ownership shares of the other owners. The more dispersed the ownership structure is, the higher the influence of the given owner is. To calculate this concentration-corrected measure of influence, Glattfelder (2013) is using a version of the Herfindahl-index in the following way:

$$H_{i,j} := \frac{W_{i,j}^2}{\sum\limits_{l \in P_j^{in}} W_{l,j}^2} \tag{1}$$

where $H_{i,j}$ is the corrected ownership share of owner i in firm j, $W_{i,j}$ is the original ownership share and P_j^{in} is the set of indices of neighbors (owners) connected to j by incoming links. This measure can take values in the interval $(0, 1]$. If $H_{i,j}$ is close to one, it means that firm i has almost exclusive influence on firm j. Based on this measure we can calculate the *direct influence* of any owner by summing up all of its influence scores:

$$h_i := \sum_{j \in P_i^{out}} H_{i,j} \tag{2}$$

where P_i^{out} is the set of indices of neighbors connected to i by outgoing links.

CEU eTD Collection

2.3.2 Direct and indirect ownership

An obvious shortcoming of Equation 2 is that it only considers direct ownership links. To account for indirect connections Brioschi et al. (1989) proposed a method called the *integrated model*. The main component of their approach can be written in the form of a recursive computation:

$$\widetilde{H_{i,j}} := H_{i,j} + \sum_{n \in P_i^{out}} H_{i,n} \widetilde{H_{n,j}} \tag{3}$$

where $\widetilde{H}$ denotes *integrated influence*. The interpretation of this formula is that the actual influence of owner A on a firm B consists of two elements: the direct influence of owner A on firm B and the integrated influence on firm B by other firms owned by owner A. This expression can be written in matrix form as well:

$$\widetilde{H} = H + H\widetilde{H} \tag{4}$$

which gives the following solution:

$$\widetilde{H} = (I - H)^{-1} H. \tag{5}$$

Although Brioschi et al. (1989) showed that the mathematical requirements to conduct this calculation are always satisfied in an ownership network, there still can be computational constraints if the matrix representing the ownership network is large. In our case the inversion of the $(I - H)$ matrix was prohibitive, therefore, we calculated its Neumann-series approximation:

$$(I - H)^{-1} = I + H + H^2 + H^3 + \dots \tag{6}$$

This method is intuitively interpretable since the k^{th} power of an adjacency matrix gives us the number of walks with length k between two nodes. If we add up all the powers, we will cover all the indirect links in the network in the end. Due to the large memory requirement of storing large matrices, we could compute the approximation only up to the 6th power. This happened because although the original matrix is very sparse, it is not necessarily true for its inverse, or even for the higher powers of it. However, as the average length of the shortest paths is relatively short in any component of the network, six steps can cover the vast majority of the relevant ownership links. Consequently, the elements of the resulting matrix $(\widetilde{H})$ can be interpreted as the *total influence* of owner A on firm B.

CEU eTD Collection

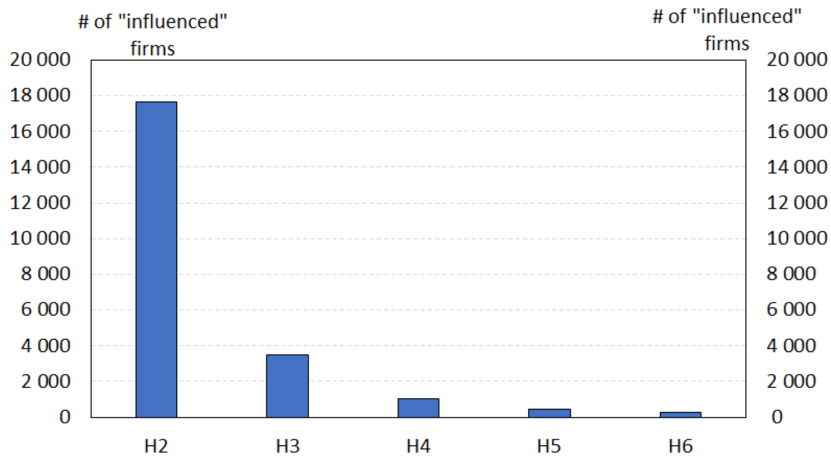


Figure 4: Additional indirect influence by the number of steps in the ownership network. ($H\#$ refers to a given power of the original influence matrix. The "# of influenced firms" is equal to the sum of the elements of a given matrix.)

Another limitation of this calculation is that we cannot observe global ultimate beneficiary owners (UBOs) as we only see such ownership relations, for which at least one of the endpoints of the links is a Hungarian firm. As ownership ties between foreign entities would be necessary to trace the exact paths of more convoluted offshore activities, we cannot take these into account in the investigation of the owners' indirect influence over the Hungarian economy.

Despite the relatively small role of higher order indirect influence indicated by Figure 4, Figure 5 shows that the total degree and the total influence are only loosely associated. There are several observations where the total influence can be high even with small degree values, while high degree is not a guarantee for high influence.

CEU eTD Collection

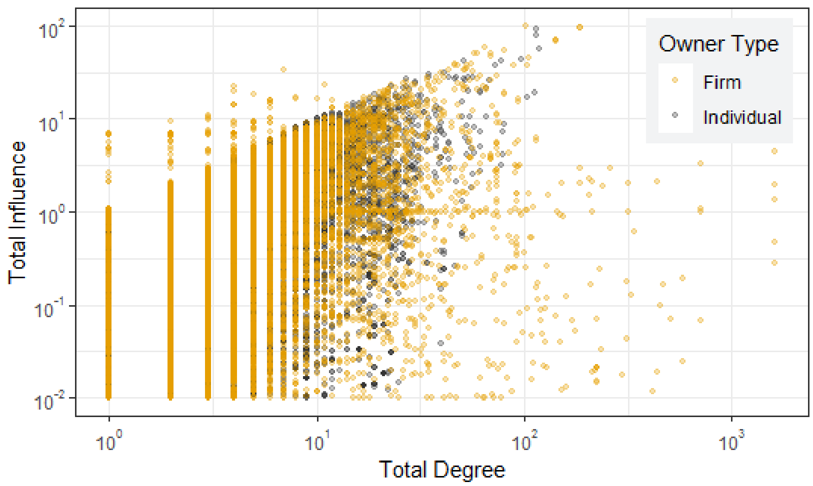


Figure 5: Total influence and total degree of firms and individuals in the ownership network.

2.3.3 Weighting of influence

Our measures so far did not consider any information regarding the significance of the owned firms. By adding a non-topological node attribute to correct for this could enhance the precision of the assessment of owners' influence considerably. We decided to use the simple approach to multiply the matrix of *total influences* ($\widetilde{H}$) by the vector of some approximation of firms' economic significance. While recognizing the depth of the methodologies in corporate valuation, the amount of firms in this exercise grants justification for opting for the simplest possible option to evaluate firms. Some of the obvious candidates as proxies for firms' weight could be e.g. capitalization (for listed companies), or total asset value (for smaller firms). The resulting *total controlled value* measure could be formulated like this:

$$\tilde{c}_i := \sum_{j \in P_i^{out}} \tilde{H}_{i,j} v_j \tag{7}$$

where v_i denotes the i's firm value.

However, the OPTEN data does not contain any variable which we could use as a proxy of firms' significance, therefore we had to join another firm-level dataset coming from the Hungarian Tax Authority containing the balance sheets and profit and loss statements of firms. Unfortunately, the overlap between these datasets is not perfect, i.e. we cannot match the required firm characteristics to almost 22% of the firms in the ownership network.

A more serious caveat of this approach is the multiplication of firms' weight

CEU eTD Collection

when one conducts the aforementioned computation. The problem arises due to the fact that the weight of a given firm contains the proportional part of the weight of the companies owned by this firm as well. (A more detailed illustration of this problem can be found in Appendix B.)

To solve this issue, we wanted to find a node attribute, which is independent from the ownership structure, but conveys some information about firms' importance in the economy. A suitable candidate to meet these requirements can be the value added of firms (denoted by rva_j for firm j.), which can be directly applied as a replacement for the previously mentioned proxies. With this solution we only have to make a slight modification on Equation 7 by replacing total assets with the real value added of firms:

$$\tilde{c_i} := \sum_{j \in P_i^{out}} \tilde{H}_{i,j} rva_j \tag{8}$$

Although this measure is clearly not an ideal proxy for firms' value, it gives more accurate results if one wants to compare the control of different economic actors than the naive approach of using traditional firm size variables.

Based on this measure, we could calculate the empirical cumulative distribution of the total control of owners, which can be seen in Figure 6. It is important to note, that neither the total influence nor the total control of a given firm include itself. The interpretation of this plot is then the following: the top right corner of the diagram represents 100% of the owners controlling 100% of the economy's value added, and the first data point in the lower left-hand corner denotes the most important owner. The red lines indicate that the top 10 owners control more than 17%, and the top 100 owners control more than 40% of the economy (measured by the real value added of firms).[14]

[14] In this calculation we did not assume any strategic cooperation between owners to control firms, and we did not take into account the fact, that having $50\% + \varepsilon$ ownership share can often be sufficient to fully control a company.

CEU eTD Collection

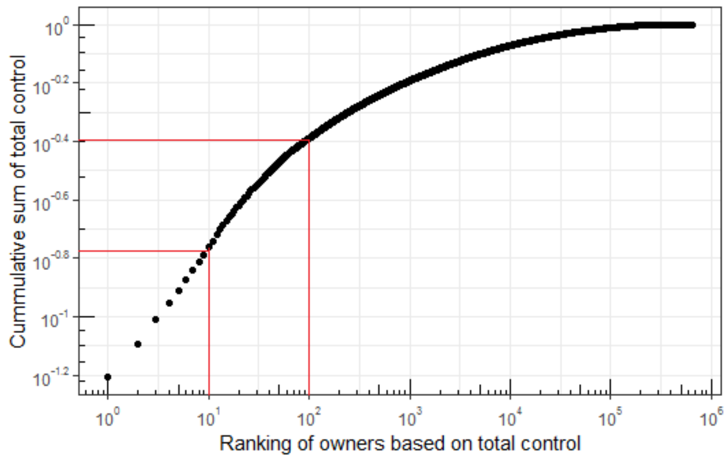


Figure 6: Cumulative distribution of owners' total control in 2017

2.4 Influence and control based on the owners' attributes

Besides the ownership links, the OPTEN data contains some attributes of the owners as well. Most importantly, we can see whether the owner is a firm or an individual as well as the country level location of its headquarter. We calculated the *direct* and *total influence* and *control* measures aggregated along the dimensions of Hungarian/foreign and firm/individual owners. As it is shown in Figure 7, there is only a small gap between the *direct* and *total* versions, however, the difference is way more pronounced between the *influence* and *control* results. As the average value added of companies owned by foreign owners and by firms is much higher, their significance is heavily underestimated in the case of the unweighted influence measures. Moreover, Table 2 reveals that foreign firms have the biggest role among these categories by controlling around 37% of the value added in the Hungarian economy. Hungarian firms and individuals have almost the same amount of *total control* (31%), while foreign individuals have much smaller significance by exercising less then 1% control.

CEU eTD Collection

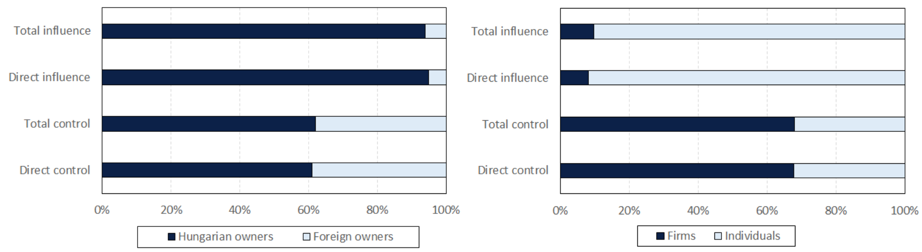


Figure 7: Total/direct control/influence of owners

Table 2: Total control (in 1000 billion HUF) based on the owners' attributes

	Hungarian owners	**Foreign owners**	**Sum**
Firms	8.88 (31%)	10.58 (37%)	19.47 (68%)
Individuals	8.91 (31%)	0.26 (1%)	9.17 (32%)
Sum	17.79 (62%)	10.85 (38%)	28.65 (100%)

We can make similar analysis on a more disaggregated level concerning the significance of foreign countries in the Hungarian economy (Figure 8)[15]. As holdings and special purpose firms designed for tax optimization might distort the results especially in the case of the *control* measures, *direct* and *total influence* might be a better indicator for foreign countries' importance in the Hungarian economy. For example The Netherlands is generally not as important economic partner for Hungary as Germany or Austria, but there are several large companies which control their Hungarian subsidiaries through holding entities with headquarters in The Netherlands. As the organizational structure of these transnational companies can change frequently based on their strategic decisions, we can observe significant changes in the control measures of countries, while their influence remains relatively stable over the examined years (as it can be seen if one compares Figure 8 and Figure 9).

[15] If a firm operating in Hungary is a foreign-owned firm and it owns other firms, then foreign influence/control includes not only these owned firms but also the foreign-owned firm itself.

CEU eTD Collection

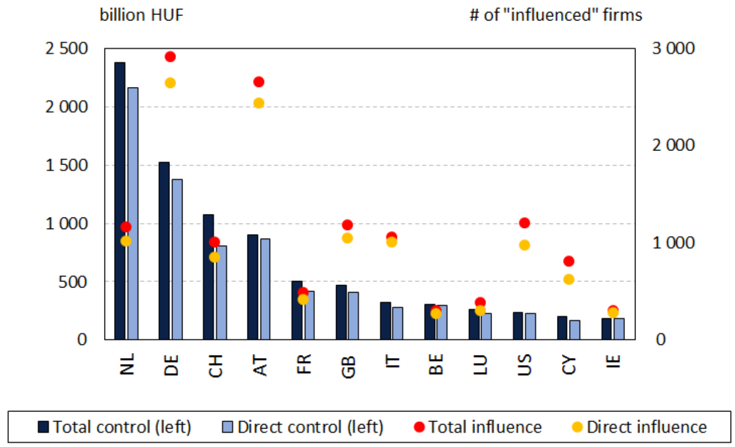


Figure 8: Total/direct control/influence of countries in 2017. (The "# of influenced firms" is equal to the sum of the total influence of owners belonging to a given country.)

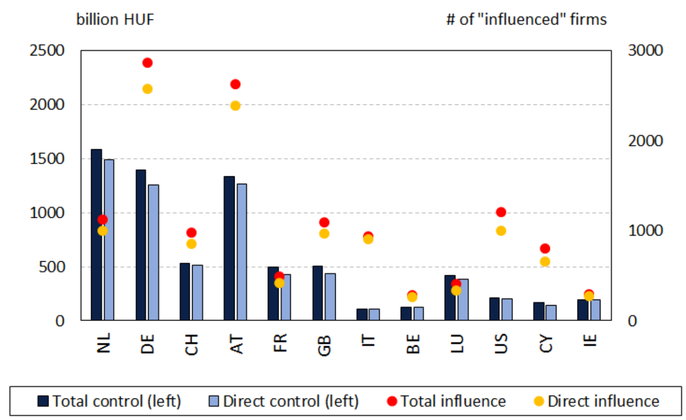


Figure 9: Total/direct control/influence of countries in 2016. (The "# of influenced firms" is equal to the sum of the total influence of owners belonging to a given country.)

If we focus only on the owners belonging to the *Hungarian firm* category, we can examine more disaggregated levels by adding further attributes from our additional firm dataset. Figure 10 and 11 shows our measures of significance of owner firms aggregated based on their head quarters' location at the level of counties (NUTS 3) and regions (NUTS 2) of Hungary. Although these diagrams are calculated based on partial data without considering the role of individuals and foreign entities, the results are in line with intuition that the more developed areas such as the capital and the counties with major towns play a more important role in the ownership network. (E.g. Fejer county is a traditional hub

CEU eTD Collection

for large Hungarian industrial companies, such as Videoton, Dunaferr, Kofem.)

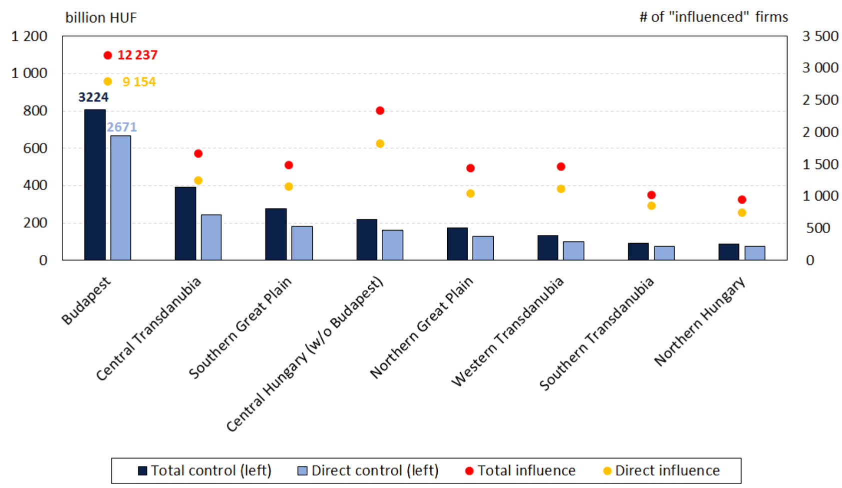


Figure 10: Total/direct control/influence of regions in 2017. (The "# of influenced firms" is equal to the sum of the total influence of owners belonging to a given region.)

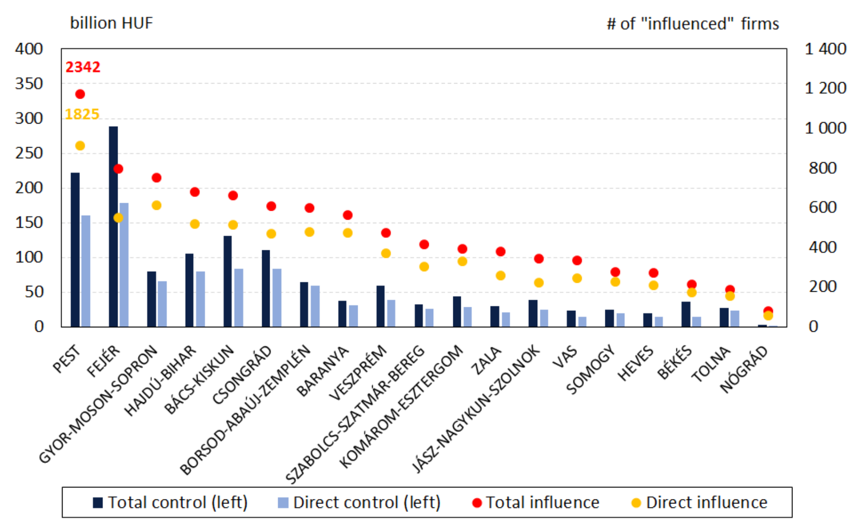


Figure 11: Total/direct control/influence of counties in 2017. (The "# of influenced firms" is equal to the sum of the total influence of owners belonging to a given county.)

We can also use firm size categories as an alternative aggregation dimension (Table 3). Although micro-enterprises have the largest role based on every

CEU eTD Collection

measure, it can be misleading to rely on only one type of metrics as *influence* greatly underestimates the significance of large companies.

Table 3: Total/direct control/influence based on size of the owner firm

Firm size category	Total control (billion HUF)	Direct control (billion HUF)	Total influence	Direct influence
Micro	2 090	1 418	16 690	12 279
Small	398	325	3 214	2 691
Medium	286	226	1 448	1 236
Large	1 052	956	719	545

We carried out this analysis also based on the NACE industry categories of the Hungarian owner firms (Figure 12). We can see the dominance of the finance and insurance industries in the *control* measures, however, the *construction* and the *professional, scientific and technical activities*[16] industries are even more influential based on their *influence*.

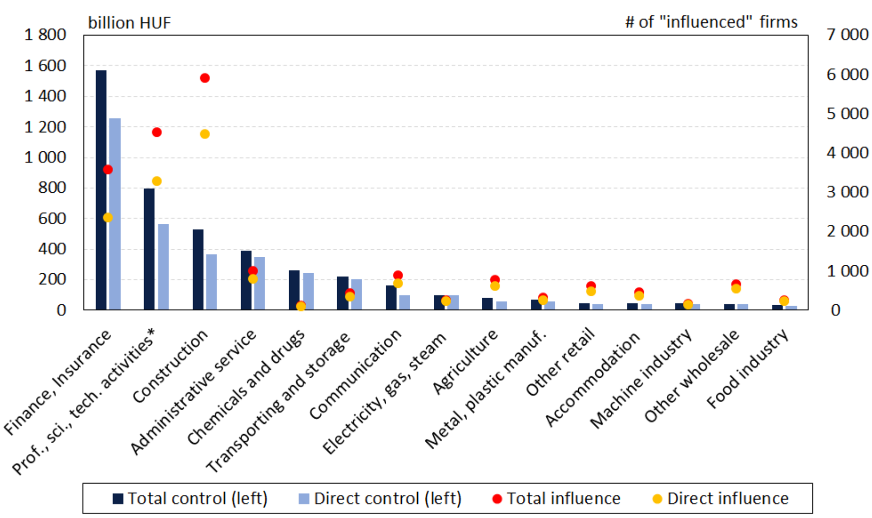


Figure 12: Total/direct control/influence of the top 15 industries in 2017. (The "# of influenced firms" is equal to the sum of the total influence of owners belonging to a given industry.)

[16]The "Prof., sci., tech. activities" category refers to professional, scientific and technical activities, which contains legal, auditing, consulting services as well as scientific and technical (e.g. architectural) services.

CEU eTD Collection

3 The supplier network of Hungarian firms

We obtained access to firm-level supplier connection data containing trade links among Hungarian firms where the tax content of the transactions between two firms exceeds EUR 3000 in the given year (Figure 13). This data is available between 2014-2017, and it is coming from Hungarian firms' VAT reports collected by the National Tax and Customs Administration of Hungary.

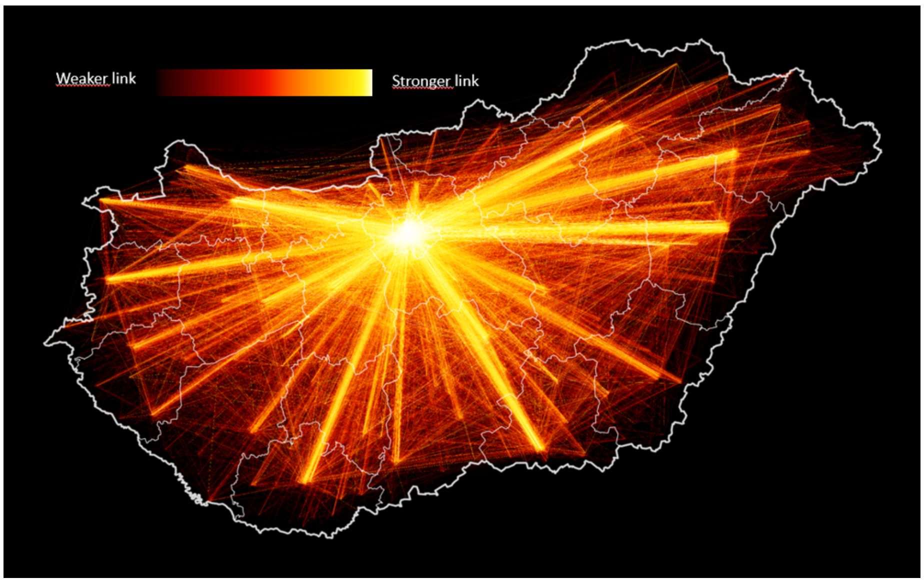


Figure 13: Supplier connections among Hungarian firms in 2017

This chapter is based on the paper titled *Unfolding the hidden structure of the Hungarian multi-layer firm network (forthcoming)* by Andras Borsos and Martin Stancsics.

3.1 Preparation of the supplier network data

Similarly to the ownership data, the most intuitive way to handle this network is to think about it as an *adjacency matrix* (A or in the case of weighted networks W), where each cell corresponds to the purchased value of the firm in the row dimension from the firm in the column dimension ($A_{i,j}$ or $W_{i,j}$). That is, an outward link starting from a given node denotes that it buys from the firm toward which the arrow points.

It is important to emphasize that this data is directly not comparable to the typical industry-level, symmetric input-output tables due to some fundamental conceptual differences. The firm level data contains information only about

CEU eTD Collection

trade relationships above the regulatory threshold, and only for those products and services which are subject to the VAT. (Albeit it also ignores the reverse VAT situations.) However, the data includes transactions between firms which are not necessarily residents in Hungary if the fulfillment of the transaction happened in Hungary. In the I/O table of the Hungarian Central Statistical Office (HCSO) these trades often belong to the foreign trade category (although in some cases they would not be part of the I/O table statistics at all). Moreover, we cannot see in the granular data any further information about the products and services, thus, it is impossible to know if a purchase happened for investment reasons, which would be handled differently in the I/O table than a purchase due trade purposes. A further problem can arise when the invoice comes from a trader firm, in which cases the source of the product is unknown. If it was imported, then it will be categorized as foreign trade in the HCSO I/O tables. There are some differences between the industry classifications as well, and also in the calculation of the prices. The HCSO I/O table uses basic prices, which are different from the market price as they do not include margins, transportation costs and the net position of the taxes and allowances. Due to all these factors, the industry-level aggregation of our granular data results in a completely different table than the I/O matrix produced by the HCSO.

The links of the network change significantly from one year to another because there are a lot of one-off, incidental transactions. More than 50% of the links disappear between the observed periods, and new links emerge in a similar extent. As these relationships are not particularly relevant from the point of view of spreading processes and they increase the noise in our measurements, we filtered the network to contain only long-term supplier connections. We consider a link *long-term connection* (i) if there were at least two transactions between the parties, and (ii) if there is at least one quarter time difference between the first and the last transaction between the two firms (Figure 14). Even with these mild requirements, only 54% of the links are long-term, however, these cover 93% of the aggregate trade volume in the network.

CEU eTD Collection

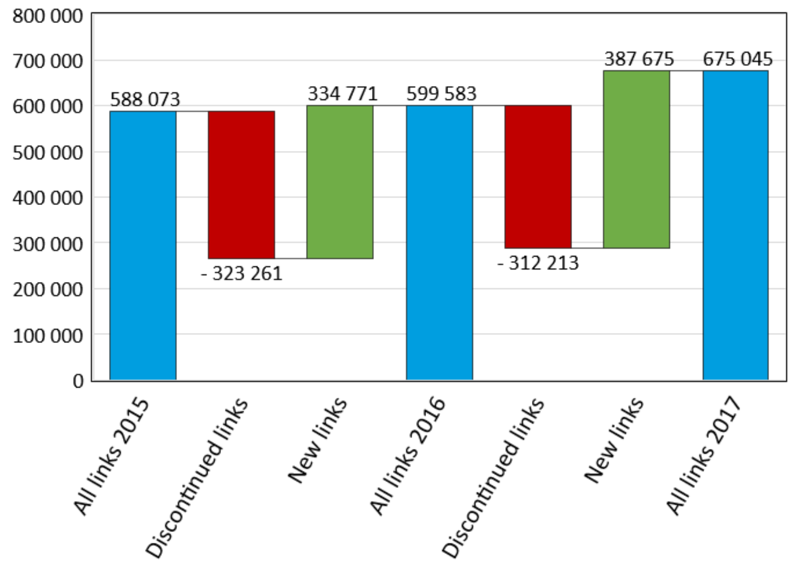


Figure 14: Temporal stability of the supplier connections

Another source of distortion we have to deal with is that there is no general rule for VAT reporting in the case of firms belonging to the same ownership-based group. Sometimes they file their VAT reports collectively, but as it is only optional, there are many firms belonging to a group which report individually. To handle this difficulty we can utilize the ownership layer of the firm connections by applying the following procedure[17]:

1. In the case of every ownership link when the *total influence* exceeds 50%, we combined the link's endpoints into a group.
2. If some firms in the supplier network belonged to the same group, we replaced them by a new node representing the group.
3. We added the links of the original firms to the new "group" node.
4. We eliminated the resulting self-loops (within-group links) (Figure 15).

[17] Although we did not consider ownership links with influence weights under the 50% threshold, and we cannot see global ultimate beneficiary owners, we could still cover probably the vast majority of the ordinary intertwinings among firms.

CEU eTD Collection

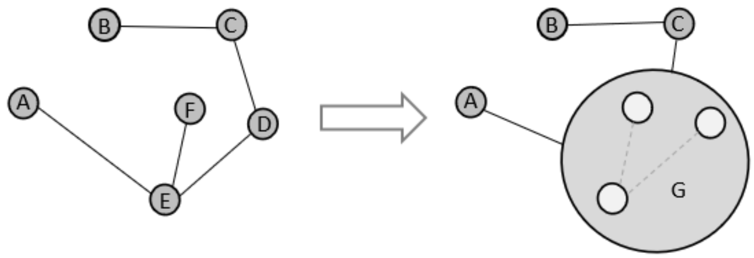


Figure 15: During the correction of the supplier network we combined together firms belonging to the same ownership group, and we eliminated the links within the groups.

Further details about the features, quality and cleaning of the data can be found in Appendix C.

3.2 Topological analysis of the supplier network

Similarly to the section about the ownership network, we present only the results of one year (2017) as it is sufficiently representative for the whole examined period. (The basic description of the network in other years can be found in Appendix C.)

As this network consists of Hungarian firms only, the number of nodes is much lower than in the case of the ownership network which contained individuals and foreign actors as well. However, the density of this network layer is higher (even after filtering for long-term connections and correcting for ownership groups), which contributes to the emergence of a giant component covering around 94% of the nodes. In this case we decided to focus only on this component as it represents credibly the whole network while all the other isolated parts are negligible in size.

The resulting network consists of 89 778 nodes and 235 913 links. The average total degree in this giant component is around 5.2 (which implies that the average in/out-degree is the half of this value, 2.6), which can be interpreted as the average number of long-term supplier and buyer relationships for firms (Table 4). This result is difficult to compare to any other datasets in the literature, as other papers usually consider all the transactions (with different thresholds and without filtering for long-term links) among (often only a sample of larger) firms.

CEU eTD Collection

Table 4: Basic description of the 2017 supplier network's giant component

Number of nodes	89 778
Number of edges	235 913
Density	2.93e-5
Average in-/outdegree	2.63
Shortest path lengths (avg.)	4.92
Shortest path lengths (st.dev.)	1.1
Local clustering (avg.)	0.078
Reciprocity	0.11

Regarding the degree distribution of the graph we can see slightly different figures depending on whether we consider the *total, in-, or outdegree* of the nodes. For all the three measures the distributions have fat tails, however, we do not encounter as many extreme values in the case of the number of suppliers of firms (*outdegree*) as in the case of the number of buyers (*indegree*) (Figure 16). Despite of this disparity, we could ascertain that there are firms in the network which can be considered *hubs.* These agents can play a special role in any spreading process for which the production network is a relevant medium, thus, the identification of these firms and the assessment of their importance can be vastly important.

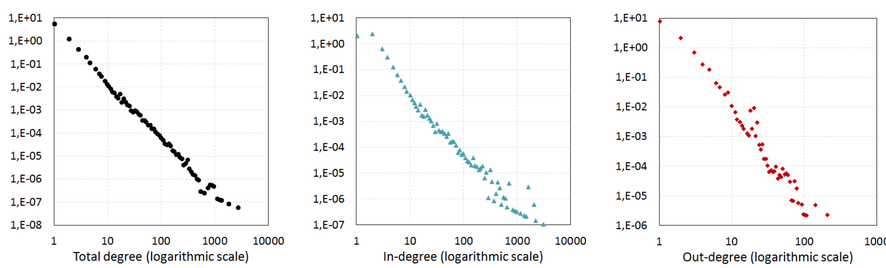


Figure 16: Degree distribution of the supplier network in 2017

We can also see that the average shortest path length is below 5 with a standard error as low as 1.1, which implies that shocks can be transmitted easily between firms via hubs in the network. Furthermore, Table 4 also shows measures of local feedback loops: there is an unexpectedly high probability of reciprocal dyads (11%) [18] and closed triplets (7.8%), which can further amplify shocks and promote contagions.

[18] In the case os reciprocity we cannot filter for repurchases, which can somewhat inflate this measure.

CEU eTD Collection

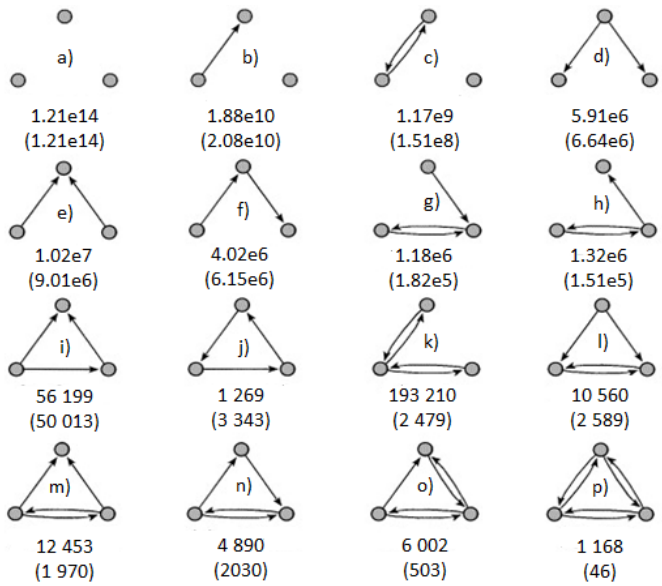


Figure 17: Motif statistics of the 2017 supplier network's giant component. (The numbers in parenthesis show the corresponding statistics for a directed configuration model generated using the empirical degree sequences.)

As it can be seen on Figure 17, we assessed the frequency of the different triadic motifs also in the supplier network. In line with the observed difference in the distribution of in- and outdegrees, we can see based on motifs d) and e) that there are much more instances of firms having multiple buyers than having multiple suppliers. (Although the numbers are close in both cases to the null model's.) While in the sample there are only a few observations of the simplest triadic loop formation shown by motif j), there are several appearances of loops hidden in the more convoluted motifs, such as n), o) and p). These more complicated motifs are often observed e.g. among wholesale trader firms and manufacturers operating on the upstream part of supply chains. (E.g. two wholesale trader of chemical materials are trading with each other to both directions, and both of them are connected to a chemical material producer.) Based on this, the network is far from being acyclic which observation can be in juxtaposition with the results of McNerney (2009), in which paper the authors examined input-output economic systems, and found that economies tend to be acyclic at the scale of triadic patterns on industry level.

3.3 Supplier connections based on firm characteristics

By connecting this data to the locations, balance sheets and profit and loss statements of firms we could carry out analysis not only based on topological

CEU eTD Collection

information but also using several node and link attributes.[19] We examined several firm characteristics, but we present here only those cases where we found meaningful patterns.

We examined the supplier relationship between firms based on their (labor) productivity. We assigned each firm into groups formed based on productivity deciles, and then we collapsed the network to this aggregation level. This way, the resulting 10 by 10 matrix shows the flows in the supplier network among firms grouped based on they labor productivity. However, although the number of firms will be the same in every group, the size of the firms, and consequently the number (and weight) of their supplier connections can be different. To control for this effect, we could compare the observed flows to a null model which gives us the expected flow between any two groups if the links were formed randomly in the network. To perform this calculation we divided the observed flows by the product of the *outdegree* of the group in the row dimension and the *indegree* of the group in the column dimension (or in the case of weighted networks we can use the *strengths*)[20]:

$$\tilde{W}_{p,q} = \frac{W_{p,q}}{S_p^{out} S_q^{in}} \tag{9}$$

where $W_{p,q}$ denotes the original and $\tilde{W}_{p,q}$ denotes the normalized flow between productivity groups p and q. S_p^{out} is the sum of the strengths of outgoing links for group p and S_q^{in} is the sum of the strengths of incoming links for group q.

Figure 18 shows some homophily based on productivity. The cells near the lower-left corner, but at some level also near the upper-right corner are darker, indicating stronger linkages between firms with similar productivity levels. These observations can have several connections to the existing literature, however, in this paper we do not try to identify the factors leading to this pattern or assess the potential consequences of this network structure.

[19] As some of these characteristics of the firms are not trivial to consolidate based on their ownership background, in this section we considered the firms as they were present in the data originally.

[20] This normalization is in the spirit of the *configuration model* which generates uncorrelated random networks with a given degree sequence.

CEU eTD Collection

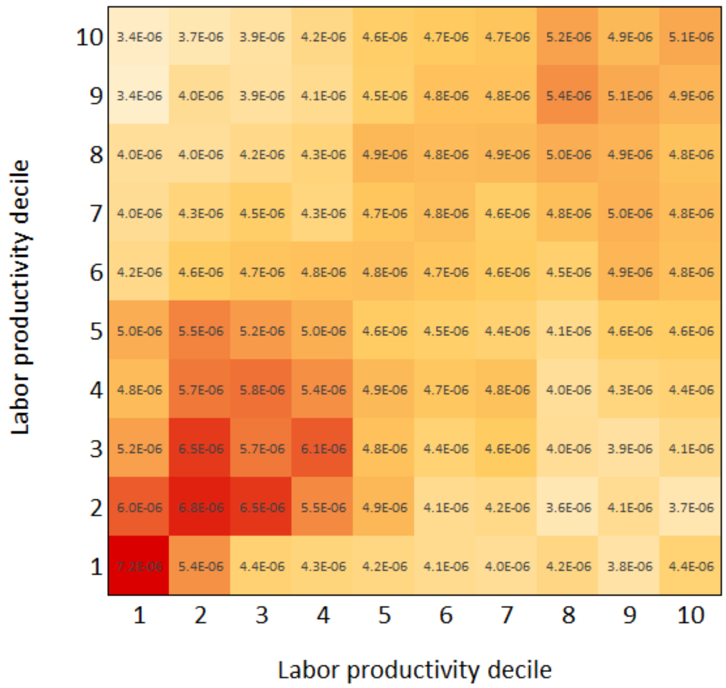


Figure 18: Flows in the supplier network between firms belonging to different productivity-based deciles in 2017. Labor productivity increases from group 1 to 10. Darker coloring indicates stronger trade connection.

Trade connections between groups with very different productivity are also very polarized: Productive firms sell much more to the less productive firms than the other way around. We quantified this polarization between the groups using the following formula which is based on Iino and Iyetomi (2012):

$$P_{p,q} = \frac{W_{p,q} - W_{q,p}}{W_{p,q} + W_{q,p}} \tag{10}$$

where $P_{p,q}$ denotes the polarization ratio.

This formulation of the polarization shows the typical direction of trade between groups in the row and in the column dimensions. The polarization matrix is antisymmetric, i.e. $P_{p,q} = -P_{q,p}$. If the relationship between groups p and q is only one-directional, $P_{p,q}$ will be ± 1 (the sign depends on the direction), while if the flow is the same to both directions, $P_{p,q} = 0$. On Figure 19 we can see that the more productive a group is the larger its dominance is in the trade relationships with less productive groups.

CEU eTD Collection

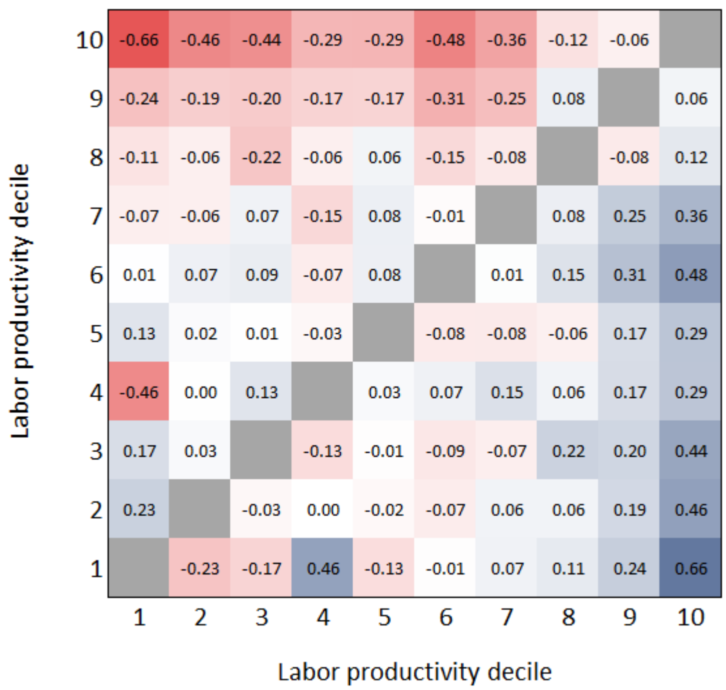


Figure 19: Polarization among firms in different labor productivity deciles in 2017. Labor productivity increases from group 1 to 10. Red color means that typically the group in the row dimension supplied to the group in the column dimension, while blue indicates the inverse situation. Darker coloring indicates stronger polarization.

On Figure 20, we can see a weaker but similar pattern if we use return of assets (ROA) instead of labor productivity (as it can be expected due to the high correlation between productivity and profitability measures).

CEU eTD Collection

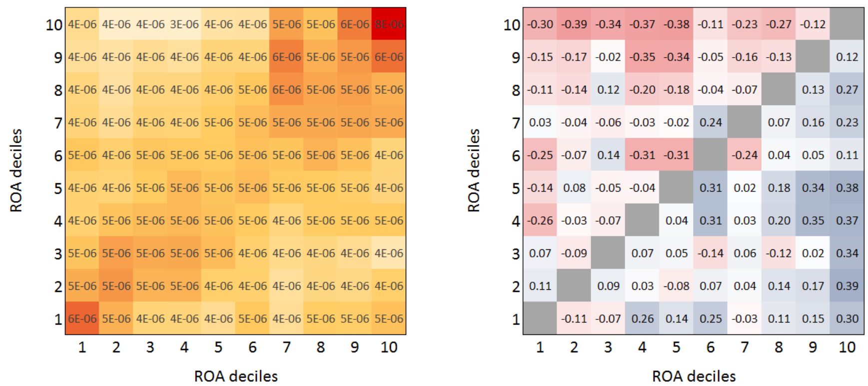


Figure 20: Flows and polarization in the supplier network between firms belonging to different ROA-based deciles in 2017

We can calculate similar measures to assess geographical clustering as well (Figure 21). The diagonal elements of the matrix are clearly darker than the off-diagonals, which shows that trade connection within regions are stronger than between regions, indicating the presence of location-based preferences in link formation.

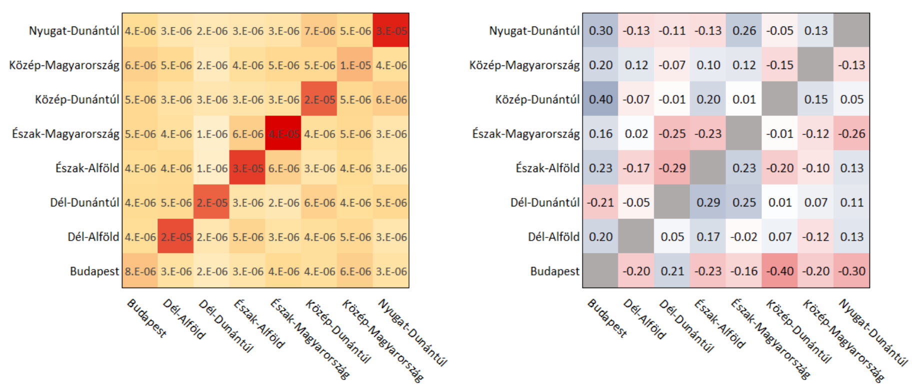


Figure 21: Flows and polarization in the supplier network between firms in different regions in 2017

As in this case it would be also informative to see the absolute magnitude of the trade connections (i.e. without comparing them to a null model) to assess the dominance of different regions, we can simply consider the number of links between the different regions as well. According to Figure 22 the dominance of

CEU eTD Collection

Budapest is apparent: Firms in the capital have way more connection than in any other region, and they have a lot of trade links to all the other regions as well. Furthermore, firms in Budapest typically supplied more to firms in other regions than the other way around.

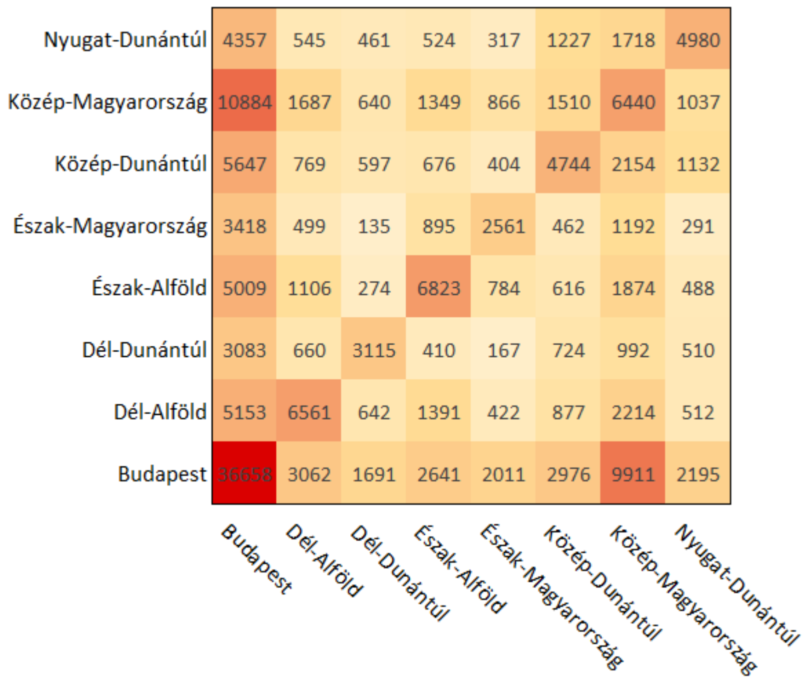


Figure 22: Trade connections in the supplier network between firms in different regions in 2017. Darker coloring indicates stronger trade connection.

3.4 Exploring the modular structure of the supplier network

An often observed characteristics of real-world social and economic networks is that they have a mesoscopic structure which can be best described by the concept of *communities*. A network is regarded to have a community structure if its nodes can be grouped into internally densely connected sets (which potentially overlap), i.e. members of a community are relatively densely connected within their group, but there are only sparser connections between the groups (Yang et al., 2010) (Figure 23). Identifying community structures can be very revealing about a complex system, as the observation of the grouping of the actors based on this dimension is only possible through the examination of the

CEU eTD Collection

whole network on granular level.

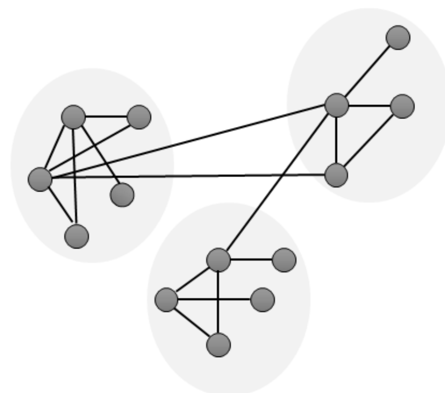


Figure 23: Schematic picture of the community structure of a small network

In the case of our production network, community detection can result in the identification of blocks within the economy, in which the coherence is provided by the intricate supplier relations among the constituent firms. As local shocks usually propagate more unimpededly within the surrounding community than between the separated communities, this segmentation of the network can help us tremendously in the more detailed understanding of spreading processes on the supplier network.

There are many algorithms (coming from different disciplines e.g. computer science, biology, mathematics, physics and sociology) which have been developed for identifying communities. (An excellent review of community detection algorithms can be found in Javed et al. (2018).) In our analysis we opted to use a widely-used technique called the "Louvain-method" which is based on the *modularity* of the network (M. E. Newman, 2006). As these methodologies are not common in economics, we describe the applied algorithms in more details in Appendix D.

3.4.1 Describing the modular structure of the Hungarian supplier network

Our procedure detected 249 communities, however, the ten largest already contain almost 80% of the firms, so we concentrate only on these in the more detailed analysis. The groups formed based on the community detection results are much more separated than in the case of any of the former grouping variables. Figure 24 shows that the diagonal of the matrix is clearly outstanding compared to the other cells indicating that connections within communities are much stronger than connections between communities.

CEU eTD Collection

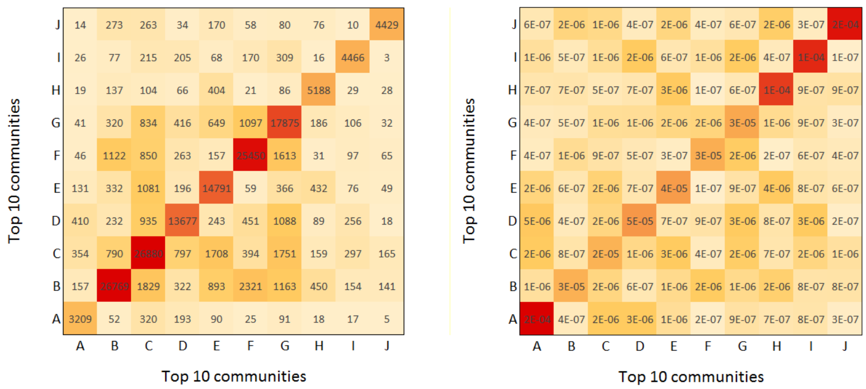


Figure 24: Strength of connection between communities in the supplier network in 2017 (The figure on the left is based on the number of links between communities, while the figure on the right shows the connections compared to the randomly expected number of links between the groups.)

Similarly to the results of Fujiwara and Aoyama (2010), the most intuitive variable we can use to interpret the communities is their sectoral composition. The largest communities all can be interpreted as a production chain of certain product categories within the economy. E.g. the first group on Figure 25 consists of firms belonging mainly to the food industry, food wholesale and food retail sectors. The second group contains firms from the machine and electronics industry; metal and plastic manufacturers; electricity, gas and steam suppliers. All the other groups can be similarly well interpreted as blocks containing chains of production of a well defined product category.

CEU eTD Collection

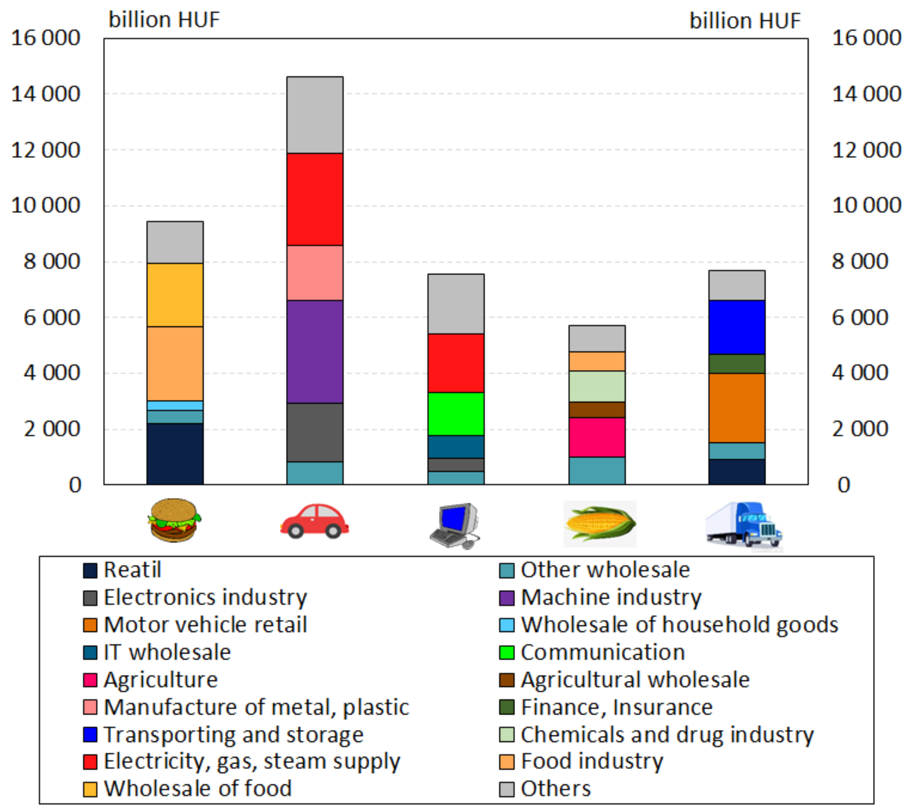


Figure 25: Industrial composition based on the size (total assets) of the firms belonging to the top 5 communities in 2017 (The pictograms indicate the main profile of a given community. E.g. the first group on Figure 25 consists of firms belonging mainly to the food industry, food wholesale and food retail sectors. The second group contains firms from the machine and electronics industry; metal and plastic manufacturers; electricity, gas and steam suppliers. All the other groups can be similarly well interpreted as blocks containing chains of production of a well defined product category.)

If we examine the polarization on Figure 26, we can see clear patterns only for two out of the ten largest groups. In the case of Group B (which corresponds to the machine and electronics production chain) we can see that they buy a lot of intermediate inputs from other blocks, but they do not supply to them in similar extent. In the case of Group J (which corresponds mainly to logistics, insurance and motor vehicle retail) the polarization is exactly the opposite: this block supplies to all the other segments of the economy way more than it buys from them as inputs. A natural explanation contributing to these result can be the unobserved export and import activities of firms in these segments of the economy.

CEU eTD Collection

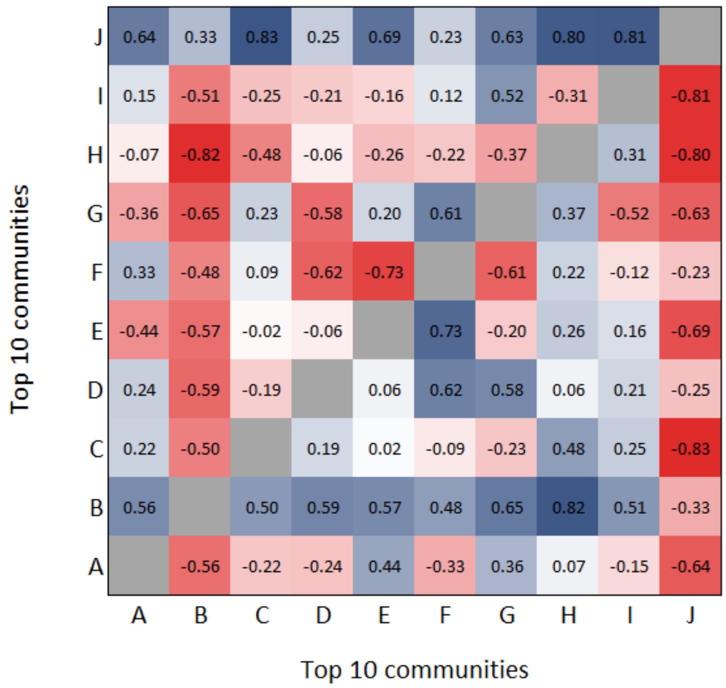


Figure 26: Polarization among firms in different communities in 2017. Red color means that typically the group in the row dimension supplied to the group in the column dimension, while blue indicates the inverse situation. Darker coloring indicates stronger polarization.

3.4.2 Bridges between communities of the Hungarian supplier network

Although we saw that the communities are highly separated, from the point of view of shock propagation it is still crucial to examine how these large blocks of the economy are connected to each other. Firms having supplier partners belonging to other communities create bridges between distant parts of the network, and therefore, propagate the spreading of contagious processes in the whole system.

The simplest measure to capture firms' shock transmitting ability is to consider the number of links of a node which are pointing to other communities. As it is shown on Figure 27, although the degree of a node is correlated with the number of links pointing to other communities, there is still a large variation which is not explained just by the degree. To further investigate the firm characteristics associated with our dependant variable, we used a simple regression analysis. We found that firms in the chemicals and drug industry, furthermore in the transportation and infrastructure sectors, and firms with high value added

CEU eTD Collection

and high export sales rate have particularly many outside connections (Table 5).

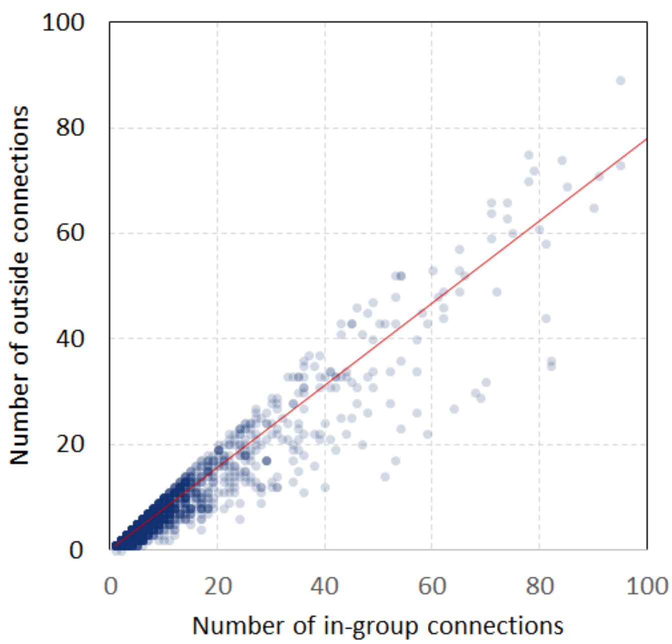


Figure 27: The number of links within the firms' communities and the number of connections pointing to other communities (The visualization shows only 10% of the nodes.)

We considered another approach as well to assess firms' intercommunity shock spreading ability. Firms having many connection with only a few communities can be less important from the point of view of connecting the different blocks of the economy than those firms which have connections to many communities. Our community detection methodology was so far incapable to identify overlaps between the communities. However, using a different representation of our network based on Evans and Lambiotte (2010) and Ahn et al. (2010) we are able to take into account this feature as well. This would make it possible to see if a firm is part of multiple communities which would be an indication for its increased ability to transmit shocks.

As a first step to obtain this measure, we have to transform our graph into a so called *line graph*. Nodes of the *line graph* are the links of the original graph, and two nodes (links) are connected if they had a shared endpoint(in the original graph where they were links) (Figure 28).

CEU eTD Collection

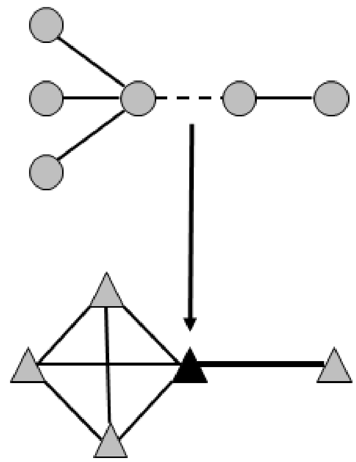


Figure 28: Creating Line graph from a traditional graph. Nodes of the line graph are the links of the original graph, and two nodes (links) are connected if they had a shared endpoint (in the original graph where they were links).

Using our standard community detection method on the line graph we can assign the links of the original graph into communities. As links of a node in the original graph can belong to more than one communities now, we can count for every node how many communities its links belong to. We can use this number of group memberships of a firm to measure its potential to transmit shocks.

By putting this measure into the same specification as before, we got somewhat different results compared to the first regression (Table 5). Based on the overlapping community approach, the firms which are typically associated with multiple groups often belong to sectors which provide non-essential products and services which are not directly related to the production activity of their buyers (e.g. insurance, catering, administration, etc.). This way, they pose only limited threat to the proper functioning of the production processes (at least in the short run). This result suggests that these firms are not necessarily very influential in spreading shocks, which finding might suggest a more optimistic interpretation about the resilience of the economy: Although the different blocks of the firm network are accessible to each other, but the firms which are truly influential from the point of view of shock propagation are usually only connected to a few communities.

CEU eTD Collection

Table 5: Regression results (Among the NACE categories only industries with positive, significant coefficients are listed.)

	Dependent variable:	
	# of outside links	# of memberships
	(1)	(2)
degree	0.028***	0.032***
	(0.0005)	(0.001)
log(balance sheet total)	0.225***	0.229***
	(0.008)	(0.017)
log(value added)	0.091***	0.049***
	(0.008)	(0.017)
exporter	0.046***	0.026
	(0.013)	(0.027)
government owned	0.099	−0.381
	(0.166)	(0.276)
foreign owned	−0.161***	0.006
	(0.019)	(0.033)
Wood, paper, printing industry	0.353***	0.111
	(0.080)	(0.089)
Finance, Insurance	−0.036	0.747***
	(0.152)	(0.168)
Manufacture of wearing apparel	−0.049***	0.598***
	(0.142)	(0.157)
Accommodation	−0.054	0.409***
	(0.090)	(0.100)
Transporting and storage	0.179***	−0.110
	(0.061)	(0.068)
Chemicals and drug industry	0.992***	−0.813***
	(0.160)	(0.177)
Electricity, gas, steam supply	1.303***	0.131
	(0.173)	(0.191)
Water supply; sewerage	1.145***	−0.411***
	(0.121)	(0.133)
Constant	0.152	−0.865**
	(0.333)	(0.369)
Observations	57,407	57,407
Region FE	✓	✓
Industry FE	✓	✓
SME classification FE	✓	✓
R^2	0.710	0.892
Adjusted R^2	0.710	0.892
Residual Std. Error (df = 57356)	2.463	2.724
F Statistic (df = 50; 57356)	2,807.892***	9,457.715***
Note:		*p<0.1; **p<0.05; ***p<0.01

Since there are many firms without any connections outside of its community, we also considered the possibility that the excess zeros in the distribution are generated separately from the data generating process of the count values. As our dependent variables are heavily dispersed, we used a zero-inflated negative binomial regression to explore this alternative approach. Table 6 shows the results for the count model. When we consider the model with the number of

CEU eTD Collection

links pointing to other communities as the dependent variable we can see that the firm size and the export activity are the most influential factors, however, if we use the number of group memberships, also firms operating in wholesale industries seem to have a higher level of embeddedness in other communities.

Regarding the process governing the presence of the excess zeros, in the case of the number of outside links the degree and the export activity of firms are negatively associated with the probability of excess zeros, while operating in the food, agriculture, household goods retail and household goods wholesale industries have positive coefficients. In the case of the number of community memberships, the firm size and the variables indicating the geographic region of firms seem to have larger influence on predicting excess zeros. (Table 7)

Table 6: Count model coefficients (negbin with log link). Among the NACE categories only industries with positive, significant coefficients are listed.

	Dependent variable:	
	# of outside links	# of memberships
	(1)	(2)
degree	0.028***	0.032***
	(0.0005)	(0.001)
log(balance sheet total)	0.225***	0.229***
	(0.008)	(0.017)
log(value added)	0.091***	0.049***
	(0.008)	(0.017)
exporter	0.046***	0.026
	(0.013)	(0.027)
government owned	0.099	−0.381
	(0.166)	(0.276)
foreign owned	−0.161***	0.006
	(0.019)	(0.033)
Other wholesale	−0.127***	0.259***
	(0.031)	(0.068)
Food industry wholesale	−0.358***	0.376***
	(0.040)	(0.077)
Agricultural wholesale	−0.417***	0.523***
	(0.061)	(0.105)
Constant	−3.122***	−2.874***
	(0.184)	(0.338)
Observations	57,407	57,407
Region FE	✓	✓
Industry FE	✓	✓
SME classification FE	✓	✓
Log Likelihood	-58,968.150	-28,877.760
Note:		*p<0.1; **p<0.05; ***p<0.01

CEU eTD Collection

Table 7: Zero-inflated model coefficients (binomial with logit link). Among the NACE categories only industries with positive, significant coefficients are listed.

	Dependent variable:	
	# of outside links	# of memberships
	(1)	(2)
degree	−2.611***	−0.545***
	(0.022)	(0.012)
log(balance sheet total)	0.001	−0.213***
	(0.033)	(0.037)
log(value added)	0.121**	0.063
	(0.039)	(0.039)
exporter	−0.354***	−0.053
	(0.072)	(0.065)
government owned	1.519	0.407
	(1.140)	(0.940)
foreign owned	0.062	0.216*
	(0.133)	(0.105)
Southern Great Plain	0.018	0.092
	(0.098)	(0.099)
Southern Transdanubia	0.076	0.325*
	(0.127)	(0.129)
Northern Great Plain	0.067	0.403***
	(0.100)	(0.104)
Northern Hungary	0.104	0.271*
	(0.122)	(0.128)
Central Transdanubia	0.053	0.415***
	(0.105)	(0.109)
Central Hungary	−0.190*	0.131
	(0.087)	(0.088)
Western Transdanubia	0.075	0.512***
	(0.106)	(0.112)
Food industry retail	1.338***	−0.207
	(0.314)	(0.423)
Food industry wholesale	0.709**	−0.239
	(0.240)	(0.191)
Construction, real estate	−0.505**	−0.009
	(0.129)	(0.144)
Agriculture	0.682***	0.308·
	(0.195)	(0.186)
Household goods retail	0.724***	−0.015
	(0.200)	(0.219)
Household goods wholesale	0.583**	0.67
	(0.225)	(0.208)
Manufacture of metal, plastic	0.441**	−0.188
	(0.166)	(0.162)
Constant	1.551	4.832***
	(1.156)	(0.995)
Observations	57,407	57,407
Industry FE	✓	✓
SME classification FE	✓	✓
Log Likelihood	-58,968.150	-28,877.760
Note:		*p<0.1; **p<0.05; ***p<0.01

CEU eTD Collection

4 Measuring Systemic Risk in the Hungarian Firm-Level Production Network

The data described in the previous chapters enabled us to build a microsimulation model of shock propagation to quantify short-term damages after supply chain disruptions in the production network. Using this model, we show that industry level analysis is not sufficient from the point of view of shock propagation. Our granular approach makes it possible to consider the heterogeneity in the production processes of firms by allowing us to introduce differentiation in their production functions and in the importance of different input types. As an application, we quantify the systemic risk of firms by simulating how much of the production network is affected through distinct upstream and downstream spreading mechanisms. Due to our high-resolution approach we can also explore the most influential companies in more detail. Finally, we will demonstrate that the different allocations of an industry-level shock among the firms in the given sector can lead to a very wide potential damage range in the system, which would remain hidden if one did not consider granular data.

This chapter is based on the paper titled *Systemic Risk and Shock Propagation in Firm Level Production Networks (in preparation)* by Chsristian Diem, Andras Borsos, Tobias Reisch, Janos Kertesz and Stefan Thurner.

4.1 Advantages of firm-level analysis

In this section I provide justification why using firm-level data for analyzing production networks is more beneficial than the traditional industry-based input-output literature.

Firstly, the traditional argument of Lucas Jr (1977) claimed that firm-level idiosyncratic shocks cancel each other out on the aggregate level based on the law of large numbers, hence, it is also satisfactory to only consider factors which have influence on entire industries. However, results obtained in the last decade (e.g. Gabaix (2011) and Acemoglu et al. (2012)) showed, that distributions of several relevant firm characteristics follows power-low functional forms, i.e. there is a relatively high probability of extremely large observations, which can (depending on the exponent of the distribution) violate the assumptions of the law of large numbers. This implies that only more fine-grained data is suitable to assess properly the consequences of shocks

Secondly, our analysis of the data showed that even within fine grained industry classifications firms tend to have very heterogeneous inputs. This can lead to inaccurate results when using them for assessing shock propagation in production networks. Figure 29 shows an illustration for this using the example of NACE sector C25 "Manufacture of fabricated metal products except machinery and equipment“. This category contains 178 firms, resulting in 15753

CEU eTD Collection

possible pairwise combinations. We calculated the Jaccard similarity for every pair based on the NACE4 industry classification of their suppliers, and created a histogram based on these similarity values. We found a Jaccard overlap of 0 for 69.1% of all pairs, meaning that if we compare two arbitrary firms from C25, mostly they don't have any input in common.

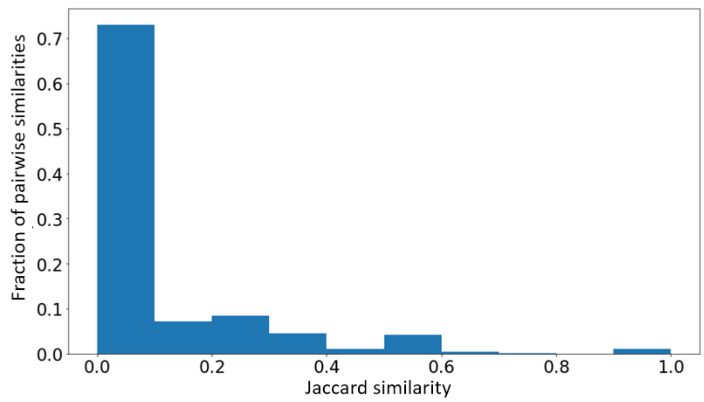


Figure 29: Pairwise similarity of production functions in NACE sector C25 Manufacture of fabricated metal products except machinery and equipment. (Sector 'C25 - Manufacture of fabricated metal products, machinery and equipment' contains 178 firms, resulting in 15753 possible pairwise combinations. We calculated the Jaccard similarity for every pair based on the NACE4 industry classification of their suppliers, and created a histogram based on these similarity values. We found a Jaccard overlap of 0 for 69.1% of all pairs, meaning that if we compare two arbitrary firms from C25, mostly they don't have any input in common.)

Thirdly, heterogeneity is especially important if the crisis scenario does not affect all firms within a sector to the same extent, which is usually the case in reality. There are indefinitely many possibilities to translate a sector level shock to the firm level for the given sector. (For instance a 10% shock in an industry can mean a 10% shock to all firms, but also a 100% shock to 10% of the firms in the industry.) Our firm-level simulation takes this into account and yields different cascades for shocks which would appear to be the same at industry level, but are distributed differently among firms within an industry. This makes it possible to apply more clear-cut shocks, which gives immense versatility in applications. (E.g. it is possible to investigate even a single firms default.)

Fourthly, heterogeneity matters also in the production processes of firms as the functional form of the production function controls how shocks propagate between firms in two ways: (i) It determines the new output level when inputs are not available (downstream spreading), (ii) and the new level of required inputs if the demand for the output drops (upstream spreading). The main considerations regarding firm specific production functions is discussed in the next section.

CEU eTD Collection

4.2 Features of firm specific production functions

4.2.1 Substitutability of inputs

The standard textbook examples of production functions are the linear-, Cobb-Douglas- and Leontief functions. The three production functions for the case of two inputs x_1, x_2 converted into output x_3 are represented respectively as

$$x_3 = \alpha_1 x_1 + \alpha_2 x_2, \qquad x_3 = A x_1^{\alpha_1} x_2^{\alpha_2}, \qquad x_3 = \min[\alpha_1 x_1, \alpha_2 x_2], \tag{11}$$

where α_i are the technical coefficients determining by which rates the conversion takes place, and A denotes total factor productivity[21].

The main effect of choosing the production function is the degree to which firms can substitute between different types of inputs to reach the same level of output. However, substitution of inputs has a different interpretation in the short term and in the long term. Regarding long run perspectives, it usually refers to the substitution between different factors of production (typically between labor and capital, but sometimes also between inputs from different sectors). This has the interpretation of fundamentally changing the way how things are produced. E.g. a table can be made by using mostly manual labor but also by machines, however switching between these two ways of production takes significant time and investment. Similarly, a table can be made out of wood or out of steel while having roughly the same functionality, but different machines are required to switch from one to the other. In the short term however, having one type of machine only, it is not possible to substitute wood with metal, or manual labor with machines. Thus, in the short run, a more relevant question is how much goods and services a firm can still produce if one of its suppliers is not able to deliver the usually provided inputs. In this model, we focus only on this shorter time horizon where the mode of production cannot be changed drastically.

In the case of the linear production function, the loss of a particular input does not play a drastic role since it is assumed that the inputs do not depend on each other. Although the output level will drop, but only relative to the amount of the input. The Leontief type production function represents the other extreme, as in this case all inputs are required in exact proportions to produce the output. It encodes a parts list for a product, and if one of the inputs on the list is not available, the product cannot be "assembled". Thus, the default (non-availability) of a supplier can have the drastic consequence of halting the production of the affected products (or services) completely. The Cobb-Douglas production function can be more flexible. In this case, losing 50%

[21] Note that all three production functions are special cases of the more general constant elasticity of substitution (CES) production function (McFadden (1963)).

CEU eTD Collection

of one input does not translate directly to a loss of 50% in output, but depends on the exponent of the input in the production function.

In economic models with high levels of aggregation and with long term focus the common practice is to use Cobb-Douglas production functions which can capture substitution between the factors of production, typically labour and capital. However, this aspect becomes less and less suitable with a short term scope and at finer resolutions of the production network, where substitution among production factors is very limited. In contrast with sector level models, each firm in our data can be assigned a specific production function based on the given firm's input vector and industry classification. More specifically, we are using a combination of linear and Leontief production functions. The Leontief function is appropriate for firms with physical production processes since it is based on the idea of parts lists and recipes according to which modern production management works. However, the linear production function is more realistic for sectors like distribution (whole sale, retail) and services, where the loss of a single input cannot influence the whole production, instead, the output is directly proportional to the respective input. Fortunately, the NACE industry classification naturally distinguishes between production-based industries and service industries: NACE2 codes up to 45 are related to physical production processes, whereas the codes from 45 to 99 are related to wholesale, retail and different types of services.

4.2.2 Criticality of inputs

Another influential aspect of production processes is the varying criticality of inputs, i.e. which inputs (suppliers) are actually relevant for production in the short term. It is unrealistic to assume that a firm – even with Leontief production function – needs all of the different inputs so crucially such that it has to shut down production if they are not available in the short term[22]. The non-availability of services (like marketing, accounting, etc.) does not cause physical production problems, but might still have an adverse effect on output. However, without iron ore the production of steel seems to be rather impossible. Thus, a model for supply chain shock propagation should also consider how crucial inputs are for the short term functioning of the firm.

In our model, we assume that some sectors produce crucial inputs for other sectors, whereas other sectors produce services whose non-availability does not lead to immediate production problems. A similar approach is taken by (Pichler et al., 2020) which paper uses an industry analyst survey to assess which inputs are crucial for 55 industries of the world input output database. However, the approach of dealing with each industry separately becomes increasingly costly when the granularity of sectors is increasing to NACE2 (88 categories)

[22]For example consider a steel producer requiring services from the hotel industry. It is strong to assume that production is shut down if the respective hotel closes. However, it might still have an impact on the firms production if the service is not available.

CEU eTD Collection

or NACE4 (615 categories), let alone to assess it on the firm level. Even on the industry level it requires 377,610 (=615*614) potential links between industries to be assessed quantitatively. Instead, we resort to a straightforward remedy and categorize each NACE4 level sector as either crucial input (sectors producing physical output), or not crucial input (typically services). We alter the Leontief production function into having two types of inputs, crucial ones and non crucial ones. We treat the crucial inputs as before, but the non-crucial ones are treated as having the same effect as in the linear production function.

From the point of view of contagion along the supply chain, it is also important to differentiate between investment goods like new machines and preproducts. As long as a new machine (investment good) is not an immediate replacement for a broken one (which cannot be repaired) the non-availability of this good causes a missed increase in productivity or capacity, but not a direct disruption in production[23]. On a more abstract level, this is reflected conceptually in production functions where the means of production (capital) are responsible for how efficiently the inputs are transformed into outputs. We address this issue by keeping only those supplier-customer links which occur frequently, and thus, are unlikely to be investment goods. (See the definition for long-term connections in Chapter 3.1.)

4.3 Replaceability of suppliers

As discussed in Section 4.2.1., on shorter time horizons the mode of production can only be changed in relatively subtle ways, e.g. by using a given input of a different supplier. Hence, the last major issue for shock propagation which needs to be addressed before formalizing our model is the problem of replacing a supplier by another who can deliver a sufficiently similar input.

Modeling in a detailed way how different companies can be replaced with each other (as suppliers) would require detailed knowledge of inventory levels, production capacities of suppliers, and detailed product information (e.g. if two suppliers in the same industry can even produce the same good).[24] There are only a few strategies in the firm-level supply chain contagion literature to tackle this challenge, but generally all attempts suffer from data availability shortcomings. Inoue and Todo (2019b) does consider replaceability only between existing suppliers. Wu (2016) uses a different strategy and creates a measure of replaceability based on the weight of a given supplier in its costumers' production-related cost. This measure assumes that if a supplier is an important part of a firm's production, then it is harder to find a substitute for it. Another often used solution is to distinguish between standardized goods (goods with a clear reference price listed in trade publications) and differentiated goods (goods

[23]Investment products can still matter for upstream contagion. As a possible extension of the model, two separate matrices could be used for downstream and upstream contagions.

[24]Furthermore, a consistent modeling of supplier replacement would include a dynamic rewiring of the network, which is extremely difficult to model in a realistic way.

CEU eTD Collection

with multidimensional characteristics) based on Rauch (1999). Although this distinction can be used as a categorical variable in econometric estimations e.g. (Giannetti et al. (2011)), it is not informative about the extent of replaceability even for standardized products. Barrot and Sauvagnat (2016) uses two other proxies as well to measure the specificity of suppliers: the level of R&D expenditures and the number of patents held by a firm. Unfortunately these pieces of information are only relevant for a tiny fraction of companies, and not at all applicable for the whole network of Hungarian firms.

In our model we propose a different strategy, which employs a straightforward, data driven way to construct a short-term supplier replaceability index based on intra-industry market shares. The basic intuition is that a supplier having a small market share within its industry should be on average relatively easy to replace by a small increase of the production of its competitors to cover the additional demand for their products. However, a supplier producing a considerable share of the goods in a given industry is more difficult to replace, as it is unlikely that its competitors can increase their production immediately to cover the additional demand[25].

Additionally, the alternative suppliers in the given industry might also have experienced shocks during the contagion processes in the model, consequently, to make this approach more realistic, we took into account also the deteriorations in their production capacity. Nevertheless, it is important to distinguish at this point between two different sources which could cause drawbacks in the production level of these potential alternative suppliers. On the one hand, one should account for downstream shocks, i.e. shocks coming from the suppliers, which is a truly limiting disruption in their production. On the other hand, one should *disregard* the upstream shocks experienced by them, because demand shocks coming from costumers are not actual restraints on their production (at least from the point of view of replacing their competitors). This is an important distinction, because this way the alternative suppliers can be compensated for their previously experienced demand side shocks, furthermore, we can also take into account that in the case of a system-wide crisis it might not be possible to find alternative suppliers.

4.4 Description of the model

Based on the available data we formulate a newly proposed shock propagation model for firm level production networks.

We consider n companies with company i having production function f_i and a weighted supplier-buyer network $W \in \mathbb{R}_+^{n \times n}$, where W_{ij} denotes the value of goods i supplies to j. Furthermore, we have a product category vector

[25] The available data does not make it possible to identify crucial suppliers which might have small market share in their industry, but they provide unique inputs. This can happen for example because of a special technology or a tailor-made product.

CEU eTD Collection

$p \in \{1,\ldots,m\}$, where $p_i = k$ indicates that company i is producing products of type k. Since actual product information is not available, we have to use the NACE four-digit industry classification of company i instead. Consequently, we also have to assume that all the links W_{ij}, starting from company i correspond to the product type p_i. The vector $y \in \mathbb{R}^n_+$, with y_i indicates the value of products of type p_i, firm i is producing. The element Π_{ik} of the input matrix $\Pi \in \mathbb{R}^{n\times m}_+$ indicates the value of input of type k firm i uses for production. The production equation for company i is then

$$y_i = f_i(\Pi_{i1}, \Pi_{i2}, \ldots, \Pi_{im}) \tag{12}$$

From the matrix W and the vector p we can reconstruct the input vector $\Pi_{i.}$ of company i. The ith column of W contains the supplier vector of company i, i.e. $W_{.i}$, indicating the amount other companies sold to i. For each input k we sum over all suppliers j belonging to this category, i.e.,

$$\Pi_{ik} = \sum_{j=1}^{n} W_{ji}\delta_{k,p_j} \quad , \tag{13}$$

where $\delta_{ij} = 1$ is the Kronecker Delta. We substitute this expression into Equation 12, and receive the network dependent production equation

$$y_i = f_i\Big(\sum_{j=1}^{n} W_{ji}\delta_{1,p_j}, \sum_{j=1}^{n} W_{ji}\delta_{2,p_j}, \ldots, \sum_{j=1}^{n} W_{ji}\delta_{m,p_j}\Big) \tag{14}$$

The second equation that we assume to hold is the output supply equation:

$$y_i = \sum_{l=1}^{n} W_{il} \tag{15}$$

From Equation (14) it is obvious that a loss of a supplier has an implication on output y_i. This loss in output spreads further *downstream* to the buyers of firm i. Similarly, if Equation (15) holds, the loss of a customer will translate into decreased demand of y_i and will spread further *upstream* to the suppliers of i.

The actual specification of f_i is probably the most important determinant for downstream shock propagation, since it determines how strong the effects on output are if a supplier cannot deliver. Additionally, we need to specify a rationing mechanism for supplying buyers with a reduced amount of output. We choose to simply impose a proportional rationing mechanism, i.e. each company gets the same fraction of its original demand. In financial contagion modeling this is a common assumption (see, for example, Eisenberg and Noe (2001)).

CEU eTD Collection

In contrast to downstream contagion, upstream contagion can be modelled independently of the production function. For this we have to assume that (i) the one product per company assumption holds, (ii) input proportions stay the same for different levels of output, (iii) and that a reduction in the output affects all inputs proportionally to their original levels[26]. Similarly to the downstream contagion, we assume a rationing mechanism with the same logic: if there is more than one supplier for a product, we assume proportionality based on their initial importance.

To define a contagion process on a network, firstly we describe the initial state, and then we formulate the dynamics taking place on this system.

We assume that the initial state corresponds to $t = 0$ and that the initial shock occurs at $t = 1$. We at time $t = T$ observe a new intermediary stable state [27] of the system, when the effects of the initial shock are incorporated into the production levels of all firms.

To do this, we first introduce the state variable corresponding to the production level of a company relative to its initial production level before the shock. Let $y_i(0)$ be the initial amount of products p_i produced by company i and $y_i(t)$ is the amount company i produces at time t after adjustments to the initial shock. Then we define the state variable as

$$h_i(t) = \frac{y_i(t)}{y_i(0)} \tag{16}$$

Given that the initial shock is negative, and assuming that in the short term firms cannot increase their production level, $h_i(t) \in [0, 1]$ with initial value $h_i(0) = 1$. Thus, the value $h_i(t)$ quantifies the fraction of the production level at time t compared to the original level at time $t = 0$ before a shock occurred. We initialize the shock at time $t = 1$ by setting $h_i(1) = \psi_i$ where $\psi_i \in [0, 1]$ represents the severity of the initial shock faced by company i. ψ_i is interpreted as the percentage of the original production $y_i(0)$ that is lost due to this shock. This abstract specification of the shock is flexible enough to represent the mere failure of a single company, but also a system wide event. The shock experienced at time $t = 1$ unfolds according the dynamics we specify with a recursive update for $h_i(t)$.

The production level of a company $h_i(t)$ embedded into the production network W can be affected in two different ways. First, $h_i(t)$ can drop due to a downstream shock, which is transmitted from suppliers to buyers. Second, $h_i(t)$ can drop due to an upstream shock, which is transmitted from customers to suppliers. Downstream and upstream shocks both affect the level of produc-

[26]For a physical production process these assumptions seem justified, but for the creation of services the situation is less clear.

[27]This stable state is in fact just a hypothetical intermediary state before the system recovers again. In that sense it is a worst case scenario if no mitigating action is taken by any of the involved agents.

CEU eTD Collection

tion, $h_i(t)$, but spread differently from i onward to other neighboring nodes. Thus, we introduce two state variables, both closely related to h. The variable $h_i^d(t)$ keeps track of the production level firm i can maintain after considering all downstream shocks it faces up to t. The variable $h_i^u(t)$ keeps track of the production level firm i can maintain after considering all upstream shocks it faces up to t. (The separate treatment of the up- and downstream effects has its limitations. These are discussed in Section 4.6.)

In the following subsections we describe the details of these contagion process. Firstly, we will describe the applied production functions. Then we derive the update rules for $h_i^d(t)$, modeling the spreading of downstream shocks (with mixed production functions to account for firms having different types of productions). Thirdly, we will introduce for each company a class of non-crucial production inputs, then we propose an extension to take into account the replaceability of suppliers. Finally, we will describe the update rules for $h_i^u(t)$, representing the spreading of upstream shocks (independent of the production function).

4.4.1 Formalization of the production functions

In our application, we consider two types of production functions f_i: the Leontief production function for physical production companies and the Linear production function for services and wholesale / retail industries. We assign each company one of the two production functions based on their sector affiliations as discussed in Section 4.2.1.

We define the Leontief production function for company i as

$$y_i = \min(\frac{1}{\alpha_{i1}}\sum_{j=1}^{n} W_{ji}\delta_{1,p_j}, \frac{1}{\alpha_{i2}}\sum_{j=1}^{n} W_{ji}\delta_{2,p_j}, \ldots, \frac{1}{\alpha_{im}}\sum_{j=1}^{n} W_{ji}\delta_{m,p_j}) \tag{17}$$

We set the corresponding parameters α_k for company i to

$$\alpha_{ik} = \frac{\sum_{j=1}^{n} W_{ji}\delta_{p_j,k}}{\sum_{l=1}^{n} W_{il}} \tag{18}$$

Note that with this definition, effectively every company has a different production function.

Equation 18 corresponds to the definition of technical coefficients in the input-output literature. The value $1/\alpha_{ik}$ specifies the amount of product group k that is required by company i to produce one unit of product group p_i. α_{ik} is simply the weight of inputs from industry i relative to output the company produces. Since W represents monetary flows between companies, a firm needs

CEU eTD Collection

to spend α_{ik} monetary units on inputs of type k to produce one monetary unit of p_i. (Note that for a source node $\alpha_{i.}$ would be 0, and for sink nodes it would be infinity.)

We define the linear production function as

$$y_i = \frac{1}{\alpha_i}\sum_{j=1}^{n} W_{ji} \quad , \tag{19}$$

with

$$\alpha_i = \frac{\sum_{j=1}^{n} W_{ji}}{\sum_{l=1}^{n} W_{il}} \tag{20}$$

It can be seen immediately that this formulation implies that Equation 19 is a special case of Equation 17 when there is only one product type considered.

4.4.2 Downstream shock propagation

The output which company i, with production function according to Equation (17) can produce at $t+1$ depends on the available inputs at time t and the level of the exogenous shock ψ_i. The available amount of input k at time t depends on the current production level of the suppliers $h_j(t) = y_j(t)/y_j(0)$ and is $\sum_{j=1}^{n} W_{ji}\delta_{k,p_j} h_j^d(t)$. Thus, we have the following recursion for the output:

$$y_i(t+1) = \min\Big(\frac{1}{\alpha_{i1}}\sum_{j=1}^{n} W_{ji}\delta_{1,p_j} h_j^d(t),\, \dots,\, \frac{1}{\alpha_{im}}\sum_{j=1}^{n} W_{ji}\delta_{m,p_j} h_j^d(t), \psi_i y_i(0)\Big) \tag{21}$$

To formulate the recursion relation in terms of $h_i^d(t)$ we divide both sides by the level of initial output $y_i(0)$.

$$\begin{aligned} h_i^d(t+1) = \min\Big(&\frac{1}{\alpha_{i1}}\frac{1}{y_i(0)}\sum_{j=1}^{n} W_{ji}\delta_{1,p_j} h_j^d(t), \\ &\dots, \frac{1}{\alpha_{im}}\frac{1}{y_i(0)}\sum_{j=1}^{n} W_{ji}\delta_{m,p_j} h_j^d(t), \psi_i\Big) \end{aligned} \tag{22}$$

For practical reasons, we apply the respective production function on the updated relative input availability matrix $\bar{\Pi}(t)$, which shows the percentage of input k still available to firm i at time t[28]. For firms with Leontief production

[28] The detailed justification for this, as well as the exact derivation of the formulas can be found in the Supplementary Materials of Diem et al. (in press).

CEU eTD Collection

function the recursion becomes:

$$h_i^d(t+1) \quad = \quad \min\left(\bar{\Pi}_{i1}(t), \bar{\Pi}_{i2}(t) \ldots, \bar{\Pi}_{im}(t), \psi_i\right) \quad , \tag{23}$$

where

$$\bar{\Pi}_{ik}(t) = 1 - \sum_{j=1}^{n} \Lambda_{ji}^d \delta_{k,p_j} (1 - h_j^d(t)) \tag{24}$$

with Λ_{ji}^d showing the weight of supplier j within a product category among firm i's suppliers:

$$\Lambda_{ji}^d = \begin{cases} \frac{W_{ji}}{\sum_{l=1} W_{li}\delta_{p_l,p_j}} & \text{if } W_{ij} > 0 \quad , \\ 0 & \text{else} \end{cases} \tag{25}$$

For firms with linear production function the recursion is:

$$h_i^d(t+1) = \min\Big(\sum_{k=1}^{n} \bar{\Pi}_{ik}(t) \frac{\Pi_{ik}(0)}{\sum_{k=1}^{n} \Pi_{ik}(0)}, \psi_i\Big) \tag{26}$$

4.4.3 Extensions of the algorithm

Now, we can extend this basic recursion for downstream contagion to account for the fact that not all inputs are equally crucial for a company with Leontief production function, and that some suppliers can be replaced by others.

Firstly, we introduce the following changes in Eq (23). We define a new set $\mathcal{O}^1$ containing the indices of all industries k, which are not crucial for firms with Leontief production function, and a set $\mathcal{O}^2$ containing the ones which are crucial. The crucial inputs will be handled as before, but we treat the non-crucial ones as having the same effect as in the linear production function. This way, we can write the update equation as

$$h_i^d(t+1) = \min\Big(\min_{k\in\mathcal{O}^2} \left(\bar{\Pi}_{ik}(t)\right), \sum_{k\in\mathcal{O}^1} \bar{\Pi}_{ik}(t) \frac{\Pi_{ik}(0)}{\sum_{k=1}^{n} \Pi_{ik}(0)}, \psi_i\Big) \tag{27}$$

Finally, to take into consideration the replaceability of the suppliers, we modify Equation 24 by adding a new term to the formula, $\sigma_i(t)$ representing the market share of a supplier within its industry (NACE4) with the corrections described in Section 4.3.:

CEU eTD Collection

$$\bar{\Pi}_{ik}(t) = 1 - \sum_{j=1}^{n} \sigma_j(t) \Lambda^d_{ji} \delta_{k,p_j} (1 - h^d_j(t)) \tag{28}$$

where

$$\sigma_i(t) = \frac{y_i(0)}{\sum_{j=1}^{n} y_j(t) h^d_j(t) \delta_{p_j,p_i}} \tag{29}$$

The interpretation of the replaceability of suppliers in the model can be illustrated by the following example. Let us assume that a supplier with a 10% market share in its industry (after considering also its competitors' state) is responsible for 50% of the input required by one of its partners in a given product category. If the production level of this supplier drops to 80%, then the production of its partner will decrease (based on Equation 28 by 10% $\times$ 50% $\times$ (1 $-$ 80%) = 1%. If we disregarded the possibility of replacing this supplier, the corresponding decrease in the firm's output would be 10%. This way, we enable in the model to replace missing supplies by taking into account the market conditions in a twofold way. On the one hand, $\sigma_j(t)$ reflects the fact that the given input in this example can be bought from the remaining 90% of the market. On the other hand, we also acknowledge that this replacement might not be possible entirely (or it entails some costs), which we can also proxy by the market share of the firm in consideration[29].

4.4.4 Upstream shock propagation

Now we turn to the formulation of upstream contagion. As explained above, $h^u_i(t)$ keeps track of the production level firm i can maintain after considering all demand reductions it faces from its customers up to t. We formulate a recursion for $h^u_i(t)$ based on Eq (15). The level of production of company i in response to demand reductions from its customers is

$$y_i(t+1) = \min\Big(\sum_{j=1}^{n} W_{ij} h^u_j(t), \psi_i y_i(0)\Big) \tag{30}$$

From this we can formulate the recursion relation in terms of $h^u_i(t+1)$:

[29] In the replaceability factor we do not model the new links which are formed as substitutes for missing inputs. This also means that we do not assume that replacing Hungarian suppliers can only happen with other Hungarian suppliers. This way we allow firms to find replacement even from abroad. The only restriction is that the replacement potential is proportional to the market share of the distressed supplier, and this calculation is based on the observable Hungarian market share.

CEU eTD Collection

$$h_i^u(t+1) = \min\left(1 - \sum_{j=1}^{n} \Lambda_{ji}^u (1 - h_j^u(t)), \psi_i\right) \tag{31}$$

with Λ_{ji}^u showing the weight of supplier j within a product category among firm i's suppliers:

$$\Lambda_{ji}^u = \begin{cases} \frac{W_{ij}}{\sum_{l=1}^{n} W_{il}} & \text{if } W_{ij} > 0 \\ 0 & \text{else} \end{cases} \tag{32}$$

The update equations are iterated until the algorithm converges at time

$$T := \min_t \{t \in \mathbb{N} | \max\left(h^d(t) - h^d(t+1), h^u(t) - h^u(t+1)\right) \leq \epsilon\} + 1 \quad , \tag{33}$$

where $\epsilon = 10^{-2}$ is chosen as convergence threshold. Thus, we assume that the propagation stops when all shocks are smaller than ϵ, and the corresponding time point of convergence is T.

The final production level after convergence for each company is defined to be $h_j(T) = \min(h_j^d(T), h_j^u(T))$, i.e. the minimum of the demand and the supply constrained production capacity of the company.

However, before generating the results, it is necessary to make one last correction. We know that the available data in the matrix W does not cover every transaction due to the reporting threshold and the missing outbound transactions. Thus, the out-strength does not match exactly the revenue of a company and the in-degree does potentially not represent the total amount of material costs. This leads to an overestimation of the importance of the observed links, W_{ij}. To correct for this, the matrices used for the shock propagation Λ^d and Λ^u can be adjusted with the available income statement data. In the case of upstream contagion where the elements of Λ^u are row standardized, the revenue can be used instead: $\Lambda_{ij}^u = \Lambda_{ij}^u \frac{s_i^{out}}{revenue_i}$, where $revenue_i$ is the revenue of company i in the respective time period. For the downstream direction, we can do the same procedure with material costs: $\Lambda_{ij}^d = \Lambda_{ij}^d \frac{s_i^{in}}{cost_i}$.

4.5 Results

To derive a measure for the severity of the initial shock, the production level $h_j(T)$ needs to be weighted by a factor v_j determining the importance of the company. The weight vector v can be chosen according to different aspects of interest. Obvious choices are the number of employees, the value added, or the revenue of the company. The in- or out-strength can also be interesting choices from a network perspective.

CEU eTD Collection

The production network loss for a given initial exogenous production constraint vector ψ is calculated as the weighted loss of production:

$$c_\psi = \sum_{j=1}^{n} v_j (h_j(1) - h_j(T)) \tag{34}$$

In order to measure the systemic risk a single company i imposes on the rest of the system, we conduct the following analysis. We set $\psi_i = 0$ for firm i and $\psi_j = 1 \forall j \neq i$ for all other firms j. Thus, we look at the hypothetical loss of production the rest of the system faces if firm i were to be taken out of the system. Then, we can call this the systemic production risk index c_i:

$$c_i = \sum_{j=1}^{n} v_j (h_j(1) - h_j(T)) \tag{35}$$

By calculating this for all the firms in the network, we can compose the systemic risk profile of the Hungarian production network for the year 2017. Figure 30 shows this by putting the individual firm's systemic risk measure in decreasing order. We present here the results generated by eight different versions of our model. This way one can compare the effect of applying firm specific linear and Leontief production functions with differentiation between crucial and non-crucial inputs ('Mixing 2)' to alternative cases where we disable the crucial - non-crucial distinction ('Mixing 1'), or we use uniformly linear or Leontief production functions for each firm. Furthermore, these four types can also be examined with and without the supplier replaceability feature.

These plots underpin the intuitions that the general assumption of Leontief production functions result in implausible collapses in the economy. On the other hand, using only the linear functional form can most likely give only a lower bound estimation. The two mixed versions of the production functions in the network gave rather similar outcomes, which indicates a limited importance for our input criticality assumption. However, the mitigating effect coming from the replaceability of suppliers is much more significant. Depending on the assumed production functions this factor can alleviate the cascades to a considerable extent.

CEU eTD Collection

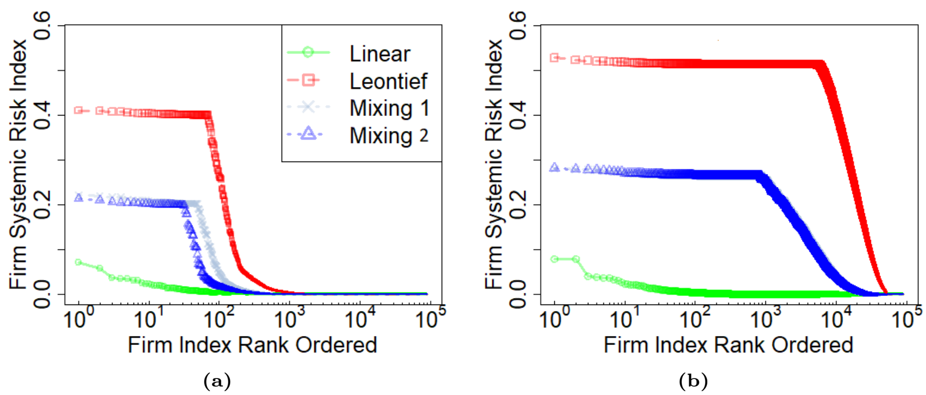


Figure 30: Systemic risk profile of the Hungarian firms in the cases of four different production function assumptions. Panel (a) shows the results when allowing for 'replaceability of suppliers', while panel (b) shows the results when this feature is disabled.

Our simulations show that only less than 100 firms have the potential to destroy more than five percent of the Hungarian national production network, and hence, pose a threat to the overall economy. In the case of the default of even only one of these companies, up to 21% of all production can be affected, however, the vast majority of the firms have only very limited impact on the production network. As Figure 31 demonstrates, these results are robust to different weighting strategies as well. (We performed the same simulations also for the year 2016, and the results proved to be very similar to the 2017 numbers.)

CEU eTD Collection

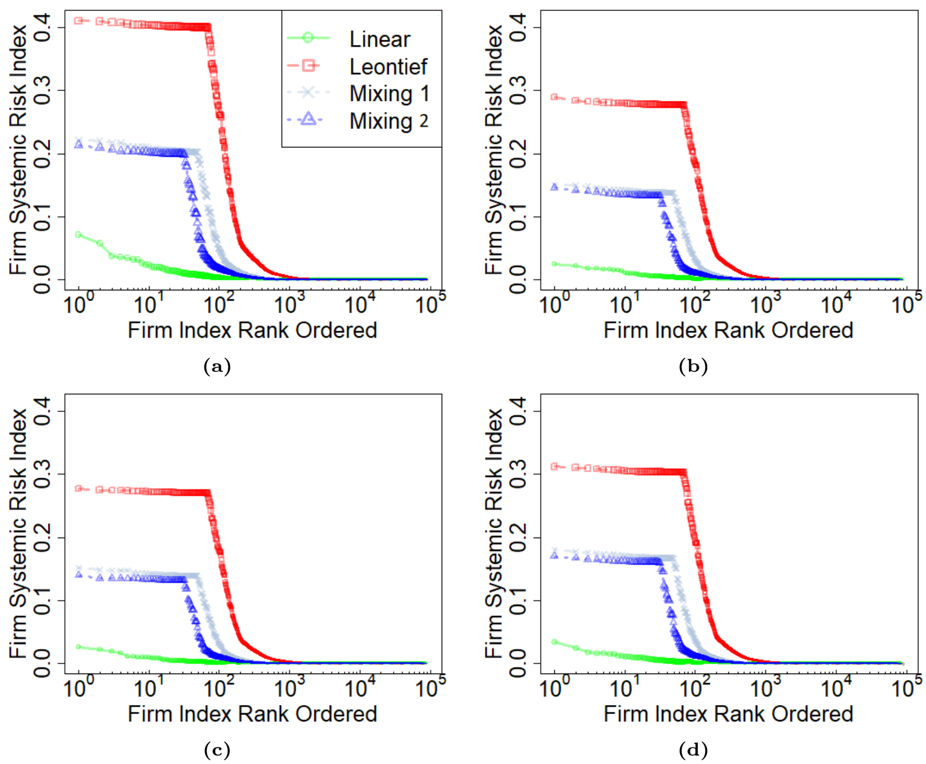
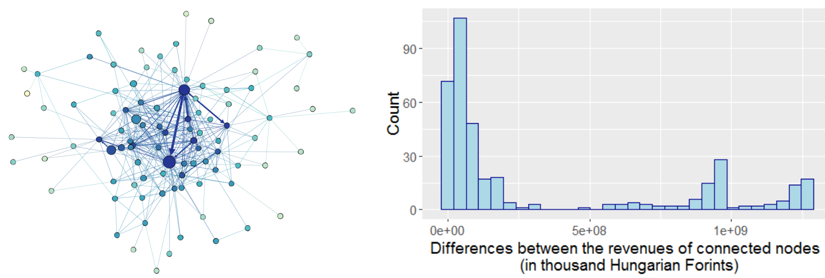


Figure 31: Systemic risk profile of the Hungarian firms with different weights. Panel (a) shows the results with out-strength (i.e. sales based on the network data) as weights, panel (b) shows the results with the revenue of firms, panel (c) was created using the number of employees, while panel (d) uses the value added of firms as weights.

Due to our high-resolution approach, we could further analyse the results to explore the most influential companies in more detail. We found that the list of the top 100 most systemically important firms contains not only intuitively expected large companies, but also quite a few SMEs. After examining the subgraph of the top 100 companied, one can observe many links where the difference between the sizes (based on firms' revenue) of the connected nodes is extremely high. (Figure 32) This indicates that the smaller firms on the top list are probably crucial suppliers to the largest Hungarian companies, consequently their default can have similar system-wide implications than that of the largest firms. (The same pattern can be observed in the 2016 network as well.) This demonstrates that the knowledge of the systemic riskiness of single firms is crucial for understanding and preventing potentially large failure cascades in these networks.

CEU eTD Collection



(a) **(b)**

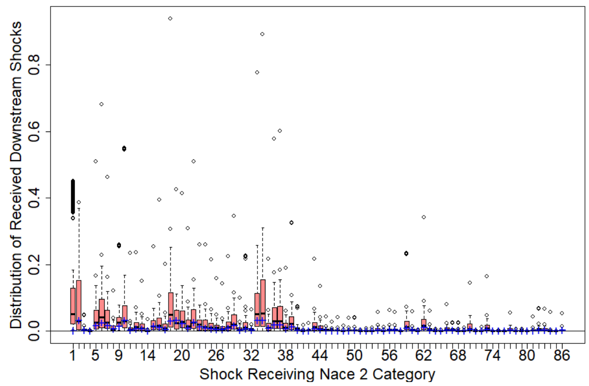
Figure 32: Connections between the top 100 systemically most important Hungarian firms. Panel (a) shows the visualization of the subgraph of the most systemically important companies. Darker coloring of the nodes means higher revenues, the size of the nodes corresponds to their strength, and the thickness of the links reflects the value of the trade connection between the nodes in the given year. Panel (b) shows the histogram of the differences between the revenues of the connected firm pairs.

We could also use our pipeline to illustrate the implications of the potential heterogeneity in the distribution of shocks among firms in a given sector. To demonstrate this, we selected a single industry for which we simulated a 20% shock the following way. We generated 1000 different realizations of this 20% shock by distributing the shock across companies in this sector differently in each scenario. We aggregated the losses in each of the 1000 realizations to NACE 2 industry-level to highlight how different the consequences are in the distinct parts of the economy depending on the allocation of a shock which would seem to be identical in a sectoral level analysis. Figure 33 shows the results of this exercise in the case of shocking the 'Crop and animal production, hunting and related service activities' sector (which corresponds to the 1st NACE 2 category).

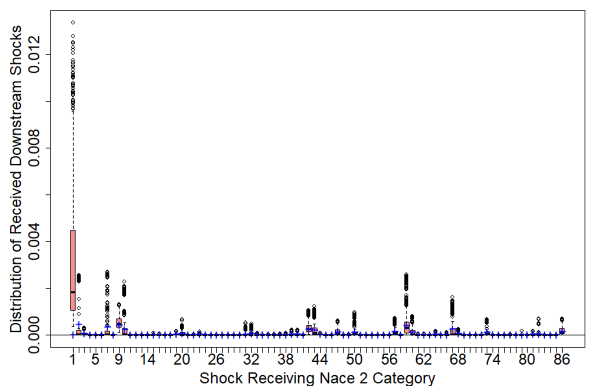
The x axis shows the NACE2 categories which are affected by the shock propagation process, while the y axis shows the empirical shock distribution for the 1000 shock scenarios. To offer an even more detailed view, we separated the downstream (Panel 'a') and the upstream effects (Panel 'c'). (Panel 'b' shows the downstream effects under the assumption of linear production function for all firms.) The blue '+' symbols highlight the scenario in which all the firms in the initially shocked sector are affected uniformly (i.e. each company in the sector is hit by a 20% shock).

The heterogeneity can be illustrated for example with the upstream affectedness of 'Veterinary activities', which corresponds to the 75th NACE 2 category. When the initial shock hits more animal production firms, the veterinary sector is strongly affected. However, if rather firms dealing with growing plants are affected, the shock to this sector becomes way weaker.

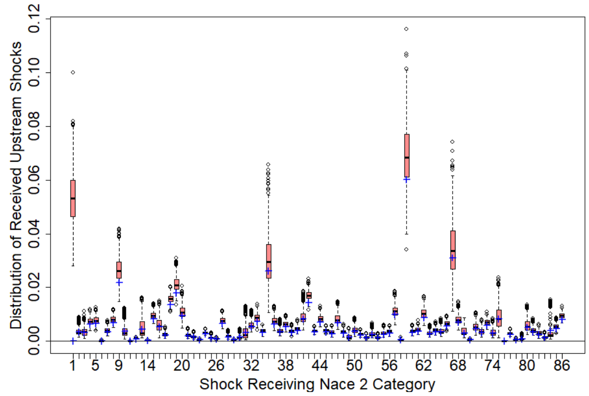
CEU eTD Collection



(a)



(b)



(c)

Figure 33: Distribution of industry-level (NACE2) losses in the case of 1000 different firm-level allocation of a 20% shock to the 'Crop and animal production, hunting and related service activities' sector. Panel (a) shows the results for downstream shock propagation with mixed production functions and differentiation between crucial and non-crucial inputs. Panel (b) shows the results for downstream shock propagation only with linear production function assumed to all firms. Panel (c) shows the results for upstream shock propagation.

CEU eTD Collection

4.6 Limitations and potential extensions

Although one of the key feature of our model is its extensive data orientation, there are still room for improvement in this direction as well. In Section 3, I have already described the limitations of the supplier network data, but there is also another area in which we lack an important piece of information. Since we cannot see product-level transactions, we could only use the industries of firms to label their products. This simplification diffuses the shock propagation effects by spreading contagion to a wider set of firms and decreases impacts on single firms. (For example, if a trader company is facing with a demand shock, we do not know which goods are affected, so we will reduce all the inputs proportionally, which blurs the channels of upstream contagion.) Additionally, if we observed product-level transactions, the replacement of inputs could also become more realistic. However, even if this information was available, we would need to create a mapping between specific inputs to the different outputs produced, i.e. a product recipe for each product. This is probably the next frontier for data driven (contagion) modeling of firm-level production networks, which we do not consider to be obtainable in the short run.

Another aspect in which our work suffers from limitations is the application of a relatively simple, proportional rationing algorithm. Although this solution helps keeping the model tractable, it also entails some compromises. For instance, if a firm buys less from an input, and it has more than one suppliers in that category, they all suffer a reduction of demand. In reality companies sometimes manage suppliers by introducing some kind of pecking order, e.g. by acknowledging that less vulnerable suppliers could bear more demand reduction. A further problem, that the independent updating of h^d and h^u can lead to the companies being unnecessarily constrained. A firm with two suppliers could request from one supplier more than it can deliver, and from the other less then it could deliver. Similarly, a firm with two buyers could send to the first less than it requests and to the other more than it requests. One possible remedy could be if we made the rationing mechanism not only proportional to the initial shock, but to the current level of upstream and downstream constraints as well. However, this dynamic adjustment would be computationally too expensive for our available infrastructure, and can also lead to the potential loss of monotonicity of h^d and h^u, hence, convergence would not be guaranteed anymore.

This leads to another area which offers a way to improve the plausibility of the model, namely, the elaboration of the interactions between the upstream and downstream shock propagation channels. Currently, when firms replace their suppliers (as a response to a downstream shock), we do not model the formation of new links, thus, we ignore the improvement in the new supplier's h^u. Furthermore, in reality not only suppliers, but also buyers might be replaceable. If a costumer decreases its demand toward a supplier (negative upstream shock), the supplier might be able to sell more to another costumer (positive downstream

CEU eTD Collection

shock). To include these mechanisms in an accurate way, we would need to also take into account inventory levels, and unused production capacities.

With these improvements we could also utilize buffers to soften shocks by allowing the continuation of production even if crucial supplies become unavailable. Furthermore, if one can estimate the time period for which a firm can keep up its production without the supply sources, it would also offer a way to introduce *time* in the modeling framework. Currently our model produces theoretical fixed points on an infinite time horizon, which are still useful for many applications, but they correspond to unrealistically severe outcomes. By adding buffers to the model, we could also give more plausible estimates about the extent of the cascades for sensible time intervals.

However, inventories can shield companies only against downstream shocks but not against upstream shock propagation. In the case of reduced demand, financial buffers can be more relevant to keep the firm operational while it generates reduced sales revenue. By modeling the liquidity and solvency positions of firms, several new applications would become attainable. Production network cascades are not confined to the real economy itself, but have financial implications for firms' and thus, banks' balance sheets as well. Failures or reductions in production lead to lost revenues in tandem with lost profits and cash-flows (Barrot and Sauvagnat (2016)). Consequently, firms' ability to repay their debt will deteriorate, and the demand for bank loans to bridge liquidity gaps will increase. Even if the affected firms initially survive the financial effects of a production shock, large losses weaken firms equity cushions, thereby also the long term probability of default can increase. Furthermore, firms could end up in a state where they are able to cover their running costs, but cannot repay the liquidity bridge loans received during the crisis. This so called 'zombification' of firms (see, for example, Caballero et al. (2008) and Blattner et al. (2019)) can further inflate the impact on financial institutions and on the overall economy.

Fortunately, since we have access to detailed balance sheet and profit and loss statements, we have the opportunity to explore several of the potential further research directions described in this section. However, in this thesis these extensions are not yet available, hence, in the next section I will present a less mechanistic, econometric approach to connect the production network to the financial sector.

CEU eTD Collection

5 Shock propagation in the banking system with real economy feedbacks

As we experienced also during the escalation of the crisis after 2008, shock events either in the financial sector or in the real economy can be easily transmitted to the other. Since in most countries banks are the main sources of firms' financing, if the banking system is hit by a shock, it can lead to financial problems for firms dependent on bank loans[30] due to the lower lending activity. In turn, if the real economy is declining, banks can suffer losses e.g. on non-performing loans. Crucially, these shocks can even be augmented not only in the banking system, but also in the production network of firms. Hence, the environment in which the underlying processes beyond the observed emergent phenomena in the financial system are taking place is not limited to the financial sector, but it interferes heavily with the realm of the real economy.

In this chapter we are taking a step towards simulating a banking system contagion model with real economy interlacement utilizing the actual observed system of the interacting network of all the firms and banks in Hungary. In the applications of this model we attempt to demonstrate that the feedback mechanisms in these coupled networks could amplify the losses in the economy beyond the shortfalls expected when we consider the subsystems in isolation. As a test for this, we embed the model into a liquidity stress testing framework of the Central Bank of Hungary. We believe, that this high resolution representation of the economy grants high validity of the simulation results, and makes the model especially suitable for policy analysis. To illustrate the versatility of our agent-based modeling framework, we present two further applications for different policy purposes: (i) We elaborate a way to use the model for SIFI identification, (ii) and we show an example of assessing the impact of shocks originated in the real economy.

This chapter is based on the paper titled *Shock Propagation in the Banking System with Real Economy Feedback (forthcoming)* by Andras Borsos and Bence Mero.

5.1 Description of the model

In this section we provide intuitive description and justification for our work, while Section 4 will show the exact formulation of our simulations.

Our model can be divided into four theoretical blocks (Figure 34):

- In the first block, we model the adjustments and contagions in the banking system after an exogenous shock.

[30] Most importantly SMEs are vulnerable to these shocks as they cannot raise capital or issue bonds so easily as listed companies.

CEU eTD Collection

- As an adjustment mechanism of banks, a credit supply shock hits the real economy, which increases firms' probability of default (PD) on their loans.
- The amplification of the shock in the production network further increases firms' PD.
- As a feedback from the firm network, banks suffer losses on their corporate loan portfolios.

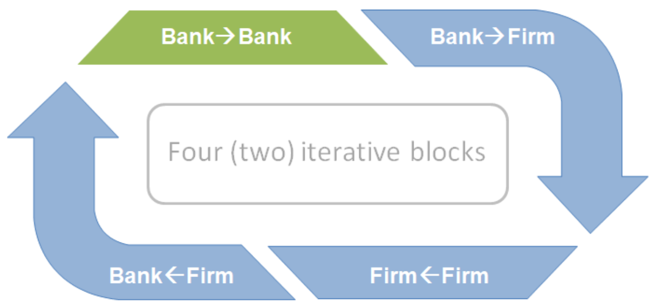


Figure 34: Our framework consists of four modeling blocks: (i) contagions in the banking sector, (ii) shock propagation from banks to firms, (iii) assessing the amplification of these shocks in firms' production network and (iv) feedback from firms to the banking system.

As the last three blocks are all parts of the process describing how credit supply shocks translate into an increased probability for firms becoming nonperforming on their loans, these can be handled together during the implementation as one modeling unit describing the real economy feedback. Before we describe the detailed formulation of the simulation steps, we introduce the four blocks separately.

5.1.1 Banking system[31]

Our model of the banking system contains two channels of contagion and several mechanisms that capture banks' adjustment. One source of contagion is happening through the interbank lending market: If a bank suffers a loss of a magnitude that results in its failure, and thus it becomes unable to repay the loans it borrowed in the interbank market, it causes losses to its partners. The second channel stems from the form of bank adjustment when a bank attempts to improve its position by selling assets whose price may change as a result of these transactions, and thus other banks also suffer losses because of the price change. (This mechanism is hereinafter referred to as fire sales.) According to

[31] The description of the banking system block is an edited version of Box 10 from the 2016 May *Financial Stability Report* of the Hungarian Central Bank. (https://www.mnb.hu/en/publications/reports/financial-stability-report/financial-stability-report-may-2016)

CEU eTD Collection

the logic of the model, contagion and adjustment mechanisms follow one another cyclically until the fixed point of the system is reached[32]. (Figure 35)

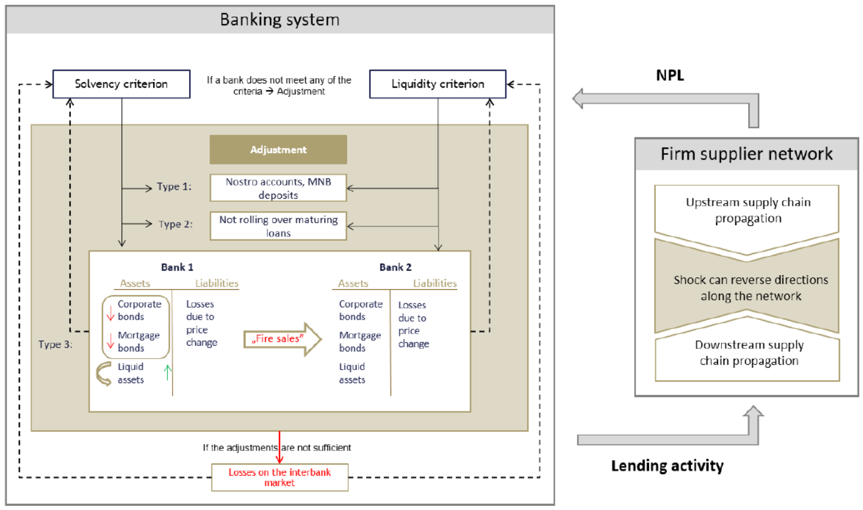


Figure 35: Schematic structure of the banking block. Regular arrows indicate adjustment options, while dashed arrows shows occurrences of losses.

During running the model, first we examine whether the given bank meets the levels of the liquidity (Liquidity Coverage Ratio - LCR) and solvency (Capital Adequacy Ratio - CAR) indicators required by the regulatory authority. If not, to meet the regulatory requirements, banks have to adjust their balance sheets. Our assumptions regarding the adjustment options are built on empirical findings in the European banking system: Brinkhoff et al. (2018) shows the results of the European Systemic Risk Board's macroprudential surveys that aim to assess banks' behaviour in macroeconomic stress scenarios. They have found that lowering credit risk exposures is the largest component of the expected reduction in their risk-weighted assets. Additionally, Behn et al. (2019) also showed that banks in danger of breaching regulatory requirements often choose socially detrimental adjustment strategies, most of all by reducing lending activity. The assumption that banks would even use balance sheet transformation which entail fire sales contagion to raise liquidity in a stress situation is supported by e.g. Allen and Carletti (2008), Adrian and Shin (2010) and Diamond and Rajan (2011). However, adjustment steps can differ between countries due to country- and bank-specific dissimilarities. As we could implement our model on Hungarian data, we fine-tuned the assumptions to the Hungarian experiences during the 2008 crisis. Furthermore, the exact adjustment opportunities can vary depending on the application as well. The assumed behaviors of banks

[32] Eisenberg and Noe (2001) showed that a unique fixed point exists in the system, however, they only considered the interbank contagion channel without fire sales.

CEU eTD Collection

in the model reflect these evidences and principles.

In order to improve the *liquidity* situation, banks in the model attempt to increase their liquid assets by liquidating those assets that cannot be taken into account in the LCR calculation or can only be taken into account with a high discount. This adjustment may take place in three stages. In the first step banks carry out operations that are feasible in a stress situation as well, do not cause a decline in reputation, do not entail large losses, and do not generate further contagion in the banking sector. Adjustment possibilities like this may include the drawing of nostro accounts (accounts that a bank holds in a foreign currency in another bank) and the non-renewal of just maturing deposits at the central bank. If no further adjustment is necessary, a given bank's reaction is evenly distributed across the above listed instruments. If carrying out the first level is not sufficient, the bank makes adjustments which do not meet the above listed considerations. In the second stage banks make the parts of the household and corporate loan portfolios which are just maturing on a cash flow basis expire. We assumed that banks make 100 per cent of the household loans maturing within 90 days and 50 per cent of the corporate loans maturing within 90 days expire (however, this time window can vary based on the assumed initial shock and the application of the model). The difference between the retail and corporate portfolio is explained by the fact that reputation loss can be more severe in the case of corporate clients. Finally, if necessary, even those assets are liquidated (corporate bonds and mortgage bonds) whose selling may result in a fire sales effect as other banks whose balance sheet also contains the given security also suffer losses through the price change. The extent of the price change depends on the type, the overall amount and the liquidated amount of the given asset.

Improving the *solvency* position takes place along similar logic, with the difference that in order to improve a bank's position, asset restructuring is possible on the basis of the risk weights (which are taken into account during the calculation of the risk-weighted asset value), instead of the LCR discount rates. Accordingly, in this case the bank transforms the assets with high risk weight into assets with risk-free rating (e.g. into cash when making assets mature). According to our model specification, in the case of a solvency problem banks have somewhat fewer options to adjust as some assets in the first stage have practically zero risk weight, so their liquidation would not improve the CAR.

If even all these adjustments are insufficient to meet the requirements (LCR and CAR), the given bank goes bankrupt, and its interbank loans become non-performing. We account simultaneously for the losses stemming from the interbank exposures and the fire sales type price losses. In the case of a default event, we differentiate in the LGD parameter based on the extent the given bank violated the requirements. After accounting for all the banks, if no change has taken place in the assets compared to the previous iteration, the process stops. Otherwise, if further loss occurred because of the contagion, some banks may have gone below the regulatory limit again, and the process restarts.

CEU eTD Collection

5.1.2 Shock transmission from the banking system to the real economy

In the model of bank-firm network relationships the main mechanisms to transfer shocks from banks towards firms is the decline in credit availability from the supply side. Ivashina and Scharfstein (2010) offers an underpinning for this mechanism by showing that firms had difficulties during the recent financial crisis in renewing their credit lines. An important factor which can modulate this kind of vulnerability is the number of connections a given firm has to the banking sector. The ability for a bank to privately observe information and maintain a close relationship with its customer enables these firms to have increased access to capital with more complex and non-standard credit needs (Von Thadden, 1995). Based on this, it can be beneficial if a firm has more than one long-term, embedded connections with financial institutions.

This embeddedness is also useful during crisis times when firms often prefer to solve their financial problems privately in a credit relationship, rather than damaging their reputation on the financial markets. Jiangli et al. (2008) showed that banks are able to smoothen out shocks to firms by rescheduling payments or by the renegotiation of the terms of the credit contract. However, this effect seems to be much weaker during systemic crisis situations. In this case, banks do not necessarily accommodate firms with new lending, rather they often refuse future lending. Puri et al. (2011) suggests that banks can smooth out idiosyncratic shocks but they amplify systemic shocks. They also showed that banks affected by a shock reject substantially more loan applications than non-affected banks.

In Hungary, the economy experienced a massive drop in lending after the 2008 crisis. Although Figure 36 and 37 do not distinguish supply and demand side factors, however, the extent of the disruption in the trends can still be considered as an obvious sign of credit retrenchment.

CEU eTD Collection

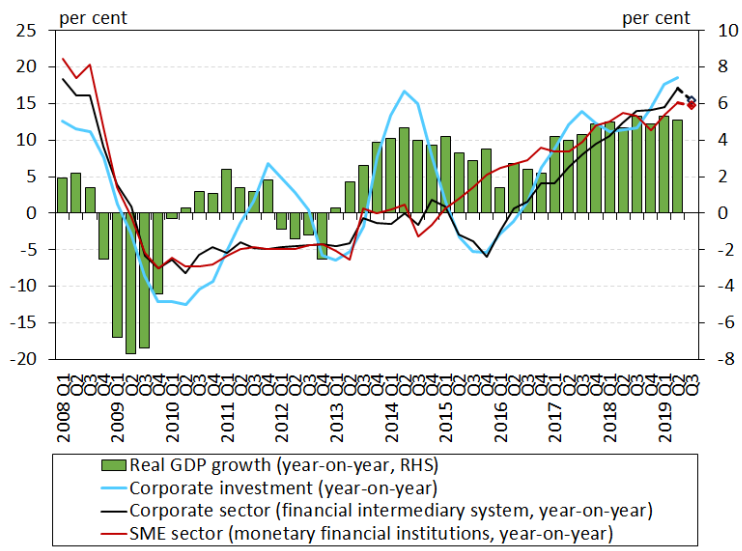


Figure 36: Growth rate of outstanding corporate and SME loans and indicators of the real economy. Source: Central Bank of Hungary.

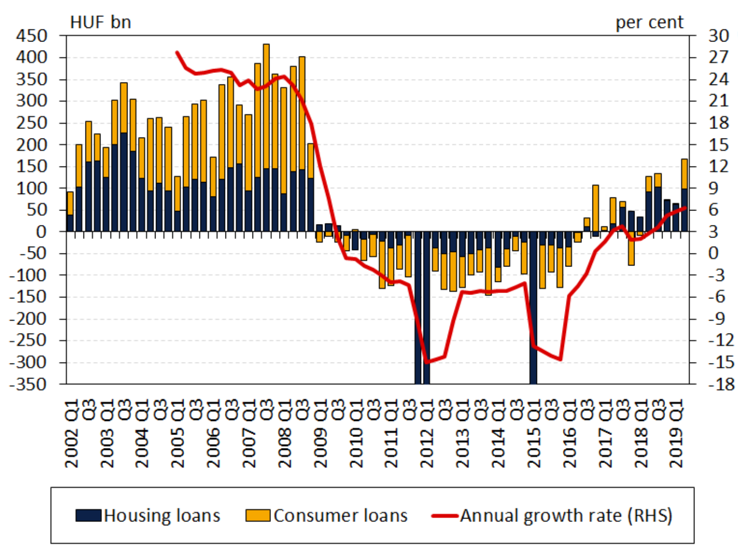


Figure 37: Household (housing and consumer) loan transactions and its annual growth rate. Source: Central Bank of Hungary.

CEU eTD Collection

5.1.3 Shock amplification in the production network

Credit supply shocks can have an impact via the supplier network even on firms which were not affected directly. In this block of the model our objective is to assess also the indirect effect of shocks coming from the banking system. Our approach to deal with this challenge is different from the mechanical modeling style we applied for the banking system. As firms are extremely heterogeneous and their operation is much less regulated than that of banks, it would be extremely burdensome to work out the details of their behaviour. Instead, we used a spatial econometric approach to estimate the increase in the probability of default of firms on their loans after a credit supply shock hits some part of the production network they are indirectly connected to[33].

This solution is connected to the literature of supply chain contagions, which gained momentum after supplier information about firms became more and more often accessible. These studies supplied ample evidence that production networks are not resilient even to firm-level idiosyncratic shocks as firms are not capable to react flexibly enough[34]. Moreover, shocks can even be amplified through supplier links. E.g. according to the results of Barrot and Sauvagnat (2016), the reduction of sales by $1 at the supplier level causes a decrease of $2.4 in sales at the customer level.

Furthermore, this stream of the economic literature distinguishes between upstream and downstream shock propagations:

- If a firm experiences a credit supply shock, its production might fall on account of the financial distress, so the shock will affect intermediate input suppliers as well. In addition, suppliers might not be able to collect money from defaulting partners. This means that the shock travels to the upstream direction on the supply chain.
- Regarding the other direction, if a supplier defaults after a credit supply shock, the intermediate inputs it produced might not be easy to replace for its costumers, hence, the shock spreads to the downstream direction.

Interestingly, shocks can reverse directions along the network, which means that in effect they can also spread horizontally. A popular example for this is the case of car manufacturing industry in the United States. In the fall of 2008, the president of Ford Motor requested government support for General Motors and Chrysler, but not for Ford. He wanted government support for his company's rivals because the failures of GM and Chrysler were predicted to result in the failure of many of the suppliers of Ford Motor. Namely, a shock to General Motors can trigger upstream shock propagation in the car-parts industry, which becomes a negative supply shock (downstream propagation) to Ford. One can

[33] The details of this estimation are discussed in Section 5.

[34] See for example Carvalho et al. (2016), Demir et al. (2018), Boehm et al. (2019).

CEU eTD Collection

imagine other scenarios for horizontal shock propagation as well. For instance, if a supplier is hit by a shock, its competitors can gain market share if the input is not too specific.

5.1.4 Shock transmission from the real economy to the banking system

The parameters estimated for the direct and indirect impact of credit supply shocks on firms' PD can be applied directly to simulate firm defaults. If a firm becomes nonperforming, banks with loan exposures towards the firm will suffer losses on their corporate loan portfolio[35]. To handle the stochastic nature of this procedure we calculated with the expected value of 1000 realizations of credit losses.

The problem of nonperforming loan portfolios became one of the most pressing issues in several European countries. Rampant NPL portfolios are not only problematic for banks, but it cuts back lending activity even further creating a negative feedback loop in the economy (Accornero et al., 2017). Figure 38 shows the devastating situation in Hungary following the 2008 crisis.

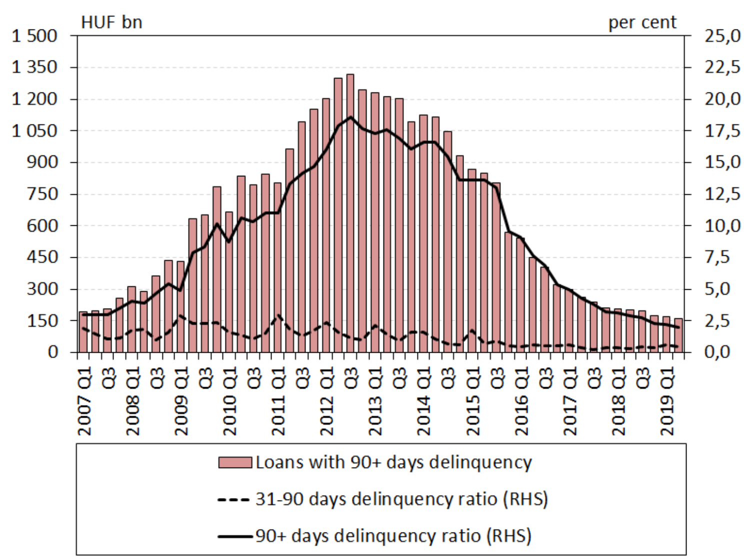


Figure 38: Ratio of non-performing corporate loans in the credit institution sector. Source: Central Bank of Hungary.

[35]During most of the simulations we used either 50% or 100% as LGD parameters, which simplification conceals the vast difficulties of estimating LGD parameters specific to several relevant bank and firm characteristics.

CEU eTD Collection

5.1.5 The time scale of the model

The processes described in this section so far must be synchronized in the time scale of the model. First of all, it is important to emphasize that even if the blocks of the simulation follow each other iteratively, this is often merely the practical representation of processes simultaneously reinforcing each other. Furthermore, the time scale of the simulations is highly dependant on the assumed initial shock and the application. As it is shown by the Figures above, the effect of the shock in 2008 was rather drastic and immediate both in the case of the plummeting lending activity and the soaring delinquency ratio, and we also experienced that the situation worsened for several quarters at an almost constant rate. However, if we apply the model within the framework of a liquidity stress test, then the relevant time scale might be only 30 or 90 days.

To address this concern, we can adjust the model by tuning two types of parameters to match the time scale of the modeled phenomena. Firstly, the window in which banks can make their loan portfolio expire should be set to the time period applicable for the given run. Secondly, the parameters governing the probabilities of firms becoming non-performing on their loans can also be adjusted to manage the mismatch between the data frequency in the estimation and the application's time scale. It might arise as an additional concern, that we consider the shock propagation based on estimates coming from yearly data, which masks the differences between short-term and medium-term dynamics. On the one hand, one would assume that the production function is more similar to the Leontief function in the short run, but the opportunities for substitution become later gradually more and more relevant. On the other hand, firms' liquidity buffer can attenuate the propagation of shocks for a while. Unfortunately, we cannot measure which one of these impacts dominates in different time windows, so we opted for not making corrections to any direction based on these considerations, so we transform the yearly estimates simply proportionally to the time frame of the application.

Of course, similarly to any other model, the reliability of the results can be lower and lower as the time window increases and less and less elements of the economy can be assumed to remain constant. As this is only a partial model, it is not suited to incorporate long-term changes in the economy. For instance, during the years following the 2008 crisis the situation of the banks was heavily influenced by several factors including capital injections, extra taxes, restructuring of some banks by the state, introduction of new regulations, etc. This way, regarding the dynamics of contagions within the banking system, we can only make plausible assumptions for relatively short time periods.

5.1.6 Data requirements of the microsimulation

To implement this microsimulation model on real data we obtained access to several detailed datasets at the Hungarian Central Bank and at the Hungarian

CEU eTD Collection

Tax Authority. While detailed information about banks and bilateral exposures at the interbank market are part of the standard data reporting towards central banks in most countries, we could also access

- the central credit information database (KHR) containing all loan contracts between banks and firms,
- firms' balance sheet and profit and loss statements from corporate tax reports, and
- transaction level data about the supply chain connections among firms from VAT reports.

Although most of these datasets have been already preprocessed and have relatively high quality, the construction of the supplier network required several corrections. VAT reporting in Hungary contains information also about the trade partners of firms, where the tax content of all the trade transactions between two companies exceeds €3000 in the given year. This information is available between 2014-2017[36], which made it possible to reconstruct the Hungarian production network with relatively high precision. By adding the location and financial reports of firms to the data we could utilize not only topological characteristics but also several node attributes. The most important shortcomings of this data are the missing observations stemming from mainly two sources: (i) international trade and (ii) connections below the value threshold. As a result of these, around 50% of the procurements is present in the observed system. The supplier network changes notably from one year to another, which is mainly due to the lot of one-off, incidental transactions. As these links are important from the point of view of shock propagation, we applied a filtering to keep only long-term supplier connections[37]. In 2017, only slightly more than half of the links are long-term, however, these cover around 93% of all the traded value.

A further distortion we had to handle is that firms belonging to the same ownership group sometimes report collectively, but very often it happens individually. To correct for this, we obtained access also to OPTEN's ownership connection database. Although we did not see global ultimate beneficiary owners, only local connections, we could still cover most of the relevant connections among firms. We also considered indirect ownership links by a calculation analogous to the Leontief inverse[38]. After all these corrections, our final network consists of yearly 80-100 thousand nodes and 200-250 thousand links.

[36] Although the quality is very poor for 2014.

[37] We classified a connection as long-term if there were at least two transaction between the firms, and if there is at least one quarter time difference between the first and the last trade occasion between them.

[38] More specifically, we computed the Neumann-series approximation (up to the fourth order) of this version of the Leontief inverse.

CEU eTD Collection

5.2 Details of the simulation

The banking system block and the real economy feedback part (which consists of the last three theoretical blocks) iteratively follow each other during the simulation. If any of the banks makes some adjustment in its lending activity (which exceeds a very low tolerance parameter in the model) the real economy feedback is triggered. If this feedback results in additional losses for the banking system (which exceeds the tolerance parameter), than the banking system's contagion mechanisms become active again. Within a "banking system block" there is a similar inner loop: If significant losses occur at any of the banks, its adjustment and/or its default can cause losses to the other banks as well, which can lead to further adjustments. Although the simulation runs in a sequential manner, this is often merely the technical representation of simultaneous events. When we denote the order of events (or states of variables) with the notation t, we refer to the iterative rounds of the simulation and not actual time. The logic of the simulation can be summarized by the following pseudocode:

Algorithm 1 The structure of the simulation

```
while additional feedback losses ≥ ε do

    Banking block:
    while additional losses ≥ ε do
        for banks do
            if (CAR_bank < CAR_regulation) then
                Solvency adjustments
            end
            if (LCR_bank < LCR_regulation) then
                Liquidity adjustments
            end
        end
        Calculating additional losses after bankruptcies
    end

    Feedback block:
    for firms do
        Calculating firm PDs after credit supply shock
        Simulating (1000 times) firms' defaults on their loans
    end
    for banks do
        Calculating additional feedback losses for banks
    end
end
```

In the following subsections we will give detailed formulation of the simula-

CEU eTD Collection

tion steps.

5.2.1 Banking system contagions

In the model we consider the nine largest Hungarian banks, which cover around 85% of the market[39]. At the Central Bank of Hungary we can observe banks' exact measures regarding their liquid assets, expected cash inflows and outflows, furthermore the equity instruments which are relevant for the CAR calculation and the risk-weighted assets.

Another crucial piece of information in the banking block is the representation of the interbank market. As the transactions here are usually very short-term, mostly overnight, a snapshot would not reflect a representative state of the market. Instead, we constructed the network by taking the average daily exposures in a month for each bank, which we then distributed in the proportion of the monthly average exposures towards the banks' partners.

Additionally, we consider further asset classes which are relevant for banks' adjustment processes. These are (1) short-term (within three month) claims towards the central bank, (2) nostro accounts, (3) government bonds, (4) corporate loans, (5) household loans, (6) corporate securities and (7) mortgage bonds. Each asset class has some parameters which govern their role during the adjustment decisions of banks (Table 8)[40]:

- *LCR haircut* indicates that to what extent a given asset should be discounted during the calculation of liquid assets for LCR.
- *Risk weight* is the discount parameter to determine the risk-weighted assets of a bank.
- The *rank* parameter determines the order in which assets are used by banks to adjust their balance sheet to be able to meet the regulatory requirements. *Rank* is determined following the principles laid out in the previous section, and it can be considered as an externally given solution of banks' optimization problem. Assets can have the same rank parameter, in which situation the required adjustment is evenly distributed between those assets.
- *Minimum Price* denotes the lowest relative price in the scenario where all the banks in the model liquidate completely the given asset category. As there are other holders of those assets on the market, the banking system can have only limited impact on the price.

[39]The inclusion of smaller institutions, which often have in some aspect special operations would only add complications to the model without any significant benefit.

[40]The risk weights and the LCR haircuts are regulated in a very detailed way, which we did not follow in the model with the same level of precision.

CEU eTD Collection

	Risk weight	Rank (solvency)	LCR haircut	Rank (Liquidity)	Minimum Price
Central bank claims	0	0	10%	1	100%
Nostro account	0	0	10%	1	100%
Government bonds	20%	1	50%	1	90%
Household loans	50%	2	100%	2	100%
Corporate loans	50%	2	100%	2	100%
Corporate securities	50%	3	100%	3	50%
Mortgage bonds	50%	3	100%	3	50%

Table 8: Adjustment parameters of the relevant asset classes.

Solvency adjustments of banks

During modeling the solvency related behavior of the banks, firstly we have to test whether a bank meets the regulatory CAR requirement in every iteration. For a given bank i this test is given by Equation 36.

$$\frac{(E_{i,t_0} - L_{i,t})}{RWA_{i,t_0} + \sum_j [\underbrace{rw_j \times p_{t,j} \times (A_{i,j,t} - A_{i,j,t_0})}_{\text{Change in the RWA due to asset liquation}} + \underbrace{rw_j \times (p_{t,j} - p_{0,j}) \times A_{i,j,t}}_{\text{Change in the RWA due to price change}}]} < CAR_{reg} \tag{36}$$

where E_{i,t_0} is bank i's original equity, $L_{i,t}$ is the cumulative loss occurred up until round t for bank i, RWA_{i,t_0} is the original risk-weighted asset of bank i, rw denotes the vector of risk weights associated with all the asset classes considered in the model, p_t is the vector of relative prices for all the asset classes[41] (the original price, p_0 is one in every case), and $A_{i,t}$ shows the assets of bank i. The change of the risk-weighed assets can be decomposed into the change due to asset liquidation ($\Delta RWA(A)$) and the change caused by price changes ($\Delta RWA(p)$). CAR_{reg} is the regulatory requirement of the capital adequacy ratio.

From this we can also calculate how much equity bank i lacks to comply with the regulation:

[41] Accounting standards vary among countries and asset classes, but for the sake of simplicity, we generally follow the principles of mark-to-market evaluation in the model. Although the implications of this approach are often debated, it reflects realistically the fair value of the assets during crisis periods.

CEU eTD Collection

$$Missing\ Equity_{i,t} = [RWA_{i,t_0} + \Delta RWA(A) + \Delta RWA(p)] \times CAR_{reg} \\ -(E_{i,t_0} - L_{i,t}) \tag{37}$$

We also have to assess how much assets are available for selling which could help to improve the solvency situation. In the case of solvency, banks first consider only one asset, sovereign bonds, with a rank parameter equal to 1 (Stage 1). Maturing household and corporate loans have a rank parameter equal to 2 (Stage 2), and corporate securities and mortgage bonds belong to Stage 3. The amount of available assets which can be used to improve solvency ($Assets\ for\ Adjustment_S$) in a stage is simply the sum of a given banks' assets in that category:

$$Assets\ for\ Adjustment_{S,i,t} = \sum_{j \in Stage_r} A_{ij,t} \tag{38}$$

where $Stage_r$ is the set of assets with rank r.

Then the actual solvency adjustment ($Adjustment_{S,i,t}$) is the minimum of the available adjustment opportunities and the necessary adjustment to reach the requirement. Even if a bank cannot meet the required CAR, it will try to approach it as much as possible.

$$Adjustment_{S,i,t} = min\Big(Assets\ for\ Adjustment_{S,i,t};\ \frac{Missing\ Equity_{i,t}}{(rw_r) \times CAR_{reg}}\Big) \tag{39}$$

where rw_r is the risk weight of assets with rank r.

If the required adjustment cannot be covered by Stage 1 assets, also Stage 2 and finally Stage 3 assets are needed. Adjustment within a given stage happens by selling the same percentage of each asset in that stage.

Liquidity adjustments of banks

During the testing of banks for their compliance with the CAR we accounted for the changes in the numerator and the denominator due to the adjustments in previous rounds. As the LCR has a more complicated formula (Equation 40) with more interactions with previous adjustments, we will present separately the alterations of LCR's components.

$$LCR = \frac{High\ Quality\ Liquid\ Assets\ (HQLA)}{Outflows - min(Inflows;\ 0.75 \times Outflows)} \tag{40}$$

Bank i's HQLA is computed as the sum of the amount of liquid assets at the current price plus the amount which was sold earlier possibly at a different price (both corrected by the vector of haircut parameters):

CEU eTD Collection

$$HQLA_{i,t} = HQLA_{i,t-1} + \underbrace{\sum_{j \in A_{LCR}} (A_{ij,t-1} - A_{ij,t}) \times (p_{j,t}) \times (1 - h_{LCR,j})}_{\substack{\text{Change in the HQLA due} \\ \text{to asset liquidation}}}$$
$$+ \underbrace{\sum_{j \in A_{LCR}} (p_{j,t} - p_{j,t-1}) \times A_{ij,t-1} \times (1 - h_{LCR,j})}_{\substack{\text{Change in the HQLA due} \\ \text{to price change}}} \quad (41)$$

where A_{LCR} is the set of assets which can be used for liquidity adjustment and $h_{LCR,j}$ is the LCR haircut parameter for asset j.

As opposed to the solvency examination, here we are calculating the difference between time t and $t-1$ instead of t_0. The reason for this is that now the time of the adjustment matters because prices can change during the simulation, and using different prices also means different change in the HQLA. The amount of cash received during liquidation has an important role for LCR (as it is part of the HQLA), but it was not relevant for the RWA as losses in the solvency block appeared in the numerator of the CAR.

Importantly, adjustments of banks aiming to improve their liquidity by increasing HQLAs can interfere with the denominator of the LCR as well. The usage of some of the adjustment options (short term central bank deposits and nostro accounts) influences the expected cash inflows as well, and this effect might distort the expression in the denominator. Additionally, losses on the interbank market also contribute to the reduction of the expected inflows:

$$\Delta Inflows_i = \Delta Nostro_i + \Delta CBclaims_i - L_{interbank,i,t-1} \quad (42)$$

where Δ refers to the change between t and $t-1$, while $L_{interbank,i,t-1}$ is the losses suffered by bank i in the previous round of the simulation.

The denominator of the LCR (LCR_{denom}) can be constructed now using the following expression:

$$LCR_{denom} = Outflows - min[max(Inflows + \Delta Inflows; 0);\ 0.75 \times Outflows] \quad (43)$$

After updating all the components of the LCR, we can also calculate the additional HQLA need if a bank is below the regulatory limit:

$$Missing\ HQLA_{i,t} = LCR_{denom,i,t} \times LCR_{reg} - HQLA_{i,t} \quad (44)$$

CEU eTD Collection

To get the required adjustment in a given stage we have to correct the $Missing\ HQLA$ with the LCR haircut parameters and with the current weighted average prices.

$$Required\ adjustment^{r}_{L,i,t} = \frac{Missing\ HQLA_{i,t}}{\sum\limits_{j \in A^{r}_{LCR}} h^{r}_{LCR,j} \times p_{j,t} \times \frac{A_{j,t}}{\sum A_{j,t}}} \tag{45}$$

where A^{r}_{LCR} is the set of assets which can be used for liquidity adjustment in stage r and h^{r}_{LCR} is the vector of LCR haircut parameters for assets with rank r.

Similarly to the solvency part, we have to assess how much assets are available for selling to improve the liquidity situation. In a given adjustment stage r it follows the same logic as Equation 38.

$$Assets\ for\ Adjustment_{L,i,t} = \sum_{j \in Stage_{r}} A_{ij,t} \tag{46}$$

Finally, the actual liquidity adjustment ($Adjustment_{L}$) is the minimum of the available adjustment opportunities and the necessary adjustment to reach the requirement. Similarly to the CAR, even if a bank cannot meet the required LCR, it will try to approach it as much as possible.

$$Adjustment_{L,i,t} = min\Big(AssetsForAdjustment_{L,i,t}; Required\ adjustment_{L,i,t}\Big) \tag{47}$$

If the required adjustment can be covered by Stage 1 assets, only these will be utilized by selling the same percentage of each of them. If also Stage 2 or 3 assets are needed, the necessary adjustment will be distributed in the same proportional manner.

Clearing of the losses in the banking system

After managing the solvency and liquidity situation of all the banks, we evaluate the state of the system. We consider a bank bankrupt, if even after all the adjustment opportunities it is unable to meet the regulatory criteria. However, we somewhat differentiate in the consequences of a default event based on the extent the given bank violated the requirements. In the case of the LCR, the loss given default (LGD) parameter was determined as 0% when the LCR is between 50-100%, and 100% for a requirement breach where the LCR goes below 50%. For the capital adequacy ratio a similar threshold is used at the 4% level of the CAR. A bank's LGD (lgd_{k}) is determined in every round based on their LCR and CAR levels:

$$lgd_{i,t} = max(lgd^{S}_{i,t}; lgd^{L}_{i,t}) \tag{48}$$

CEU eTD Collection

where $lgd^S_{i,t}$ is the LGD level which would be imposed based on the CAR of bank i, and $lgd^L_{i,t}$ is the LGD which would come from the LCR of bank i at round t.

Based on these parameters we update the interbank exposures following Equation 49.

$$W^B_t = W^B_{t-1} \times S_t \tag{49}$$

where W^B is the weighted adjacency matrix representing the exposures among K banks on the interbank lending market. A cell $w^B_{i,j}$ denotes the amount that bank i lends to bank j.

$$W^B = \begin{bmatrix} w^B_{11} & w^B_{12} & \cdots \\ \vdots & \ddots & \\ w^B_{K1} & & w^B_{KK} \end{bmatrix}$$

S is a diagonal matrix containing the surviving ratio of the interbank exposures based on the LGDs of each bank:

$$S = \begin{bmatrix} 1 - lgd_1 & & \\ & \ddots & \\ & & 1 - lgd_K \end{bmatrix}$$

The losses on the interbank exposures ($Loss_{ib}$) can be represented as the difference between the initial and the final state of the interbank matrix:

$$Loss_{ib} = W^B_t - W^B_0 \tag{50}$$

Finally, we are calculating the losses due to the change of the asset prices. The formula describing the functional form of price development is based on Georgescu (2015):

$$p_{j,t} = exp\Big(\alpha_j \sum_{i=1}^{K} s_{i,j,t}\Big) \tag{51}$$

where $p_{j,t}$ is the price of asset j at round t, $s_{i,j,t}$ is the sold amount of asset j until round t by bank i, and α controls the price elasticity. α is chosen such that when all of asset j in the system are sold, the price drops to the price level determined by the *minimum price* parameter of the given asset:

$$\alpha_j = ln(MinimumPrice_j) / \sum_{i=1}^{K} A_{i,j} \tag{52}$$

CEU eTD Collection

The losses due to fire sales ($Loss_{fs}$) can be calculated then as the difference in banks' asset values due to the price changes:

$$Loss_{fs} = (p_0 - p_t) \times A_0 \tag{53}$$

After accounting for all the banks, if the amount of the assets compared to the previous iteration changed more then the tolerance parameter ϵ, or some banks have gone below the regulatory requirements, the banking block part of the algorithm restarts. Otherwise, the banking block stops.

5.2.2 Real economy feedback

Real economy feedback is triggered if any of the banks used corporate credit retrenchment during the adjustment process (similarly to Silva et al. (2018)). Firstly, we calculate the extent of the reduction of these loans in the case of all banks:

$$\Delta Loans_{corp,i} = \frac{(Loans_{corp,i,t_0} - Loans_{corp,i,t})}{Loans_{corp,i,t_0}} \tag{54}$$

where $Loans_{corp,i,t}$ is the size of the corporate loan portfolio (which is maturing within 30 days) of bank i at round t.

As establishing new bank connections is costly (see e.g. Kim et al. (2003)), and during a crisis the credit crunch can be general, bank i's credit retrenchment ($\Delta Loans_{corp,i}$) can be interpreted as a direct credit supply shock for firm j (css_j^0) who needs (re)financing from the given bank. [42]

After determining the credit supply shock experienced by firms directly, we assess the spillover effects happening via the supplier network. The simplest – although from a computational perspective sometimes inefficient – way to represent the firm network is using an *adjacency matrix* (A^F or in the case of weighted networks W^F). In this matrix, $W^F_{m,n}$ corresponds to the traded amount supplied by firm m to firm n.

To account for shock propagation to the upstream direction, we first normalize the weighted adjacency matrix by the output (revenue+activated own

[42] If a firm is connected to more than one banks, then the credit supply shock the firm faces will be some function of the shocks coming from the banks the firm has connections with. The choice of this functional form is not trivial: using the weighted mean (where the weights are coming from the lending history between the firm and the banks) would imply that firms would have demand towards their bank connections in the same proportion as in their pre-crisis credit mix. However, firms might try to switch between the existing bank connections during a crisis, so taking the minimum of the shocks coming from the existing bank connections seems to be a more realistic assumption.

CEU eTD Collection

performance)[43] of firms in the row dimension:

$$\widetilde{W^F_{us}} = \Gamma \times W^F \tag{55}$$

where $\widetilde{W^F_{us}}$ is the row-normalized matrix representing the supplier network, and Γ is a diagonal matrix containing the reciprocal of the output of each firm.

By multiplying this row-normalized matrix with the vector of credit supply shocks experienced by each firm, we will have a vector representing the weighted sum of the credit supply shocks of the buyers (at one step distance in the network) of each firm css^1_{us}:

$$css^1_{us} = \widetilde{W^F_{us}} \times css^0 \tag{56}$$

where css^0 is the vector of direct credit supply shocks experienced by the firms.

To calculate higher order spillovers we can also determine weighted sum of the credit supply shocks of the buyers of the buyers (so at two steps distance downstream in the network) of each firm css^2_{us}:

$$css^2_{us} = \widetilde{W^F_{us}} \times css^1_{us} \tag{57}$$

We could go even further in the network, however, during the estimation of the coefficients for firms' PDs we have found only shocks coming at most from two steps distance have significant effect. However, we can consider shock propagation from two steps distance to the downstream direction as well. To calculate these terms we have to make only a slight modification. We have to normalize the weighted adjacency matrix by the output of firms in the column dimension, which can be done by multiplying with the same diagonal matrix, but this time using the transpose of W^F:

$$\widetilde{W^F_{ds}} = \Gamma \times (W^F)^T \tag{58}$$

The calculation of the weighted sum of the shocks coming from the suppliers at distance one and two happens the same way as in the upstream case:

$$css^1_{ds} = \widetilde{W^F_{ds}} \times css^0 \tag{59}$$

$$css^2_{ds} = \widetilde{W^F_{ds}} \times css^1_{ds} \tag{60}$$

[43]One could use the rowsums of the weighted adjacency matrix for normalization as well, however, using the output instead makes the interpretation of the results more intuitive.

CEU eTD Collection

As we mentioned in the previous section, shocks can reverse directions along the network, which means that they can also spread "horizontally". If we consider only two steps distance again, we have to deal with two types of horizontal shock propagation: (i) In one situation the shock can come from the suppliers of my buyers, (ii) while in the second case it can come from the buyers of my suppliers. We can account for these shocks similarly to the previous calculations. As in the first case the shock goes first downstream and then upstream, it will be denoted by $css_{ds \to us}$, while the second case is the opposite: $css_{us \to ds}$. The calculation of each of them is shown by Equations 61 and 62 respectively.

$$css_{ds \to us} = \widetilde{W_{us}^{F}} \times css_{ds}^{1} \tag{61}$$

$$css_{us \to ds} = \widetilde{W_{ds}^{F}} \times css_{us}^{1} \tag{62}$$

As a next step, we translate the direct and indirect shocks hitting a firm into additional probabilities that a given firm becomes non-performing on its loans. This step is described by Equation 63.

$$\begin{split}\Delta PD_j = css_j^0 \times \beta_{css^0} + css_{us,j}^1 \times \beta_{css_{us}^1} + css_{us,j}^2 \times \beta_{css_{us}^2} + css_{ds,j}^1 \times \beta_{css_{ds}^1} + \\ css_{ds,j}^2 \times \beta_{css_{ds}^2} + css_{ds \to us,j} \times \beta_{css_{ds \to us}} + css_{us \to ds,j} \times \beta_{css_{us \to ds}}\end{split} \tag{63}$$

where ΔPD_j is the increase in a firm's probability of becoming non-performing on its loans as a result of the direct and indirect consequences of the credit supply shocks[44]. The β parameters are coefficients showing the effects of one unit increase in the credit supply shock variables. (The estimation of these coefficients will be described in the next section.)

To complete the feedback mechanism, we simulate the default of the firms based on their ΔPD and we calculate the losses for each bank on their loans belonging to the defaulted firms. As it is a stochastic procedure, we create 1000 realizations and use the average of them as the actual losses suffered by the banks[45]. Equation 64 shows the losses suffered by bank i on its corporate loan portfolio in one realization round ($Loss_{fb,i,t}$).

$$Loss_{fb,i,t} = \sum_{v=1}^{D} OP_{v \to i} \times lgd_f \tag{64}$$

where D is the number of firms becoming non-performing in the given realization, OP is the outstanding principal amount of the loan contract between bank

[44]We concentrate now on the additional PD of banks' clients, as their base PD is accounted for during the normal operation of banks.

[45]We preferred to calculate here the average instead of the median, because the average reflects more the consequences of tail events which we did not want to ignore in the model.

CEU eTD Collection

i and firm v and lgd_f is the loss given default parameter for corporate loans.

After accounting for all the banks, if the loan losses of any of the banks exceeds the tolerance parameter ϵ, the banking block part of the algorithm is triggered again. Otherwise, the simulation ends.

5.3 Estimation of the feedback parameters

The parameters controlling how credit supply shocks influence firms' probability of becoming non-performing would be difficult to determine reliably by expert judgment or based on the experiences of past crises, hence, we attempted to estimate them independently of the model. However, this task has two main challenges: Firstly, the identification of credit supply shocks is far from being trivial, and secondly, we want to estimate not only the direct effects, but also the spillovers via the production network. In the next two subsections we will describe our approaches to deal with these difficulties.

5.3.1 Identification of credit supply shocks

Shocks can influence banks' credit supply and firms' credit demand simultaneously, thus, the observed change in lending amount cannot be considered the change of supply only. There are two typical strategies to handle this well-known endogeneity problem. When it is possible, researchers can use natural or quasi-natural experiments, such as an unexpected policy change, a nuclear accident or a natural disaster for identification. (See e.g. Khwaja and Mian (2008), A. V. Banerjee and Duflo (2014), Chodorow-Reich (2014) and Dörr et al. (2018).) The main advantage here is the strongly credible exogeneity of the shocks. However, it is often not possible to find or quantify such exogenous shocks, in which cases one can use only more indirect identification strategies. An indirect approach which gained popularity recently was developed by Amiti and Weinstein (2018). Their method uses matched firm-bank loan data, where the identification is based on the observation of firms with multiple bank connections in different time periods. Although this approach has weaker internal and external validity, it does not require to find a suitable instrument. Furthermore, by imposing adding-up constrains this procedure has the additional advantage to ensure consistency with the aggregate lending dynamics. This, or similar solutions were applied by e.g. Chava and Purnanandam (2011), Schnabl (2012), Jiménez et al. (2012), Dwenger et al. (2015), Amador and Nagengast (2016) and Degryse et al. (2017).

As the time window in which we observe both the Hungarian firm network and the loan contract data is relatively short, we had only very limited opportunities to find a suitable exogenous shock which we can use for identification. This period (2015-2017) was without major turbulences in the Hungarian banking sector, however, there were some policy measures which we attempted to

CEU eTD Collection

exploit to identify the supply side of the corporate credit market.

The Hungarian Central Bank launched a program in 2015 called Market-based Lending Scheme (MLS) to stimulate economic growth by supporting banks' lending activity[46]. Within the framework of the MLS, the central bank offered two instruments: The first incentive was that the banks could hedge their lending-related interest rate risk by an interest rate swap (LIRS) offered by the central bank to incentivize banks to grant longer-term, fixed-rate SME loans. Additionally to the LIRS, a preferential deposit facility was also introduced to support banks' liquidity management.

However, there was a condition for banks if they wanted to participate in the MLS: By having recourse to the LIRS instrument, banks had to make an implicit commitment to increase their net lending to small and medium-sized enterprises by an amount equalling one fourth of the allocated LIRS. During the programme, the central bank concluded LIRS transactions amounting to a total €2.2 billion with 17 credit institutions, which means the undertaking of an SME loan expansion of nearly €550 million by the banks participating in the programme (Figure 39). As this means an ex ante dedication to future lending, it can be interpreted as a proxy for banks' credit supply. Banks made such commitments for 2016 and 2017 as well, which makes it possible to use this as a credit supply shock indicator in our estimation[47].

[46]The description of the MLS is based on Box 5 in the 2016 May *Financial Stability Report* of the Hungarian Central Bank, where further details can also be found. (https://www.mnb.hu/en/publications/reports/financial-stability-report/financial-stability-report-may-2016)

[47]Although the MLS program created a positive loan supply shock, we are assuming that a negative shock would have similar effect to the opposite direction.

CEU eTD Collection

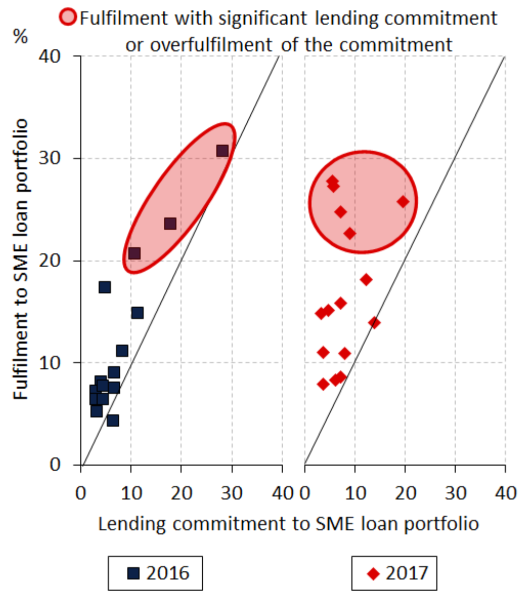


Figure 39: Banks' commitments and fulfillments in the MLS program. Source: Central Bank of Hungary.

A potential concern might arise due to the possibility that the variation in the commitment decisions of banks could be influenced to some extent by their anticipation of credit demand towards them. While this effect cannot be completely dismissed, it probably plays only a negligible role in the variation of the commitments. Although there are a few banks among the largest nine banks in Hungary (which were included in our model) which have some specialization (e.g. some banks are stronger in the household segment, others in the corporate market), however, even in their cases it is unlikely to experience very different demand from their SME clients as banks' specialization is not based on such firm characteristics (e.g. their industry) which could justify relevant differences in credit demand dynamics. Furthermore, the examined period can be considered free from serious economic turbulences in Hungary, thus, even if there were dissimilarities in banks' expectations concerning demand factors, these are more likely to be the result of the uncertainty of these kind of forecasts. However, as robustness check to the MLS shocks, we also performed the indirect method of Amiti and Weinstein (2018) following the implementation of Amador and Nagengast (2016). Further details of this methodology are described in Appendix E, where we also compare the outcomes of the regressions which are using different credit supply shock variables.

CEU eTD Collection

5.3.2 Estimation of direct and indirect effects

To represent the network-based interactions among firms, we turned to estimation techniques coming from the spatial econometrics literature. This branch traditionally deals with spatially structured data, however, the same methods can be applied to capture more abstract interaction structures, such as the production network of firms. (For a detailed review of the field see e.g. Elhorst (2014).) Spatial estimation models usually display the dependence among the observations using the so-called spatial weight matrix (W), which makes it possible to represent units affecting each other mutually. In our case the spatial weight matrix is analogous to the normalized supplier exposure matrices.

Three basic types of spatial interaction models can be distinguished: (i) the spatial autoregreesive (SAR) model, (ii) the spatial error model (SEM) and (iii) the exogenous interaction (SLX) model. As the mechanisms modeled by each of these techniques can be present simultaneously, more complicated models were also developed to combine the different spatial interactions. Equation 65 shows a general formulation containing all of these potential spatial terms in matrix form:

$$Y = \rho WY + X\beta + WX\zeta + u \tag{65}$$

where Y is the dependent variable (e.g. the default of a firm's loans), W is the supplier exposure matrix, X is the matrix of explanatory variables (most importantly for us the credit supply shock) and

$$u = \lambda W + \epsilon$$

where

$$\epsilon \sim i.i.d.$$

The term ρWY represents the SAR part for which the interpretation would be that a given firms' probability of becoming non-performing depends on its buyers' or suppliers'[48] probability of becoming non-performing on their loans. The λW is the SEM term referring to shocks which would jointly affect firms that are connected to each other in the supplier network. Finally, $WX\zeta$ is the SLX term implying that firms' probability of becoming non-performing depends on its partners' independent variables, most importantly on their credit supply shock. As this last term is exactly what we are interested in for the model, we formulated a panel logit SLX specification without including the other types of spatial interactions (Figure 40). This way we assumed that (i) in the examined period there were no significant correlated shocks affecting firms based on their supplier connections, and (ii) the credit supply shocks did not spread through any other unobserved channels.

[48] It depends on whether we are using W, or W^T.

CEU eTD Collection

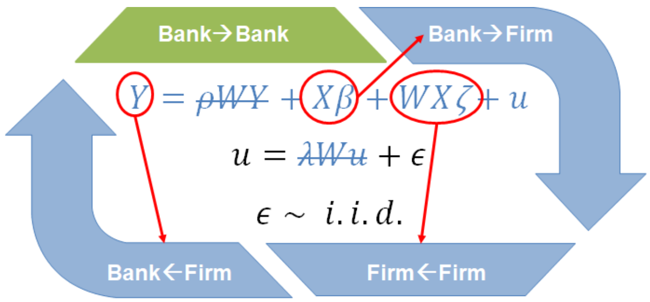


Figure 40: The different terms of the SLX model framework capture all the mechanisms which are relevant for the model. $X\beta$ refers to the direct effect of credit supply shocks, while $WX\zeta$ captures the spreading of the shock on the production network. (The term ρWY represents the SAR part for which the interpretation would be that a given firms' probability of becoming non-performing depends on its buyers' or suppliers' probability of becoming non-performing on their loans. The λW is the SEM term referring to shocks which would jointly affect firms that are connected to each other in the supplier network. As we are only interested in $WX\zeta$ in this application, we formulated a panel logit SLX specification without including the other types of spatial interactions.)

This relatively simple framework makes it possible to flexibly include further time and spatial lags, and even more than one spatial weight matrices. As $W_{i,j}$ is defined as firm i sells to firm j, than the matrix $W^k X$ would represent shock spreading to the upstream direction from distance k, while $(W^T)^k X$ would mean shock propagation to the downstream direction from distance k. By including these matrices up to $k = 4$ in the estimation[49], we can have separate coefficients for different spatial lags for upstream, downstream and horizontal contagion as well.

To avoid any concerns about the potential endogeneity of the supplier exposure matrices, we are exploiting the time dimension of the data by using the one-year lagged versions of them. As we are considering only long-term supplier connections, the usage of the lagged versions does not cause significant information loss, but it can assure that the endogenous nature of link formation will not interfere with the spreading process.

A further difficulty which needs to be addressed is the handling of firms without loans. Ignoring them completely during the estimation would also mean their removal from the supplier network. However, even if a firm does not have any bank connection, and cannot experience credit supply shocks directly, it still can have a role in propagating shocks which were originated elsewhere in the production network. In order to preserve these pieces of information, we delete these firm only after calculating all the higher order matrices. This way we can retain all the indirect pathes between firms even if we disregard firms without loans during the estimation.

[49] As the average shortest path length of the production network is 4.9 with a standard error as low as 1.1, investigating four steps in both upstream and downstream directions is sufficient to cover the vast majority of potential shock propagation.

CEU eTD Collection

After taking into account all the considerations above, we arrive at our final specification, which gives estimates for all the parameters in Equation 66:

$$NP_t = \beta_0 + \beta_{css} CSS_t + \sum_{k=1}^{4} \left[\beta_{css^k_{us}} (\widetilde{W^F_{us,t-1}})^k CSS_t + \beta_{css^k_{ds}} (\widetilde{W^F_{ds,t-1}})^k CSS_t \right] + \beta_{css_{ds \to us}} CSS_{ds \to us,t} + \beta_{css_{us \to ds}} CSS_{us \to ds,t} + X_t \beta_{controls} + \epsilon_t \quad (66)$$

where NP_t is a dummy variable indicating whether a firm became non-performing (defined as more than 90 days delinquency) in the given year. In the estimation we included as controls firms' revenue, ROA, liquidity buffer, size category, the export share of their revenue and a dummy variable indicating state owned companies. Furthermore, we added fixed effects for firms' industry, regional location and for the year.

According to the results (Table 9), the impact of credit supply shocks can be significant even two steps away in the supplier network. Although in the case of upstream propagation, p-values are a bit higher at distance two from the source of the shocks, we included even this level of spreading in the model as they are not that far away from the significance levels of the downstream case. However, in the case of distance three and four there is no indication of any effect of the credit supply shocks. Regarding the horizontal channels, our results indicate significant spreading only when the shock is firstly transmitted towards a supplier and then to another buyer of that supplier, but not for the reverse situation, so in the end we excluded the $css_{ds \to us}$ channel by setting its parameter to zero in the model.

CEU eTD Collection

Table 9: Regression Results

	Dependent variable:					
	Probability of default on loans					
	Logit (1)	Logit (2)	Logit (3)	Logit (4)	Logit (5)	Logit (6)
css^0	-0.030^{**}	-0.028^{**}	-0.029^{**}	-0.028^{**}	-0.028^{**}	-0.028^{**}
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
css^1_{us}		-0.070^{***}	-0.055^{**}	-0.058^{**}	-0.057^{**}	-0.060^{**}
		(0.022)	(0.023)	(0.024)	(0.024)	(0.025)
css^1_{ds}		-0.121^{***}	-0.068^{*}	-0.073^{*}	-0.065^{*}	-0.035
		(0.033)	(0.038)	(0.039)	(0.039)	(0.042)
css^2_{us}			-0.079	-0.105	-0.111	-0.110
			(0.057)	(0.068)	(0.070)	(0.072)
css^2_{ds}			-0.245^{**}	-0.307^{**}	-0.352^{**}	-0.306^{**}
			(0.105)	(0.143)	(0.145)	(0.147)
css^3_{us}				0.099	0.113	0.098
				(0.143)	(0.193)	(0.193)
css^3_{ds}				0.219	-0.153	-0.164
				(0.336)	(0.426)	(0.431)
css^4_{us}					-0.058	-0.037
					(0.469)	(0.469)
css^4_{ds}					1.088	1.137
					(0.737)	(0.739)
$css_{ds\to us}$						0.030
						(0.052)
$css_{us\to ds}$						-0.104^{*}
						(0.055)
Constant	-5.124^{***}	-5.119^{***}	-5.125^{***}	-5.126^{***}	-5.128^{***}	-5.117^{***}
	(0.405)	(0.407)	(0.408)	(0.408)	(0.408)	(0.409)
Year dummies	✓	✓	✓	✓	✓	✓
Industry dummies	✓	✓	✓	✓	✓	✓
Location dummies	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	91,528	91,528	91,528	91,528	91,528	91,528
Log Likelihood	-7,863.209	-7,847.722	-7,843.375	-7,842.913	-7,841.987	-7,839.985
Akaike Inf. Crit.	15,796.420	15,769.440	15,764.750	15,767.830	15,769.980	15,769.970
Note:	Robust standard errors are clustered by banks. *p<0.1; **p<0.05; ***p<0.01					

Since the coefficients of our estimation are odds ratios which cannot be used directly as parameters in the model, we had to calculate the marginal effects to obtain interpretable results. After this step, we arrive at the final feedback

CEU eTD Collection

parameters (Table 10):

css^0	css^1_{us}	css^1_{ds}	css^2_{us}	css^2_{us}	$css_{us\to ds}$	$css_{ds\to us}$
0.0008	0.0013	0.0007	0.0010	0.0051	0.0900	0.0000

Table 10: Average marginal effects of the estimated feedback parameters.

As a robustness check, we performed the same estimation using the indirect credit supply shock variable as well. Although in this case we had significant results only for the one-step downstream shock spreading with the indirect shocks, and the marginal effects were somewhat different as well, the overall impact of the credit supply shocks were similar to that of the main specification. We show this in Appendix E using one of the applications as an illustration.

A further important consideration could be that the parameter values in Table 10 were estimated using data on yearly frequencies, however, some applications of the model might require a shorter time scale for the simulation. In these situations we adjusted the parameters proportionally; e.g. if we considered only a three months time window (for instance in a liquidity stress test), we divided the parameters by four to handle the mismatch with the estimation.

5.4 Applications

Since the primary objective of this model is to offer a versatile tool for various policy analyses, we present here three potential applications: (i) Firstly, we embedded the model into a liquidity stress testing framework, (ii) then we elaborated a way to use it for SIFI identification, and lastly (iii), we show an example of assessing the impact of shocks originated in the real economy. (Additionally, we offer a brief comparison of our model with the DebtRank algorithm in Appendix F.)

5.4.1 Embedding the model into a liquidity stress test

As one of the first applications, we embedded the model to the liquidity stress testing framework of the Hungarian Central Bank. This liquidity stress test has been featuring contagion channels in the banking system since 2016, however, we could add now a unified shock propagation modeling block with feedback mechanisms from the real economy. During the implementation we used the standard stress scenario of the liquidity stress test (presented in the central banks' biannual Financial Stability Reports), which is a complex exogenous shock calibrated to the 2008 crisis (Table 11).

CEU eTD Collection

Assets			Liabilities		
Item	**Degree**	**Currencies affected**	**Item**	**Degree**	**Currencies affected**
Exchange rate shock on derivatives	15 per cent	FX	Withdrawals in household deposits	10 per cent	HUF/FX
Interest rate shock on interest rate sensitive items	300 basis points	HUF	Withdrawals in corporate deposits	15 per cent	HUF/FX
Calls in household lines of credit	20 per cent	HUF/FX	Withdrawals in debt from owners	30 per cent	HUF/FX
Calls in corporate lines of credit	30 per cent	HUF/FX			

Table 11: Components of the liquidity stress scenario of the Hungarian central bank.

When we ran the stress test simulation using only a limited version of the framework which did not contain any contagion mechanisms, only one out of the nine largest banks was unable to comply with the LCR during the stress scenario. If we enabled for contagion channels in the banking block only, two out of the nine largest banks have become unable to comply with LCR even with using adjustment opportunities. In this case, an additional €258 million fire sales loss and €5 million interbank loss occurred in the banking system. After enabling the real economy feedback channels as well, 0.5% of the firms in the model went bankrupt causing €184 million loss for banks on defaulting loans. Furthermore, losses due to fire sales further increased by €41 million, and a third bank went below the regulatory requirement, but this time it happened due to solvency insufficiency. Although it is still the fire sales channel which is responsible for the largest chunk of banks' losses in the simulation, the real economy feedback contributes by almost the same extent. We also noticed that the loss-based ranking of the banks has changed as well after we enabled the feedback mechanisms. (Figure 41)

CEU eTD Collection

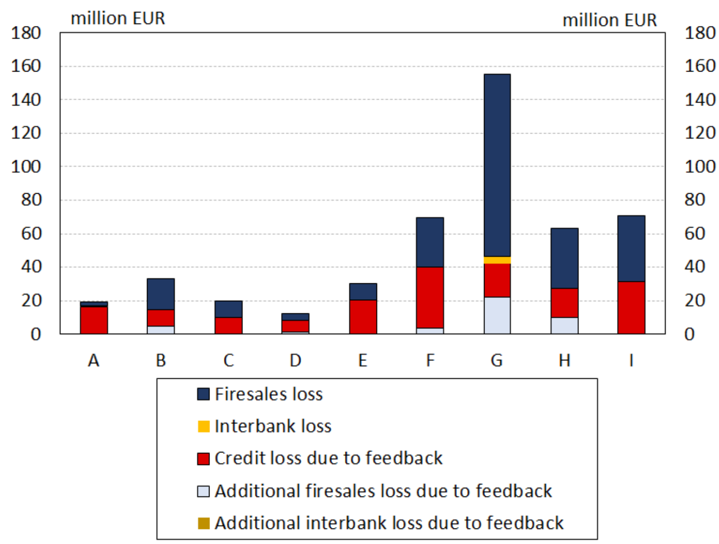


Figure 41: Results for the nine largest banks in Hungary (based on 2017 data).

From the point of view of systemic risks and financial stability, it is clear that ignoring the feedback mechanisms can lead to the severe underestimation of risks and potential losses in this shock scenario. Furthermore, while the interlacing between the liquidity and the solvency problems was largely hidden in the reduced stress testing frameworks, the real economy feedbacks made this aspect also more pronounced. Additionally, by including the feedback channels we can gain some insight into the impacts of a banking sector liquidity shock on the non-financial firms as well. (Although we do not claim that the model is capable of giving a full picture about all the consequences of the stress scenario on the real economy.)

5.4.2 SIFI identification based on Shapley value

The problem of identifying systemically important financial institutions (SIFIs) has been dealt with by numerous papers, among which we relied in this exercise on those using the concept of Shapley value (Tarashev et al. (2011), Bluhm et al. (2014), Aldasoro et al. (2017)). Shapley value is a concept originated in game theory, and it was developed to allocate the outputs generated in cooperative games among agents (Shapley, 1953).

The typical technique how the Shapley value is applied for SIFI identification is to calculate the difference between the system-wide losses occurring after a shock event with and without the participation of a given bank in the simulation of the banking system. We calculated this difference for the idiosyncratic default of each bank, however, Shapley value in its original form would require

CEU eTD Collection

to repeat this calculation for all the possible subsystems of the banking system $(f(N^{SUB}) - f(N^{SUB} - i))$. The actual Shapley value would be then the average of the additional losses that a bank generates by participating in any subsystem of the bank network:

$$Shapley_i = \frac{1}{n} \sum_{n_s=1}^{n} \frac{1}{c(n_s)} \sum_{N^{SUB} \supset i} \left[f(N^{SUB}) - f(N^{SUB} - i) \right] \tag{67}$$

where $Shapley_i$ is the Shapley value of bank i, $N^{SUB} \supset i$ denotes all the subsystems that contains bank i, n_s means the number of banks in a given subsystem and $c(n_s) = \frac{(n-1)!}{(n-n_s)!}(n_s - 1)!$ is the number of subsystems containing bank i and are comprised of n_s banks.

Due to computational constraints, we did not perform the calculations for all the subsystems, only for the whole bank network[50]. This way, our Shapley-based measure for bank i is the average difference between the aggregate system-wide losses (caused by the idiosyncratic default of each bank occurring one by one) with and without the presence of the bank in interest[51]:

$$SIFI_i = \sum_{\substack{m,n \\ m \neq n}}^{N} S_{m,n} - \sum_{\substack{p,q \\ p \neq q}}^{N-1} S^{-i}_{p,q} \tag{68}$$

where S is an $N \times N$ matrix, in which N denotes the number of banks, and $S_{m,n}$ is the losses suffered by bank n after the exogenous default of bank m. S^{-i} is an $(N-1) \times (N-1)$ matrix which contains the losses occurring without the participation of bank i in the system. The main diagonals of these matrices are ignored in this application.

In order to gain more detailed insight in the sources of systemic risk for each bank, we present our SIFI measure decomposed into three factors:

- System-wide losses due to a given bank's default:

$$DamagingPotential_i = \sum_{n \neq i}^{N} S_{i,n} \tag{69}$$

[50]Castro et al. (2017) proposed a polynomial method using stratified random sampling with optimum allocation to estimate the Shapley value, however, for our purposes it is more advantageous to simply ignore the subsystems since we are more interested in the importance of institutions when the whole system is present.

[51]When a bank is deleted from the system, all the links attached to it will be removed as well. To avoid interference with the simulation of the model, we assumed that the banks which borrowed from the removed institution can replace their interbank funding with other financing sources offering the same conditions, while the assets of the removed bank are reallocated to agents outside of the model.

CEU eTD Collection

- A given bank's losses due to other banks' defaults:

$$Vulnerability_i = \sum_{m \neq i}^{N} S_{m,i} \tag{70}$$

- Other banks' extra losses due to the amplification of the impact of other banks' defaults by bank i:

$$Amplification_i = \sum_{m \neq i}^{N} \sum_{n \neq i}^{N} S_{m,n} - \sum_{p=1}^{N-1} \sum_{q=1}^{N-1} S_{p,q}^{-i} \tag{71}$$

The importance of these factors can vary across the examined banks. (Figure 42) There are banks, whose systemic importance comes from their vulnerability to shocks. Other banks might be resilient from this aspect, but their default can cause severe damage in the banking system. The amplification component has notable role only in the case of one examined bank, which indicates that either the complexity of the Hungarian bank network was not high enough (in 2017) to make it possible for a bank to cause severe damage only by transmitting losses, or at least the assumed idiosyncratic shocks were too weak to trigger cascading failures.

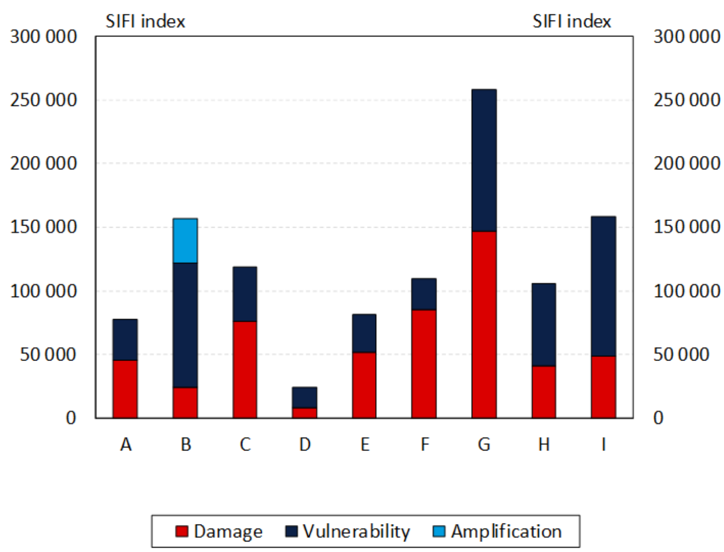


Figure 42: Decomposed SIFI index of the nine largest banks (2017). (Although the units of this SIFI index are expressed in Hungarian Forint, they would be difficult to interpret as amounts of money since they are the sums of the differences between aggregate losses in the case of multiple scenarios.)

CEU eTD Collection

5.4.3 Impact assessment of real economy shocks

Our model contains elaborated details only for the banking system, but not about any other sector of the economy. However, in a limited form it might still be possible to examine the effects of shocks coming from the real economy if we keep in mind that in this framework, shocks have to be translated into the change in firms' probability of becoming non-performing. This way we can capture only the credit loss and the supply chain contagion aspects of real economy shocks, which is far from being a complete assessment. With this caution in mind, we attempted to assess the consequences of shocks originated in certain industries on the banking system. A recent example of an unexpected stress event can be the COVID-19 pandemic, which had very severe impact on some industries whose firms could transmit the shock to other industries, and to the banks as well.

As a first step in this analysis we identified the most vulnerable industries to this shock using four-digit NACE categories. Most of the affected sectors in Hungary belong to the manufacturing, wholesale and retail trade, transporting, storage, accommodation, food service activities, real estate activities, administrative and support service activities, arts, entertainment, recreation and other services activities. (A detailed table about the affected sectors can be found in Appendix G.) We assumed that these directly affected firms have 100% susceptibility for being hit by the shock, which means the maximum exposure to the shock.

After the identification of the most involved sectors, we calculated the indirect exposures (up to four steps) to these industries in each firm's revenue. (E.g. if 20% of firm A's revenue comes from buyers belonging to the directly affected sectors, then firm A's exposure will be 20%. If there is another buyer of this firm, which is responsible for another 20% revenue and it has 50% exposure, then the vulnerability of firm A will be $20\% + 10\% = 30\%$.) During this procedure we did not include the directly affected firms as they have reached already the maximum level of involvement with the crisis[52]. To acknowledge some heterogeneity among firms, we corrected their exposure with firms' potential liquidity buffers[53]. We calculated these buffers in the proportion of their revenue as well, so we could simply subtract it from the exposure measure.

As we estimated only the parameters governing shock spreading and feedback in the case of credit supply shocks (which would not be applicable here), we had to make some assumptions about the connection between this shock and firms' probability of becoming non-performing. If the final value of the exposure was 100%, or a firm operates directly in some of the affected sectors, we increased

[52] It also means that directly affected firms cannot amplify shocks further. E.g. if a firm has a buyer belonging to one of the directly hit industries, and this buyer is responsible for 10% of the firm's revenue, then there cannot be second, or higher order contagions through the same buyer, as the whole 10% exposure has been already taken into account as vulnerability.

[53] We calculated basically the quick liquidity ratio with a slight modification: We took the difference between the numerator and the denominator from the original formula.

CEU eTD Collection

the probability of becoming non-performing by ΔPD percentage points. If the vulnerability was below 100%, we decreased the ΔPD parameter proportionally. These PD values could be directly fed into the model as inputs to simulate the effects of this shock. As we do not know the exact value of ΔPD, we ran the simulation ten times increasing it by five percentage points each time.

Figure 43 shows the number of lost jobs due to the defaulting firms, and the losses of the banking sector in the case of different values of the ΔPD parameter. Figure 44 illustrates the losses separately for the nine largest Hungarian banks[54]. If one considers the direct scenario, the banking system could suffer a loss of more than €1.1 billion, which is equivalent to almost 13% of the equity in the banking system.

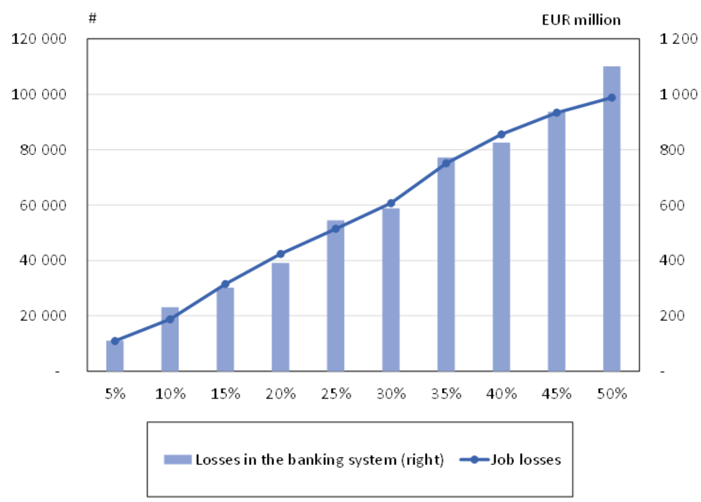


Figure 43: The number of lost jobs and the losses of the banking sector in the case of different values of the ΔPD parameter.

[54]Since our data are about 2017, the results should also be interpreted as if the shock had happened in 2017.

CEU eTD Collection

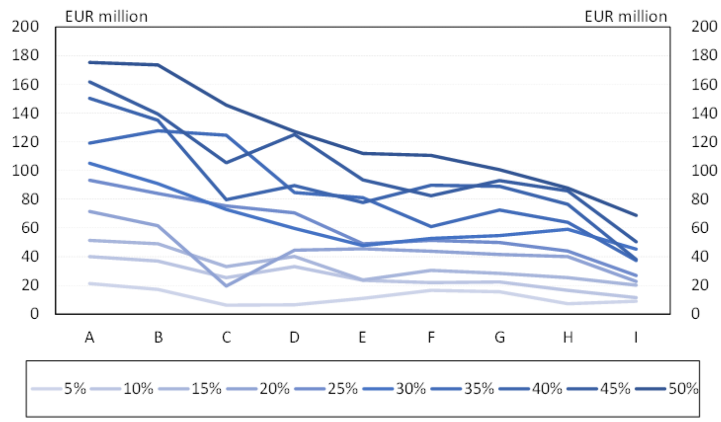


Figure 44: Losses of the nine largest Hungarian banks in the case of different values of the ΔPD parameter.

Other shocks coming from the real-economy could be included in a similar fashion, however, for the sake of reliable interpretation of the results, it is necessary to thoroughly explore the connection between firms' exposure to the shock and their probability of becoming non-performing on their loans. Although we do not have yet the necessary statistics to calibrate all the parameters in the case of the COVID-19 crisis, the results can still indicate a plausible range for the expected consequences.

CEU eTD Collection

6 Discussion and conclusion

There is an increasing interest in several areas of economics towards the inclusion of networks not only in theoretical models but also in empirical analyses. This thesis contributes to these endeavours in a threefold way: (i) it provides insights into the basic structure and the unique traits of micro-level firm network data; (ii) secondly, it introduces a model of shock spreading in firm-level production networks, which makes it possible to rectify several shortcomings of industry-level supply chain analyses; (iii) and lastly, it offers a novel way of modeling feedback channels between the financial sector and the real economy in the context of interacting economic networks.

In order to make it possible to investigate firm networks, I obtained access to sensitive datasets about the ownership links and the supplier connections of Hungarian firms. Using these sources it has become possible to build the multi-layer representation of the Hungarian firm network which enabled us to gain insight into its previously unobserved structure. Although this data is almost unmatched in the literature, it is very important to acknowledge that it still has limitations regarding its completeness; furthermore, that it requires careful preparations which is highly dependant on the application.

In the case of ownership networks, three issues needed to be addressed to avoid serious biases: the distribution of ownership shares, the computation of indirect links and the weighting based on the size of the firms. The resulting ownership system proved to be very sparse and disconnected, however, it still revealed some topological characteristics and the typical motifs at the micro-level. It was also possible to assess the significance of economic entities regarding the extent to which they can influence and control the economy via their ownership relations. The same investigation has been conducted also at more aggregated levels revealing the role of different groups formed based on several characteristics, such as the nationality, the legal category or the HQ location of the owners.

In the case of the supplier layer, the network has been collapsed to the level of ownership groups, where it became possible to accurately define long-term connections. The analysis of this system identified several topological characteristics, which can be responsible for facilitating contagious processes. This network has high enough density to allow the emergence of a giant component, within which one can reach any firm with only a few steps. This is due to the presence of *hub nodes* having very high degree and *bridge nodes* which connect the otherwise isolated blocks of the economy. These blocks were identified by community detection methods and they were shown to represent the different production chains of certain product categories within the economy. In addition to these topological traits, the network also demonstrated strong homophily based on several firm attributes.

These pieces of information enabled the building of a microsimulation model of shock propagation to quantify short-term damages after supply chain disrup-

CEU eTD Collection

tions in the production network. It turned out that from the point of view of shock propagation, I-O tables offer an unreliable grouping due to the fact that the core input structure of firms is very heterogeneous within even fine grained (NACE 4 level) industry classes. The granular approach proposed in this thesis makes it possible to consider this heterogeneity in the production processes of firms by allowing to differentiate in their production functions and in the importance of different input types.

The first application of this model quantified the systemic risk of firms by simulating the effects of distinct upstream and downstream spreading mechanisms on the production network. These simulations showed that only less than 100 firms have the potential to destroy more than five percent of the Hungarian national production network, and hence, pose a significant threat to the overall economy. In the case of the default of even only one of these companies, up to 21% of all production can be affected, however, the vast majority of the firms have only very limited impact on the production network. Additionally, the analysis offered a lower bound of firm-level systemic risk by assuming only linear production functions, and upper bounds by assuming only Leontief production functions. Due to the high-resolution approach the most influential companies could be explored in more detail. The list of the top 100 most systemically important firms contains not only intuitively expected large companies, but also quite a few SMEs. This means that the default of relatively small, but crucial suppliers can have similar system-wide implications just as in the case of the largest companies. Hence, the knowledge of the systemic riskiness of single firms has a pivotal role in understanding and preventing potentially large failure cascades in these networks. Finally, it was shown that the different allocations of an industry-level shock among firms in a given sector can lead to a wide distribution of cascade sizes. This heterogeneity remains hidden if one cannot see the fine-grained structure of the supplier network. This way, the proposed model is suitable to evaluate more accurately defined shocks, and to assess the potential damage range of an industry-level shock considering different affectedness for individual firms.

The last chapter described a novel way to analyse the financial stability of an economy in a microsimulation environment which is suitable to capture contagious mechanisms in an interconnected system of economic networks. More specifically, it considered the interactions between the network of banks (exhibiting contagious mechanisms among them) and the network of firms (transmitting shocks to each other along the supply chain) which systems are linked together primarily via loan-contracts. As the previously described production network model does not consider the financial situation of firms, it cannot accommodate shocks coming from the banking sector. To circumvent this problem, a data-driven, econometric approach has been applied to determine (i) the effect of credit supply shocks on firms, (ii) the extent to which these shocks are augmented in the production network, and (iii) the implications on firms' probability of becoming non-performing on their loans.

CEU eTD Collection

The results of this model confirmed that the feedback mechanisms between the coupled financial and real economy networks can amplify the losses in the economy beyond the shortfalls expected when considering the interacting subsystems in isolation. As a first test for this, the model has been embedded into the liquidity stress testing framework of the Central Bank of Hungary, and the results proved the importance of the real economy feedback channel, without which systemic risks could potentially be severely underestimated. The inclusion of this feature did not only doubled the system-wide losses, but it also made the connection between liquidity and solvency problems more pronounced. To illustrate the versatility of this modeling framework, two further applications were presented for different policy purposes. Firstly, the model was used for SIFI identification, which showed that the source of the systemic importance of banks can greatly vary between the damaging potential of their default and their vulnerability to shocks coming from other banks. Secondly, the example of the COVID-19 pandemic was used to illustrate how the impact of shocks originated in the real economy can be assessed by the model.

Given the wide range of potential further applications, a more elaborated embedding of the financial system in the real economy would be certainly desirable. The presented framework could be extended in several directions from this point of view. Regarding the financial sector, at the moment only banks are included but no other financial institutions (such as the insurance sector, investment funds or central clearing counterparties), which can contribute greatly to the complexity of the economy. However, the current representation of the real economy was even more simplified. A significant upgrade would be to model the operation of firms more comprehensively. This would make it possible to reflect on the now missing credit demand component, and one could also include shocks coming from outside the financial sector more realistically. In an even more general model it would also be possible to generate endogenous shocks. Nowadays researchers usually impose an exogenous stress scenario calibrated to a crisis event to see how the modelled mechanisms amplify the initial shock. However, in reality, these mechanisms are also responsible for the shocks growing to the initially observed extents.

In parallel with these opportunities one should also be aware of some pitfalls during the elaboration of more and more details of the economy in a data-driven simulation environment. This line of research would lead to the territory of agent-based macroeconomic modeling, for which one of the greatest challenges is to create detailed models, but preserve their tractability to avoid becoming unfathomable "black boxes". Furthermore, as it is apparent in this work, the development of these models should go hand in hand with the advancement of the empirical literature which produces vital information for the key parameters. With sufficient awareness of these limitations, I believe that microsimulations and computational models in economics are extremely useful and versatile tools, which can be even further improved if the concepts of network science become more embedded in the way we think about the economy.

CEU eTD Collection

7 Appendix

7.1 Appendix A – Further description of the ownership data

Depending on the year, we can observe around 800 000 ownership links among almost 1 million actors covering the ownership structure of around 400 000 firms (Table A.1).

Table A.1: Basic description of the ownership data

Year	Number of firms	Number of links
2015	367 857	774 944
2016	385 341	800 112
2017	405 823	824 293
2018	428 092	850 132
2019	454 540	884 059

The majority of the owners are individuals, while there are only yearly 51 000 – 66 000 links where the owner is a firm. According to Table A.2, the ratio of firms is much higher in the case of foreign owners than in the case of Hungarian owners (40% as opposed to 6%).

Table A.2: Number of links with foreign/HUN and individual/firm owners

Year	Foreign individual	Foreign firm	HUN individual	HUN firm
2015	18 908	13 815	704 626	37 595
2016	20 055	14 259	725 618	40 180
2017	21 224	14 686	745 363	43 020
2018	23 157	15 132	765 885	45 958
2019	26 691	16 413	791 120	49 835

As it is depicted on Figure A.1, the links of the network do not change significantly from one year to another. Around 75% of the links are present in all the five observed years.

CEU eTD Collection

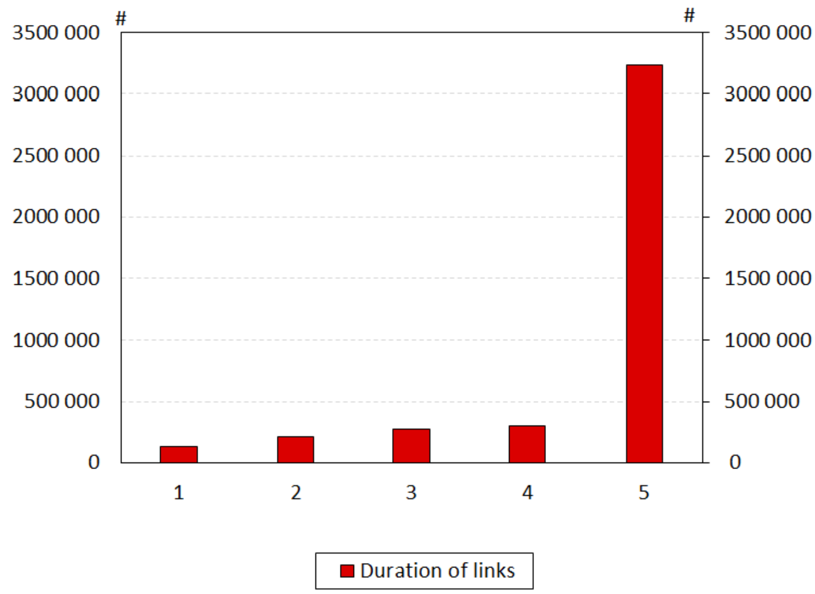


Figure A.1: Duration of links (in years) in the Hungarian firm ownership network

Furthermore, in the case of 35-40% of the links we can see even the extent of the influence (expressed in percentages). However, sometimes (in 16 000-19 000 cases depending on the year) we found firms with ownership links where the sum of the overall influences exceeded 100%. In these cases we corrected the influence for each owner proportionally to make the sum equal to 100%. For links where the influence information was missing, we used a simple imputation method by dividing the missing amount of influence among the remaining owners. (For example if we observed a link with 50% share, and we could also see that there are two more owners associated with the same firm without influence information, we simply divided the missing 50% between the two remaining owners equally.)

Using firms' anonymized tax numbers as keys, we could connect the ownership data to another dataset coming from the Hungarian Tax Authority, which contains several firms characteristics. Although the overall quality of the data is quite good, it is far from being complete: we cannot see tax numbers in 25-30% of the Hungarian owner firms, and the dataset is even more incomplete in the case of listed companies, for which we rarely observe the ownership structure.

7.2 Appendix B – Weighting of firms' influence

As it is depicted in Figure B.1.a), the *total controlled value* of firm A would be the sum of the following elements:

CEU eTD Collection

- 100% of firm B's value via a direct link between A and B;
- 50% of firm C's value via a direct link between A and C; and
- 50% of firm C's value via the indirect link between A and C through B.

However, as B's whole value is originated from its ownership over 50% of C, the direct link between A and B and the indirect link between A and C through B overlap. As a result, we inflate A's *total controlled value* stemming from this system of ownership ties.

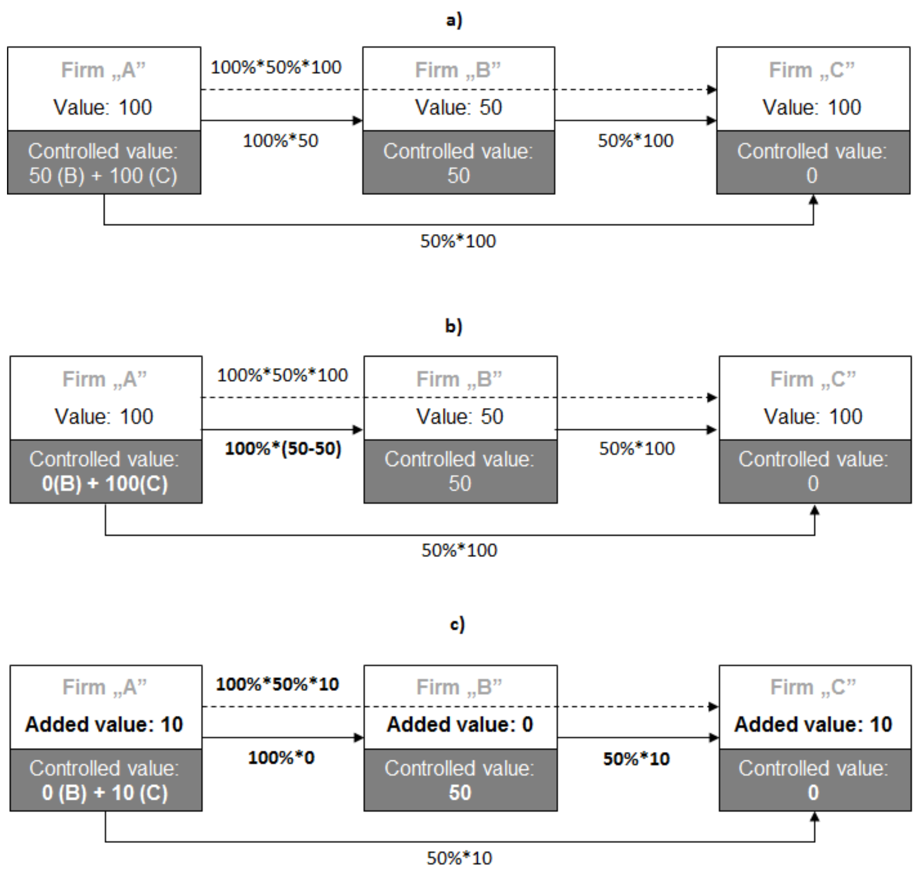


Figure B.1: The calculation of indirect control in the ownership network

Direct ownership links are indicated by solid lines, while indirect links represented by dashed lines. Bold font shows changes between a), b) and c).

Figure B.1.b) shows a possible correction by reducing firm B's value by

CEU eTD Collection

the part which comes from its ownership over firm C. This method might seem simple, however, if we generalize it for the whole network, it becomes very demanding to implement. We would have to examine every paths in the network and correct all the members of these chains (except the endpoints). As it is computationally infeasible, we turned to a simplification of the problem. We wanted to find a node attribute, which is independent from the ownership structure, but conveys some information about firms' value. The best candidate meeting these requirements is the value added of firms, which can be directly applied as a replacement for the previous value proxies (as it is shown in Figure B.1.c)).

7.3 Appendix C – Further description of the supplier data

We observed the supplier system between 2014-2017, however, the quality in the first year was very poor probably due to the inexperience of both the authorities and the firms in the new reporting requirements. Because of this reason we considered only the period between 2015-2017 in our analysis. The basic description of the network properties of these years can be seen in Table C.1.

Table C.1: Description of the supplier network between 2015-2017

	2015	**2016**	**2017**
Number of nodes	63 772	79 049	89 778
Number of edges	153 006	205 105	235 913
Density	3.76e-5	3.28e-5	2.93e-5
Average in-/outdegree	2.39	2.59	2.63
Shortest path lengths (avg.)	5.05	4.7	4.92
Shortest path lengths (st.dev.)	1.2	1.1	1.1
Local clustering (avg.)	0.071	0.077	0.078
Reciprocity	0.13	0.11	0.11

Although the general quality in these years is high, we still had to make a few correction. We filtered for situations where the VAT rate calculated from the supplier network data deviates from the official rate (which is 27% in the examined period). We also corrected, if the tax amount and the purchase value was mixed up. Furthermore, we checked the consistency between the sales revenue of firms (coming from corporate tax declaration data) and the sum of the purchases of a given firm's products and services observed in the supplier network.

Besides the EUR 3000 reporting threshold there is another limitation of the analysis of the supplier network, which is coming from the lack of international trade links in the data. To assess the overall significance of missing links, we connected the supplier network to another dataset (also coming from the Hungarian Tax Authority), which is collected as part of the tax declaration of firms. Due to the sensitivity of these pieces of information, we had to use firms'

CEU eTD Collection

anonymized tax numbers as keys to merge the different data sources (similarly to the case of the ownership network). This data contains several characteristics related to firms' balance sheets and profit and loss statements, among which we can observe their yearly material costs as well. Although the comparability of these datasets is not perfect, it was possible to compare the sum of the reported supply transactions for a given firm to its aggregate material costs. This exercise revealed that around 50% of the material costs cannot be matched to the supplier network (Figure C.1). In the case of the larger firms, the main reason for this disparity is probably the unobserved import, while for smaller firms the value threshold is more likely to be the dominant constrain.

While the lack of international trade data is an important limitation (especially because Hungary is a small, open economy vulnerable to cross-border shock spreading), it should be noted that even by observing the direct import-export links, one would still gain only a limited coverage of international exposures. Shocks affecting Hungary can originate way further away in the international trade network, hence, to be able to fully resolve this caveat one would have to be able to observe the whole global supplier network.

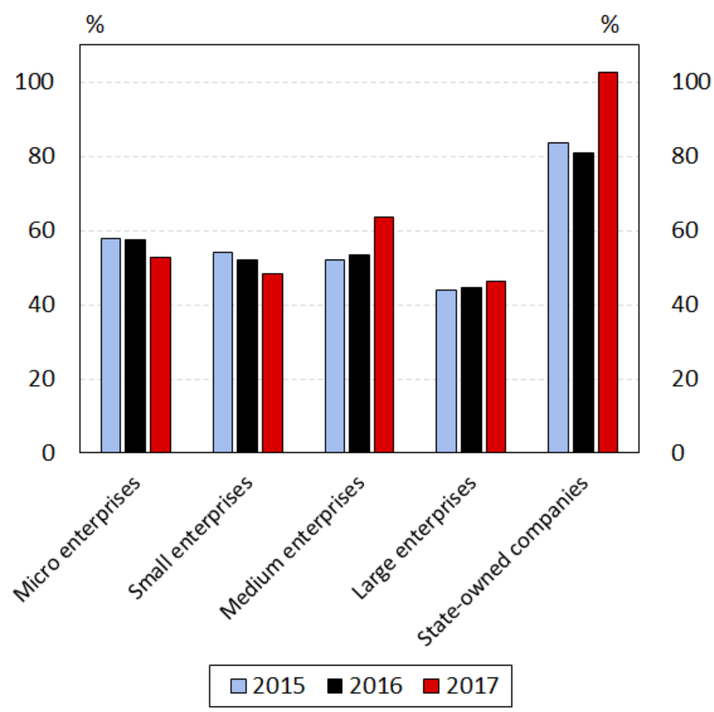


Figure C.1: Ratio of the sum of supplier transactions and material costs of firms by company sizes.

According to the regulation, the frequency of the reporting can be yearly,

CEU eTD Collection

quarterly or monthly depending on the size of the firms and on the weights of the firms' supplier links. As it is shown on Figure C.2, the vast majority of the firms report monthly, however, we used yearly aggregation even in these cases for two reasons: (i) we need longer periods to define long-term relationships; and (ii) the other datasets with which we want to connect this network are also on yearly frequency.

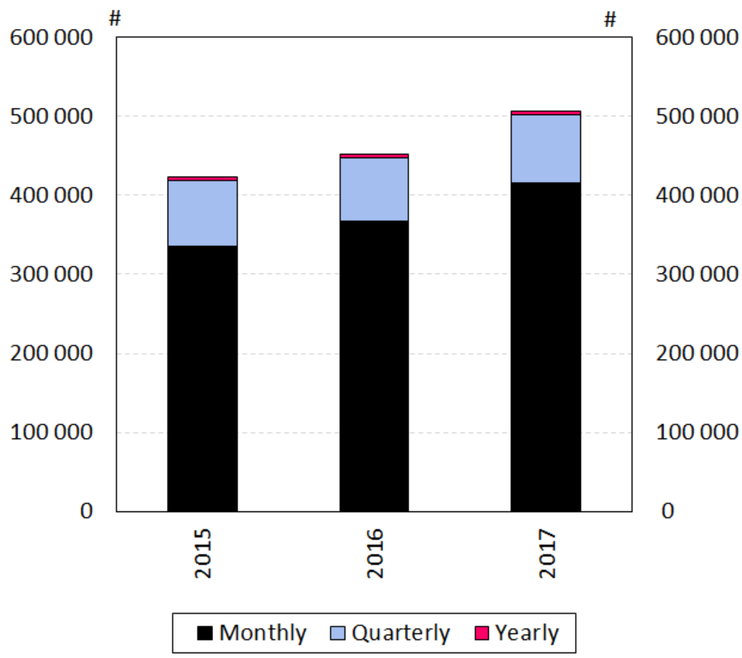


Figure C.2: The frequency of firms' reporting about their supplier relationships

A further interesting feature of this data reporting is that firms are required to submit information not only about their partners whom they are buying from, but also about their buyers. This made it possible to build the supplier network from both directions, and examine their differences. If we construct the network from connections where the subject firm reports its suppliers, we have to face the problem, that there are suppliers not subject to the VAT, which results in missing information. However, when one uses the connections where the subject firm reports its buyers, another problem can emerge, namely that firms are less motivated to report as these transactions entail VAT obligations for them. As both approaches has different shortcomings, and even the reporting periods are not guaranteed to match for the two sides of a given transaction, the overlap is – as expected – far from being perfect. Although it is impossible to precisely assess the extent of these biases, we decided to use the former approach because irregularities connected to VAT declarations are certainly not

CEU eTD Collection

randomly distributed, therefore, they can distort the results more seriously. This issue offers another research direction connected to the detection of fraudulent activities, however, this topic is outside of the scope of our analysis for now.

7.4 Appendix D – Description of the community detection methodology

The task of detecting communities is closely related to the well-known clustering problems. The reason why there are so many sophisticated network related techniques in the literature is the concern of computational feasibility (coming from the often enormous size and the complexity of the analyzed systems). In almost every approaches of this challenge there are two pressing questions: (i) how to measure how well the network is separated given a particular division, and (ii) what algorithm to use to find the best grouping of the nodes[55].

Our choice of the function describing the fitness of a given partition is a widely-used measure called *modularity* (M. E. Newman, 2006). The intuition behind this approach is that a good division of a network is not merely one in which there are dense connections between nodes within modules and sparse connections between nodes in different modules. A better formulation would be to say that we are looking for a partition in which there are *fewer than expected* links between communities and *more than expected* links within communities. Corresponding to this idea, the modularity is the number of links present within communities minus the expected number of links placed at random:[56]

$$Q = \frac{1}{2L}\sum_{i \neq j}\Big(A_{i,j} - \frac{k_i k_j}{2L}\Big)\delta(C_i, C_j) \tag{D.1}$$

where Q denotes modularity, L is the number of links in the graph, and the Kronecker delta indicates whether node i and j belong to the same community.

The higher the probability of a link is, the smaller its contribution to the modularity score is. If the sum of the increments in the end is positive, that indicates the possible presence of a community structure. Therefore, our goal is to find the divisions of a network with the largest modularity score. Even in our case with less than 100 thousand nodes, it would be obviously impossible to go over all the possible partitions and calculate the modularity for all of them. Blondel et al. (2008) proposed an agglomerative, multi-level modularity optimization algorithm called the "Louvain-method" which is based on a hierarchial approach:

1. Initially, each node represents a community with a single member.

[55]It should be noted that in some cases a partition is not the best approach as there are often overlapping communities.

[56]Random placement in this case means the randomization of the links with the preservation of the degree distribution.

CEU eTD Collection

2. Every vertex is moved one-by-one to a community where the modularity is increased in the largest extent by the reallocation of the given node.

3. In the second phase, each community is considered a new node on their own, and the process goes back to Step 1.

4. The algorithm stops either when all the nodes are assigned in one all-encompassing community, or when we cannot increase the modularity anymore.

Although this type of modularity maximizations has some shortcomings, e.g. it has a resolution limit and problems in detecting overlapping communities or hierarchical structures (Javed et al., 2018), it is a widely accepted method which can be relatively simply and reliably implemented even by using R's Igraph package. However, as the "Louvain-method" is based on a stochastic algorithm, it can give different results for different realizations. To overcome this problem Tandon et al. (2019) proposed a method in which we aggregate the different realizations using the *fast consensus* procedure:

1. Building the consensus graph based on existing links only to avoid too high computational cost. (In the consensus graph two nodes are connected if they belong to the same community.)

2. As two nodes do not necessarily belong to the same group in every realization, we can use a threshold, below which we ignore the link. (For instance, if we observe them together in less than 20% of the realizations.)

3. Adding triadic closure links (because nodes sharing neighbors belong to the same community with higher probability).

4. Run our standard community detection algorithm on the consensus graph.

5. Iterate the procedure until we reach our tolerance level.

7.5 Appendix E – Indirect credit supply shock identification

For the sake of tractability we start the description of this estimation by showing a general modeling specification for the problem. If one assumes that the credit demand of a multi-bank firm changes the same way towards all of its partner banks, than the percentage difference between the changes in the amounts of credits can be attributable to supply side factors (Figure E.1).

CEU eTD Collection

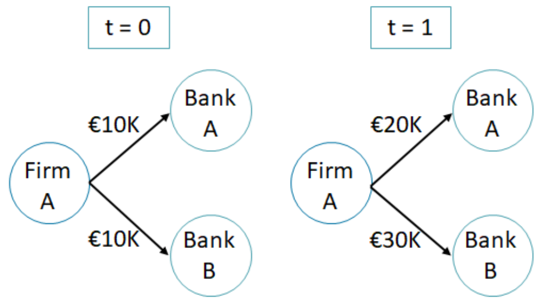


Figure E.1: The amount of credits between firm A and bank B increases more than towards bank A. If one assumes that the credit demand of the firm changes uniformly towards both of its partner banks, then this difference can be attributed to supply side factors.

In this case, the lending Ψ_{fbt} between bank b and firm f at time t can be decomposed into supply (β_{bt}) and demand factors (α_{ft}):

$$\frac{\Psi_{fbt} - \Psi_{fb,t-1}}{\Psi_{fb,t-1}} = \alpha_{ft} + \beta_{bt} + \epsilon_{fbt} \tag{E.1}$$

where we assume that the expected value of the error term is zero, $\mathbf{E}[\epsilon_{fbt}] = 0$. α_{ft} captures all firm-specific characteristics and shocks which can affect its borrowing, while β_{bt} comprises all the bank-specific factors which can have an impact on the credit supply of a given bank. Although Equation E.1 could be directly estimated on our data coming from the credit registry, Amiti and Weinstein (2018) highlighted that this formula ignores the aggregate equilibrium on the lending market. That is, firms can only obtain new loans if a bank is willing to provide that credit; and similarly, banks can increase their lending activity only if there are firms soaking up the additional supply. They offer an alternative formulation which corrects for this inefficiency and allows us to consider newly formed loan contracts as well. According to this, the growth in a given bank's lending D^B_{bt} can be expressed as the supply of the bank plus the weighted sum of its client firms' demand, where the weights are the share a given firm had in the bank's lending in the previous period:

$$\begin{aligned} D^B_{bt} &= \sum_f \left(\frac{\Psi_{fbt} - \Psi_{fb,t-1}}{\Psi_{fb,t-1}} \right) \times \frac{\Psi_{fb,t-1}}{\sum_f \Psi_{fb,t-1}} \\ &= \beta_{bt} + \sum_f \phi_{fb,t-1} \times \alpha_{ft} + \sum_f \phi_{fb,t-1} \times \epsilon_{fbt} \end{aligned} \tag{E.2}$$

where

$$\phi_{fb,t-1} = \frac{\Psi_{fb,t-1}}{\sum_f \Psi_{fb,t-1}} \tag{E.3}$$

CEU eTD Collection

Analogously, the growth in a given firm's borrowing D^F_{ft} is the composition of its own demand and the weighted sum of the supply of its partner banks:

$$\begin{aligned} D^F_{ft} &= \sum_b \Big(\frac{\Psi_{fbt} - \Psi_{fb,t-1}}{\Psi_{fb,t-1}} \Big) \times \frac{\Psi_{fb,t-1}}{\sum_b \Psi_{fb,t-1}} \\ &= \alpha_{ft} + \sum_b \theta_{fb,t-1} \times \beta_{bt} + \sum_b \theta_{fb,t-1} \times \epsilon_{fbt} \end{aligned} \tag{E.4}$$

where

$$\theta_{fb,t-1} = \frac{\Psi_{fb,t-1}}{\sum_b \Psi_{fb,t-1}} \tag{E.5}$$

As $\phi_{fb,t-1}$ and $\theta_{fb,t-1}$ are determined directly from the data, we can make similar assumptions about the error terms as before: $\mathbf{E}\big[\sum_f \phi_{fb,t-1} \times \epsilon_{fbt}\big] = 0$ and $\mathbf{E}\big[\sum_b \theta_{fb,t-1} \times \epsilon_{fbt}\big] = 0$. With these moment conditions we arrive at a system of linear equations with α_{ft} and β_{bt} as unknowns:

$$D^B_{bt} = \beta_{bt} + \sum_f \phi_{fb,t-1} \times \alpha_{ft} \tag{E.6}$$

$$D^F_{ft} = \alpha_{ft} + \sum_b \theta_{fb,t-1} \times \beta_{bt} \tag{E.7}$$

Although this system consists of the same number of equations and unknowns (which is equal to the number of banks plus the number of firms) in every year, the system is still under-determined as the sum of the shares in lending are equal to one ($\sum_f \phi_{fb,t-1} = 1$ and $\sum_b \theta_{fb,t-1} = 1$). To be able to find a unique solution, we have to impose an additional constraint, which can be handled analogously to the dummy variable trap problem by choosing a reference category. To obtain economically interpretable results, we transformed α_{ft} and β_{bt} by subtracting their median respectively in every year. (This implies, that banks' credit supply shocks can only be compared to each other within the given year. However, since we also include time fixed effects, this concern is not problematic as the time-specific components are removed from the banks' shocks.) The transformation gives us the following expression for the banks:

$$D^B_t = (\bar{A}_t + \bar{B}_t)\iota_B + \Phi_{t-1} N_t + \Phi_{t-1}\tilde{A}_t + \tilde{B}_t \tag{E.8}$$

where D^B_t is a vector containing the loan growth rates of banks at time t, $(\bar{A}_t + \bar{B}_t)$ are the median firm and bank common shocks, which would affect all firm-bank pairs the same way in year t. ι_B is a vector of ones, N_t is the vector of the average industry-level shock for all the firms, and Φ_t is the matrix

CEU eTD Collection

of weights of all the loans of every borrowers:

$$\Phi_t = \begin{bmatrix} \phi_{11,t} & \cdots & \phi_{F1,t} \\ \vdots & \ddots & \\ \phi_{1B,t} & & \phi_{FB,t} \end{bmatrix}$$

The first term in Equation E.8 represents common shocks, e.g. a change in the key interest rates by the central bank, which would affect all lending connections. The second term shows industry-level shocks to a given banks' clients. It captures changes in a bank's lending coming from its specialization to some industries, which can make its lending activity differ from the general trend. The third term can be interpreted as the change in the bank's lending due to idiosyncratic firm-level demand shock. Lastly, the fourth term represents the credit supply shock of a bank which is independent from all the above listed influences, so we can use it as a credit supply shock variable in our estimation of feedback effects. Since this term was expressed as the deviation from the median bank's supply shock in year t, its interpretation is also relative to this median. This way, the zero value of the credit supply shock does not mean unchanged lending activity, but rather the median change in the system in a given year. If a bank decreases its lending by 20%, but all the other banks' lending drops only by 15%, than the credit supply shock of the given bank will be 5%.

The described methodology of Amiti and Weinstein (2018) is based on firms with multiple bank connections, regarding which we made a slight modification following Degryse et al. (2017). As only a small portion of Hungarian firms have multiple bank connections (Figure E.2), we wanted to enhance the external validity by including also firms with only one bank link. If the vast majority of the firms were excluded from the estimation, β_{bt} might not reflect the representative credit supply shocks of banks, but only those experienced by firms with more bank connections. Since Hungarian firms show strong heterogeneity especially along the dichotomy of large, productive foreign-owned companies and small, inefficient SMEs, representativeness might be essential in gaining correct estimates.

CEU eTD Collection

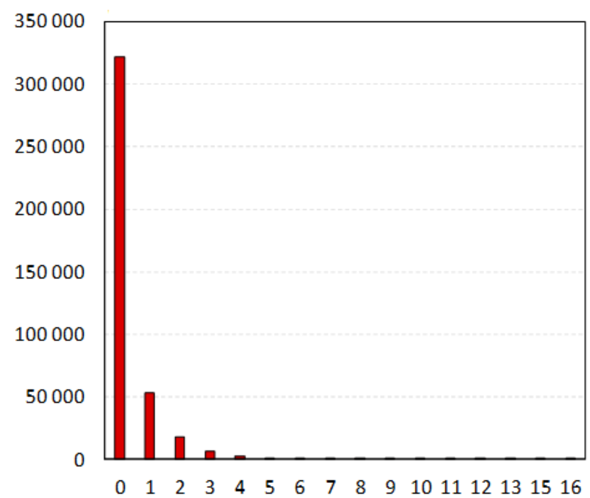


Figure E.2: Distribution of Hungarian firms based on the number of bank connections. Bank connections are defined by credit contracts or financial leasing. (Based on 2017 data.)

The main idea of Degryse et al. (2017) is that firms with similar size, operating in the same region and in the same industry can have similar dynamics in their credit demand as well. To exploit this information we replaced the $Time \times Firm$ fixed effects with $Location \times Industry \times Size \times Time$ fixed effects as control to demand-side factors in a given a year[57].

The results of the parameter estimation using this indirect credit supply shock variable are summarized below:

[57]The industry classifications are based on the two-digit NACE categories, location is determined by the town of the headquarters of firms, while size categories are given by the Hungarian XXXIV. SME regulation.

CEU eTD Collection

Table E.1: Regression Results

	Dependent variable:					
	Probability of default on loans					
	Logit (1)	Logit (2)	Logit (3)	Logit (4)	Logit (5)	Logit (6)
css^0	−0.047	−0.031	−0.032	−0.032	−0.033	−0.007
	(0.121)	(0.121)	(0.121)	(0.121)	(0.121)	(0.125)
css^1_{us}		−0.957**	−0.928**	−0.910**	−0.905**	−0.846*
		(0.424)	(0.427)	(0.429)	(0.429)	(0.433)
css^1_{ds}		−0.819	−0.892	−0.607	−0.596	−0.563
		(0.706)	(0.716)	(0.740)	(0.741)	(0.747)
css^2_{us}			−1.057	−1.088	−1.057	−0.966
			(1.128)	(1.145)	(1.187)	(1.187)
css^2_{ds}			1.697	2.424	2.192	2.340
			(2.162)	(2.139)	(2.324)	(2.320)
css^3_{us}				1.032	1.180	1.270
				(2.993)	(3.082)	(3.052)
css^3_{ds}				−8.499	−9.214	−9.415
				(5.441)	(5.945)	(5.938)
css^4_{us}					−2.380	−1.576
					(10.274)	(10.153)
css^4_{ds}					4.982	6.801
					(15.479)	(15.720)
$css_{ds\rightarrow us}$						−1.244
						(1.358)
$css_{us\rightarrow ds}$						−0.393
						(1.297)
Constant	−4.491***	−4.487***	−4.487***	−4.485***	−4.485***	−4.485***
	(0.230)	(0.230)	(0.230)	(0.230)	(0.230)	(0.230)
Year dummies	✓	✓	✓	✓	✓	✓
Industry dummies	✓	✓	✓	✓	✓	✓
Location dummies	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	156,110	156,110	156,110	156,110	156,110	156,110
Log Likelihood	-16,639.930	-16,636.460	-16,635.750	-16,634.680	-16,634.620	-16,634.140
Akaike Inf. Crit.	33,351.870	33,348.930	33,351.500	33,353.360	33,357.250	33,360.270
Note:	Robust standard errors are clustered by banks. *p<0.1; **p<0.05; ***p<0.01					

CEU eTD Collection

css^0	css^1_{us}	css^1_{ds}	css^2_{us}	css^2_{us}	$css_{us \to ds}$	$css_{ds \to us}$
0.0002	0.0190	0.0126	0.0216	0.0000	0.0090	0.0280

Table E.2: Marginal effects of the estimated feedback parameters.

To assess the sensitivity of our model to the differences between the feedback parameters estimated using direct and indirect credit supply shocks, we used the application in Section 6.1 as an illustration. After enabling the real economy feedback channels in the model, 0.51% of the firms in the model went bankrupt (as opposed to 0.53% in the main specification) causing €175 million loss for banks on defaulting loans (which is only slightly differ from the €184 million in the original results). Furthermore, losses due to fire sales further increased by €48 million (instead of €41 million), and a third bank went below the regulatory requirement the same way due to solvency insufficiency.

Based on these results, the main difference between the two specifications seems to be that in the case of the indirect credit supply shock estimates the role of the direct effect of the shocks is somewhat weaker, and the role of the contagion among firms is stronger. However, the overall impact is basically identical from the point of view of the losses in the banking system.

7.6 Appendix F – Comparison with the DebtRank algorithm

One of the pioneering methods in the financial contagion literature was developed by Battiston et al. (2012). In their proposed model shocks can propagate among banks through the interbank vulnerability matrix V:

$$V_{i,j} = \frac{A_{i,j}}{e_i} \tag{F.1}$$

where V is a $N \times N$ matrix, (where N is the number of banks), $A_{i,j}$ denotes the unsecured exposures of bank i towards bank j, and e_i is the capital buffer of bank i.

The DebtRank algorithm starts with initializing a $N \times 1$ vector $b(0)$ containing the initial shocks to banks' capital. If $b_i(0) = 1$, then bank i lost all of its capital, so the bank is in default. If $0 < b_i(0) < 1$, then the bank is distressed, and finally if $b_i(0) = 0$, then bank i is undistressed. Shocks propagate in the banking network based on the following update rule:

$$b_i(t) = min[1, b_i(t-1) + \sum_{j \in D_t} V_{i,j} b_j(t-1)], \tag{F.2}$$

CEU eTD Collection

$$s_i(t) = \begin{cases} D, & \text{if } b_i(t) > 0 \textit{ and } s_i(t-1) \neq I, \\ I, & \text{if } s_i(t-1) = D, \\ s_i(t-1), & \text{otherwise.} \end{cases} \tag{F.3}$$

in which $s_i(t) \in \{U, D, I\}$ is a categorical variable showing whether a bank is undistressed(U), distressed (D) or inactive (I). After convergence, the DebtRank index DR is calculated by weighting the losses of banks by their economic value, for which several proxies can be used. Now we simply use banks' equity:

$$DR(b(0)) = \sum_i (b_i(t) - b_i(0))e_i, \tag{F.4}$$

To compare this algorithm to the model proposed in Chapter 5, a straightforward strategy is to simulate the default of each of the 9 largest banks in Hungary one by one, and compare the system-wide losses predicted by DebtRank and also by our model. The result of this exercise is shown in Figure F.1. It can be seen that there is a considerable difference in the outcome of basically every bank default simulation. This dissimilarity can have many sources. Debtrank does not consider several of the contagion and adjustment channels which are included in our model, but even higher order rounds of shock spreading are ignored[58].

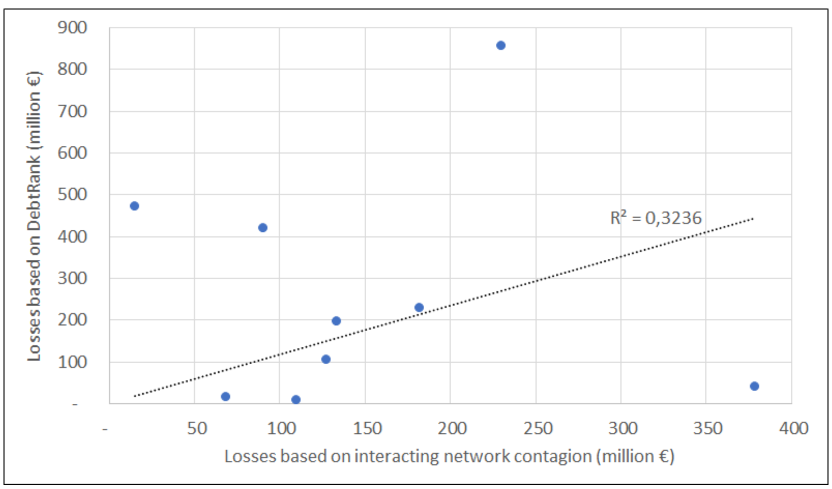


Figure F.1: System-wide losses after the initial default of each of the 9 largest Hungarian banks calculated (i) using DebtRank (ii) and our interacting network contagion model.

In the case of complex systems, even a slight variation of the initial conditions can lead to strikingly different outcomes, so it is not surprising to see very

[58] This shortcoming was corrected by Bardoscia et al. (2015).

CEU eTD Collection

different systemic impacts in this comparison. While now we do not compare the results to any grand truth observation, it seems plausible to claim that the extent to which a model captures the actual economic processes can matter a great deal in accurately determining systemic risks in a complex financial system.

CEU eTD Collection

7.7 Appendix G – List of directly affected industries

Manufacturing	22xx, 25xx, 28xx, 29xx, 30xx
Wholesale and retail trade	4511, 4719, 4751, 4752, 4753, 4754, 4759, 4761, 4762, 4763, 4764, 4765, 4771, 4772, 4775, 4777, 4778, 4779, 4782, 4789, 4799
Transporting and storage	4930, 4932, 4939, 5010, 5030, 5110, 5223
Accommodation and food service activities	55xx, 56xx
Real estate activities	6810, 6820, 6831, 6832
Administrative and support service activities	7711, 7721, 7722, 7729, 7911, 7912, 7990
Arts, entertainment and recreation	9001, 9002, 9311, 9321
Other services activities	9511, 9512, 9521, 9522, 9523, 9524, 9525, 9529, 9601, 9602, 9604, 9609

Table G.1: List of directly affected industries.

CEU eTD Collection

References

Accornero, M., Alessandri, P., Carpinelli, L., & Sorrentino, A. M. (2017). Non-performing loans and the supply of bank credit: evidence from italy. *Bank of Italy Occasional Paper*(374).

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., & Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, *80*(5), 1977–2016.

Acemoglu, D., Ozdaglar, A., & Tahbaz-Salehi, A. (2015). Systemic risk and stability in financial networks. *American Economic Review*, *105*(2), 564–608.

Adrian, T., & Shin, H. S. (2010). Liquidity and leverage. *Journal of financial intermediation*, *19*(3), 418–437.

Ahn, Y.-Y., Bagrow, J. P., & Lehmann, S. (2010). Link communities reveal multiscale complexity in networks. *nature*, *466*(7307), 761.

Aldasoro, I., Gatti, D. D., & Faia, E. (2017). Bank networks: Contagion, systemic risk and prudential policy. *Journal of Economic Behavior & Organization*, *142*, 164–188.

Allen, F., & Carletti, E. (2008). Mark-to-market accounting and liquidity pricing. *Journal of accounting and economics*, *45*(2-3), 358–378.

Amador, J., & Nagengast, A. (2016). The effect of bank shocks on firm-level and aggregate investment.

Amiti, M., & Weinstein, D. E. (2018). How much do idiosyncratic bank shocks affect investment? evidence from matched bank-firm loan data. *Journal of Political Economy*, *126*(2), 525–587.

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

Ballester, C., Calvó-Armengol, A., & Zenou, Y. (2006). Who's who in networks. wanted: The key player. *Econometrica*, *74*(5), 1403–1417.

Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2013). The diffusion of microfinance. *Science*, *341*(6144).

Banerjee, A. V., & Duflo, E. (2014). Do firms want to borrow more? testing credit constraints using a directed lending program. *Review of Economic Studies*, *81*(2), 572–607.

Baqaee, D. R. (2018). Cascading failures in production networks. *Econometrica*, *86*(5), 1819–1838.

Baqaee, D. R., & Farhi, E. (2019). The macroeconomic impact of microeconomic shocks: beyond hulten's theorem. *Econometrica*, *87*(4), 1155–1203.

Baqaee, D. R., & Farhi, E. (2020). Productivity and misallocation in general equilibrium. *The Quarterly Journal of Economics*, *135*(1), 105–163.

Barabási, A.-L., et al. (2016). *Network science.* Cambridge university press.

Baral, P., & Fique, J. P. (2012). Estimation of bilateral exposures-a copula approach. In *Technical report.* mimeo.

Bardoscia, M., Battiston, S., Caccioli, F., & Caldarelli, G. (2015). Debtrank: A microscopic foundation for shock propagation. *PloS one*, *10*(6).

CEU eTD Collection

Bargigli, L., Di Iasio, G., Infante, L., Lillo, F., & Pierobon, F. (2015). The multiplex structure of interbank networks. *Quantitative Finance*, *15*(4), 673–691.

Barrot, J.-N., & Sauvagnat, J. (2016). Input specificity and the propagation of idiosyncratic shocks in production networks. *The Quarterly Journal of Economics*, *131*(3), 1543–1592.

Barucca, P., Bardoscia, M., Caccioli, F., D'Errico, M., Visentin, G., Battiston, S., & Caldarelli, G. (2016). Network valuation in financial systems. *Available at SSRN 2795583*.

Battiston, S., Puliga, M., Kaushik, R., Tasca, P., & Caldarelli, G. (2012). Debtrank: Too central to fail? financial networks, the fed and systemic risk. *Scientific reports*, *2*, 541.

Battiston, S., Weisbuch, G., & Bonabeau, E. (2003). Decision spread in the corporate board network. *Advances in Complex Systems*, *6*(04), 631–644.

Beaman, L. A. (2012). Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the us. *The Review of Economic Studies*, *79*(1), 128–161.

Behn, M., Daminato, C., & Salleo, C. (2019). A dynamic model of bank behaviour under multiple regulatory constraints.

Benhabib, J., Bisin, A., & Jackson, M. O. (2010). *Handbook of social economics.* Elsevier.

Bernard, A. B., & Moxnes, A. (2018). Networks and trade. *Annual Review of Economics*, *10*, 65–85.

Bimpikis, K., Candogan, O., & Ehsani, S. (2019). Supply disruptions and optimal network structures. *Management Science*, *65*(12), 5504–5517.

Bimpikis, K., Fearing, D., & Tahbaz-Salehi, A. (2018). Multisourcing and miscoordination in supply chain networks. *Operations Research*, *66*(4), 1023–1039.

Blattner, L., Farinha, L., & Rebelo, F. (2019). When losses turn into loans: The cost of undercapitalized banks. *ECB Working Paper*.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, *2008*(10), P10008.

Bluhm, M., Faia, E., & Krahnen, J. P. (2014). Endogenous banks' networks, cascades and systemic risk.

Boehm, C. E., Flaaen, A., & Pandalai-Nayar, N. (2019). Input linkages and the transmission of shocks: Firm-level evidence from the 2011 tōhoku earthquake. *Review of Economics and Statistics*, *101*(1), 60–75.

Borsos, A., Stancsics, M., et al. (2020). *Unfolding the hidden structure of the hungarian multi-layer firm network* (Tech. Rep.). Magyar Nemzeti Bank (Central Bank of Hungary).

Brinkhoff, J., Langfield, S., & Weeken, O. (2018). *From the horses mouth: surveying responses to stress by banks and insurers* (Tech. Rep.). ESRB Occasional Paper Series.

Brioschi, F., Buzzacchi, L., & Colombo, M. G. (1989). Risk capital financing and the separation of ownership and control in business groups. *Journal*

CEU eTD Collection

of Banking & Finance, *13*(4-5), 747–772.

Broer, P., Antony, J., et al. (2010). *Linkages between the financial and the real sector of the economy: A literature survey* (Tech. Rep.). CPB Netherlands Bureau for Economic Policy Analysis.

Budnik, K. B., Balatti, M., Covi, G., Dimitrov, I., Groß, J., Hansen, I., ... others (2019). Macroprudential stress test of the euro area banking system.

Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., & Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature*, *464*(7291), 1025–1028.

Caballero, R. J., Hoshi, T., & Kashyap, A. K. (2008). Zombie lending and depressed restructuring in Japan. *American Economic Review*, *98*(5), 1943–77.

Caballero, R. J., & Simsek, A. (2013). Fire sales in a model of complexity. *The Journal of Finance*, *68*(6), 2549–2587.

Caccioli, F., Farmer, J. D., Foti, N., & Rockmore, D. (2015). Overlapping portfolios, contagion, and financial stability. *Journal of Economic Dynamics and Control*, *51*, 50–63.

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

Calvo-Armengol, A., & Jackson, M. O. (2004). The effects of social networks on employment and inequality. *American economic review*, *94*(3), 426–454.

Carvalho, V. M., Nirei, M., Saito, Y., & Tahbaz-Salehi, A. (2016). Supply chain disruptions: Evidence from the great east japan earthquake. *Columbia Business School Research Paper*(17-5).

Carvalho, V. M., & Tahbaz-Salehi, A. (2018). Production networks: A primer.

Castro, J., Gómez, D., Molina, E., & Tejada, J. (2017). Improving polynomial estimation of the shapley value by stratified random sampling with optimum allocation. *Computers & Operations Research*, *82*, 180–188.

Chava, S., & Purnanandam, A. (2011). The effect of banking crisis on bank-dependent borrowers. *Journal of Financial Economics*, *99*(1), 116–135.

Chodorow-Reich, G. (2014). The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis. *The Quarterly Journal of Economics*, *129*(1), 1–59.

Choi, S., Galeotti, A., & Goyal, S. (2017). Trading in networks: Theory and experiments. *Journal of the European Economic Association*, *15*(4), 784–817.

Coen, J., Lepore, C., & Schaanning, E. (2019). Taking regulation seriously: fire sales under solvency and liquidity constraints.

Cont, R., & Schaanning, E. (2019). Monitoring indirect contagion. *Journal of Banking & Finance*, *104*, 85–102.

Costello, A. M. (2020). Credit market disruptions and liquidity spillover effects in the supply chain. *Journal of Political Economy*, *128*(9), 000–000.

Covi, G., Gorpe, M., & Kok, C. (2019). Comap: mapping contagion in the euro area banking sector.

CEU eTD Collection

Csóka, P., & Herings, P. J.-J. (2018). Decentralized clearing in financial networks. *Management science*, *64* (10), 4681–4699.

Csóka, P., & Herings, P. J.-J. (2020). An axiomatization of the proportional rule in financial networks. *Management Science*.

Davis, G. F., Yoo, M., & Baker, W. E. (2003). The small world of the american corporate elite, 1982-2001. *Strategic organization*, *1* (3), 301–326.

Davis, J. A., & Leinhardt, S. (1967). The structure of positive interpersonal relations in small groups.

Dees, S., & Henry, J. (2017). Stress-test analytics for macroprudential purposes: Introducing stamp. *Satellite Models*, *13*.

Degryse, H., De Jonghe, O., Jakovljević, S., Mulier, K., & Schepens, G. (2017). The impact of bank shocks on firm-level outcomes and bank risk-taking. *Available at SSRN*.

de Jeude, J. v. L., Aste, T., & Caldarelli, G. (2019). The multilayer structure of corporate networks. *New Journal of Physics*, *21* (2), 025002.

Demir, B., Javorcik, B., Michalski, T. K., & Ors, E. (2018). Financial constraints and propagation of shocks in production networks. *Work. Pap., Univ. Oxford, UK*.

de Silanes, F. L., La Porta, R., & Shleifer, A. (1999). Corporate ownership around the world. *Journal of Finance*, *54* (2), 471–517.

Dhyne, E., Magerman, G., & Rubínová, S. (2015). *The belgian production network 2002-2012* (Tech. Rep.). NBB Working Paper.

Diamond, D. W., & Rajan, R. G. (2011). Fear of fire sales, illiquidity seeking, and credit freezes. *The Quarterly Journal of Economics*, *126* (2), 557–591.

Diem, C., Borsos, A., Reisch, T., Kertesz, J., & Thurner, S. (in press). Systemic risk and shock propagation in firm level production networks. *Unpublished working paper*.

Dietrich, D., & Hauck, A. (2020). Interbank borrowing and lending between financially constrained banks. *Economic Theory*, *70* (2), 347–385.

Dörr, S., Raissi, M., & Weber, A. (2018). Credit-supply shocks and firm productivity in italy. *Journal of International Money and Finance*, *87*, 155–171.

Drehmann, M., & Tarashev, N. (2013). Measuring the systemic importance of interconnected banks. *Journal of Financial Intermediation*, *22* (4), 586–607.

Duarte, F., & Eisenbach, T. M. (2018). Fire-sale spillovers and systemic risk. *FRB of New York Staff Report*(645).

Dutta, B., & Jackson, M. O. (2003). On the formation of networks and groups. In *Networks and groups* (pp. 1–15). Springer.

Dwenger, N., Fossen, F. M., & Simmler, M. (2015). From financial to real economic crisis: Evidence from individual firm-bank relationships in germany. *Freie Universität Berlin, School of Business & Economics Discussion Paper*(2015/28).

Eisenberg, L., & Noe, T. H. (2001). Systemic risk in financial systems. *Management Science*, *47* (2), 236–249.

Elhorst, J. P. (2014). *Spatial econometrics: from cross-sectional data to spatial panels* (Vol. 479). Springer.

CEU eTD Collection

Elliott, M., Golub, B., & Jackson, M. O. (2014). Financial networks and contagion. *American Economic Review*, *104*(10), 3115–53.

Elsinger, H., Lehar, A., & Summer, M. (2013). Network models and systemic risk assessment. *Handbook on Systemic Risk*, *1*, 287–305.

Evans, T. S., & Lambiotte, R. (2010). Line graphs of weighted networks for overlapping communities. *The European Physical Journal B*, *77*(2), 265–272.

Farmer, J. D., & Foley, D. (2009). The economy needs agent-based modelling. *Nature*, *460*(7256), 685–686.

Farmer, J. D., Gallegati, M., Hommes, C., Kirman, A., Ormerod, P., Cincotti, S., ... Helbing, D. (2012). A complex systems approach to constructing better models for managing financial markets and the economy. *The European Physical Journal Special Topics*, *214*(1), 295–324.

Frank, O. (1971). *Statistical inference in graphs.* Foa Repro.

Fujiwara, Y., & Aoyama, H. (2010). Large-scale structure of a nation-wide production network. *The European Physical Journal B*, *77*(4), 565–580.

Fujiwara, Y., Terai, M., Fujita, Y., & Souma, W. (2016). Debtrank analysis of financial distress propagation on a production network in japan.

Furfine, C. H. (2003). Interbank exposures: Quantifying the risk of contagion. *Journal of money, credit and banking*, 111–128.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, *79*(3), 733–772.

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

Gai, P., & Kapadia, S. (2010). Contagion in financial networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *466*(2120), 2401–2423.

Galeotti, A., Goyal, S., Jackson, M. O., Vega-Redondo, F., & Yariv, L. (2010). Network games. *The review of economic studies*, *77*(1), 218–244.

Garcia-Bernardo, J., Fichtner, J., Takes, F. W., & Heemskerk, E. M. (2017). Uncovering offshore financial centers: Conduits and sinks in the global corporate ownership network. *Scientific Reports*, *7*(1), 6246.

Gatti, D. D., Gallegati, M., Greenwald, B., Russo, A., & Stiglitz, J. E. (2010). The financial accelerator in an evolving credit network. *Journal of Economic Dynamics and Control*, *34*(9), 1627–1650.

Georgescu, O.-M. (2015). Contagion in the interbank market: Funding versus regulatory constraints. *Journal of Financial Stability*, *18*, 1–18.

Giannetti, M., Burkart, M., & Ellingsen, T. (2011). What you sell is what you lend? explaining trade credit contracts. *The Review of Financial Studies*, *24*(4), 1261–1298.

Glattfelder, J. B. (2013). Backbone of complex networks of corporations: The flow of control. In *Decoding complexity* (pp. 67–93). Springer.

Golan, A., Judge, G., & Robinson, S. (1994). Recovering information from incomplete or partial multisectoral economic data. *The Review of Economics and Statistics*, 541–549.

CEU eTD Collection

Gorg, H., & Strobl, E. (2001). Multinational companies and productivity spillovers: A meta-analysis. *The economic journal*, *111*(475), F723–F739.
Goyal, S. (2012). *Connections: an introduction to the economics of networks.* Princeton University Press.
Granovetter, M. (1976). Network sampling: Some first steps. *American journal of sociology*, *81*(6), 1287–1303.
Gross, C., & Siklos, P. L. (2020). Analyzing credit risk transmission to the nonfinancial sector in europe: A network approach. *Journal of Applied Econometrics*, *35*(1), 61–81.
Gurgone, A., Iori, G., & Jafarey, S. (2018). The effects of interbank networks on efficiency and stability in a macroeconomic agent-based model. *Journal of Economic Dynamics and Control*, *91*, 257–288.
Haldane, A. G., & May, R. M. (2011). Systemic risk in banking ecosystems. *Nature*, *469*(7330), 351–355.
Hallegatte, S. (2008). An adaptive regional input-output model and its application to the assessment of the economic cost of katrina. *Risk Analysis: An International Journal*, *28*(3), 779–799.
Heemskerk, E. M., & Takes, F. W. (2016). The corporate elite community structure of global capitalism. *New Political Economy*, *21*(1), 90–118.
Hulten, C. R. (1978). Growth accounting with intermediate inputs. *The Review of Economic Studies*, *45*(3), 511–518.
Idier, J., & Piquard, T. (2017). Pandemic crises in financial systems: a simulation-model to complement stress-testing frameworks.
Iino, T., & Iyetomi, H. (2012). Directional bias between communities of a production network in japan. In *Intelligent decision technologies* (pp. 273–280). Springer.
Inoue, H., & Todo, Y. (2019a). Firm-level propagation of shocks through supply-chain networks. *Nature Sustainability*, *2*(9), 841–847.
Inoue, H., & Todo, Y. (2019b). Propagation of negative shocks across nationwide firm networks. *PloS one*, *14*(3), e0213648.
Inoue, H., & Todo, Y. (2020). The propagation of the economic impact through supply chains: The case of a mega-city lockdown against the spread of covid-19. *Available at SSRN 3564898*.
Ivashina, V., & Scharfstein, D. (2010). Bank lending during the financial crisis of 2008. *Journal of Financial economics*, *97*(3), 319–338.
Jackson, M. O. (2010). *Social and economic networks.* Princeton university press.
Jackson, M. O., & Pernoud, A. (2020). Systemic risk in financial networks: A survey. *Available at SSRN*.
Jackson, M. O., & Zenou, Y. (2015). Games on networks. In *Handbook of game theory with economic applications* (Vol. 4, pp. 95–163). Elsevier.
Javed, M. A., Younis, M. S., Latif, S., Qadir, J., & Baig, A. (2018). Community detection in networks: A multidisciplinary review. *Journal of Network and Computer Applications*, *108*, 87–111.
Jiangli, W., Unal, H., & Yom, C. (2008). Relationship lending, accounting disclosure, and credit availability during the asian financial crisis. *Journal*

CEU eTD Collection

of Money, Credit and Banking, *40*(1), 25–55.
Jiménez, G., Ongena, S., Peydró, J.-L., & Saurina, J. (2012). Credit supply and monetary policy: Identifying the bank balance-sheet channel with loan applications. *American Economic Review*, *102*(5), 2301–26.
Khwaja, A. I., & Mian, A. (2008). Tracing the impact of bank liquidity shocks: Evidence from an emerging market. *American Economic Review*, *98*(4), 1413–42.
Kikkawa, K., Magerman, G., & Dhyne, E. (2019). Imperfect competition in firm-to-firm trade. *Available at SSRN 3389836*.
Kim, M., Kliger, D., & Vale, B. (2003). Estimating switching costs: the case of banking. *Journal of Financial Intermediation*, *12*(1), 25–56.
Kranton, R. E., & Minehart, D. F. (2001). A theory of buyer-seller networks. *American economic review*, *91*(3), 485–508.
Kumar, A., Chakrabarti, A. S., Chakraborti, A., & Nandi, T. (2020). Distress propagation on production networks: Coarse-graining and modularity of linkages. *arXiv preprint arXiv:2004.14485*.
Leech, D. (2002). Computation of power indices.
Leicht, E., & D'Souza, R. M. (2009). Percolation on interacting networks. *arXiv preprint arXiv:0907.0894*.
Leontief, W. W. (1951). *The structure of american economy, 1919-1939: an empirical application of equilibrium analysis.* (Tech. Rep.).
Liu, X., Siler, P., Wang, C., & Wei, Y. (2000). Productivity spillovers from foreign direct investment: Evidence from uk industry level panel data. *Journal of International Business Studies*, *31*(3), 407–425.
Lucas Jr, R. E. (1977). Understanding business cycles. In *Carnegie-rochester conference series on public policy* (Vol. 5, pp. 7–29).
Luo, S. (2019). Propagation of financial shocks in an input-output economy with trade and financial linkages of firms. *Review of Economic Dynamics*.
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.
Magerman, G., De Bruyne, K., Dhyne, E., & Van Hove, J. (2016). *Heterogeneous firms and the micro origins of aggregate fluctuations* (Tech. Rep.). NBB Working Paper.
Mastrandrea, R., Squartini, T., Fagiolo, G., & Garlaschelli, D. (2014). Enhanced reconstruction of weighted networks from strengths and degrees. *New Journal of Physics*, *16*(4), 043022.
McFadden, D. (1963). Constant elasticity of substitution production functions. *The Review of Economic Studies*, *30*(2), 73–83.
McNerney, J. (2009). Network properties of economic input-output networks.
Montagna, M., & Kok, C. (2016). Multi-layered interbank model for assessing systemic risk.
Newman, M. (2018). *Networks*. Oxford university press.
Newman, M. E. (2006). Modularity and community structure in networks. *Proceedings of the national academy of sciences*, *103*(23), 8577–8582.

CEU eTD Collection

Oberfield, E. (2012). Business networks, production chains, and productivity: A theory of input-output architecture.

Pichler, A., & Farmer, J. D. (2021). Modeling simultaneous supply and demand shocks in input-output networks. *arXiv preprint arXiv:2101.07818*.

Pichler, A., Pangallo, M., del Rio-Chanona, R. M., Lafond, F., & Farmer, J. D. (2020). Production networks and epidemic spreading: How to restart the uk economy? *Production Networks and Epidemic Spreading: How to Restart the UK Economy*.

Pichler, A., Pangallo, M., del Rio-Chanona, R. M., Lafond, F., & Farmer, J. D. (2021). In and out of lockdown: Propagation of supply and demand shocks in a dynamic input-output model. *and out of Lockdown: Propagation of Supply and Demand Shocks in a Dynamic Input-Output Model (February 18, 2021)*.

Poledna, S., Molina-Borboa, J. L., Martínez-Jaramillo, S., Van Der Leij, M., & Thurner, S. (2015). The multi-layer network nature of systemic risk and its implications for the costs of financial crises. *Journal of Financial Stability*, *20*, 70–81.

Popoyan, L., Napoletano, M., & Roventini, A. (2020). Winter is possibly not coming: Mitigating financial instability in an agent-based model with interbank market. *Journal of Economic Dynamics and Control*, 103937.

Puri, M., Rocholl, J., & Steffen, S. (2011). Global retail lending in the aftermath of the us financial crisis: Distinguishing between supply and demand effects. *Journal of Financial Economics*, *100*(3), 556–578.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of international Economics*, *48*(1), 7–35.

Riccetti, L., Russo, A., & Gallegati, M. (2013). Leveraged network-based financial accelerator. *Journal of Economic Dynamics and Control*, *37*(8), 1626–1640.

Rogers, L. C., & Veraart, L. A. (2013). Failure and rescue in an interbank network. *Management Science*, *59*(4), 882–898.

Roncoroni, A., Battiston, S., Hałaj, G., & Kok, C. (2019). Interconnected banks and systemically important exposures.

Schnabl, P. (2012). The international transmission of bank liquidity shocks: Evidence from an emerging market. *The Journal of Finance*, *67*(3), 897–932.

Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games*, *2*(28), 307–317.

Silva, T. C., da Silva Alexandre, M., & Tabak, B. M. (2018). Bank lending and systemic risk: A financial-real sector network approach with feedback. *Journal of Financial Stability*, *38*, 98–118.

Tandon, A., Albeshri, A., Thayananthan, V., Alhalabi, W., & Fortunato, S. (2019). Fast consensus clustering in complex networks. *Physical Review E*, *99*(4), 042301.

Tarashev, N., Drehmann, M., et al. (2011). *Measuring the systemic importance of interconnected banks* (Tech. Rep.). Bank for International Settlements.

CEU eTD Collection

Thurner, S., Hanel, R., & Klimek, P. (2018). *Introduction to the theory of complex systems*. Oxford University Press.

Tintelnot, F., Kikkawa, A. K., Mogstad, M., & Dhyne, E. (2018). *Trade and domestic production networks* (Tech. Rep.). National Bureau of Economic Research.

Tomasello, M. V., Napoletano, M., Garas, A., & Schweitzer, F. (2017). The rise and fall of r&d networks. *Industrial and corporate change*, *26*(4), 617–646.

Tumminello, M., Lillo, F., & Mantegna, R. N. (2010). Correlation, hierarchies, and networks in financial markets. *Journal of Economic Behavior & Organization*, *75*(1), 40–58.

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

Upper, C., & Worms, A. (2004). Estimating bilateral exposures in the german interbank market: Is there a danger of contagion? *European economic review*, *48*(4), 827–849.

Vitali, S., & Battiston, S. (2011). Geography versus topology in the european ownership network. *New Journal of Physics*, *13*(6), 063021.

Vitali, S., & Battiston, S. (2014). The community structure of the global corporate network. *PloS one*, *9*(8), e104655.

Vitali, S., Battiston, S., & Gallegati, M. (2016). Financial fragility and distress propagation in a network of regions. *Journal of Economic Dynamics and Control*, *62*, 56–75.

Vitali, S., Glattfelder, J. B., & Battiston, S. (2011). The network of global corporate control. *PloS one*, *6*(10), e25995.

Von Thadden, E.-L. (1995). Long-term contracts, short-term investment and monitoring. *The Review of Economic Studies*, *62*(4), 557–575.

Watanabe, T., Uesugi, I., & Ono, A. (2015). *The economics of interfirm networks* (Vol. 4). Springer.

Wu, D. (2016). Shock spillover and financial response in supply chain networks: Evidence from firm-level data. *Unpublished working paper*.

Yang, B., Liu, D., & Liu, J. (2010). Discovering communities from social networks: methodologies and applications. In *Handbook of social network technologies and applications* (pp. 331–346). Springer.

Zajac, E. J., & Westphal, J. D. (1996). Director reputation, ceo/board power, and the dynamics of board interlocks. In *Academy of management proceedings* (Vol. 1996, pp. 254–258).

CEU eTD Collection