Essays on Aggregate Fluctuations, Network Dynamics and Statistical Regularities in Economics

Inauguraldissertation
zur
Erlangung der Doktorwürde
der
Fakultät Sozial- und Wirtschaftswissenschaften
der
Otto-Friedrich-Universität Bamberg

Vorgelegt von
Alexander Hempfing

Disputation
Bamberg, 22. Februar 2021

URN: urn:nbn:de:bvb:473-irb-495690
DOI: https://doi.org/10.20378/irb-49569

Promotionskommission

Professor Mishael Milaković, PhD	(Erstgutachter und Sprecher)
Professor Dr. Christian Proaño	(Zweitgutachter)
Professor Dr. Frank Westerhoff	(Drittgutachter)

PK 045
PROMOTIONSKOLLEG

Gefördert durch die

Hans Böckler Stiftung
Mitbestimmung · Forschung · Stipendien

Diese Dissertation wurde im Rahmen eines strukturierten Programms an der Otto-Friedrich-Universität Bamberg verfasst und materiell wie ideell durch die Hans-Böckler-Stiftung gefördert. Hierfür bin ich sehr dankbar.

In memory of my grandparents, in gratitude to those who remained: Mama, Line, and Sebastian.

Acknowledgements

"Learning is like rowing against the current. As soon as you stop, you drift back again."
Xúnzǐ

This dissertation would not have been possible without the support and sustained assistance of many. I am infinitely grateful to have received such continuous encouragement.

First of all, I would like to thank my first supervisor, Mishael Milaković. Being introduced to the economic sciences under his guidance was the breeding ground that made this dissertation possible. His advice, his ideas and his far-reaching knowledge were as valuable as the critical questioning of thoughts, which always pursued the goal of improving the presented findings, reflecting on them, becoming better and thus endeavoring to make a lasting contribution to a field of science that could hardly be more complex in its underlying structure. I am especially grateful for this support.

Furthermore, I would like to thank Christian Proaño and Frank Westerhoff. As part of my dissertation committee and as co-founders of the Bamberg Research Group on Behavioral Macroeconomics (BaGBeM), they played a decisive role in making this PhD possible. The fact that they always had an open ear and provided valuable comments made the past years as instructive as they were pleasant. In this context, I would like to thank the Hans Böckler Foundation, which has supported me both materially and ideologically over the past three years. Without their financial support, it would not have been possible for me to undertake this journey.

I would also like to explicitly thank my co-authors, without whom the content of this dissertation would not have been possible. Jan Schulz, co-author of two of the four papers presented here, especially deserves my sincere thanks. His expertise, vision and friendship have always been a pillar of support. It is also thanks to Philipp Mundt that I was able to develop significantly in my last year. His professionalism, his know-how and his eye for detail were both a role model and a source of inspiration for me. I would also like to take this opportunity to thank all other members and associates of the BaGBeM and the Department of Economics, who made the past three years a unique experience.

Finally, I would like to thank my family for their continuous encouragement, and for having supported me to the best of their ability from day one. Last but not least, I would like to thank my fiancée, who has stood by me all these years, bringing me down to earth when necessary, and encouraging me when things were not going well. For all this, I am hugely grateful.

Preface

The individual behavior of economic entities, such as agents, firms or industries, and their interaction in markets may create complex economic dynamics and regularities in the aggregate. This doctoral thesis seeks to demonstrate how, and under which conditions, such dynamics may arise, leading to and explaining certain cyclicities in macroeconomic fluctuations, tie formation processes in economic networks and persistent distributional properties in foreign exchange and virtual currency markets. With this aim, this dissertation is divided into three parts.

Part I of this thesis consists of two papers in which we analyze how aggregate fluctuations arise, and the role that can be attributed to idiosyncratic shocks at the microeconomic and mesoeconomic level as well as higher-order linkages in the supply and demand interdependencies driving macroeconomic outcomes. As such, the first paper in *Chapter 1* considers recent advances in the modeling of aggregate fluctuations, which have emphasized the importance of heavily skewed and fat-tailed distributions in firm size (Gabaix, 2011) or connections in production networks (Acemoğlu et al., 2012), arguing that idiosyncratic shocks on the microeconomic level can have considerable macroeconomic effects. We illustrate that both the size of economic entities, measured as the granularity of sales shares, and the heavily skewed distribution of connections in production networks play a significant role in the description of aggregate fluctuations, and that stable results can only be achieved if both concepts are taken into account. In particular, fluctuations at the aggregate level can best be described when explicitly accounting for the size of an entity and its position within the production network when facing idiosyncratic shocks. *Chapter 2*, containing the second paper, aims to shed light on the remarkable degree of synchronicity displayed by business cycles. Here it is shown that the large degree of comovement of GDP time series can mainly be attributed to aggregate regularities in the final goods markets, and that shocks faced by these markets seem to propagate through higher-order connections rather than direct pairwise linkages. Both papers in this part were co-written with Jan Schulz and Mishael Milaković.

In spite of the importance of interacting heterogeneous firms and industries in the research of international trade and production networks, leading to aggregate fluctuations, there has been a surprising lack of attention to the fact that trading partners are embedded in an evolutionary process of production and exchange across borders globally, which develops over time. This is because trade relationships provide feedback on the underlying network structure, and the structure itself influences trade link formation processes. This makes it difficult to distinguish endogenous network processes from exogenous events, and to attribute them to theoretical or empirical considerations. Part II contains one paper, represented by *Chapter 3*, which attempts to tackle this issue. The contribution seeks to improve our understanding of network dynamics and tie formation processes within and between economies over time. The goal of this paper is therefore to develop a stochastic actor-oriented model that is able to capture established stylized facts in the literature and examine reasons for the evolution of production as well as import and export trade preference networks. The empirical value of the model is demonstrated by combining industry-level OECD data with several covariates such as total factor productivity, labor cost intensity, economic complexity, geographic distance or regional proximity. This allows us to evaluate the influence of these variables on the process of network link formation over time. We identify various important features for production networks in international trade, such as network proximity, geographical distance or technological complements, being crucial drivers for decision processes in trade link formation. Moreover, we find that OECD industries have a tendency to share trade activities, following their peers, a feature that points to herding behavior in the selection of partners outside the OECD or the adaptation of particularly profitable import and export portfolios. Furthermore, we provide initial insights into the dynamics of tie formation processes, where inter-industry OECD trade affects decisions made towards import and export preferences, and vice versa. We find that triadic dependencies exists, allowing certain industries to exploit sourcing differentials of different suppliers, e.g. by cooperation or through competition between two suppliers with partially overlapping capabilities. The content of this part was co-written with Philipp Mundt.

Part III contains *Chapter 4*, which provides a statistical evaluation of the distributional properties and statistical regularities of virtual, intra-virtual and traditional currency exchange rates. The last part of this thesis arose from my own research interest in modern means of payment, following the debate as well as the rise and fall of Bitcoins and similar virtual currencies in recent years. This aspect intersected with the research interests of my first supervisor on statistical equilibrium economic modeling, where interactions in markets, regardless of their mode of operation, give rise to aggregate regularities in the form of statistical distributions, which are neither anticipated nor intended by individual market participants. The analysis at hand shows that, in spite of their differing mode of formation, daily log-returns of all currency types share peculiar properties that have also been examined thoroughly in other fields of economic literature. Unlike the previously mentioned papers, this chapter was written alone and has already been published in an international double-blinded peer-reviewed journal*.

Each of the four papers are independent from each other, and can be read without any prior knowledge of each other, as they are written in a format suitable for publication in professional journals. Despite the papers mentioned in this preface sharing different specifics, what they have in common is that they seek a better understanding of economic processes within a complex and dynamic world.

*The original version of this article has already been published and can be found online at the following source: Hempfing A (2019) What's left after the hype? An empirical approach comparing the distributional properties of traditional and virtual currency exchange rates. PLoS ONE 14(7):e0220070. https://doi.org/10.1371/journal.pone.0220070.

Contents

List of Figures

List of Tables

Part I

Essays on Aggregate Fluctuations in Economics

Chapter 1

A Mesoeconomic Aggregation Rule for Microeconomic Shocks*

Co-written by Jan Schulz and Mishael Milaković

Abstract

Latest theoretical advances have called the invariance of sales shares with respect to productivity differences into question, arguing that sizable higher-order macroeconomic effects may emerge when sales shares react to productivity shocks. Here we propose a parsimonious model for the adjustment of sales shares in response to productivity shocks, and operationalize it to readily quantify the impact of idiosyncratic shocks on aggregate fluctuations. Using input-output data for the European (EU28) economy, we find that our model significantly outperforms recently suggested specifications that rely either on the granularity of sales shares or the heavily skewed distribution of connections in production networks. While our results confirm earlier findings in the sense that microeconomic shocks are an important driver of macroeconomic fluctuations, we demonstrate that previous approaches substantially underestimate their relative impact because they fail to account for what we term the mesoeconomic interaction of granularity and network effects. Our main empirical finding is that idiosyncratic shocks to less than two percent of industries already explain almost eighty percent of the business cycle once we account for this interaction.

1.1 Introduction

Recent advances in the modeling of aggregate fluctuations have emphasized the importance of heavily skewed and fat-tailed distributions in firm size (Gabaix, 2011) or in the connectivity of production networks (Acemoğlu et al., 2012), arguing that id-

*The article in this chapter has been presented to participants of the 1st and 2nd Behavioral Macroeconomics Workshop in Bamberg, the 7th Workshop on Networks in Economics and Finance in Lucca, the 6th Annual GENED Meeting in Kiel, the 22nd Forum for Macroeconomics and Macroeconomic Policies in Berlin, and the 24th Workshop on Economic Science with Heterogeneous Interacting Agents in London.

iosyncratic shocks on the microeconomic level can have considerable macroeconomic effects. These approaches essentially rest on the foundational theorem for efficient economies put forward by Hulten (1978), and mark an important deviation from the diversification argument of Lucas (1977), which claims that microeconomic shocks will cancel out in the aggregation process. In light of Hulten's theorem, whereby the aggregate productivity of an economy can be expressed as a sales share-weighted sum of idiosyncratic productivity levels, empirical studies have typically assumed sales shares to be constant between periods of observation. Baqaee and Farhi (2019) call this invariance of sales shares with respect to productivity differences into question and show that sizeable higher-order effects can emerge when sales shares do respond to productivity shocks, and that the magnitude of such propagation effects will depend on the structure of the underlying production network. Inspired by their theoretical results, we propose a parsimonious model for the perpetual adjustment of sales shares through productivity shocks between discrete periods that is readily estimated from available data. In this regard our contribution lies in an operationalization of Baqaee and Farhi's theoretical considerations that explicitly accounts for the dynamic adjustment of sales shares. Moreover, while our results confirm earlier findings in the sense that microeconomic shocks are an important driver of macroeconomic fluctuations, we also demonstrate that previous empirical approaches substantially underestimate the relative impact of idiosyncratic shocks because they fail to account for what we term the mesoeconomic interaction of granularity and network effects. Using input-output data for the European (EU28) economy on the industry level, we find that shocks to less than two percent of industries already explain almost eighty percent of the business cycle once we account for this interaction.

The aggregation from microeconomic shocks to macroeconomic behavior is conceptually far from trivial because the interactions among economic agents can amplify or attenuate initial shocks. Consequently, borrowing a term from statistical physics, we adopt a *mesoscopic* perspective on aggregation (see, e.g., Imry, 2008) and understand both the size distribution as well as the network structure as mesoeconomic objects that we use as intermediate layers for aggregation. Mesoeconomic objects are still subject to microeconomic forces, in our case the microeconomic elasticities of substitution, yet the properties emerging from the interaction of industries with each other are irreducible to their individual micro behavior. This irreducibility stems from the fact that relative sizes and network structures are not individual properties of the considered micro entities but can only be defined in relation to all other micro entities at once. While our model is obviously not the only conceivable approach to mesoeconomic aggregation, its empirical performance merits serious consideration and avoids many of the problems that direct aggregation from the micro to the macro level brings about.[1] Intuitively, shocks to firms or industries that are both large in size and central in the production network will have a substantial macroeconomic impact whereas shocks to small and peripheral entities will not. The mesoeconomic interaction of granularity and network effects is less intuitive when the two effects work in opposite directions. To gauge the relative importance of the two effects, we also derive the polar cases of an economy that is homogeneous in its network structure yet heterogeneous in its size distribution (the purely granular case), and of an economy that is heterogeneous in its network structure yet homogeneous in its size distribution (the purely networked case). The commonly used approximation by Riemann sums that leads to the purely granular case is thus a special case of our model when network effects are irrelevant. Empirically it is harder to find cases of large but peripheral industries than it is to find relatively small but highly central industries,[2] so it is perhaps not overly surprising that the explanatory power with respect to aggregate fluctuations is greater for the purely networked case than for the purely granular case. More importantly, though, the interaction of granularity and network structure substantially increases explanatory power compared to these two polar cases.

Our paper relates to at least three strands of previous research. Hulten's theorem and Domar (1961) aggregation of idiosyncratic productivity shocks are the basis of modern growth accounting and its empirical applications, suggesting a linear impact of productivity shocks on aggregate fluctuations via sales shares (or Domar weights). The contributions by Gabaix (2011) and Acemoğlu et al. (2012) are important because they demonstrate that even in a world of linear shock responses idiosyncratic shocks will have a significant impact on the business cycle if Domar weights are fat-tailed. Viewed from this perspective, their approaches are two sides of the same coin (Gabaix, 2016), showing that idiosyncratic shocks will generally not cancel out in the process of aggregation. Second, from a microeconomic point of view, it is only the network perspective that sheds light on comovement and shock transmission among producers of intermediate goods, and microeconomic studies indeed find a large degree of sectoral comovement (Saint-Paul, 1993; Foerster et al., 2011; Di Giovanni et al., 2014; Stella,

[1]Kirman (1992) argues that traditional aggregation rules have typically imposed an unrealistically large degree of homogeneity on economic agents, for instance the Gorman (1961) polar form, resulting in the inability of such models to explain the moments of aggregate output time series (Ascari et al., 2015).

[2]A popular example for the latter is the production of computer hard drives by Hitachi, Seagate, Toshiba, and Western Digital, who were servicing around one third of global demand from their Thai facilities when devastating floods hit the country in the summer and fall of 2011. This resulted in global supply chain interruptions with massive production delays in the global information and communications technology sector, yet Western Digital, which at the time was the leading manufacturer of hard drives with a market share of about fifty percent, had revenues (total assets) on the order of around 10 (8.5) billion US dollars, hardly qualifying them as a very large firm in the global production network.

2015), a property that we can also reproduce in our model. Third, Baqaee and Farhi (2019) challenge and qualify Hulten's aggregation rule, extending it beyond first-order terms to allow for nonlinearities that can emerge when sales shares vary with productivity shocks. They argue that Hulten's aggregation rule is an artifact of the constant returns to scale (CRS) property assumed for every micro production function, and recent sectoral level studies estimate elasticities of substitution that differ from the CRS benchmark (Atalay, 2017; Boehm et al., 2018; Barrot and Sauvagnat, 2016). This is also what casual empiricism suggests, since relative sales shares vary on all levels of aggregation from plants to firms to industries or sectors.

To paraphrase Sims (1980), we face a "wilderness" of potential alternatives for the aggregation of production once we depart from the knife-edge Cobb-Douglas case, apparently preventing the estimation of the correct structural model. To circumvent this problem, we opt for a small set of nonparametric assumptions that imply restrictions on the functional form of the considered production function, resulting in a reduced-form model based on the elasticity of substitution among supplier relationships. Aggregate shocks are then given by the mesoeconomically weighted sum of idiosyncratic productivity shocks, allowing us to decompose the aggregate shock into its constituent parts in order to isolate the impact of each industry on aggregate fluctuations. Our mesoeconomic measure robustly outperforms previous approximations in the literature that assume an invariance of sales shares between periods or rely exclusively on comovement. It is worthwhile to point out that we achieve this performance using raw data only, that is without resorting to any arbitrary manipulation of the productivity growth rate distributions such as *winsorizing*, which considerably distorts the performance of granular approaches (Dosi et al., 2019).[3] Finally, our approach can also be thought of as a short-run complement to the complexity index proposed by Hidalgo and Hausmann (2009), who argue that the relative degree of complexity of national production structures is a useful predictor for national levels of production in the long run. Notably, though, our short-run perspective is transnational and emphasizes instead that aggregate short-run fluctuations depend on industry level characteristics that are not confined to national boundaries.

We introduce our reduced-form model in the next section, followed by a description of the data and the econometric identification strategy for testing our model predictions in section 1.3. Our main empirical results, including several robustness checks,

[3] Since purely granular approaches to aggregate fluctuations are very sensitive to winsorizing, we quantify its impact for productivity shocks that have a normal or a double-exponential distribution in appendix 1.A. We show analytically that in the empirically relevant leptokurtic case the bias from winsorizing is substantial.

are presented in section 1.4, and we conclude with a discussion of the implications and limitations of our results in section 1.5.

1.2 Idiosyncratic Shocks and Mesoeconomic Aggregation

Hulten's theorem states that for any efficient economy in general equilibrium with N producers (firms or industries) indexed by i

$$d\,log\,TFP = \sum_{i=1}^{N} \frac{S_i}{Y} d\,log\,TFP_i, \tag{1.1}$$

where TFP denotes the total factor productivity and Y the GDP of the economy, while S_i are the sales and TFP_i the productivity levels of entity i. Following Gabaix (2011) and assuming that GDP and TFP growth rates are related through a proportionality constant μ reflecting factor usage, we have

$$d\,log\,Y = \mu \sum_{i=1}^{N} \frac{S_i}{Y} d\,log\,TFP_i. \tag{1.2}$$

This implies that for any given time period the rate of change of GDP is determined by the Domar weighted sum of idiosyncratic productivity shocks. Hulten's theorem holds for any efficient economy, irrespective of the existence or particular shape of input-output linkages in the economy, but merely up to a first-order approximation, or equivalently when the Domar weights S_i/Y are time invariant with respect to idiosyncratic productivity TFP_i. Whenever sales shares do respond to productivity shocks, however, potentially sizeable higher-order effects can emerge from these changes in sales shares. Baqaee and Farhi (2019) consider a second-order approximation and show that the impact of time-varying sales shares on GDP growth is uniquely determined by two effects: the derivatives of sales shares and of the input-output multiplier with respect to productivity shocks. While the former describes the change of sales shares with respect to productivity differences, the latter describes how these changes affect the ratio of total sales to GDP. It turns out that all higher order terms vanish only in the Cobb-Douglas CRS case. Whenever the production functions of some producers deviate from this knife-edge scenario, second-order terms can become quantitatively important, depending on the degree of factor reallocation that is assumed.

The operationalization of these theoretical insights is complicated by at least two concerns. First, the estimation of a complete structural model depends on unobservable general equilibrium elasticities of substitution. Second, since sales shares and productivity shocks are observable, we could construct a measure of the total change in TFP using (1.2) by integrating over a discrete time interval, where now both productivity levels TFP_i and sales shares S_i/Y can vary over time,

$$\frac{\Delta Y_t}{Y_{t-1}} = \mu \sum_{i=1}^{N} \int_{t-1}^{t} \frac{S_i}{Y}(t) d\, log\, TFP_i(t), \tag{1.3}$$

yet a good approximation of the integral in continuous time necessitates reasonably high observational frequencies. The discrete approximation by (left) Riemann sums that is used by both Gabaix (2011) and Baqaee and Farhi (2019) reads

$$\frac{\Delta Y_t}{Y_{t-1}} \approx \mu \sum_{i=1}^{N} \frac{S_{i,t-1}}{Y_{t-1}} \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}, \tag{1.4}$$

and becomes more accurate the higher the frequency of observation. This approximation, however, presupposes that sales shares stay constant between $t-1$ and t and jump to their new level precisely at time t. Given quarterly or annual frequencies, it seems rather implausible for sales shares to stay (even approximately) constant between periods.[4] Hence we want to improve and generalize approximation (1.4) in order to account for the adjustment in sales shares within a discrete time interval using observations only at $t-1$ and t. To do so, we employ the following parsimonious set of assumptions that recovers the approximation by Riemann sums in (1.4) as a special case. Since our data are on the industry level, the exposition will refer to industries, yet in principle it applies equally to firms. One assumption per level of aggregation fully determines our model, providing a convenient modular structure:

1. On the *micro* level, we follow standard procedure and assume that idiosyncratic productivity shocks are exogenous and that each industry's revenue and TFP growth rates are related by a finite proportionality constant $\mu > 1$.

2. On the *meso* level, we assume a constant and finite elasticity of substitution $\omega > 1$ between supplying and supplied industries, implying that intermediate inputs are *micro substitutes* (Blackorby and Russell, 1989).[5] If two industries do not have a supplier-recipient relationship, there is no direct effect from the growth of revenue in one to the other, but there will generally be an indirect effect that is mediated by the elasticity of substitution.

3. On the *macro* level, we aggregate from end-of-period sales levels to GDP via a

[4]While the input-output multiplier in our European data is remarkably stable over time, with an average value of around 0.6 and a coefficient of variation (CoV) of 0.01, the average CoV for Domar weights is almost an order of magnitude larger at 0.09.

[5]We deliberately refrain from decomposing this substitutability in its partial price and quantity effect, since this would require us to define the degree of factor mobility, in particular labor mobility, which depends on the considered time-horizon (Baqaee and Farhi, 2019).

constant input-output multiplier $\phi \in (0, 1]$ that measures the ratio of GDP to total sales, so $\phi < 1$ implies the existence of intermediate goods, whereas $\phi = 1$ corresponds to an islands economy without intermediate consumption.

Since we assume linear relations both at the micro and macro level, nonlinear effects have to originate at the meso level and it turns out that end-of-period sales levels are determined by the *interaction* of the size distribution with the topology of the production network, mediated by the elasticity of substitution.

Heuristically, the idea is to trace how a productivity shock to one industry propagates through the production network and then to aggregate this effect over all industries experiencing simultaneous shocks. This will lead to an expression that resembles the centrality measure first proposed by Katz (1953), and it is also instrumental in deriving the purely granular and the purely networked economy as special cases. The linkages in the production network (or graph) G are described by the binary adjacency matrix A, where $A_{ij} = 1$ if j is a supplier to i and $A_{ij} = 0$ otherwise. Let d denote the graph distance between two industries and consider what we term the *order effect* of a productivity shock to an industry j as a function of the graph distance to an industry i that is (possibly but not necessarily) different from j. Denote the rate of change in the sales of an industry j that has a distance d to the initially affected industry i by $\Delta S^d_{j,t}/S_{j,t-1}$. From our micro assumption, the initial effect of a productivity shock on the sales of industry i is the zeroth-order effect denoted by $\Delta S^0_{i,t}/S_{i,t-1}$, and is proportional to the initial productivity shock[6]

$$\frac{\Delta S^0_{i,t}}{S_{i,t-1}} = \mu \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \tag{1.5}$$

Suppose that each industry i takes goods produced by some industry j as inputs; given our meso level assumption of a constant elasticity of substitution $\omega > 1$, the first-order effect of a change in industry i's sales on the sales growth of a supplier j to i is then given by

$$\frac{\Delta S^0_{i,t}}{S_{i,t-1}} = \omega \frac{\Delta S^1_{j,t}}{S_{j,t-1}}, \tag{1.6}$$

implying that the first-order effect on the sales growth of a supplier industry j stemming from a single productivity shock to industry i is

[6]This is a conventional assumption in real business cycle models, and sometimes the proportionality constant $\mu > 1$ is interpreted as factor usage (see Gabaix, 2011, for a more detailed discussion of this interpretation).

$$\frac{\Delta S_{j,t}^1}{S_{j,t-1}} = \mu\omega^{-1}\frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \tag{1.7}$$

Next consider an industry k that is a supplier to industry j; given a direct path from i to k through j, the second-order effect on k from the initial shock to i is therefore

$$\frac{\Delta S_{k,t}^2}{S_{k,t-1}} = \mu\omega^{-2}\frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}, \tag{1.8}$$

and so forth. It is convenient to rewrite equations (1.7) and (1.8) in terms of their order effect for an arbitrary industry j that might be but is not necessarily connected to the initially affected industry i by a distance d. This allows to generalize for an arbitrary number of paths between industries at a distance d by exploiting a property of the d-th power of the binary adjacency matrix A. For this, we note that A_{ij}^d denotes the number of paths between industry j and i through $d-1$ intermediaries. Thus A_{ij}^d gives the number of ways through which industry j is affected by a change in the productivity of industry i in the d-th round of adjustments, mediated by the constant elasticity of substitution.

If we assume that the network topology does not change during a discrete time interval (but might very well change from one period of observation to the next), the n-th order effect on an arbitrary industry j is then given by

$$\frac{\Delta S_{j,t}^n}{S_{j,t-1}} = A_{ij}^n\mu\omega^{-n}\frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \tag{1.9}$$

Equation (1.9) and $\omega > 1$ imply that the correlation between the sales growth rates of different industries has to be positive and decreasing in the network distance between industries, and this is exactly what Carvalho (2014) finds for his dataset of US industries, lending empirical support to the plausibility of this requirement. To get the total effect from the shock on i to an industry j from all orders, we sum over n from zero to infinity

$$\frac{\Delta S_{j,t}}{S_{j,t-1}} = \mu\sum_{n=0}^{\infty} A_{ij}^n\omega^{-n}\frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}, \tag{1.10}$$

and $\omega > 1$ is a sufficient condition for this sum to converge. In the next step, consider the total effect on all N industries j

$$\sum_{j=1}^{N}\frac{\Delta S_{j,t}}{S_{j,t-1}} = \mu\sum_{j=1}^{N}\sum_{n=0}^{\infty} A_{ij}^n\omega^{-n}\frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \tag{1.11}$$

The expression in (1.11) closely resembles the Katz centrality vectors frequently employed in the network literature to measure the centrality of nodes when higher-order effects are relevant.

To see this, consider the vector of Katz centrality scores k for the graph G with a binary but not necessarily symmetric adjacency matrix A depending on some attenuation parameter $\beta > 1$ and given by

$$k(\beta) = ((\mathcal{I}_{N\times N} - \beta A)^{-1} - \mathcal{I}_{N\times N})\mathcal{I}_{N\times 1} \tag{1.12}$$

$$= ((\mathcal{I}_{N\times N} + \beta^{-1}A + \beta^{-2}A^2 + ...) - \mathcal{I}_{N\times N})\mathcal{I}_{N\times 1} \tag{1.13}$$

$$= \left(\sum_{j=1}^{N}\sum_{n=0}^{\infty} \beta^{-n}A_{ij}^n - 1\right)_{i=1}^{N}, \tag{1.14}$$

where $\mathcal{I}_{N\times N}$ is the identity matrix with dimension $N \times N$ and $\mathcal{I}_{N\times 1}$ is a unit vector with length N. Bonacich (1991) shows that if β approaches the dominant (Perron) eigenvalue λ_{max} from above, the vector of Katz centrality scores approaches the vector of eigenvector centrality scores denoted e. For $\beta = \omega$ we have from equation (1.14) that $k(\omega) = (\sum_{j=1}^{N}\sum_{n=0}^{\infty} A_{ij}^n\omega^{-n} - 1)_{i=1}^{N}$, so this term corresponds to the vector of Katz centrality scores. Assuming that $\omega = \lambda_{max}$, that is the coefficient ω corresponds to the Perron eigenvalue, this implies that $k(\lambda_{max}) = e$, so the Katz centrality vector is equal to the vector of eigenvector centrality scores, and we have $\mathcal{I}_{N\times 1} + e = (\sum_{j=1}^{N}\sum_{n=0}^{\infty} A_{ij}^n\omega^{-n})_{i=1}^{N}$.

Let $\theta_{i,t} = 1 + e_{i,t}$ denote the *shock transmission vector*, capturing both the initial impact of the productivity shock as well as all meso level propagations between t and $t-1$. It follows that

$$\sum_{j=1}^{N} \frac{\Delta S_{j,t}}{S_{j,t-1}} = \mu\theta_{i,t}\frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \tag{1.15}$$

Finally, to allow for simultaneous shocks apart from the singular one considered so far, we sum the right-hand side over i

$$\sum_{j=1}^{N} \frac{\Delta S_{j,t}}{S_{j,t-1}} = \mu\sum_{i=1}^{N}\theta_{i,t}\frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \tag{1.16}$$

Since we consider the sum over a vector, $\sum_{j=1}^{N}\frac{\Delta S_{j,t}}{S_{j,t-1}} = \sum_{i=1}^{N}\frac{\Delta S_{i,t}}{S_{i,t-1}}$, we can swap indices to obtain

$$\sum_{i=1}^{N} \frac{\Delta S_{i,t}}{S_{i,t-1}} = \mu\sum_{i=1}^{N}\theta_{i,t}\frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \tag{1.17}$$

By imposing the restriction $\omega = \lambda_{max}$, the total rate of change resulting from the sum of arbitrary order-effects between t and $t-1$ will correspond to the centrality weighted sum of all productivity shocks. Hence we have characterized the end-of-period state of all industries in the economy, which completes the meso level aggregation. To close the system and to aggregate to the macro level, we recall from our macro assumption that

$$\phi \sum_{i=1}^{N} \Delta S_{i,t} = \Delta Y_t, \tag{1.18}$$

where ΔY_t denotes the change in GDP between $t-1$ and t, and the input-output multiplier $\phi < 1$ indicates the presence of intermediate goods. From equations (1.17) and (1.18) we have

$$\Delta Y_t = \phi \sum_{i=1}^{N} \Delta S_{i,t} = \phi \sum_{i=1}^{N} \frac{\Delta S_{i,t}}{S_{i,t-1}} S_{i,t-1} = \phi\mu \sum_{i=1}^{N} \theta_{i,t} \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}} S_{i,t-1}, \tag{1.19}$$

and therefore

$$\frac{\Delta Y_t}{Y_{t-1}} = \phi\mu \sum_{i=1}^{N} \theta_{i,t} \frac{S_{i,t-1}}{Y_{t-1}} \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \tag{1.20}$$

Equation (1.20) is our central result and relates the macro state of the economy to micro productivity shocks through the mesoeconomic interaction of network structure and Domar weights. We obtain equation (1.20) under the crucial assumption of micro substitutability ($\omega > 1$) and the technical requirement that the parameter ω is equal to the Perron eigenvalue λ_{max}. While the former is theoretically intuitive if we want to exclude explosive growth, the latter is certainly plausible from an empirical point of view as the Perron eigenvalue is remarkably stable over time and close to twenty, implying that about five percent of productivity shocks are transmitted to suppliers.

Notice that the Riemann sum approximation (1.4) is a special case of our model specification. We can interpret this case economically as the correct specification for an islands economy (not merely as a discrete approximation with desirable mathematical properties) because in an islands economy all higher-order terms from network interaction are identically zero so $\theta_{i,t} = (1 + e_i) = 1$ for all $i = 1, \ldots, N$. Since there is no intermediate consumption in an islands economy we have $\phi = 1$, leaving us with the formulation used by Gabaix (2011) and Baqaee and Farhi (2019),

$$\frac{\Delta Y_t}{Y_{t-1}} = \mu \sum_{i=1}^{N} \frac{S_{i,t-1}}{Y_{t-1}} \cdot \frac{TFP_{i,t}}{TFP_{i,t-1}}. \tag{1.21}$$

The invariance of sales shares is thus essentially a network irrelevance result. This alternative derivation is in line with the theoretical result in Baqaee and Farhi (2019) that the invariance of sales shares is a linearity result, since the network topology is the only possible source of nonlinearity we consider. We include this purely granular specification as a robustness check in the empirical part. Given the modular outline of the model, the purely granular case corresponds to a direct aggregation from the micro to the macro level without any meso level interaction. Another polar case, also included for empirical robustness, is the purely networked case where all industries are of equal (or representative) size, that is $S_{i,t-1}/Y_{t-1} = 1/N$ for all $i = 1, \ldots, N$ whereby equation (1.20) reduces to

$$\frac{\Delta Y_t}{Y_{t-1}} = \frac{\mu\phi}{N} \cdot \sum_{i=1}^{N} \theta_{i,t} \cdot \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \tag{1.22}$$

This formulation corresponds to a direct aggregation from the meso to the macro level such that productivity shocks are directly transmitted down the supply-chain without being impacted by the distribution of sales shares. Hence the modular structure of our model allows for testing the relevance of each aggregation step in addition to testing the performance of the generalized formulation with mesoeconomic interactions.

1.3 Data and Estimation

We use the second release of the World Input-Output Database (2017), henceforth WIOD, that presently covers the period from 2000 to 2014. A detailed breakdown of the WIOD and its various data sources, for the most part based on national accounts, is provided by Timmer et al. (2015). Macroeconomic data for GDP and GDP per capita series as well as several other macro variables used for robustness checks are taken from the World Bank's Development Indicators database (World Bank, 2018). We focus on the EU28, mainly to exploit the large degree of cross-country connectivity in an integrated free trade area with a large ensemble of industries. While the sample period is rather short, it contains at least two important crises periods (the global financial and economic crisis and the European sovereign debt crisis) with substantial variation in GDP growth rates. Thus the sample period provides a reasonably challenging testing ground for the empirical modeling of aggregate fluctuations in general, and it should also help in discriminating between the different cases we have derived in the previous section.

1.3.1 Data

The WIOD provides annual series for 56 industries per country, leaving us with 1568 distinct industries in the EU28. The size-based metrics are taken from the Socio-Economic

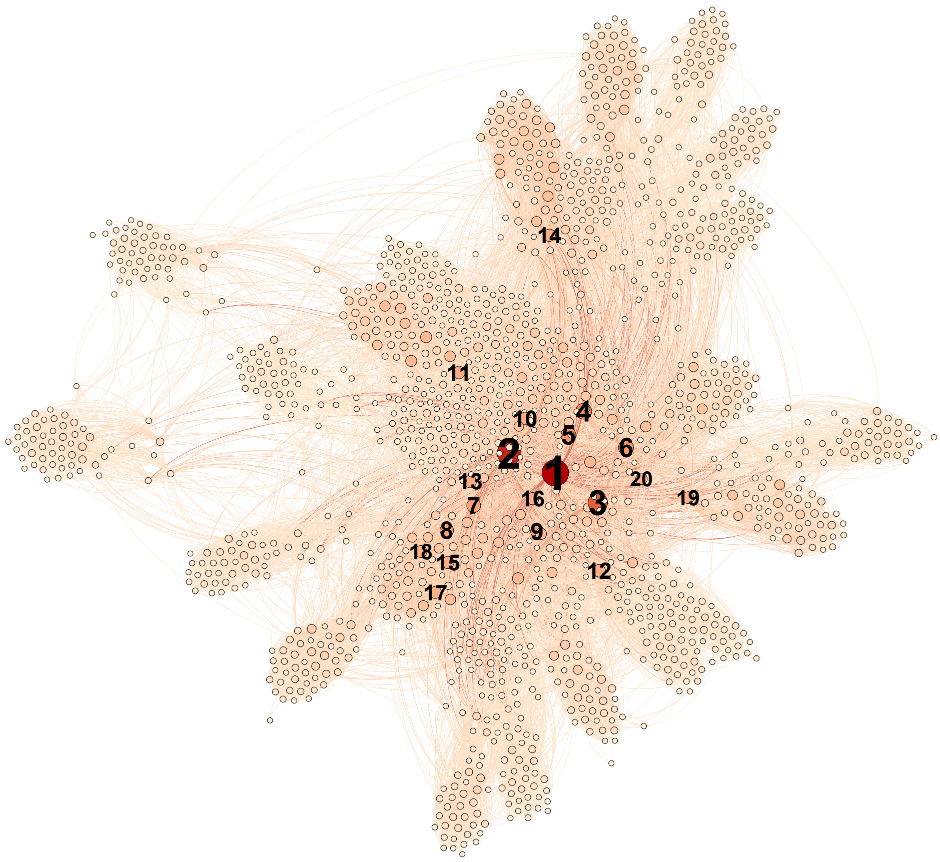
Accounts (SEA), in particular the annual volume of industry sales (gross output) and the number of employees used to calculate Domar weights and productivity shocks, respectively. The complementary cumulative distribution functions (CDFs) of industry sales, reported for each sample year in Appendix 1.A, reveal a large degree of heterogeneity in industry size, but they do not exhibit power law-like behavior. Thus aggregation from the firm to the industry level apparently prevents the straightforward application of "pure" granular theory that has so far been entirely based on the power law property of firm size distributions.

In addition to size-based metrics from the SEA, our model also requires the specification of adjacency matrices in order to account for the dynamic adjustment of sales shares through supplier relations. To construct these matrices, we binarize the World Input-Output Tables (WIOTs) from the WIOD that describe inter-industry flows for all industry pairs, leading to a 1568×1568 matrix for each year. Following Carvalho (2014), we apply a standard threshold of one percent of an industry's total input purchases for binarization. This threshold preserves about eighty percent of non-zero links from the weighted adjacency matrix in all years. The resulting networks are extremely sparse and visualized in Figure 1.1 for the year 2014, where we observe $25\,825$ non-zero edges out of a possible $1568^2 \approx 2.5 \times 10^6$, resulting in a low network density of around 0.01, with an average degree of 17 in that year. The size of each node in the figure is proportional to its centrality score,[7] and the network looks very similar in other years as well. After all, the binarization threshold does not artificially inflate connectivity and should consequently not have a spurious impact on aggregate fluctuations. Yet mesoeconomic aggregation emphasizes the importance of higher-order effects in the production network, and the figure shows that all clusters are still connected to each other via some intermediaries, corroborating the importance of shock transmission through higher-order connections in spite of the network's low density and average degree. The strongest propagators of productivity shocks are located in the German manufacturing sector, in line with the intuition that industries producing highly complex final products rely on long and diversified supply chains.

1.3.2 Estimation

To quantify the impact of industry level shocks on aggregate fluctuations we essentially adhere to the procedure of Gabaix (2011). First we construct *residuals* that represent the sum of idiosyncratic shocks to the $K < N$ largest industries, here weighted ac-

[7]In the empirical part the transmission vector corresponds to eigenvector centralities, which are much more efficiently computed by a standard power iteration method than Katz centralities. This does not bias the estimates in our dataset because the initial impact of productivity shocks rapidly vanishes in relation to the higher-order terms in equation (1.11), even for a finite-sum approximation.



1	German manufacture of motor vehicles, trailers and semi-trailers	11	British human health and social work activities
2	German manufacture of machinery and equipment n.e.c.	12	Czech manufacture of motor vehicles, trailers and semi-trailers
3	German construction	13	German manufacture of basic materials
4	French construction	14	Swedish manufacture of motor vehicles, trailers and semi-trailers
5	German manufacture of chemicals and chemical products	15	Italian manufacture of fabricated metal products, excl. machinery and equipment
6	German manufacture of food products, beverages and tobacco products	16	German electricity, gas, steam and air conditioning supply
7	German manufacture of electrical equipment	17	Italian construction
8	Italian manufacture of machinery and equipment n.e.c.	18	Italian wholesale trade, excl. motor vehicles and motorcycles
9	German manufacture of fabricated metal products, excl. machinery and equipment	19	Spanish manufacture of motor vehicles, trailers and semi-trailers
10	French manufacture of motor vehicles, trailers and semi-trailers	20	German manufacture of rubber and plastic products

Figure 1.1: The EU28 production network exhibits substantial heterogeneity in centrality scores, along with clusters that mostly coincide with national borders. The twenty most central industries are listed in descending order and node sizes are proportional to centrality scores.

cording to the theoretically motivated specifications in equations (1.20) – (1.22).[8] Then we run linear regressions on the various residuals and consider the coefficients of determination, in other words their respective (adjusted) R^2, in order to compare how much of the variation in GDP growth can be explained by each of the three model specifications. To do so, we need to measure productivity shocks and here we will use labor productivity instead of TFP.[9] Let $g_{i,t} - \bar{g}_t$ denote the productivity shock to industry i at time t, where $g_{i,t} = \ln(S_{i,t}/E_{i,t}) - \ln(S_{i,t-1}/E_{i,t-1})$ is the observed change in i's labor productivity, with $E_{i,t}$ being the number of employees in industry i at time t, while $\bar{g}_t$ measures the aggregate productivity shock. This conventional formulation assumes that the observed productivity signal $g_{i,t}$ is additively separable into an idiosyncratic and an aggregate component, but the specification of the aggregate shock $\bar{g}_t$ is not trivial, so we provide a detailed and rather careful discussion of its construction in Appendix 1.A.

Now we are in a position to compute residuals for our respective model specifications, starting with the *mesoeconomic residual* $\mathcal{M}$ that aggregates the idiosyncratic productivity shocks to the K largest industries $i = 1, \ldots, K$ in period t using the mesoeconomic weight $\theta_{i,t} \cdot S_{i,t-1}/Y_{t-1}$,

$$\mathcal{M}_t = \sum_{i=1}^{K} \theta_{i,t} \frac{S_{i,t-1}}{Y_{t-1}} (g_{i,t} - \bar{g}_t). \tag{1.23}$$

For robustness we also consider the case of purely linear shock responses in an islands economy with the *granular residual* $\mathcal{G}$, using merely Domar weights

$$\mathcal{G}_t = \sum_{i=1}^{K} \frac{S_{i,t-1}}{Y_{t-1}} (g_{i,t} - \bar{g}_t), \tag{1.24}$$

and the *network residual* $\mathcal{N}$ that aggregates directly from the meso level by ignoring the size differences between industries, using merely the transmission vector θ

$$\mathcal{N}_t = \sum_{i=1}^{K} \theta_{i,t} \, (g_{i,t} - \bar{g}_t). \tag{1.25}$$

[8]The granular case studied by Gabaix corresponds to eq. (1.21), but here for industries instead of firms.

[9]Our model is built on TFP shocks that are Hicks-neutral, while labor productivity shocks generally are not. This is at least one of the reasons why the growth of TFP and labor productivity follow different trajectories, as shown for instance by Syverson (2004) for the manufacturing industry on the plant level. Nevertheless we employ labor productivity for three reasons: from a theoretical viewpoint, the estimation of TFP presupposes precisely the CRS assumption we challenge with our model; from an empirical viewpoint, we simply do not have TFP data in the WIOD; finally, the use of labor productivities facilitates the comparison with the granular results of Gabaix, who also theorizes with TFP but then utilizes labor productivity for his empirical exercise.

With this information we can translate the mesoeconomic model and its two polar variants into the following regressions, labeled $I - III$ for easier reference, in order to investigate their coefficients of determination

$$I: \ (a) \ g_t^Y = \beta_0 + \beta_1 \mathcal{M}_t + \epsilon_t \qquad (b) \ g_t^Y = \beta_0 + \beta_1 \mathcal{M}_t + \beta_2 \mathcal{M}_{t-1} + \epsilon_t \tag{1.26}$$

$$II: \ (a) \ g_t^Y = \beta_0 + \beta_1 \mathcal{G}_t + \epsilon_t \qquad (b) \ g_t^Y = \beta_0 + \beta_1 \mathcal{G}_t + \beta_2 \mathcal{G}_{t-1} + \epsilon_t \tag{1.27}$$

$$III: \ (a) \ g_t^Y = \beta_0 + \beta_1 \mathcal{N}_t + \epsilon_t \qquad (b) \ g_t^Y = \beta_0 + \beta_1 \mathcal{N}_t + \beta_2 \mathcal{N}_{t-1} + \epsilon_t \tag{1.28}$$

where aggregate fluctuations are denoted g_t^Y and measured by the growth rates of (per capita real) GDP. The left panel, marked (a), corresponds to equations (1.20) – (1.22) while the right panel, marked (b), includes a lagged term of the respective residual to account for the possibility that the adjustment of sales shares takes time and GDP thus responds to micro productivity shocks in a delayed fashion, also addressing concerns of reverse causality.

Our primary tool of investigation are plots that show how much of the explained variation R^2 in GDP growth rates can be attributed to the largest K industries.[10] To establish a benchmark we show in Appendix 1.A that the expected R^2 in the presence of idiosyncratic shocks, denoted $E[R^2]$, for K industries with weights $w \in \{\theta\, S/Y, S/Y, \theta\}$ is given by the ratio of Hirschman-Herfindahl indices H for the weights of these K industries relative to the weights for all N industries,

$$E[R^2] = \frac{H(K)}{H(N)} = \sum_{i=1}^{K} \frac{(w_i)^2}{\sum_{j=1}^{N} (w_j)^2}. \tag{1.29}$$

Since the three considered weights exhibit considerable leptokurtosis we should observe a large initial increase in R^2 already for a small number of industries, followed by an almost monotonous yet very moderate increase for the remaining range of K, approaching a plateau while K approaches its upper bound because the size of industry weights then decreases rapidly. These three qualitative features will guide our comparison and assessment of the different specifications in the next section.

Notably, and in stark contrast to conventional granular approaches (Gabaix, 2011; Ebeke and Eklou, 2017; Blanco-Arroyo et al., 2018), we use only raw WIOD data in all our specifications and do not winsorize the empirical density of productivity growth rates.[11] In Appendix 1.A, we show that such winsorizing should actually *decrease* rather

[10] This approach is inspired by Blanco-Arroyo et al. (2018). While they are investigating different purely granular "regimes", we will compare the empirical R^2 plots to a benchmark that would be expected for the correct identification of both idiosyncratic shocks and weights.

[11] Winsorizing is typically justified by reference to extraordinary events such as large mergers. Yet

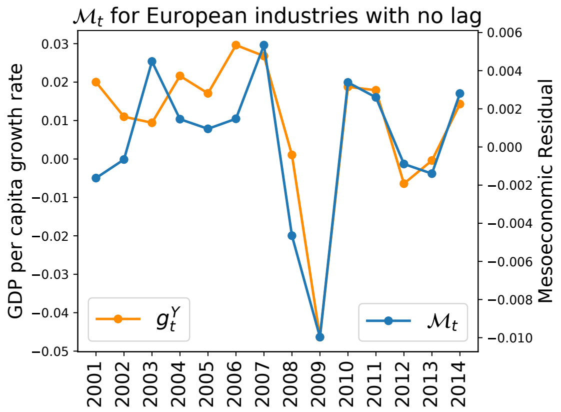
than increase explanatory power for a correct model specification, essentially due to a decrease in residual variance, and this effect is amplified by the well-documented leptokurtic nature of empirical productivity growth rate densities (Dosi et al., 2012; Yu et al., 2015; Dosi et al., 2019).[12] Explanatory power thus only increases when the winsorizing procedure generates spurious comovement between industries, inflating regression coefficients by compensating for the initial drop in variance and explanatory power. Viewed from this perspective, the fact that a purely granular approach needs to generate comovement through winsorizing to push the estimates over the brink of significance in itself indicates the crucial importance of hitherto neglected network interactions.

1.4 Results

The mesoeconomic model specification predicts a linear relationship of the meso residual and the empirical growth rates of GDP which we test by a linear regression. To assess the empirical performance of the full mesoeconomic specification, we estimate the regression equations in (1.26) I (a) and the variant with one lagged term I (b). The table on the right panel of Figure 1.2 shows the OLS estimates for both I (a) and I (b), evaluated at the snapshot of K^*, estimated for the time series shown on the left panel. The choice of K^* is the result of careful calibration that is detailed in Appendix 1.A, in contrast to the usual practice within the granular literature with rather arbitrary K^*. However, we also find our estimates to be robust for various different K. We find strong support for the mesoeconomic hypothesis with the contemporaneous coefficient being highly significant at the 1% level in both specifications. The explanatory power of the mesoeconomic specification is sizeable, with about 69% for the purely contemporaneous case and almost 80% for the specification with an additional lagged term, reflecting the fact that the time series of the meso residual $\mathcal{M}$ tracks the trajectory of empirical GDP growth rates remarkably closely. The coefficient of the lagged term for regression equation I (b) is insignificant and almost one order of magnitude smaller than the contemporaneous one. In light of the mesoeconomic hypothesis, this indicates that the vast majority of adjustments of sales shares is picked up by the contemporaneous coefficient and occurs within the interval between yearly observations. The same is true regarding the estimated intercepts: Even though significant, their size seems to

Bottazzi et al. (2019) show in their appendix that for each of the idiosyncratic events causing the four minimum and maximum growth rates in the Gabaix (2011) sample the notion of "extraordinary events" can be reasonably challenged.

[12]We show that symmetric winsorizing at merely one percent diminishes the explanatory power by more than five percent when productivity growth rates follow (as the empirical literature strongly suggests) a Laplace distribution. Conversely, the loss in explanatory power is considerably smaller for a Gaussian distribution at about 1.8 percent. We discuss the details of these computations in Appendices 1.A and 1.A.



	I	
	(a)	(b)
$\mathcal{M}_t$	0.3975***	0.4120***
	(0.0775)	(0.0687)
$\mathcal{M}_{t-1}$		0.0963
		(0.0692)
Intercept	0.0009**	0.0007**
	(0.0002)	(0.0030)
T	14	13
R^2	0.687	0.794
Adjusted R^2	0.660	0.753
F Statistic	26.29***	19.25***

Figure 1.2: Both visual inspection and the estimated regression equations indicate that the residual time series is a good predictor of GDP growth rates. The left panel shows the time series of the mesoeconomic residual $\mathcal{M}$ against g_t^Y (GDP per capita growth rate) at $K^* = 224$ between 2000 and 2014. The right panel reports the results for the time series of $\mathcal{M}$ being regressed on g_t^Y with the purely contemporaneous specification in I (a) and including one lagged term in I (b). Standard errors are given in parentheses. Significance at the *** 1%, ** 5%, * 10% level.

be economically negligible and estimated at the order of 10^{-4}. The model performance is therefore apparently driven by movements in the contemporaneous meso residual and decidedly not a general trend in the GDP series the intercepts might pick up.

The results shown in Figure 1.2 indicate strong support for the mesoeconomic hypothesis. The snapshot results for the regression equations evaluated at a single K^* might, however, be misleading when the explanatory power is very peaked in the neighborhood of this particular value and thus sensitive to incremental variation in K. Both panels of Figure 1.3 demonstrate that this fear is unsubstantiated and that R^2 is exceptionally stable for K close to K^*. Apart from sensitivity analyses, the functional form of R^2 for both specifications is perfectly in line with the analytical predictions lending further support to the mesoeconomic hypothesis: For both specifications, the estimated R^2 exhibits all three predicted qualitative features of i) a large initial increase for the industries with very large meso weights that gets ii) more gradual with higher K and iii) diminishes on a plateau for K approaching its maximum. Most notably, Figure 1.3 reveals that a mere 2% of industries is able to explain almost 80 % of variation in GDP growth rates. Such further decomposition allows therefore to identify a much smaller subset of about 20 to 30 industries of systemic relevance than the snapshot evaluated at the maximum number of considered industries might suggest.

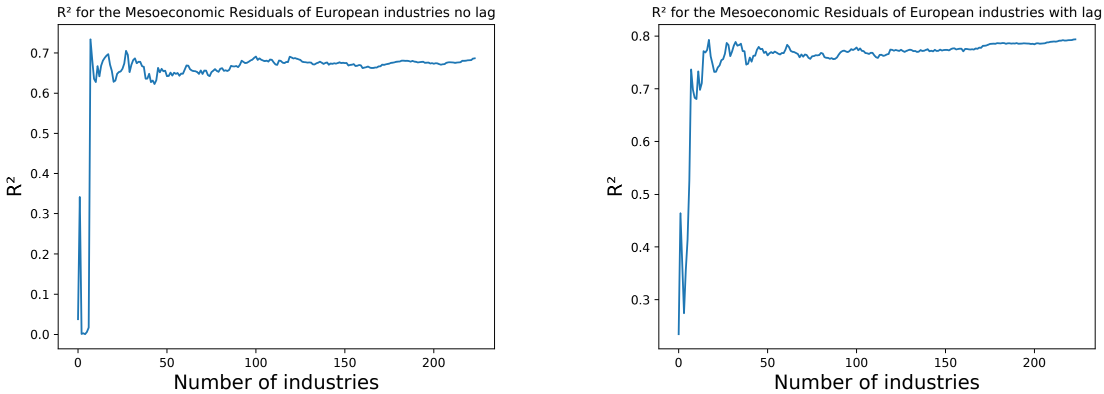
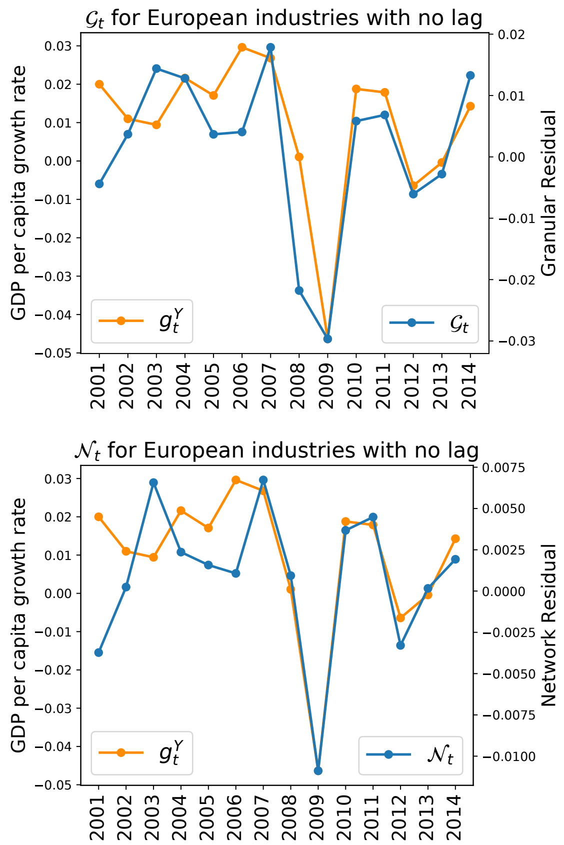


Figure 1.3: The incremental contribution of industries to explained variation of GDP varies substantially with the number K of considered industries with the largest meso weights, where almost all variation is already explained by the largest 30 industries. The left panel shows the R^2 for the purely contemporaneous regression specification I (a) with the mesoeconomic residuals calculated for an increasing number of K largest industries, while the R^2 plot on the right panel is based on specification I (b) with an additional lagged term.

1.4.1 Robustness: Limit Cases

For further robustness, we repeat the estimation exercise for the two special cases of the mesoeconomic residual in exactly the same way to isolate the effects of the size- and network-based channel. Figure 1.4 shows the time series for the granular and network residuals evaluated at K^*, with the right panels summarizing the OLS estimates for regression equations II (a) to III (b) computed for these series. For the two special cases, the contemporaneous coefficients are highly significant, indicating that both the size and network-based channels are of importance within the meso specification. Across specifications, however, the meso residual is outperforming the granular and network special cases with respect to explanatory power, lending robust support for the crucial relevance of the interaction of both channels implied by the mesoeconomic hypothesis. The explanatory power for the two special cases might even be an overestimate compared to the full meso model, since the estimated intercepts are consistently one order of magnitude larger than for the meso specification. Therefore, a higher fraction of model performance is attributable to the intercept picking up a general trend in the macro series and not to comovement with the residual series. Visual inspection of both time series is quite instructive with regards to the possible reasons for the underperformance of a particular special case. The granular residual seems to outperform the pure network specification in times of rather low aggregate volatility at the first half of the series, while the network residual performs relatively better for the other half with much larger fluctuations. This is in line with the intuitive notion that large fluctuations are due to cascading failures for which only the network residual can properly account. This is most clearly visible in the outstanding performance of the network

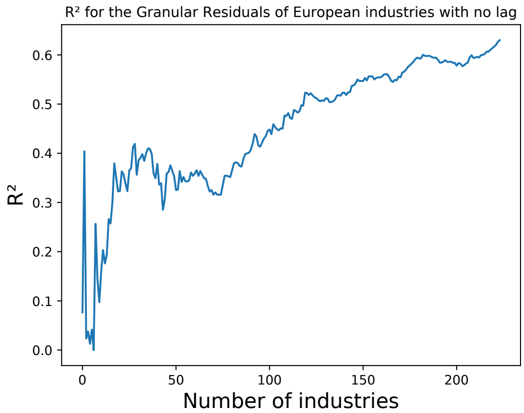


	II	
	(a)	(b)
$\mathcal{G}_t$	1.118***	1.104***
	(0.247)	(0.230)
$\mathcal{G}_{t-1}$		0.329
		(0.237)
Intercept	0.008**	0.006*
	(0.003)	(0.003)
T	14	13
R^2	0.630	0.738
Adjusted R^2	0.599	0.685
F Statistic	20.43***	14.05***
	III	
	(a)	(b)
$\mathcal{N}_t$	3.231***	3.744***
	(0.773)	(0.696)
$\mathcal{N}_{t-1}$		0.101
		(0.668)
Intercept	0.006*	0.004
	(0.003)	(0.003)
T	14	13
R^2	0.593	0.744
Adjusted R^2	0.559	0.693
F Statistic	17.47***	14.54***

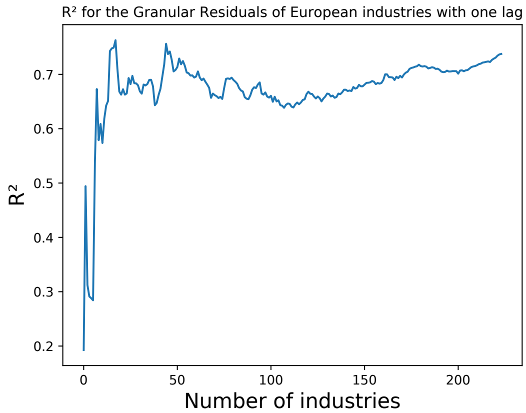
Figure 1.4: The empirical performance of the two special cases of the mesoeconomic specification is considerably worse than for the general formulation across all variants. The left panel shows the time series of the network and granular residual against g_t^Y (GDP per capita growth rate) between 2000 and 2014. The right panel reports the results for both time series being regressed on g_t^Y with the purely contemporaneous specifications II (a) and III (a) and including one lagged term in II (b) and III (b). Standard errors are given in parentheses. Significance at the *** 1%, ** 5%, * 10% level.

residual around the global financial crisis as the epitome of an event driven by systemic risk. In the regime of moderate volatility in the first half of the series, shock propagation might be of lesser importance and the failure of the network residual to take size differences properly into account becomes more apparent when contrasted to the granular variant that is designed around size heterogeneity for an island economy.

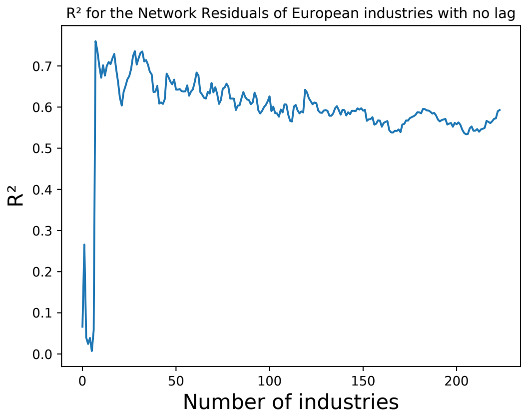
In Appendix 1.A, we show that for a correctly specified model, the estimated $\hat{\beta}_1$ regression coefficient is equivalent to the proportionality constant pertaining to it, that is, μ for the granular and $(\mu\phi)/N$ for the network case. Interpreted in this way, the estimated coefficients suggest that the granular model overestimates the impact of idiosyncratic shocks, while the pure network residual underestimates them. Since $\hat{\beta}_1$ is equivalent to μ for the granular variant, the estimated $\hat{\beta}_1$ close to unity would suggest that sales shares react almost one-to-one to productivity shocks. The implied sensi-



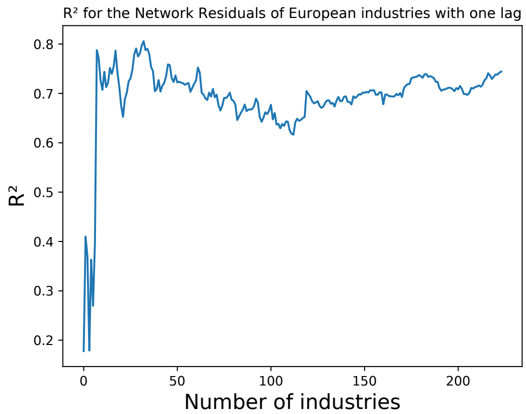
(a) Specification II (a)



(b) Specification II (b)



(c) Specification III (a)



(d) Specification III (b)

Figure 1.5: All specifications for the limit cases fail to produce at least one predicted property, when the achieved explanatory power is plotted against the number of industries with highest weight K. The upper panels show the results for the granular specifications II (a) and (b) and the lower ones for the pure network case in variants III (a) and (b).

tivity of sales shares is therefore implausibly high and exceeds even the one for the limit case of a long-run neoclassical growth model with $\mu \approx 1.5$ (Gabaix, 2011). Intuitively, the impact of direct productivity shocks is overestimated, because the initial impact of a shock has to account for the whole variation in sales shares for an islands economy without any other shock being transmitted from supplier industries. In contrast to that, the interpretation of $\hat{\beta}_1$ for the network residual as $(\mu\phi)/N$ suggests that its OLS estimates are heavily inflated and, consequently, the explanatory power is upwards biased.[13] Note that the proportionality constant implied by the full meso model is $\mu \cdot \phi$ and therefore equivalent to the constant of the network model up to a factor $1/N$ which is at the order of 10^{-3}. In fact, $\hat{\beta}_1$ is even one order of magnitude larger for the network specification and therefore overestimated with respect to the full meso model by four orders of magnitude. The downward biased estimate for the impact of productivity shocks results from the counterfactual assumption of homogeneity in size. For all considered industries in all years, the imposed homogeneous Domar weight of $1/N$ is an underestimate of their actual size.[14] Since a productivity shock to an industry does not affect the industry's suppliers directly but is mediated by the adjustment of its sales shares, underestimating initial sales shares leads to an underestimation of the initial shock impact which then cumulates at all orders of transmission. These back-of-the-envelope calculations thus suggest that the supposed empirical success of both residual variants might be partially caused by distorted $\hat{\beta}_1$ estimates contradicting considerations of economic plausibility.

Even with limit cases potentially favored by distorted $\hat{\beta}_1$ estimates, the full meso model apparently outperforms them for the whole considered range for $K = 1, \ldots, K^*$.[15] For all variants with regression equations II (a) to III (b), the empirical R^2 plots exhibit at least one feature that contradicts the analytical predictions, as Figure 1.5 shows. The purely contemporaneous granular specification displayed in the upper left panel shows much too strong oscillations with the increasing trend being also much too gradual to correspond to the extremely heterogeneous distribution of Domar weights. Its variant II (b) with an additional lagged term is probably closest to the ideal functional form out of all considered variants but still exhibits too much volatility for K approach-

[13] For algorithmic stability, the power iteration method employed to estimate transmission centralities normalizes scores in every iteration, leaving us with scores that correspond to the true Katz centralities only up to a common factor. In contrast to the purely granular case, we therefore cannot infer μ in absolute terms here but are able to compare the coefficients for the full meso model and the pure network limit case, as both use the same equally normalized centralities.

[14] This is due to the fact that Domar weights are moderately correlated with centrality scores ($r \approx 0.5$) and industries with high meso weight therefore simultaneously tend to exhibit large centrality scores and Domar weights.

[15] Notice that we take the "composite K largest industries", that is e.g., for $K = 10$, we take the 10 largest industries in each given year separately, whose identities might change over time.

ing its maximum, where diminishing Domar weights would imply that the plot should flatten out. The plots for the network residuals in both variants at the bottom panels show a large initial increase in line with theory but also exhibit lengthy intervals of K with diminishing R^2 which contradict theoretical predictions: Intuitively, taking the idiosyncratic trajectories of a higher number of subunits of a given aggregate system into account should improve predictions for the system, not worsen them. While the rather high explanatory power might thus show both residuals to capture important features of the true generating process that links micro productivity shocks to macro outcomes, both the snapshot estimates at K^* as well as the R^2 plots indicate them to be misspecified in at least some regards. Indeed, the time series point to the granular residual being more applicable for regimes of rather low aggregate volatility but fail to for high amplitudes which is exactly where the network residual apparently shines. The mesoeconomic residual robustly outperforms its granular and network limit cases in terms of explanatory power, plausibility and consistency with analytical predictions, as its application is not limited to a particular volatility regime.

1.4.2 Robustness: Macro Controls

As a second set of robustness checks, we include several controls to examine the possibility of omitted confounding variables causally influencing both the productivity patterns on the micro level and the GDP growth patterns on the macro level, causing the observed comovement between both. Candidates for such confounding variables should be conceptually distinct from the mesoeconomic residual and thus pertain to the state of the aggregate system. These include e.g. the rate of change in the oil price, as oil is close to a universal input, the ECB policy rate, affecting lending conditions independent of industry and trade with non-EU countries that are per construction not captured in our underlying input-output tables. In particular, we consider three groups of controls for supply-side, demand-side and monetary-policy shocks, namely:

1. Demand-side shocks: growth rate of total investment (denoted $investment$, as percentage of GDP, EU28); growth rate of public expenditure (denoted $pubexp$, as percentage of GDP, EU28)

2. Supply-side shocks: growth rate of the real crude oil price (denoted oil, Brent Europe); growth rate of real trade activity (denoted $trade$, measured as the sum of total imports and exports, Euro area)

3. Monetary-policy shocks: ECB policy rate (denoted $ECBpolicyrate$, Euro area)

To test for the impact of these variable groups, we include them in our baseline linear regression specification I (a). Since we cannot rule out the possibility that the specific choice of variable groups matters, we include every possible permutation of variable groups in specifications (A) to (G) to test against the benchmark model.

$$
\begin{aligned}
(A):\ g_t^Y &= \beta_0 + \beta_1 \mathcal{M}_t + \beta_2 ECBpolicyrate_t + \epsilon_t \\
(B):\ g_t^Y &= \beta_0 + \beta_1 \mathcal{M}_t + \beta_2 investment_t + pubexp_t + \epsilon_t \\
(C):\ g_t^Y &= \beta_0 + \beta_1 \mathcal{M}_t + \beta_2 oil_t + \beta_3 trade_t + \epsilon_t \\
(D):\ g_t^Y &= \beta_0 + \beta_1 \mathcal{M}_t + \beta_2 oil_t + \beta_3 trade_t + \beta_4 ECBpolicyrate_t + \epsilon_t \\
(E):\ g_t^Y &= \beta_0 + \beta_1 \mathcal{M}_t + \beta_2 investment_t + \beta_3 pubexp_t + \beta_4 ECBpolicyrate_t + \epsilon_t \\
(F):\ g_t^Y &= \beta_0 + \beta_1 \mathcal{M}_t + \beta_2 investment_t + \beta_3 pubexp_t + \beta_4 oil_t + \beta_5 trade_t + \epsilon_t \\
(G):\ g_t^Y &= \beta_0 + \beta_1 \mathcal{M}_t + \beta_2 oil_t + \beta_3 trade_t + \beta_4 investment_t + \beta_5 pubexp_t + \beta_6 ECBpolicyrate_t + \epsilon_t
\end{aligned}
$$

Specification	(A)	(B)	(C)	(D)	(E)	(F)	(G)	Benchmark
Controls	M	D	S	S, M	D, M	D, S	S,D,M	-
$\hat{\beta}_1$ Significance	$***$	$***$	$***$	$***$	$***$	$***$	$***$	$***$
AIC	-87.60	-81.95	-82.03	-84.03	-84.67	-78.78	-81.40	-84.38
BIC	-85.68	-79.40	-79.48	-80.84	-81.47	-74.95	-76.93	-83.10

Table 1.1: The mesoeconomic specification is robustly significant after the inclusion of a plethora of different macro controls. The AIC and BIC are calculated for regression specifications (A) to (G) with the mesoeconomic residual $\mathcal{M}$ for $Q = K^* = 224$. M refers to ECB policy rate, D to the group of demand-side shocks and S to the group of supply-side shocks.

Each specification (A) to (G) is estimated by OLS. The results for these estimations are summarized in Table 1.1. We report the significance level of the estimated $\hat{\beta}_1$ coefficient as well as the associated Akaike and Bayesian Information Criteria (AIC and BIC) for a consistent comparison of the specifications with a different number of regressors (Akaike, 1973; Schwarz, 1978).[16] The $\hat{\beta}_1$ coefficient stays highly significant for all different groups of controls we consider. This testifies to the robustness of our benchmark model and is evidence that there are no confounding third variables which cause idiosyncratic productivity and GDP to co-move.

In terms of penalized gains in explanatory power, the inclusion of supply-side effects with variable group S does not seem to increase performance against the benchmark model, as both measures show for (C), (D), (F) and (G). This seems to imply that the mesoeconomic specification represents supply-side developments rather well, with even inclusion of the canonical oil price shocks unable to increase performance (Summers, 1986). Since our model with its initial RBC assumption is essentially a supply-

[16] As Kuha (2004) notes, both criteria select for different types of data-generating processes. The BIC is more suitable for processes with a small number of strong effects, while the AIC selects better for models with a collection of many effects with rather small effect sizes. Model selection without such prior information should therefore be based on both criteria in agreement.

side model, this is arguably to be expected but even more so evidence for the absence of omitted supply-side variables in the regression specification. Intriguingly, however, performance apparently also does not increase when including demand-side shocks, as specifications (B), (F) and (G) show with ambiguous results for (E), even though we are not targeting demand-side developments in our theoretical model at all. The robustness checks therefore seem to leave little room for those traditional real macro shocks, when microeconomic productivity shocks are properly accounted for by the mesoeconomic residual.

Only the specification in which we include a simple proxy for monetary policy (A) would be preferred to our benchmark model by both AIC and BIC.[17] This result, however, is not surprising since the mesoeconomic residual is build on the assumption of RBC models where monetary policy is ineffective in generating real effects (Snowdon et al., 1994). As, for example, Galí (2008) shows, however, transmission of interest rate shocks through the amount of available credit, inflation or asset prices have potentially strong effects on GDP fluctuations under price-stickiness. In this sense, complementing the real-side analysis with simple proxies for monetary policy shocks can be expected to outperform a model built only on real-side effects. The fact that the mesoeconomic residual benchmark alone outperforms almost all other specifications, on the other hand, provides evidence that the real side of the European economy is well represented by the benchmark model.

1.5 Discussion

The general agreement in macroeconomics was that microeconomic shocks to firms or disaggregated industries cannot cause significant aggregate fluctuations as idiosyncratic shocks would wash out, based on the diversification argument brought forward by Lucas (1977). This consensus was shaken by seminal contributions from Gabaix (2011) and Acemoğlu et al. (2012).

This paper illustrates that both, the granularity of sales shares (Gabaix, 2011) and the heavily skewed distribution of connections in production networks (Acemoğlu et al., 2012), play a significant role in the description of aggregate fluctuations and a stable result can only be achieved if both concepts are taken into account. In particular, fluctuations at the aggregate level can best be described when explicitly accounting for an entities size and its position within the production network when facing idiosyn-

[17] This proxy can, by construction, not account for the (increasingly relevant) unconventional policy measures by the ECB and does only apply to the subset of the 19 Euro zone countries for our EU28 sample. The superior performance from including this imperfect proxy thus points to the relevance of the monetary policy channel and highlights its negligence in RBC-type models even more.

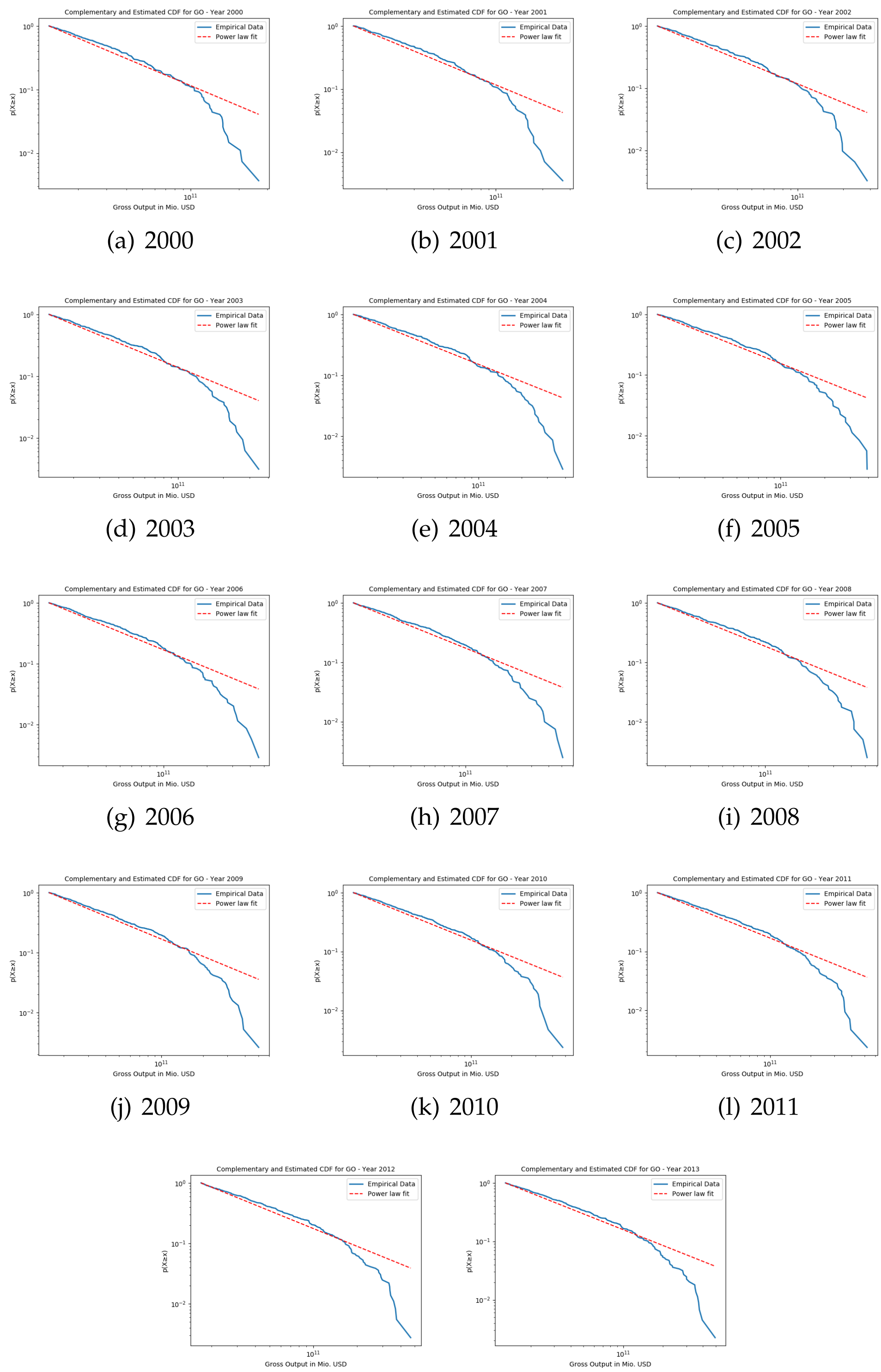
cratic shocks. The main insight suggested by this paper is that a precise and stable perception of business cycle fluctuations can only be achieved if the significant asymmetries in entity characteristics as well as their interactions are taken into account. With theory, calibration, and direct empirical evidence, this paper makes the case that, when accounting for the interactions of idiosyncratic shocks correctly, less than two percent of industries already explain almost eighty percent of aggregate fluctuations within Europe. Thus, the analysis provides a fairly good answer to the question raised by Acemoğlu et al. (2012), asking for a more systematic investigation of the quantitative importance of the mechanisms stressed in their paper.

With this, our mesoeconomic aggregation rule suggests a number of directions for future research. First, as homogeneity is assumed within industries, which is a limiting assumption in any industry level input–output table, the (technology) heterogeneity existing within industries cannot be fully capture. For this the application of our rule to firm data would allow for a more systematic analysis and behavioral assessment at the micro level. Considering size and network dynamics at a firm level has, for example, been addressed by Bernard et al. (2019a) with first intriguing results. Second, since the given model is non-parametric, we cannot investigate the impact of specific structural characteristics such as the geographic and intersectoral mobility of labour, market frictions and non-technological shocks in general. Ongoing foundational work in particular by Baqaee and Farhi (2019) addresses these concerns theoretically and asks for further development. Third, another important area for future research, also in the context of our robustness checks, is a systematic analysis of the relationships between financial networks or monetary policy and the extent of contagion and cascading failures in production networks. Ozdagli and Weber (2017) provide a first set of results indicating that production networks might be an important propagation mechanism of monetary policy to the real economy.

In sum, we anticipate our paper to be a valuable contribution to the current discussion on idiosyncratic shocks and aggregate outcomes and think that accounting for both, the granularity of economic entities and their entanglement in input-output networks, may lead to a better understanding of business cycle fluctuations, trade interdependencies and comovement in complex economic systems.

1.A Appendix

Complementary CDFs



(a) 2000 (b) 2001 (c) 2002

(d) 2003 (e) 2004 (f) 2005

(g) 2006 (h) 2007 (i) 2008

(j) 2009 (k) 2010 (l) 2011

(m) 2012 (n) 2013

Figure 1.6: Complementary and Estimated CDF for Gross Output of all years. The estimation is carried out following the procedure by Clauset et al. (2009).

Winsorizing, Standard Deviations and Variance Ratios

Assume a Subbotin initial distribution with a probability density function given by

$$f(x,\kappa,\sigma,\mu) = \frac{1}{2\sigma\kappa^{1/\kappa}\Gamma(1+1/\kappa)}\exp\left(-\frac{1}{\kappa}\left|\frac{x-\mu}{\sigma}\right|^{\kappa}\right).$$

with support over the real line and where $\kappa,\sigma \in \mathbb{R}^+$, $\mu \in \mathbb{R}$ and $\Gamma(\cdot)$ denotes the Gamma function. As can be readily verified, the Subbotin distribution includes the Gaussian for $\kappa = 2$ and the Laplacian for $\kappa = 1$ as special cases. The standard deviation of a p-winsorized distribution is equivalent to the standard deviation of a distribution censored at the pth and $1-p$th quantile, that is, a distribution where the cumulative weight from outside the pth and $1-p$th quantile is placed at the ends of the censoring interval. Thus, the impact of winsorizing on the standard deviation of a Laplacian initial distribution can be deduced from the standard deviation of a Subbotin density with $\kappa = 1$ censored at the pth and $1-p$th quantiles. The uncensored Laplacian is the Subbotin density with $\kappa = 1$, that is,

$$f(x,\mu,\sigma) = \frac{1}{2\sigma}e^{\left|\frac{x-\mu}{\sigma}\right|},$$

where $\sigma \in \mathbb{R}^+$ and $m \in \mathbb{R}$ and with support over the real line. Its pth quantile is given by $Q_L(p) = \mu + \sigma\ln(2p)$ and its $1-pth$ quantile by $Q_L(1-p) = \mu - \sigma\ln(2p)$ for $p \in (0, 0.5)$.

The standard deviation of this distribution censored at the pth and $1-p$th quantile is thus given by

$$SD_L(\mu,\sigma,p) = \sqrt{\int_{\mu+\sigma\ln(2p)}^{\mu-\sigma\ln(2p)}(x-\mu)^2dx + 2\cdot(\mu+\sigma\ln(2p))^2\cdot p},$$

as, by symmetry, $Q_L(p) - \mu = -Q_L(1-p) + \mu$. Integrating yields

$$SD_L(\sigma,p) = \sqrt{2}\sigma\sqrt{(2p\ln(p) + p(\ln(4)-2)+1)},$$

now independent of μ. Define $\lambda_L(p) = 1 - \frac{SD_L(\sigma,p)}{SD_L(\sigma)}$ as a measure of bias for the standard deviation of a winsorized Laplacian distribution compared to the uncensored initial distribution with $SD_L(\sigma) = \sqrt{2}\cdot\sigma$ as

$$\lambda_L(p) = 1 - \sqrt{(1+p(-2+\ln(4)+2p\ln(p))},$$

that is independent of σ. Thus, we can analytically determine the decrease of the standard deviation for p-winsorizing relative to the initial standard deviation for any given initial Laplacian distribution and the whole permitted range of $p \in (0, 0.5)$. In particular, $\lambda_L(0.01) = 0.0503898$. This implies that p-winsorizing a Laplacian initial distribution by a cumulative censored probability mass of $2p = 0.02$ or 2% leads to a overproportionate (and large) decrease of the standard deviation by about 5.03%.

To examine the effects of winsorizing on a Gaussian distribution, the same method as for the Laplacian can be applied. Unfortunately, however, the quantile values for the Subbotin density with $\kappa = 2$ depend on the inverse generalized Gamma function $\Gamma^{-1}(\cdot)$ that is not analytically tractable.

As a way round, a generic quantile function is considered. Assume without loss of generality that $\mu = 0$, since the standard deviation does not depend on the location of the distribution in question. The quantile function $Q(\kappa, \sigma, p)$ is now defined as $Q(\sigma, \kappa, p) = -\sigma \left(\kappa \Gamma^{-1}\left(\frac{1}{\kappa}, 2p\right)\right)^{\frac{1}{\kappa}}$,where $\Gamma^{-1}(\cdot)$ is the inverse generalized Gamma function. By $\mu = 0$ and by the symmetry of the distribution, $Q(\sigma, \kappa, p) = -Q(\sigma, \kappa, 1 - p)$.

For these quantile functions, the standard deviation of the censored Subbotin density $SD(\kappa, \sigma, p)$ at p and $1 - p$ is thus given by

$$SD(\kappa, \sigma, p) = \sqrt{\int_{-Q(\sigma,\kappa,p)}^{Q(\sigma,\kappa,p)} x^2 f(x, \kappa, \sigma)\, dx + 2Q(\sigma, \kappa, p)^2 \cdot F(Q(\sigma, \kappa, p))}\,,$$

where $F(\kappa, \sigma)$ is the cumulative distribution function of the (uncensored) Subbotin density.

To simplify notation, let $\Gamma^{-1}\left(\frac{1}{2}, 2p\right) = G(p)$. Setting $\kappa = 2$ for a Gaussian density, substituting the actual function for the generic quantile function and integrating yields $SD_G(\sigma, p)$ for the standard deviation of a Gaussian distribution censored at at the pth and $1 - p$th quantiles as a function of p and σ by

$$SD_G(\sigma, p) = \sigma \cdot \sqrt{\operatorname{erf}\left(\sqrt{G(p)}\right) - \frac{2\exp(-\,G(p))\sqrt{G(p)}}{\sqrt{\pi}} + 2G(p)\,\Gamma\left(\frac{1}{2}, G(p)\right)},$$

where $\operatorname{erf}(\cdot)$ is the error function.

Define finally $\lambda_G(p) = 1 - \frac{SD_G(\sigma,p)}{SD_G(\sigma)}$ as the ratio of standard deviation of the p-winsorized Gaussian to the standard deviation of the uncensored Gaussian distribution given by $SD_G(\sigma) = \sigma$ as

$$\lambda_G(p) = 1 - \sqrt{\operatorname{erf}\left(\sqrt{G(p)}\right) - \frac{2\exp(-G(p))\sqrt{G(p)}}{\sqrt{\pi}} + 2G(p)\,\Gamma\left(\frac{1}{2}, G(p)\right)}.$$

While $\lambda_G(p)$ is still dependent on $G(p)$ and thus, on the inverse Gamma function that is not analytically tractable, the function is only dependent on p as its sole argument. Thus, p can be easily analyzed numerically for its whole domain $p \in (0, 0.5)$. In particular, $\lambda_G(0.01) \approx 0.018046$. Thus, censoring a Gaussian distribution by a cumulative probability mass $2p = 0.02$ or 2% leads to a less than proportional decrease of the standard deviation by only about 1.8%.

Figure 1.7 plots both functions $\lambda_G(p)$ and $\lambda_L(p)$ in their domain $p \in (0, 0.5)$ as well as the cumulative censored probability mass resulting from p-winsorizing, that is, $2p$. Both functions $\lambda_G(p)$ and $\lambda_L(p)$ are expectedly bounded between 0 and unity, as with a winsorizing probability of $p = 0$, the winsorized distribution corresponds to the uncensored initial distribution with the same standard deviation and therefore $\lim_{p\to 0} \lambda_G(p) = \lim_{p\to 0} \lambda_L(p) = 0$. For p approaching 0.5, the distribution approaches the degenerate Dirac-Delta case, where the standard deviation and all higher-order moments are 0. Both functions thus approach unity for p going to 0.5.

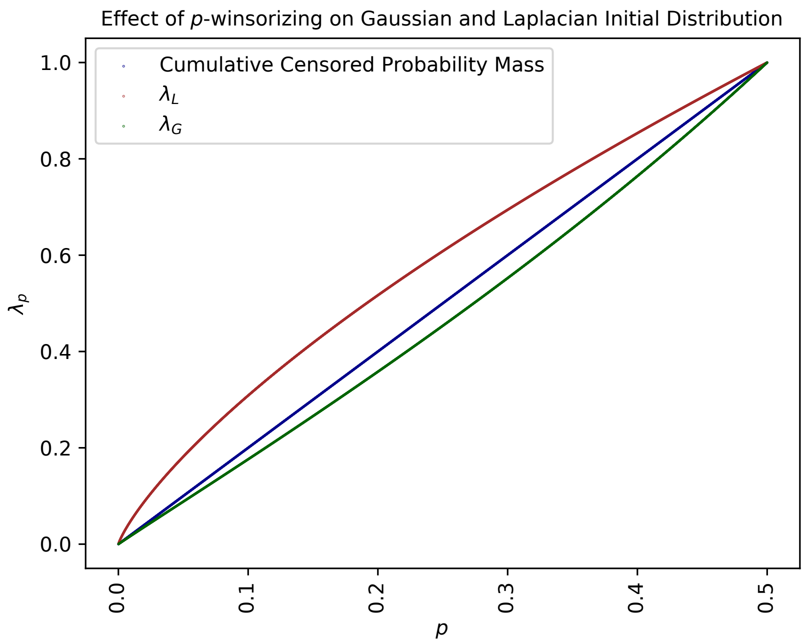


Figure 1.7: *Percentage decrease λ of the Standard Deviation for a Gaussian and Laplacian initial distribution from p-winsorizing.*

Apart from these limit cases, however, it can easily be seen that p-winsorizing leads to a overproportionate decrease of the standard deviation for any Laplacian initial distribution that is especially large for small values of p relative to the cumulative censored probability mass. In contrast to that, the decrease for the Gaussian is always less than proportional for the whole domain $p \in (0, 0.5)$. This illustrates the possibly hugely distorting effect of winsorizing distributions with fat tails, such as the Laplacian, while the distortion would be much less problematic for Gaussian distributions.

Consider finally the percentage loss in terms of variance $\Lambda_L(p) = 1 - \frac{var_L(\sigma,p)}{var_L(\sigma)}$ for the Laplacian and $\Lambda_G(p) = 1 - \frac{var_G(\sigma,p)}{var_G(\sigma)}$. Since by the definition of the standard deviation, $(SD(\sigma,p)/SD(\sigma))^2 = var(\sigma,p)/var(\sigma)$ it follows from $SD(\sigma,p)/SD(\sigma) < 1$ that $\lambda(p) < \Lambda(p)$ for all underlying distributions and for any given p in the permitted range $p \in (0, 0.5)$. Figure 1.8 plots these percentage decreases Λ for a Gaussian and Laplacian initial distribution. Indeed, the percentage loss is always more than proportionate with respect to the cumulative censored probability mass for both initial distributions.

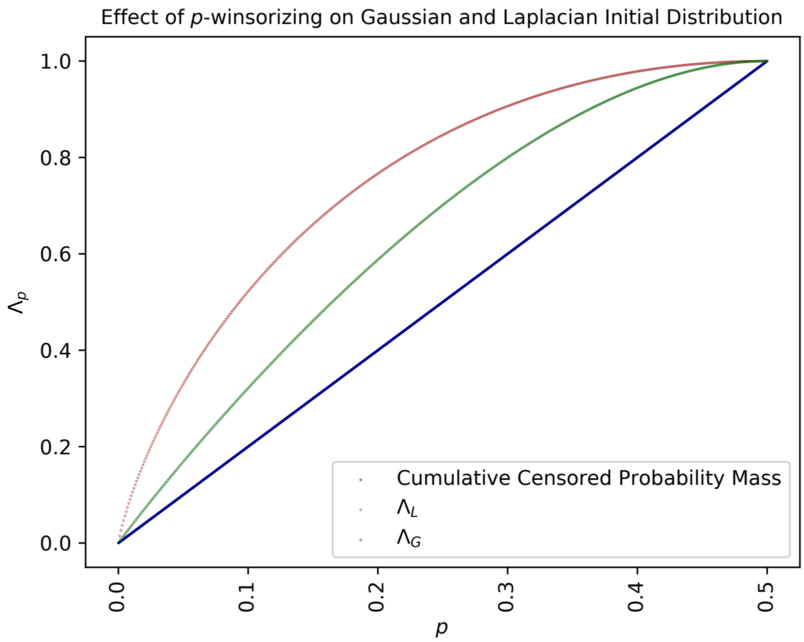


Figure 1.8: *Percentage decrease Λ of the Variance for a Gaussian and Laplacian initial distribution from p-winsorizing.*

However, given the disproportionate response to p winsorizing of the standard deviation SD_L for a Laplacian initial distribution, this effect is also much larger for the variance for a more leptokurtic distribution. For example, $\Lambda_G(0.01) \approx 0.0357664$ while $\Lambda_L(0.01) \approx 0.0982405$ which highlights the stark difference even for low degrees of winsorizing.

Generic Residuals and Simple Linear Regression

Consider a generic weight w_i for industry i. For our three specifications, this weight could be the Domar weight for the granular case, S_i/Y, the transmission centrality, θ_i, for the pure network case and $S_i/Y \cdot \theta_i$ for the idiosyncratic case. We assume that the TFP growth rate of the total economy is a weighted sum of idiosyncratic productivity shocks to N industries which implies

$$\frac{dTFP_t}{TFP_t} = \sum_{i=1}^{N} w_{i,t} d\pi_{i,t},$$

where $\pi_{i,t}$ is the productivity of firm i in period t, $w_{i,t}$ its weight at t and TFP_t is the Total Factor Productivity in period t. We assume that the idiosyncratic productivity shocks $d\pi_i$ are independent of the weights w_i which, for the granular case, would imply Gibrat's law. In the following, we call this condition *Generalized Gibrat's Law*. Assume that that productivity shocks are Hicks-neutral and that GDP growth y_t is proportional to TFP growth. For these assumptions, the cumulative idiosyncratic shocks multiplied by a factor ψ yield the GDP growth rate y_t which is given by

$$y_t = \psi \cdot \sum_{i=1}^{N} w_{i,t} d\pi_{i,t}.$$

It follows that the variance of GDP growth rates is given by

$$var(y) = \psi^2 \cdot \sigma_\pi^2 \cdot H(N),$$

where σ_π^2 is the variance of productivity growth rates and H the Hirschman-Herfindahl index (HHI) of the weights w. For generic weights w, $H(N)$ is therefore just the sum of N squared weights, that is,

$$H(N) = \sum_{i=1}^{N} (w_i)^2.$$

Consider now a generic residual $\mathcal{X}_t$ based purely on observables given by

$$\mathcal{X}_t = \sum_{i=1}^{K} w_{i,t} d\pi_{i,t},$$

where K denotes the K largest industry weights and assuming that we identified the correct productivity differences $d\pi_{i,t}$ for all K firms. If we assume the Generalized Gibrat's law to hold, that is, productivity growth rates to be independent of weights, σ_π has to be independent of $H(K)$ as it is defined by the weight distribution of industries.

Also, this implies that $E[\sigma_\pi(K)] = \sigma_\pi(N)$, that is, the variance of productivity growth rates of any subsample of K industries and $K > 1$ is expectationally equivalent to the variance of productivity growth rates for the whole sample.

The variance of generic residual $\mathcal{X}$ for the K largest industries in terms of weights is given by

$$var(\mathcal{X}^K) = \sigma_\pi^2(K) \cdot H(K).$$

Taking expectations to define the expected induced variance yields

$$\begin{aligned} E[var(\mathcal{X}^K)] &= E[\sigma_\pi^2(K)] \cdot H(K) \\ &= \sigma_\pi^2 \cdot H(K), \end{aligned}$$

as $H(K)$ is independent of σ_π by the generalized version of Gibrat's law.

Consider a simple linear regression with

$$y_t = \beta_0 + \beta_1 \cdot \mathcal{X}_t + \epsilon_t.$$

An OLS estimation would yield for this specification the estimated $\hat{\beta}_1$

$$\hat{\beta}_1 = \frac{cov(y, \mathcal{X})}{var(\mathcal{X})}.$$

Rewriting the initial specification gives

$$y_t = \psi \cdot \mathcal{X}_t^N = \psi \cdot (\mathcal{X}_t^K + \mathcal{X}_t^{N-K}),$$

for all considered periods t, where $\mathcal{X}^N$ denotes the generic residual for all N entities, $\mathcal{X}^K$ for the largest K ones and $\mathcal{X}^{N-K}$ for the remaining smallest $N - K$ industries or firms. From this definition, $\hat{\beta}_1$ can be expressed as

$$\hat{\beta}_1 = \frac{\sum_t^T (\mathcal{X}_t^K - \bar{\mathcal{X}}^K) \cdot (y_t - \bar{y})}{\sum_t^T (\mathcal{X}_t^K - \bar{\mathcal{X}}^K)^2}.$$

From $\psi \cdot (\mathcal{X}_t^K + \mathcal{X}_t^{N-K}) = y_t$ and therefore $\psi \cdot (\bar{\mathcal{X}}^K + \bar{\mathcal{X}}^{N-K}) = \bar{y}$, we get

$$= \frac{\sum_t^T (\mathcal{X}_t^K - \bar{\mathcal{X}}^K) \cdot \psi \cdot ((\mathcal{X}_t^K + \mathcal{X}_t^{N-K}) - (\bar{\mathcal{X}}^K + \bar{\mathcal{X}}^{N-K}))}{\sum_t^T (\mathcal{X}_t^K - \bar{\mathcal{X}}^K)^2}.$$

Rearranging gives

$$\begin{aligned}
&= \psi \cdot \frac{\sum_t^T (\mathcal{X}_t^K - \bar{\mathcal{X}}^K) \cdot ((\mathcal{X}_t^K - \bar{\mathcal{X}^K}) + (\mathcal{X}_t^{N-K} - \bar{\mathcal{X}}^{N-K}))}{\sum_t^T (\mathcal{X}_t^K - \bar{\mathcal{X}}^K)^2} \\
&= \psi \cdot \frac{var(\mathcal{X}^K) + cov(\mathcal{X}^K, \mathcal{X}^{N-K})}{var(\mathcal{X}^K)} \\
&= \psi \cdot (1 + \frac{cov(\mathcal{X}^K, \mathcal{X}^{N-K})}{var(\mathcal{X}^K)}).
\end{aligned}$$

Since we assume that all sectors draw independently from productivity distributions with equal first and second moments, $cov(\mathcal{X}^K, \mathcal{X}^{N-K})$ and $var(\mathcal{X}^K)$ are independent.

Also, by the same independence assumption, it has to hold that $E[cov(\mathcal{X}^K, \mathcal{X}^{N-K})]$ $= 0$. Taking expectations yields

$$\begin{aligned}
E[\hat{\beta}_1] &= E[\psi \cdot (1 + \frac{cov(\mathcal{X}^K, \mathcal{X}^{N-K})}{var(\mathcal{X}^K)}) \\
&= \psi + \psi \cdot E[\frac{cov(\mathcal{X}^K, \mathcal{X}^{N-K})}{var(\mathcal{X}^K)}]
\end{aligned}$$

and by invoking the independence of $cov(\mathcal{X}^K, \mathcal{X}^{N-K})$ and $var(\mathcal{X}^K) = \psi + \psi \cdot (E[cov(\mathcal{X}^K, \mathcal{X}^{N-K})] \cdot E[var(\mathcal{X}^K)^{-1}])$ which yields with $E[cov(\mathcal{X}^K, \mathcal{X}^{N-K})] = 0$ again from independence = ψ.

Consider now the coefficient of determination for the given simple linear regression with

$$R^2 = \hat{\beta}_1^2 \cdot \frac{var(\mathcal{X}^K)}{var(y)}.$$

Taking expectations yields

$$E[R^2] = E[\hat{\beta}_1^2 \cdot \frac{var(\mathcal{X}^K)}{var(y)}]$$

and since $var(y)$ is observable, $E[var(y)] = var(y)$ which implies $= var(y)^{-1} \cdot E[\hat{\beta}_1^2] \cdot E[var(\mathcal{X}^K)]$.

Substituting gives

$$\begin{aligned}
&= \frac{\psi^2 \cdot \sigma_\pi^2 \cdot H(K)}{\psi^2 \cdot \sigma_\pi^2 \cdot H(N)} \\
&= \frac{H(K)}{H(N)} \\
&= \sum_{i=1}^{K} \frac{(w_i)^2}{\sum_{j=1}^{N}(w_j)^2}.
\end{aligned}$$

Thus, under the generalized version of Gibrat's law and assuming that we correctly identified the idiosyncratic productivity shocks to each industry or firm, the expected R^2 for the granular residual for K industries should be given by the ratio of the weight Herfindahl for the K largest firms or industries to the weight Herfindahl for the total economy.

To illustrate this result, we perform a simple Monte Carlo simulation. The weight distribution is given by a power-law with tail exponent $\alpha = 1.9$. There are $N = 100$ industries in total and the idiosyncratic productivity shocks are drawn from a Laplacian distribution with location parameter $\mu = 0$ and scale parameter $\sigma = 0.05$. We consider 100 time periods each and repeat the simulation 50 times with different realizations for the idiosyncratic shocks. Figure 1.9 shows the mean R^2 for these 50 simulation runs and a simple linear regression model as presented above as well as the expected R^2 by the results below. This is given by $\frac{H(K)}{H(N)} = \frac{\zeta_K(\frac{2}{\alpha})}{\zeta_N(\frac{2}{\alpha})}$ where $\zeta_N(\alpha)$ is a truncated Zeta function at N with $\zeta_N(\alpha) = \sum_{k=1}^{N} k^{-\alpha}$ and with $\alpha = 1.9$.

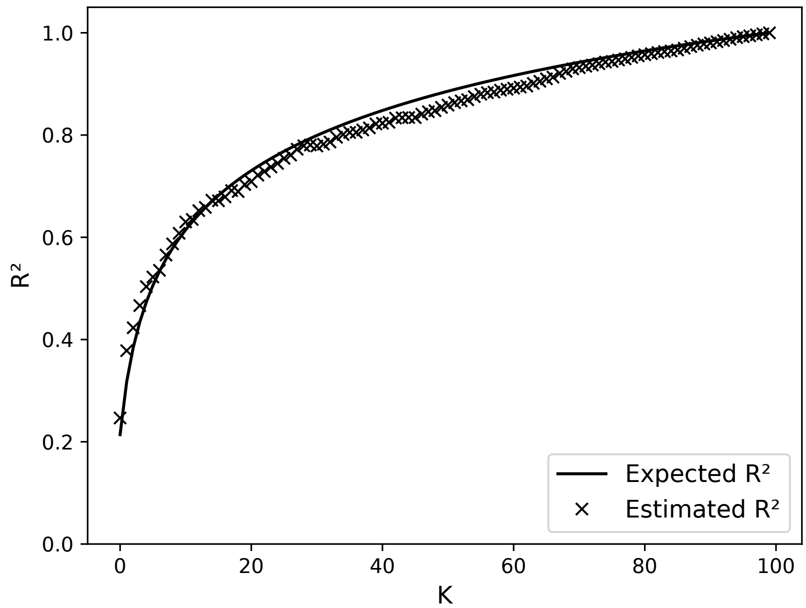


Figure 1.9: Mean R^2 for Monte Carlo simulation exercise and expected R^2 derived from the weight distribution.

As can easily be seen, the expected R^2 from the ratio of Herfindahls is remarkably close to the mean R^2 from the simulation exercise. Notice that our results are not materially sensitive to the specific choice of the distribution for idiosyncratic productivity shocks. The results are robust for several different parameter values of μ, σ and κ for a (symmetrical) Subbotin distribution that includes the outlined Laplacian as a special case for $\kappa = 1$. The approximation only breaks down for very small shape parameters $\kappa < 0.3$, for which the variance and all higher moments explode and thus, 100 different realizations for the idiosyncratic productivity shocks are not sufficient for the mean value to converge to the expected value by the law of large numbers. The result is also robust to various specifications for the weight distribution (measured by the tail exponent α).[18]

Of course, this is only the expected or mean functional form of the R^2. Indeed, depending on the specific realization of draws from the distribution for idiosyncratic shocks, the realized R^2 might differ, as Figure 1.10 shows. However, given a sufficiently heterogeneous weight distribution, this R^2 plot cannot deviate too far from the expected R^2 as shown in Figure 1.9 as it should increase almost monotonically, excluding therefore huge downturns in the empirical R^2 plot.

[18] Material available upon request.

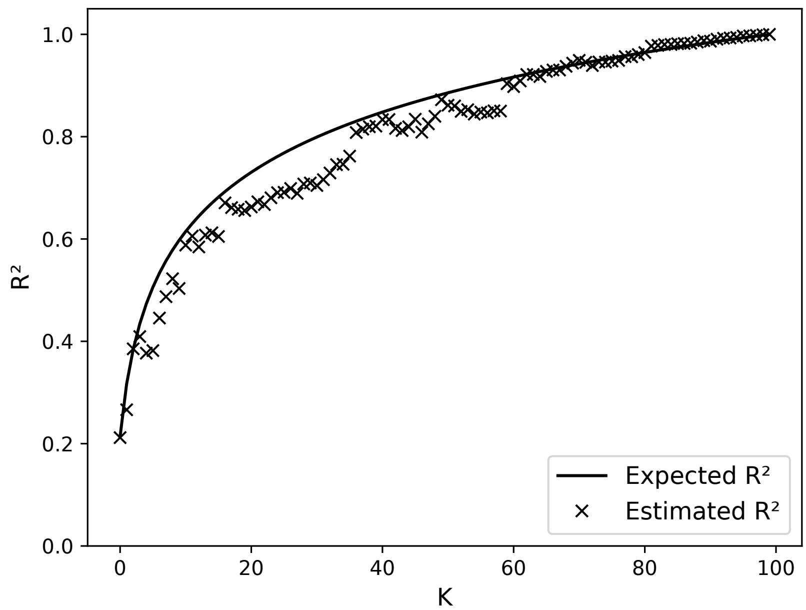


Figure 1.10: R^2 for a specific realization of idiosyncratic productivity shocks and expected R^2 derived from the weight distribution.

Finally, consider the boundary condition Gabaix (2011) gives for granular firms to exist - they have to be distributed by a power law with a tail exponent α at most equal to 2.[19] This implies that if all K weights are granular, $H(K) \geq \zeta_K(1)$. This, however, is just the harmonic series. Thus, for any given $H(N) = H$ for the total economy, the ratio $\frac{H(K)}{H} \approx \frac{1}{H(N)}(ln(K) + \gamma)$, where γ is the Euler-Mascheroni constant. Thus, the R^2 plot can be expected to increase at least proportional to the natural logarithm of K.

[19]See Appendix 1.A for further details.

Winsorizing, β estimates and R^2

Consider the same framework as in Appendix 1.A. However, assume now that the productivity growth rate distribution of the K largest industries was winsorized at the pth and $1-pth$ quantile. We define $\nu(p) = \tilde{\sigma}^2_\pi(K,p)/\sigma^2_\pi(K)$, that is, ν gives the ratio of variances for the p-winsorized and initial distribution of productivity growth rates for the K largest industries with $\nu(p) \in (0,1)$ for $p \in (0, 0.5)$. It has to hold that $\lim_{p\to 0}\nu(p) = 1$ and $\lim_{p\to 0.5}\nu(p) = 0$. If we assume the generalized version of Gibrat's law to hold, it follows that $E[\sigma^2_\pi(K)] = E[\sigma^2_\pi(N)] = \sigma^2_\pi$ and therefore $E[\tilde{\sigma}^2_\pi(K,p)] = \nu(p)\cdot\sigma^2_\pi$.

The expected, winsorized variance of $\tilde{\mathcal{X}}^K(p)$ of the generic residual defined in the section 1.A is then

$$E[var(\tilde{\mathcal{X}}^K(p))] = \nu(p)\cdot\sigma^2_\pi\cdot H(K).$$

For analytical tractability, we assume that the winsorized generic residual is affected by some time-invariant scalar $\delta^K(p) \in \mathbb{R}$ such that $\mathcal{X}^K_t = \tilde{\mathcal{X}}^K_t(p) + \delta^K(p)\forall t$, which might vary with the winsorizing probability p and with the number of considered entities K. Following the discussion in Appendix 1.A, we thus get

$$\begin{aligned} y_t &= \psi\cdot\mathcal{X}^N_t \\ &= \psi\cdot(\mathcal{X}^K_t + \mathcal{X}^{N-K}_t) \\ &= \psi\cdot(\tilde{\mathcal{X}}^K_t(p) + \delta^K(p) + \mathcal{X}^{N-K}_t) \end{aligned}$$

and analogously for the time-series averages, where $\bar{\mathcal{X}}^K(p)$ is the time-series average of the generic residual winsorized at p

$$\bar{y} = \psi\cdot(\bar{\mathcal{X}}^K(p) + \delta^K(p) + \bar{\mathcal{X}}^{N-K}).$$

For the regression coefficient, it follows that

$$\begin{aligned} \hat{\beta}_1(p) &= \frac{\sum_t^T(\tilde{\mathcal{X}}^K_t(p) - \bar{\mathcal{X}}^K(p))\cdot(y_t-\bar{y})}{\sum_t^T(\tilde{\mathcal{X}}^K_t(p) - \bar{\mathcal{X}}^K(p))^2} \\ &= \frac{\sum_t^T(\tilde{\mathcal{X}}^K_t(p) - \bar{\mathcal{X}}^K(p))\cdot\psi\cdot((\tilde{\mathcal{X}}^K_t(p) + \delta^K(p) + \mathcal{X}^{N-K}_t) - (\bar{\mathcal{X}}^K(p) + \delta^K(p) + \bar{\mathcal{X}}^{N-K}))}{\sum_t^T(\tilde{\mathcal{X}}^K_t(p) - \bar{\mathcal{X}}^K(p))^2} \\ &= \psi\cdot\frac{\sum_t^T(\tilde{\mathcal{X}}^K_t(p) - \bar{\mathcal{X}}^K(p))\cdot(\tilde{\mathcal{X}}^K_t(p) + \mathcal{X}^{N-K}_t - \bar{\mathcal{X}}^K(p) + \bar{\mathcal{X}}^{N-K})}{\sum_t^T(\tilde{\mathcal{X}}^K_t(p) - \bar{\mathcal{X}}^K(p))^2} \end{aligned}$$

and by rearranging

$$
\begin{aligned}
&= \psi \cdot \frac{\sum_t^T (\tilde{\mathcal{X}}_t^K(p) - \bar{\mathcal{X}}^K(p)) \cdot (\tilde{\mathcal{X}}_t^K(p) - \bar{\mathcal{X}}^K(p)) + (\mathcal{X}_t^{N-K} - \bar{\mathcal{X}}^{N-K}))}{\sum_t^T (\tilde{\mathcal{X}}_t^K(p) - \bar{\mathcal{X}}^K(p))^2} \\
&= \psi \cdot \frac{var(\tilde{\mathcal{X}}^K(p)) + cov(\tilde{\mathcal{X}}^K(p), \mathcal{X}^{N-K})}{var(\tilde{\mathcal{X}}^K(p))} \\
&= \psi \cdot (1 + \frac{cov(\tilde{\mathcal{X}}^K(p), \mathcal{X}^{N-K})}{var(\tilde{\mathcal{X}}^K(p))}).
\end{aligned}
$$

If the winsorizing procedure did not generate spurious dependence between industries, the terms $cov(\tilde{\mathcal{X}}^K(p), \mathcal{X}^{N-K})$ and $var(\tilde{\mathcal{X}}^K(p))$ remain independent and $E[cov(\tilde{\mathcal{X}}^K(p), \mathcal{X}^{N-K})] = 0$. Thus, as shown in Appendix 1.A, the estimate for the β_1 coefficient in a simple linear regression model remains unbiased, and $E[\hat{\beta}_1(K)|p = 0] = E[\hat{\beta}_1(K)|p] = \psi$. However, the expected R^2 as the explanatory power for K industries should for any degree of winsorizing p be biased downwards. This follows purely from the fact that the variance of the winsorized generic residual is downwards biased in expectations. To see this, consider $E[R^2(K)|p]$ as

$$
\begin{aligned}
E[R^2(K)|p] &= \psi^2 \frac{E[var(\tilde{\mathcal{X}}^K(p)]}{var(y)} \\
&= \frac{\psi^2 \cdot \nu(p) \cdot \sigma_\pi^2 \cdot H(K)}{\psi^2 \cdot \sigma_\pi^2 \cdot H(N)} \\
&= \nu(p) \cdot \frac{H(K)}{H(N)} \\
&= \nu(p) \cdot \sum_{i=1}^{K} \frac{(w_i)^2}{\sum_{j=1}^{N} (w_j)^2}
\end{aligned}
$$

and that $\nu(p) < 1$ for all $p \in (0, 0.5)$. It follows that, compared to the benchmark expected R^2 without winsorizing, the estimate is downwards biased by $\Lambda(p)$ per-cent given by

$$
\Lambda(p) = \frac{E[R^2(K)|p = 0] - E[R^2(K)|p]}{E[R^2(K)|p = 0]} = 1 - \nu(p).
$$

Calibrating Aggregate Shocks

The specification for the idiosyncratic shock follows the traditional assumption in the literature that the composite shock $g_{i,t}$ to industry i we observe at time t is composed of two independent components, an aggregate and an idiosyncratic one. Following this assumption, the idiosyncratic shock can be expressed as $(g_{i,t} - \bar{g}_t)$ with $\bar{g}_t$ denoting the aggregate shock. The aggregate shock to an industry i is calibrated by averaging over the productivity growth rates of the set of industries that are i) in the granular set of size Q and ii) in the same industry class I as industry i. If industry i is, for example, the German automotive industry, we average over all EU28 automotive industries as industry class I that are in the granular set. By the first condition on granularity, we aim to include only industries that can be expected to exhibit a non-vanishing impact on aggregate movements by their relative size. The latter one is to capture industry-specific disturbances such as industry-wide price movements better and is theoretically shown to outperform the residual variant without industry controls in the online appendix to Gabaix (2011). The results without industry controls are generally in line with our baseline and reported at the end of this appendix.

We set $\bar{g}_t = \bar{g}_{I,t}^{Q}$ for an arbitrary industry i as the median growth rate of the largest Q industries within the same industry class I as i. Deviating from convention in the granular literature, we choose the median as the measure of location that is best suited to account for the well-established leptokurtic nature of empirical productivity growth rate distributions and not the arithmetic mean that is highly dependent on extreme observations. Q as the number of largest industries in terms of size is a free parameter of the specification that has unfortunately received rather limited attention in the literature and is typically arbitrarily fixed (see also Dosi et al., 2019, for a criticism of this practice). We estimate Q in a non-arbitrary way as the set of industries that can, by granular theory, be expected to influence aggregate movements and thus contribute to the aggregate shock. By this definition, we also anticipate the criticism that our variant of the granular hypothesis at the industry level is ill-posed and its measured performance therefore distorted since industry sizes clearly do not follow the power law functional form on which granular theory is based (Gabaix, 2011, Proposition 2). This is apparent from the complementary CDFs in Appendix 1.A.

In particular, the theorem in Gabaix (2011) establishes sufficient conditions for entities to be granular in terms of the parameters of this power law that typically approximates the upper tail of firm size distributions, namely, the tail exponent α being between unity and two. This is due to the fact that a power law with parameters in this range has infinite variance and idiosyncratic shocks to entities distributed in such a way do not cancel out in the aggregate by the Central Limit Theorem.[20] To translate the parametric restrictions for a particular distributional form into non-parametric restrictions applicable to our case at the industry level, we calibrate with respect to the measured level of concentration a power law under those parametric restrictions would imply.

As we show in Appendix 1.A, the square root of the Hirschman-Herfindahl-Index (HHI) is the correct measure of concentration to calculate the relative influence of a subset of granular entities on the standard deviation of fluctuations of the aggregate system. This allows to establish boundary conditions in terms of measured concentration that are equivalent to the parameter restrictions for power law distributions without requiring the distribution to be well approximated by a power law. To translate the parametric restrictions of a power law into the implied HHI, we derive the square-root of the HHI for a discretized power-law with arbitrary $\alpha \in \mathbb{R}^+$ that is given by

$$S(k) = \frac{S_{min} \cdot N^{\frac{1}{\alpha}}}{k^{\frac{1}{\alpha}}}$$

with $k = 1, 2, ..., N$ as the respective ranks of a given S in a descending order and N as the number of values with $N \in \mathbb{N}^+$.

The HHI is defined as the total sum of N squared market shares of sales S. Thus, h as the square-root of the HHI is given by

$$h = \sqrt{\sum_{k=1}^{N} \left(\frac{S_k}{\sum_{l=1}^{N} S_l} \right)^2}.$$

With $S_k = S(k)$, this resulting h is given by

$$h = \sqrt{\sum_{k=1}^{N} \left(\frac{\frac{S_{min} \cdot N^{\frac{1}{\alpha}}}{k^{\frac{1}{\alpha}}}}{\sum_{l=1}^{N} \frac{S_{min} \cdot N^{\frac{1}{\alpha}}}{l^{\frac{1}{\alpha}}}} \right)^2} = \sqrt{\frac{\zeta_N\left(\frac{2}{\alpha}\right)}{\zeta_N^2\left(\frac{1}{\alpha}\right)}},$$

[20]The aggregation of values drawn from a distribution with finite variance converges, by the Central Limit Theorem, asymptotically to a Gaussian distribution. A power law with a tail exponent above 2 has finite variance. Thus, by the Central Limit Theorem, the aforementioned condition implies vanishing fluctuations at the aggregate level. Conversely, for power laws with tail exponents between unity and two and thus infinite variance, there exist non-vanishing aggregate fluctuations even for a very large number of firms, industries or other affected entities.

where $\zeta_N(\alpha)$ is a truncated Zeta function at N with $\zeta_N(\alpha) = \sum_{k=1}^{N} k^{-\alpha}$. Thus, h under the assumption of a discretized power law is a function of the tail exponent α and the number of Gross Outputs N, that is, $h = h^{PL}(\alpha, N)$.

This enables us to express the parametric restrictions for a power law by a non-parametric measure without any distributional assumptions. We require the empirical concentration measure h^{emp} for the N largest industries to be within the interval of concentration measures derived for a theoretical power law parametrized by the boundary conditions in Gabaix's theorem. Formally, we require for the number of industries N the inequalities $h^{PL}(1, N) \geq h_t^{emp}(N) \geq h^{PL}(2, N)$ to hold with h^{emp} evaluated at t. Theoretically, the behavior of the Q_t largest industries should be indicative of the state of the system as a whole in t as these entities should exhibit a non-vanishing and thus significant impact on aggregate outcomes. The size of the granular set at t, Q_t, is then the minimal N for which these inequalities hold. We truncate the sample at its minimum to include as few information about the aggregate system as possible in our construction of the residual and therefore to not artificially inflate its measured performance.

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Q_t^*	219	219	214	207	205	204	204	208	224	205	219	224	221	218
Zipf	0.087	0.087	0.088	0.089	0.089	0.090	0.090	0.089	0.086	0.089	0.087	0.086	0.086	0.087
Rep. Industry	1.284	1.284	1.282	1.279	1.279	1.279	1.278	1.280	1.285	1.279	1.28	1.286	1.285	1.284

Table 1.2: The calibration shows the chosen set with size Q to be only barely more concentrated than the representative industry benchmark but two orders of magnitude less concentrated than the Zipf case. Q_t^* gives the minimal granular sample sizes as defined by the MGC, column *Zipf* compares the empirical square root of the HHI to the implied square root of the HHI for the same $N = Q_t^*$ and a Zipfian distribution by $h^{emp}(Q_t^*)/h^{PL}(1, Q_t^*)$, while column *Rep. Industry* reports the ratio of the empirical h for Q_t^* to the h for the representative industry benchmark given by $h = 1/\sqrt{Q_t^*}$.

We require a consistent Q for the whole covered time period to construct the residual. Taking the maximum over all Q_t, that is, $Q = 224$, as a first guess is sufficient, as the granularity condition is fulfilled for this Q at all t. The measured concentration is more than one order of magnitude smaller than the Zipfian benchmark typically characterizing firm-size distributions (row 3), while only being about 28% above the equal weights or representative industry benchmark (row 4). This result demonstrates the fundamental differences in levels of aggregation between the firm to the industry level, where the importance of size for shock transmission apparently greatly diminishes.

For the snapshot estimations, we fix $K^* = Q = 224$ but let K vary in a second step to test for the expected functional form of the emerging R^2 function plotted against an increasing K until K^*. Even though the estimated K^* and Q are well based in theory, we also conduct robustness tests on $K^* = Q$ and the specification with industry controls. We find that our results are robust for both different choices of Q and against dropping the specification of industry controls and just averaging over Q. Varying Q such that $Q = 20, 30, 50, 100$ does not seem to have any qualitative effect on the achieved explanatory power. This material is available upon request. When we just average over the Q granular industries without taking the industry composition into account, the results are also qualitatively in line with what we report in section 1.4, even though generally worse.

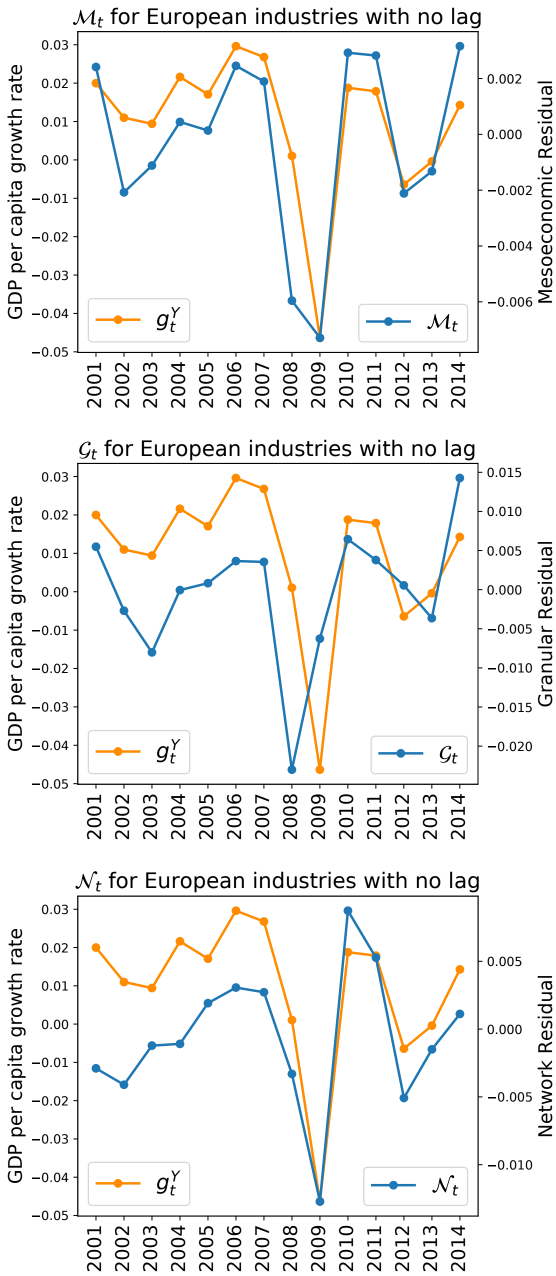
We test this for the residual variants $\tilde{\mathcal{M}}, \tilde{\mathcal{G}}$ and $\tilde{\mathcal{N}}$, where the only difference to their counterparts $\mathcal{M}, \mathcal{G}$ and $\mathcal{N}$ is that the aggregate shock is calculated without industry controls, that is, by just averaging over the largest $Q = 224$ industries in each year. For easier reference, we labeled the respective regression equations analogously to the benchmark with industry controls by IV to VI:

$$IV: \ (a) \ g_t^Y = \beta_0 + \beta_1 \tilde{\mathcal{M}}_t + \epsilon_t \qquad (b) \ g_t^Y = \beta_0 + \beta_1 \tilde{\mathcal{M}}_t + \beta_2 \tilde{\mathcal{M}}_{t-1} + \epsilon_t \tag{1.30}$$

$$V: \ (a) \ g_t^Y = \beta_0 + \beta_1 \tilde{\mathcal{G}}_t + \epsilon_t \qquad (b) \ g_t^Y = \beta_0 + \beta_1 \tilde{\mathcal{G}}_t + \beta_2 \tilde{\mathcal{G}}_{t-1} + \epsilon_t \tag{1.31}$$

$$VI: \ (a) \ g_t^Y = \beta_0 + \beta_1 \tilde{\mathcal{N}}_t + \epsilon_t \qquad (b) \ g_t^Y = \beta_0 + \beta_1 \tilde{\mathcal{N}}_t + \beta_2 \tilde{\mathcal{N}}_{t-1} + \epsilon_t \tag{1.32}$$

We show in Figures 1.11 and 1.12 that these robustness checks are in line with the results reported in section 1.4, in so far as the mesoeconomic specification outperforms its limit cases both for the snapshot evaluated at $K* = 224$ as well as for the predicted functional form for the R^2 emerging when plotted against K in the whole range of $K = 1, \ldots, K^*$. In general, all variants perform worse, indicating that the specification without industry controls might not be adequate for the estimation of aggregate shocks. Especially the granular variant experiences a sizeable loss in explanatory power with a shape of the R^2 plot completely at odds with theoretical predictions.

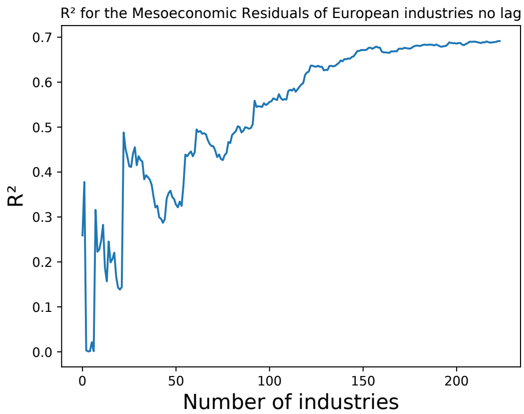


	IV (a)	IV (b)
$\tilde{\mathcal{M}}_t$	0.4810***	0.4725***
	(0.0927)	(0.0890)
$\tilde{\mathcal{M}}_{t-1}$		0.1766*
		(0.0907)
Intercept	0.0011***	0.0012***
	(0.0002)	(0.0002)
T	14	13
R^2	0.691	0.772
Adjusted R^2	0.666	0.726
F Statistic	26.88***	16.89***

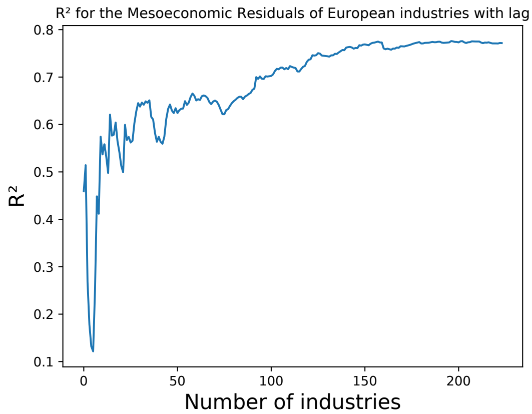
	V (a)	V (b)
$\tilde{\mathcal{G}}_t$	0.957	0.921*
	(0.576)	(0.470)
$\tilde{\mathcal{G}}_{t-1}$		1.549**
		(0.528)
Intercept	0.009*	0.011**
	(0.005)	(0.004)
T	14	13
R^2	0.187	0.555
Adjusted R^2	0.119	0.466
F Statistic	2.76	6.24**

	VI (a)	VI (b)
$\tilde{\mathcal{N}}_t$	2.981***	3.153***
	(0.627)	(0.590)
$\tilde{\mathcal{N}}_{t-1}$		0.6372
		(0.588)
Intercept	0.011***	0.010***
	(0.003)	(0.003)
T	14	13
R^2	0.653	0.744
Adjusted R^2	0.624	0.692
F Statistic	22.60***	14.51***

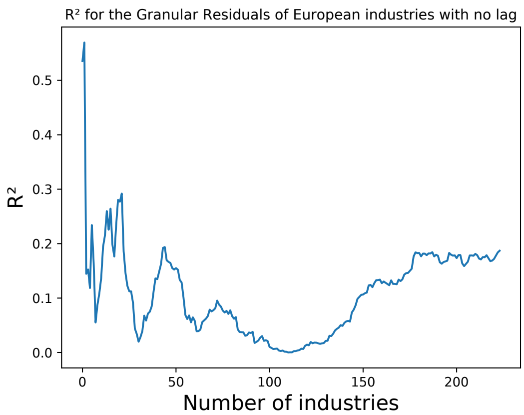
Figure 1.11: Without industry controls, the empirical performance of all residual variants $\tilde{\mathcal{M}}, \tilde{\mathcal{G}}$ and $\tilde{\mathcal{N}}$ drops considerably with the highest drop for the granular variant $\tilde{\mathcal{G}}$. The left panel shows the time series of residuals against g_t^Y (GDP per capita growth rate) between 2000 and 2014. The right panel reports the results of all residual series being regressed on g_t^Y with the purely contemporaneous specifications IV (a), V (a) and VI (a) and including one lagged term in IV (b), V (b) and VI (b). Standard errors are given in parentheses. Significance at the *** 1%, ** 5%, * 10% level.



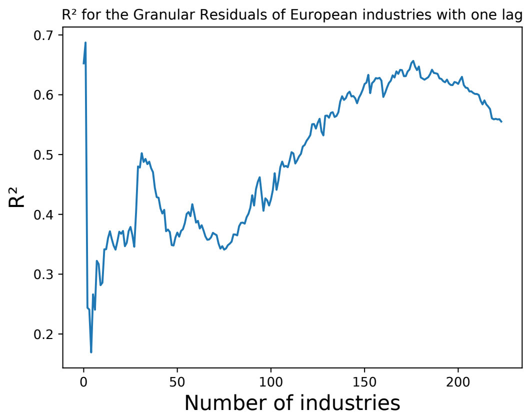
(a) Specification IV (a)



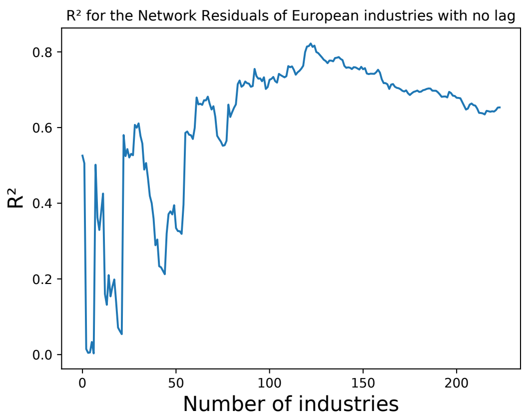
(b) Specification IV (b)



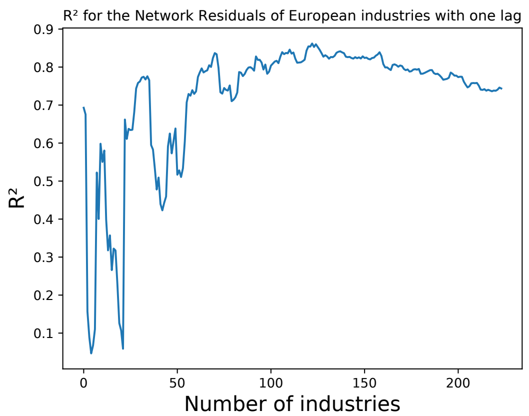
(c) Specification V (a)



(d) Specification V (b)



(e) Specification VI (a)



(f) Specification VI (b)

Figure 1.12: All variants without industry controls for the productivity growth rates perform considerably worse than there counterparts that include those. The meso specification outperforms both its limit cases and is in general closer to theoretical predictions. The upper panels show the results for the meso specifications IV (a) and V (b), the middle panels the granular specifications V (a) and V (b) and the lower ones for the pure network case in variants VI (a) and VI (b).

Chapter 2

Disentangling Growth: From Business Cycle Comovement and Higher-Order Linkages in Europe*

Co-written by Jan Schulz and Mishael Milaković

Abstract

Business cycles in Europe are highly synchronized. This comovement is of great macroeconomic interest, as it is typically seen as a necessary condition for the feasibility of a common European monetary and regulatory policy. However, there is still no consensus on the mechanisms behind such a concurrent cyclicality. The more traditional trade literature has focused on bilateral trade linkages between countries, while a more recent strand emphasizes the role of multi-sectoral production networks where idiosyncratic shocks propagate through these networks leading to comovement in the aggregate by higher-order linkages. We jointly employ a structural decomposition analysis with a principal component analysis to contribute to the discussion about first-order and higher-order linkages. Empirical evidence from the World Input-Output Database (WIOD) suggests that macroeconomic volatility in Europe tends to originate in final goods markets and to spread through higher-order rather than bilateral linkages.

2.1 Introduction

European business cycles display a remarkable degree of synchronicity (De Haan et al., 2008). Synchronous business cycles are of utmost importance for policy making within a monetary union because, for example, monetary policy gains greatly in effectiveness in the case of relatively homogeneous cycles (De Grauwe, 2020). However, the determinants of synchronization are not yet sufficiently established. While the traditional trade

*The article in this chapter has been presented to participants of the BaGBeM Brownbag Seminar in Bamberg and the International School and Conference on Network Science, NetSci-X 2020, in Tokyo.

literature, with intuitions built within two- or three-country models, focuses on bilateral relations such as trade linkages (Frankel and Rose, 1998; Baxter and Kouparitsas, 2005), a competing view focuses on the propagation of shocks through a production network with a notably multi-sectoral perspective that cannot be reduced to bilateral relations (Shea, 2002). Thereby, the literature on macroeconomic shocks suggests that aggregate fluctuations can be decomposed into (at least) two components with sector-specific relative influence. According to Garin et al. (2018), an *idiosyncratic component* affects only a single sector and an *aggregate component* jointly affects all sectors. In the jargon of this production network approach, first-order spillovers in bilateral relationships, as emphasized by the trade literature, need to be complemented by higher-order spillovers through the production network.

The idea that the production of goods and services is based on a complex web of transactions between a large number of suppliers and customers has a long tradition in the economy and goes back at least to Quesnay (1758). However, it was the seminal work of Leontief (1936) that paved the way for what is now referred to as input-output analysis, resulting in recent economic advances as the networked view of production and consumption processes attracted increasing attention. This development evolved more than 80 years after Leontief's pioneering study of the structure of the US economy because modern input–output tables are capable of mapping the complex patterns of input linkages across hundreds of industries or even firms revealing relations throughout economies in the sense of an economic network [1].

The literature on production networks has placed considerable emphasis on the amplification of exogenous shocks as they propagate through the network of producers (Acemoğlu et al., 2012; Carvalho and Tahbaz-Salehi, 2019; Baqaee and Farhi, 2019), building of two sets of assumptions in the process. The first set of empirical assumptions relates to the structural characteristics of the assumed model economies. For exogenous shocks to propagate through a network and amplify, the network topology has to be static so as not to influence aggregate volatility. This is a non-trivial assumption, since link formation for intermediate goods trade is a market outcome that may in principle adjust endogenously to variability within final goods markets, typically considered to be exogenous explanatory shocks (Atalay et al., 2011; Acemoğlu et al., 2012). In this paper, we aim to quantify the relative significance of both channels intermediate goods trade and final goods markets for aggregate comovement. Our results suggest that, as the latter exceeds the former in size, it contribution relative more to

[1] This is because the so-called technical coefficients, which form the basis for the famous *Leontief inverse*, can be interpreted as a weighted directed network, or, if binarized, as an unweighted directed network.

total value added volatility in Europe than the other.

Also, analytical results linking aggregate volatility to idiosyncratic shocks are typically derived for competitive markets within a general equilibrium setting, and fail to account for market imperfections (Acemoğlu et al., 2012; Baqaee and Farhi, 2019). To address this shortcoming, our empirical framework includes a proxy for realized market power in the form of an average cost markup over input costs. As in the case with intermediates, we find that the changes in markup contribute little to aggregate volatility. In contrast, cross-sectoral differences in the level of markup are important mediating determinants for the impact of volatility in individual final goods markets on aggregate volatility. Including (static) market distortions in network models, as pioneered by Baqaee and Farhi (2020), is therefore a major step towards closing the gap between data and model results. This also seems to imply that the apparent division of labor in the literature, with one strand focusing on aggregate volatility with static networks (e.g., Atalay et al., 2011; Acemoğlu et al., 2012; Baqaee and Farhi, 2019), and a second, distinct strand focusing on the formation of the input-output structure (e.g., Carvalho, 2014; Oberfield, 2018; García-Algarra et al., 2019), is indeed fruitful. In the process, where the more long-term perspective of the network formation literature can provide input for the assumptions of short-term models of aggregate fluctuations.

The second set of empirical assumptions relates to the propagation mechanism of shocks, rather than their origin. For networked structures to matter, shocks need to be transmitted further after the first spillover to other nodes connected in the intermediate or final goods markets. We propose a method based on principal components decomposition to identify the relative significance of such comovement caused by higher-order linkages in contrast to comovement due only to pairwise linkages for the final goods market. We find that higher-order linkages and aggregate shocks are also responsible for the vast majority of observed comovement in this market. This finding ties in with similar outcomes in the literature focusing on propagation patterns in production networks resulting from intermediate trade relationships with skewed distributions in entity size or network connectivity.

The remainder of this paper is organized as follows: Section 2.2 will introduce our methods, consisting of the joint application of a structural decomposition analysis (SDA) with a principal component analysis (PCA). Section 2.3 introduces the input-output dataset on which we base our analysis. The results of our examination are presented in Section 2.4. Section 2.5 concludes the paper and highlights potential avenues for further research.

2.2 Method

Our method is essentially based on the joint application of two established decomposition methods. We first perform a structural decomposition analysis (SDA) to identify the relative impact in size of the volatility of market power, the intermediate goods markets and the final goods markets on aggregate fluctuations. In the second part, we decompose sectoral and country-level comovement into a local and an aggregate component to quantify the impact of the underlying correlation network structure for shock transmission. To this end, we use a principal component analysis (PCA). Finally, as a robustness check, we conduct a country fixed-effects regression analysis to analyze the impact of the structural characteristics of the network, namely bilateral trade intensities, on observed comovement in the aggregate and the purely local part.

2.2.1 A Structural Decomposition of Value Added Change

We start out from the traditional input-output analysis specification following Leontief (1986) to define the input coefficients, also referred to as technical coefficients, as $a_{ij} = x_{ij}/x_j$, which is the quantity of the output of sector i used by sector j per unit of its total output x_j. A complete set of the input coefficients of all sectors of a given economy is called the structural or input-output matrix A. Using this, an economy's production network, mirrored by the Leontief inverse $L = (I - A)^{-1}$, can be derived where I is the identity matrix of appropriate dimensions.[2] The Leontief inverse represents the gross outputs generated at all stages of the production process for one unit of final demand, $\vec{f}$.[3] The vector of output levels (production) in N industries can be expressed as $\vec{x} = (I - A)^{-1}\vec{f}$. Following Timmer et al. (2015), we derive the accounting identity for the vector of sectoral value added levels $\vec{v}$ by

$$\vec{v} = F(I - A)^{-1}\vec{f}, \tag{2.1}$$

where F represents a diagonal matrix of value added to gross output ratios (v_i/x_i) in all industries of the countries under investigation. For independent F, L and $\vec{f}$, the total differential reads

$$d\vec{v} = dF \cdot L \cdot \vec{f} + F \cdot dL \cdot \vec{f} + F \cdot L \cdot d\vec{f}. \tag{2.2}$$

In line with Wolff (1985) and Dietzenbacher and Los (1998), we apply mid-point

[2] The spectral radius of A is strictly less than unity, which implies that L can be expressed as the infinite sum of the powers of input-output matrix A, illustrating that the (i, j) element of the Leontief inverse measures the importance of industry j as a direct and indirect input supplier to industry i.

[3] Ignoring empirically negligible transportation costs (Timmer et al., 2015), the country-level gross value added is equivalent to its GDP.

weights as the most accurate discrete approximation for (2.2). This allows us to additively decompose the change in value added $\Delta\vec{v}_t$ into the following four constituent SDA components with abbreviations in brackets

$$\begin{aligned}\Delta\vec{v}_t =& \Big(\frac{1}{2}(\Delta F_t \cdot L_{t-1} \cdot \vec{f}_{t-1}) + \frac{1}{2}(\Delta F_t \cdot L_t \cdot \vec{f}_t)\Big) \quad \Big\} (\Delta VAR_t) \\ &+ \Big(\frac{1}{2}(F_{t-1} \cdot \Delta L_t \cdot \vec{f}_{t-1}) + \frac{1}{2}(F_t \cdot \Delta L_t \cdot \vec{f}_t)\Big) \quad \Big\} (\Delta INTER_t) \\ &+ \Big(\frac{1}{2}(F_{t-1} \cdot L_{t-1} \cdot \Delta\vec{f}_t) + \frac{1}{2}(F_t \cdot L_t \cdot \Delta\vec{f}_t)\Big) \quad \Big\} (\Delta FIN_t) \\ &+ \epsilon_t,\end{aligned} \tag{2.3}$$

where ΔVAR_t refers to the value added share effect, ΔINTER_t to the inter-industry effect and ΔFIN_t to the final goods effect, and where ϵ_t is an error term. The error term ϵ, capturing neglected interaction effects by the discrete approximation, is typically negligible in size.

The *value added share effect* is the proportion of change in $\vec{v}$ that would occur if the value composition the share of value added to gross output changed and all other components remained constant. This is especially relevant because the entries on the diagonal matrix F can be interpreted as average cost markups μ over input costs. To see this, consider the following accounting relation for the assumption of a homogeneous good by sector i: $v_i/x_i = (p_i \cdot q_i - u_i \cdot q_i)/(p_i \cdot q_i) = (p_i - u_i)/p_i = \mu_i$, where p_i is the unit price, q_i the quantity of i goods sold which cancels out, and u_i the unit price of sector i. This notion of a markup over average costs is quite established in the more heterodox literature (Puty, 2018), especially within Kaleckian macroeconomics (Feiwel, 1975). However, it is distinct from the more mainstream use, where a markup over marginal costs is typically defined (Syverson, 2019). The evident advantage of the average cost markup is that it is based purely on observables. In contrast, marginal cost markups need to be estimated, usually via a pre-specified production function, and are therefore conditional on restrictive assumptions about the returns to scale of this function. Puty (2018) also shows that marginal cost markups may lead to counterintuitive results with regards to their cyclicality. ΔVAR_t therefore relates to a change in realized markups of a given sector, where increases in market power lead to a higher value added achieved per unit of gross output, and thus to the market power of a sector in its final goods market.

The *inter-industry effect* captures the effect of a change in the matrix of technological coefficients, and thus the entries of the Leontief inverse, which can be viewed as a change in the necessary inputs to generate the output of a given sector. Since the entries of the Leontief inverse represent market outcomes for the intermediate goods used in the production of a specific sector, changes may be due, among other things, to substitution, innovation or reorganization of production (Wolff, 1985). As with changes within the final goods sector, we remain agnostic about the precise reasons for change, and interpret ΔINTER_t as the observed realization of the combined change in intermediate market conditions.

The *final goods effect* describes the proportion of change that would occur if only final demand changed and all other components remain constant. This component therefore accounts for the direct impact of increases or decreases in sales of products for final use, especially consumption and investment goods. In contrast to the inter-industry effect, this component thus accounts for everything on the external market for goods used for consumption in contrast to the internal markets for input goods. For a later analysis, we need to decompose Equation (2.3) even further, using the same algebraic decomposition technique, to carve out the different drivers of value added change on a more fine-grained level with respect to different categories of final goods constituting ΔFIN_t such as goods for household consumption and investment goods. This results in a more detailed decomposition of fifteen components, which can be found in Equation (2.11) in the Appendix.

2.2.2 Distinguishing Aggregate and Local Comovement

This section builds on the difference between first-order and higher-order linkages, introduced by Acemoğlu et al. (2012). First-order linkages refer to direct links between sectors, causing the comovement of pairs of them. However, volatility may spill over, even due to higher-order linkages. In other words, interdependent reactions may spill over to directly connected industries, as well as to other sectors that are not connected directly, but are intermediaries in the intermediate goods market or have strong correlation prerequisites in the final goods market. Baqaee and Farhi (2019) show that these higher-order linkages can theoretically be crucial for aggregate fluctuations. To assess which fraction of variation per component is driven by higher-order rather than first-order bilateral behavior, we apply a PCA to both relevant SDA component time series and the GDP series.

Garin et al. (2018) show that the first principal component can loosely be interpreted as an aggregate shock. This is due to the fact that, by definition, the first principal component is the linear combination of time series that is able to explain the largest fraction of aggregate variance, and thus extracts the comoving component of the time series in question. Foerster et al. (2011), however, demonstrate that estimating the impact of idiosyncratic events on aggregate volatility by PCA leads to downwards biased estimates, since shocks may propagate through trade and input linkages, creating spurious comovement in sectoral time series. We therefore refer to the first principal component as aggregate behavior, including truly aggregate behavior due to aspects such as aggregate sentiments and idiosyncratic shocks propagating through the whole network due to higher-order linkages.

Importantly, the first principal component excludes any behavior that is local and confined to only a subset of the 1680 industries in our sample. We call all movement that is not due to this first principal component "localized comovement", because it relates to idiosyncratic shocks that do not impact the whole networked structure. Economically plausible candidates for local pairwise comovement between sectors are either linkages between pairs of sectors, or common aggregate developments affecting all sectors simultaneously. For local comovement, and therefore first-order linkages, a myriad of possible causes, including foreign affiliates of multinational corporations (Kleinert et al., 2015; Tweedle, 2018; Di Giovanni et al., 2018) and demand complementarities (Drozd et al., 2017) are discussed in the literature. One probable cause of aggregate movements for all sectors is sentiment spillovers across countries, which was shown theoretically by Hohnisch and Westerhoff (2008). Importantly, shock propagation through higher-order linkages causing a large subset of industries to be affected by a shock in final goods markets, may also induce aggregate movements, as opposed to first-order pairwise linkages (Acemoğlu et al., 2012).

We employ the method and notation outlined in Livan et al. (2015), which allows us to differentiate between synthetic local and aggregate time series per SDA component. Let y_{it} be the realization for one SDA component at time t for sector i with N industries and a time series length per sector T. Here, y_{it} is mapped onto a set of orthogonal principal components e_l by:

$$z_{it} = \frac{y_{it} - \mu_i}{\sigma_i} = \sum_{l=1}^{T} \sqrt{\lambda_l} V_i{}^{(l)} e_{lt}, \tag{2.4}$$

where μ_i and σ_i denote the mean and standard deviation of the N different component time series, $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_T \geq 0$ denote the nonzero variance-covariance eigenvalues, and $V_i^{(l)}$ refers to the ith component of the lth normalized eigenvector. This allows us to isolate the impact of each principal component by inspecting the eigenvalues and deducing the variation explained by it. In addition, we can construct a synthetic mean-centered time series of structural components from Equation (2.4) by

$$\tilde{x}_{it} = \sigma_i \sum_{l=2}^{T} \sqrt{\lambda_l} V_i^{(l)} e_{lt}, \tag{2.5}$$

where we drop the first principal component to adjust for the impact of aggregate shocks common to all sectors within our sample. Thus, the resulting synthetic time series can be interpreted as purely local (mean-centered) time series per industry.

2.3 Data

The empirical application of our method relies on data from the World Input-Output Database (WIOD). The second release issued in 2016 covers 43 countries from 2000 to 2014. Here, we focus on the 30 available European countries, all 28 countries of the European Union[4], plus Norway and Switzerland (EU30). The countries seem to be ideal candidates for our analysis. After all, geographically and economically speaking, they represent a large and closely connected zone that, considering the recent crises, was characterized by high macroeconomic volatility. One main reason for the relevance of the interconnectedness of the countries in our sample is that some of them have a common currency. Indeed, membership of the European Monetary Union (EMU) seems to be correlated with increasing openness with EMU and non-EMU countries alike, as Larch et al. (2018) shows, also using the WIOD. Nonetheless, these countries are also reasonably heterogeneous in their complexity of production and growth model (Gräbner et al., 2020). We are therefore confident that our findings reflect general characteristics, and are not biased due to special features of a skewed sample.

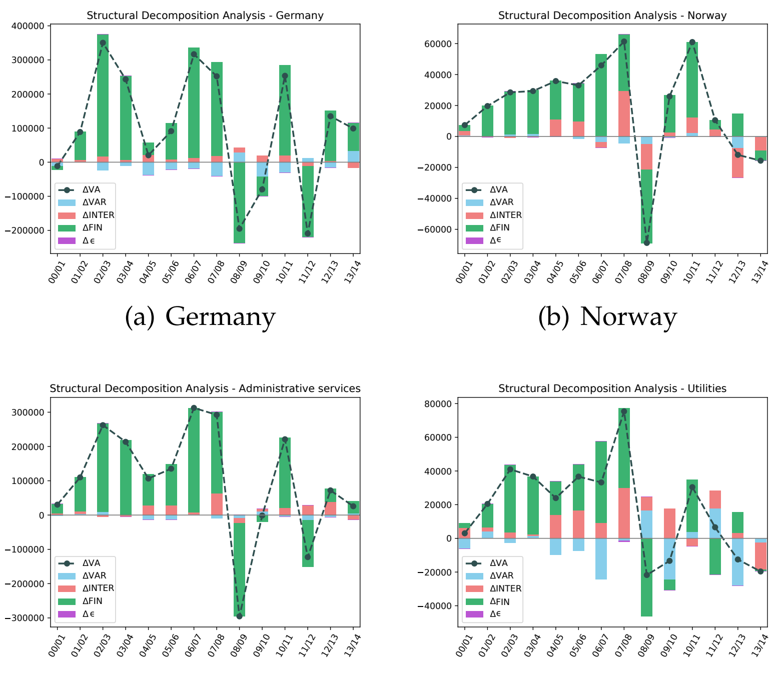
The World Input-Output Tables (WIOTs) have an industry-by-industry format, reflecting economic linkages across industries. These tables provide details for 56 industries, mostly at the two-digit ISIC rev. 4 level, or groups thereof. The columns of WIOTs contain information on production processes, while rows show the distribution of the output of industries over user categories. Products within this framework can be used as intermediates by other industries or as final goods by households, governments or firms. A critical accounting identity of the WIOTs is that the gross output of each industry (last element of each column) is equal to the sum of all users of the output from that industry (last element of each row). The specific structure is described in Timmer et al. (2015). One major limitation of the data, however, is the underlying fiction that each sector produces only a single undifferentiated good, which assumes homogeneity within industries.

[4]Including the UK, which is officially no longer a member of the EU.

2.4 Results

We begin our analysis with by testing the validity of the SDA method. To this end, we examine whether decomposition plausibly picks up the peculiarities of countries and sectors. After establishing this through illustrative examples, we conduct our main analysis and apply both the SDA and PCA to the dataset, as described in Section 2.2.1.

Figures 2.1 (a) and (b) show the time series of all components and the value added series on a country-level aggregation. Visual examination reveals that changes in ΔFIN are quantitatively the most important drivers of value added changes over time in both Germany and Norway. This finding is independent of categorization, and holds across all sectors and countries. It may therefore provide initial evidence that the assumption of a constant Leontief inverse or static production network may perform rather well. However, the relative contribution of each component to value added changes depends on the relative industry composition per country, with Norway and Germany providing an instructive example.



(a) Germany (b) Norway

(c) Administrative Sector (d) Utilities Sector

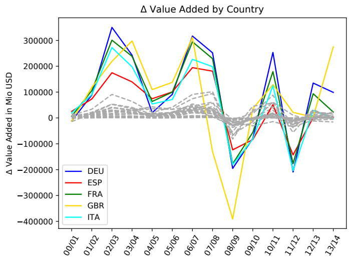
Figure 2.1: Country-level and sector-level SDA plots between 2000 and 2014, showing Germany and Norway at the country level and the EU30 administrative and utilities industry on the sector level. The relative importance of components differs widely across countries and sectors, where aggregate volatility in Norway is much more dependent on volatility in intermediate goods markets, and the utilities sector depends to a greater extent on the changes in realized markup and volatility in intermediate goods markets than the rather typical administrative sector.

In Norway, the impact of inter-industry changes on aggregate volatility is relatively high. This may be due to the fact that production in Norway is strongly dependent on oil as an input, as shown recently by Bergholt et al. (2019), rendering inter-industry markets susceptible to rather volatile world market oil prices. In the sectoral aggregation, Figures 2.1 (c) and (d), the administrative sector is a rather typical sector,being highly dependent on final goods markets, whereas in the utilities sector, both inter-industry effects and fluctuations in realized market power contribute relatively strongly to sectoral fluctuations. Again, this seems plausible, as the utilities industry is dependent on multiple inputs, rendering it susceptible to volatile world market energy prices. This could explain the relatively strong impact of both inter-industry effects and markups. Thus, for networked structures to matter, it seems plausible to evaluate all three components, depending on the country or industry setting. However, intermediate and final goods market structures particularly need to be taken into account when considering the source and propagation of aggregate volatility.

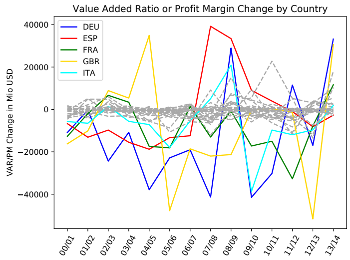
2.4.1 Counterfactual Volatilities

The additive functional form of the SDA enables us to construct counterfactual scenarios for evaluating the impact of each component on country-level volatility. Figure 2.2 shows the time series of all countries' components (b) to (d) in the sample that add to the ΔVA series in (a). To perform the counterfactual analysis, we simply subtract the relevant components from the empirical time series to compare these counterfactual cases, see Figure 2.3, with benchmark volatility, measured as the standard deviation of value added growth rates.

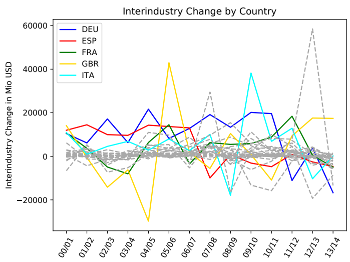
Figure 2.3 shows that subcontracting the volatility in final goods markets heavily influences the volatility of value added growth, represented as $\sigma(g(\text{VAR} + \text{INTER}))$, compared to the benchmark of $\sigma(g(\text{GVA}))$. This is not surprising when comparing only the time series of all three components and the benchmark in Figure 2.2, where the final goods component ΔFIN exhibits by far the largest amplitude on the same order of magnitude as value added differences, whereas the other components do not. This exercise seems to suggest that volatility in final goods markets, rather than in intermediate goods markets (ΔINTER) or changes in market structure (ΔVAR), is mainly behind aggregate movements. Here, it is important to note that we consider the implicit assumptions of the production network literature to have generally been met, and that intermediate goods markets per se, their actual size in terms of market volume, do not contribute much to aggregate volatility.



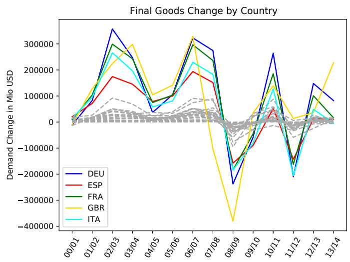
(a) ΔVA



(b) ΔVAR



(c) ΔINTER



(d) ΔFIN

Figure 2.2: First-difference time series of main components for all European countries with (a) changes in GDP, (b) changes in realized markups, (c) changes in intermediate goods markets and (d) changes in final goods markets. Changes in intermediate goods markets and markups are rather idiosyncratic, while the series for GDP and final goods markets are strongly synchronized. Subfigure (a) and (d) also exhibit by far the largest fluctuations with amplitude of equal order of magnitude, while changes in markup and intermediate goods markets exhibit much smaller fluctuations in level.

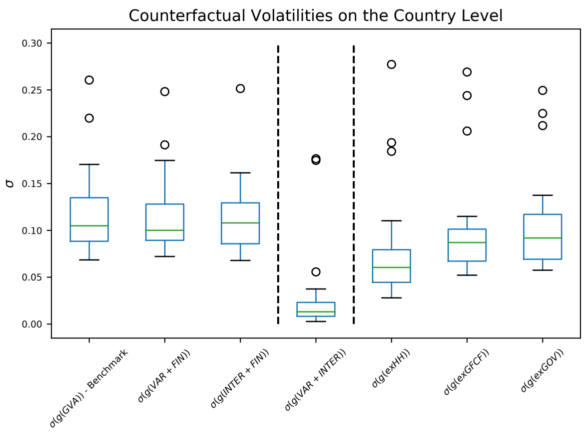


Figure 2.3: Counterfactual scenarios to test the relevance of each component for aggregate volatility at the country level. The benchmark is calculated as the standard deviation σ for growth rates of GDP per country, $\sigma(g(VAR + FIN))$ shows the volatility, excluding the volatility of intermediate goods markets, $\sigma(g(INTER + FIN))$ excluding volatility in realized markups, and the highlighted $\sigma(g(VAR + INTER))$, excluding the volatility of final goods markets, which has by far the greatest impact across countries. The three charts to the right each exclude the volatility of a specific type of final good, with $\sigma(exHH)$ excluding final goods for household consumption, $\sigma(exGFCF)$ investment goods and $\sigma(exGOV)$ final goods for the government.

Since we do not assess the network topology of the underlying structure or the effect of shocks, e.g. to the size of certain nodes within the heavy skewed distribution of connectivity within the network, we are unable to draw conclusions about their effect on the aggregate level in this setup. Although this effect has been established and reported in Chapter 1, for example, it is not the subject of this investigation. The fact that developments in the sheer size of final goods markets largely drive aggregate volatility as well as the synchronization of business cycles invites a more detailed breakdown of the final goods component in (2.3). We are able to identify household consumption and investment goods markets as the central drivers, where the other components are negligible in magnitude. The impact of these components can also be seen when performing the counterfactual analysis on the right panel of the Box-Whisker charts in Figure 2.3. Excluding household demand $\sigma(g(exHH))$ as a single source of aggregate volatility drives down volatility in all countries. The same is true, albeit with a lesser impact, for gross fixed capital formation (GFCF) or volatility in investment goods markets $\sigma(g(exGFCF))$. Intriguingly, excluding government spending in $\sigma(g(exGOV))$ from the analysis as the rightmost Box-Whisker chart has essentially no effect on measured country-specific volatility, contrasting with Keynesian prescriptions for countercyclical economic policy.

2.4.2 Local and Aggregate Comovement

A natural follow-up question from these two results is thus how much of the impact on country-level volatility do these components owe to aggregate behavior across all 30 countries in Europe, and how much is due to local developments, i.e. comovement caused by something other than the production network literature refers to as higher-order linkages. By addressing this aspect, we seek to establish a link to the existing production network literature, where variability within final goods markets are typically considered to be exogenous explanatory circumstances. We are keen to establish the fact that, also in the narrow setting at hand, higher-order linkages play a decisive role in distinguishing aggregate and local behavior in GDP time series and final goods market, represented as correlation networks. To achieve this, we perform the analysis described in Section 2.2.2.

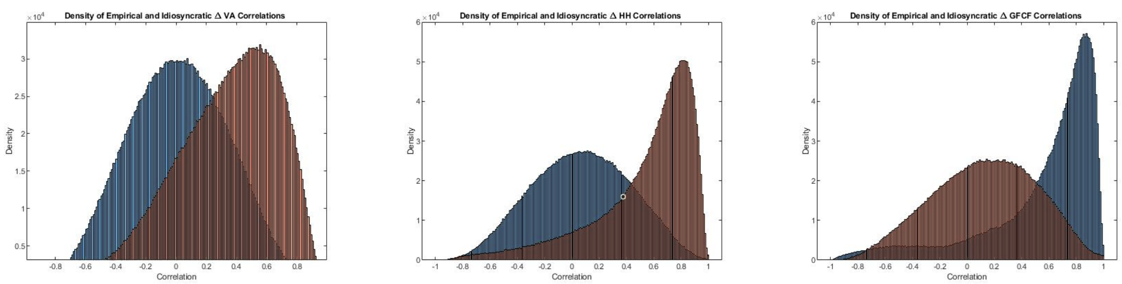
As Table 2.1 shows, we find that most of the variation in EU30 GDP time series is due to aggregate shocks, affecting the 1680 sectors in all EU30 countries at once and explaining more than 60% of volatility. This is a most unexpected finding, as it shows that comovement of the whole sample across sector- and country-level peculiarities explains the synchronicity in European business cycles. Thus, a common European business cycle seems to be behind most country-level volatility in our dataset.

Component	Var(Δ)	Due to 1st Component	Residual Variance
ΔVA	$9.161 \cdot 10^9$	$5.7835 \cdot 10^9$ (63.1)	$3.3776 \cdot 10^9$ (36.8)
ΔHH	$2.5457 \cdot 10^9$	$1.6850 \cdot 10^9$ (66.21)	$8.5998 \cdot 10^8$ (33.8)
ΔGFCF	$1.4452 \cdot 10^9$	$1.0957 \cdot 10^9$ (75.8)	$0.3496 \cdot 10^9$ (24.2)

Table 2.1: Results from the principal component analysis per SDA component, with the first column reporting the explained variance due to the first aggregate component and the right-most column showing the residual variance left unexplained by the first component. Fraction of variance explained in brackets in percent. The results show that the aggregate component is responsible for most of the observed comovement in value added and in the most important categories of final goods.

This European business cycle, in turn, is primarily a common business cycle of investment and household demand which, taken on their own, are even more dependent on aggregate comovement by about 75% and 66%, respectively. It comes as no surprise that these components exhibit such a large dependence on aggregate movement, since the value added time series are also simultaneously affected by the rather idiosyncratic ΔVAR and ΔINTER components, mitigating the impact of final goods market volatility on value added.

Our approach of combining the two methods also enables us to trace the origins of strong comovement between value added time series, as exemplified by the high Pearson correlation coefficients. As described in Equation (2.5), we construct synthetic time series to counter the impact of aggregate shocks, and inspect the correlation patterns before and after correction. This allows us to isolate the impact of locally confined comovement relative to the benchmark of both aggregate and local comovement.



(a) ΔVA (b) ΔHH (c) $\Delta GFCF$

Figure 2.4: Densities of empirical (red) and purely local (blue) Pearson correlations between about 1.3 million distinct pairs of distinct sectors for value added, volatility in final goods for household consumption and for investment goods. All empirical densities exhibit a left-skewed distribution, and thus a systematic tendency for comovement that is greatly reduced when the underlying time series are only constructed for the local component. There, distributions are almost Gaussian, with expected value close to zero, indicating almost no systematic tendency for comovement in the system.

As we show in Figure 2.4, the distribution of the correlation coefficients between distinct sectors is strongly left-skewed, indicating a systematic tendency of comovement within the sample in terms of value added, GFCF and HH demand. However, this tendency is merely due to the aggregate developments within all sectors. If we extract only the local time series, the density is well approximated by a Gaussian for all three cases, with mean close to zero for the value added series. This shows that comovement due to local comovement is indeed tiny, as the resulting distribution is close to the benchmark of correlations for iid random variables. The Maximum Likelihood Estimation (MLE) for a Gaussian shows that the mean of correlations is significantly different from 0 at the 5% significance level. This is mainly due to the rather tight bounds around the estimate from the large number of observations, where the effect size of around 0.01 (mean correlation of local ΔVA series) is still tiny. The mean Pearson correlation is also strongly reduced for household demand (to about 0.08) and GFCF (to about 0.15), showing that local comovement is not completely irrelevant, but much less important than the impression that uncorrected correlation coefficients would suggest. Our findings thus indicate, that the overwhelming role of aggregate behavior extending to the whole networked structure due to higher-order connections is remarkable. The relevance of local comovement is still greater for investment demand in line with the intuition that pairwise investment coordination between affiliates of multinational corporations should matter more for the GFCF component than demand complementarities across households. Nonetheless, when aggregate movements are properly accounted for, comovement is greatly reduced even in GFCF and HH demand.

2.4.3 First- and Higher-Order Linkages

Uncorrected comovement is very homogeneous, whereas differences only arise for the local component. Structural differences such as trade intensities thus only culminate in differences in comovement for the local component, whereas the correlation networks seems to create a tendency for common comovement across all sectors. In principle, however, the local component could still be heavily dependent on higher-order connections, leading to comovement only within a subset of industries. To assess if the local component indeed picks up on the impact of pairwise linkages, we use an indirect approach that is well established within the trade literature. We conjecture that, if the local component indeed picks up pairwise comovement, pairwise trade intensities should have a stronger impact on local comovement than on the benchmark of uncorrected correlations, where measured correlations are confounded with comovement from higher-order linkages. In particular, we follow the seminal paper on the trade-comovement nexus by Frankel and Rose (1998) and analyze the impact of trade intensities on both measures of correlation by inspecting the pertaining slope coefficient. Trade intensities are defined as

$$\text{Trade}^{cd} = \frac{1}{T}\sum_t \frac{X_t^{cd} + X_t^{dc}}{Y_t^c + Y_t^d}, \tag{2.6}$$

where X_t^{cd} are the exports of country c to d at t, or vice versa, X_t^{dc}. Y_t^c, Y_t^d is the GDP of country c or d at t, and $T = 15$ is the time series length. We regress the trade intensities on comovement, measured as the correlation in first-difference value added series, both uncorrected (corr (ΔVA)) and, for the synthetic series, without the aggregate component (corr ($\Delta\widetilde{VA}$)). In line with the pertinent literature, we consider both the level of trade intensities and their logarithm, since the log-transformation may counter the disproportionate impact of outliers on the slope coefficient sometimes observed in the data (Kose and Yi, 2006; Di Giovanni and Levchenko, 2010). We include country-fixed effects to account for the unobserved heterogeneity in the model, leading to the following four regression specifications:

$$I(a): \ \text{corr}(\Delta VA)^{cd} = \beta_0 + \beta_1 \text{Trade}^{cd} + \mu_c + \mu_d + \epsilon^{cd}, \tag{2.7}$$

$$I(b): \ \text{corr}(\Delta\widetilde{VA})^{cd} = \beta_0 + \beta_1 \text{Trade}^{cd} + \mu_c + \mu_d + \epsilon^{cd}, \tag{2.8}$$

$$II(a): \ \text{corr}(\Delta VA)^{cd} = \beta_0 + \beta_1 \ln \text{Trade}^{cd} + \mu_c + \mu_d + \epsilon^{cd}, \tag{2.9}$$

$$II(b): \ \text{corr}(\Delta\widetilde{VA})^{cd} = \beta_0 + \beta_1 \ln \text{Trade}^{cd} + \mu_c + \mu_d + \epsilon^{cd}, \tag{2.10}$$

where μ_c and μ_d are vectors of country-fixed effects with dimensions $N \times 1$, and ϵ_t is an $N \times 1$ vector of disturbances assumed to be Gaussian with zero mean and variance σ_ϵ^2. We estimate regression equations for all $N = 435$ distinct pairs of countries. The results are summarized in Table 2.2. We find that the slope coefficient $\hat{\beta}_1$ is highly significant in all estimations, but differs widely between regression specifications. In the level and the log-transformed specification, we find an almost fivefold increase in the slope coefficient when going from the empirical correlation to the correlation within the local component as the dependent variable. Pairwise trade intensities thus seem to have a much greater impact on local comovement than on composite comovement. This huge increase is exactly what would be expected if the local component predominantly captures comovement caused by pairwise linkages, where the size of the pairwise transmission channel (the trade intensity) should exhibit a much greater impact on the purely local case than the case of composite comovement with confounding effects due to higher-order linkages.

	I			II	
	(a)	(b)		(a)	(b)
Trade^{cd}	36.657***	168.34***	$\ln$ Trade^{cd}	0.017517***	0.087064***
	(7.6348)	(37.468)		(0.0032708)	0.015979)
Intercept	0.8933***	0.14169	Intercept	1.0478***	0.90309***
	(0.027939)	(0.13711)		(0.037603)	(0.1837)
N	435	435	N	435	435
R^2	0.750	0.286	R^2	0.753	0.301
Adjusted R^2	0.731	0.233	Adjusted R^2	0.734	0.249
F Statistic	40.30***	5.38***	F Statistic	41.00***	5.81***

Table 2.2: The measure of trade intensity is regressed on the measure of comovement for the uncorrected series (a) and the local component (b) only. Trade intensities are defined in levels (I) or in log-transformed levels (II). For both cases, the slope coefficient increases almost fivefold when going from the uncorrected to the local case, indicating the considerable relevance of pairwise trade linkages for comovement within the local component. Standard errors are given in parentheses, with significance indicated by * at the *** 1%, ** 5%, * 10% level.

With this evidence for the relevance of pairwise linkages in the local component, together with the emergent correlation patterns, we can thus conclude that pairwise linkages play only a secondary role in the observed comovement of European business cycles, which are mostly driven mainly by aggregate behavior and higher-order interdependencies.

2.5 Discussion and Conclusion

The literature on European business cycle comovement has typically focused on bilateral measures such as the Pearson correlation. This empirical approach naturally also lends itself to focus on pairwise spillovers to explain the remarkable degree of synchronicity across European business cycles. The production network literature has recently challenged this view and highlighted the role of cascades and higher-order linkages that can go beyond the direct impact of shocks mediated by pairwise linkages creating system-wide comovement. Since many of these contributions consider variability within the final goods markets as exogenous explanatory circumstances, we explicitly account for this gap by distinguishing aggregate and local behavior in GDP time series and final goods markets.

To achieve this, we use a structural decomposition analysis combined with a principal component analysis to provide empirical evidence at the European level that this networked view may be adequate with regards to its implicit structural assumptions and the predicted relevance of higher-order linkages. Considering the market volumes of the individual components under investigation, we demonstrated that aggregate volatility does indeed seem to originate primarily in final goods markets, and then propagate as aggregate behavior through a network of correlation linkages. In the process, volatility in final goods markets seems primarily to be due to fluctuations in investment goods and goods for household consumption, while realized final demand by the government seems to be of negligible importance for aggregate volatility, in contrast to Keynesian prescriptions.

We also provided evidence that aggregate shocks simultaneously affecting all 1680 sectors, and not pairwise sectoral linkages, explain the largest proportion of volatility. This could highlight the role of sentiment spillovers and higher-order effects of shock propagation. In contrast, purely pairwise comovement has rather a negligible aggregate impact. For the two most relevant sources of aggregate volatility, aggregate movements are much more important than first-order spillovers. Pairwise linkages seem to be more relevant in both household and especially investment demand than for the aggregate value added time series, which is also affected by counteracting inter-industry and profit margin components.

Our results thus show that the bilateral measures of association, which the trade literature has long focused on, may be misguided, at least in parts, because production networks and behavioral linkages in final goods markets are by their very nature much more topologically rich than what could be deduced from their constituent bilateral relations. The remarkable degree of synchronicity in EU30 business cycles may therefore depend on the underlying networked structure, not only in production itself, but also in its final markets, contributing to total output. It follows that impairments or advantages to single industries or countries may have much greater aggregate consequences for other market participants than its direct trade linkages might indicate. Our results thus suggest that policy measures aimed at increasing economic integration should not only focus on changing trade intensities, but also take into account how a change in bilateral trade intensities may influence other participants within a network of interdependent relationships, irrespective of them accruing within the intermediate or final goods markets.

With this in mind, our findings open a number of avenues for further research. For one, a natural experiment that could be exploited is Brexit, where recent studies have indeed found evidence for the heterogeneous relevance of sectors for trade relations, which are thus also differentially affected by the post-Brexit disintegration (Giammetti et al., 2020; Giammetti, 2020). Moreover, examining the implications of the results presented in this paper on a more fine-grained firm level could be a difficult yet worthwhile endeavor to investigate and further elaborating the method presented in this paper, and its applicability. An investigation at the firm level would enable us to go beyond the limiting assumption of one unitary good produced per sector. Excluding this assumption would allow to explicitly take into account substitution effects between similar goods produced within industries or to account for changes related to the technological paradigm shift as experienced by intermediate goods and final goods markets in recent years. Another possible angle would be to explicitly account for the effects of policy on economic integration during the period under investigation.

In sum, our findings suggest that essentially all EU30 countries predominantly follow a common business cycle and that this cyclicality is mainly driven by final goods markets, when considering the sheer volume size of all components contributing to total output. Moreover, similar to findings in the literature on linkages in production networks, it seems that final goods markets experience higher-order linkages, guiding them towards an aggregate behavior, rather than relying on local, bilateral linkages leading to a common business cycle in Europe.

2.A Appendix

A Further Decomposition of Value Added

$$
\begin{aligned}
\Delta v_t = &\left(\frac{1}{2}(\Delta F_t \cdot L_{t-1} \cdot \vec{f}^{hh}_{t-1}) + \frac{1}{2}(\Delta F_t \cdot L_t \cdot \vec{f}^{hh}_t)\right) \\
&+ \left(\frac{1}{2}(F_{t-1} \cdot \Delta L_t \cdot \vec{f}^{hh}_{t-1}) + \frac{1}{2}(F_t \cdot \Delta L_t \cdot \vec{f}^{hh}_t)\right) \\
&+ \left(\frac{1}{2}(F_{t-1} \cdot L_{t-1} \cdot \Delta\vec{f}^{hh}_t) + \frac{1}{2}(F_t \cdot L_t \cdot \Delta\vec{f}^{hh}_t)\right) \\
\\
&+ \left(\frac{1}{2}(\Delta F_t \cdot L_{t-1} \cdot \vec{f}^{np}_{t-1}) + \frac{1}{2}(\Delta F_t \cdot L_t \cdot \vec{f}^{np}_t)\right) \\
&+ \left(\frac{1}{2}(F_{t-1} \cdot \Delta L_t \cdot \vec{f}^{np}_{t-1}) + \frac{1}{2}(F_t \cdot \Delta L_t \cdot \vec{f}^{np}_t)\right) \\
&+ \left(\frac{1}{2}(F_{t-1} \cdot L_{t-1} \cdot \Delta\vec{f}^{np}_t) + \frac{1}{2}(F_t \cdot L_t \cdot \Delta\vec{f}^{np}_t)\right) \\
\\
&+ \left(\frac{1}{2}(\Delta F_t \cdot L_{t-1} \cdot \vec{f}^{gov}_{t-1}) + \frac{1}{2}(\Delta F_t \cdot L_t \cdot \vec{f}^{gov}_t)\right) \\
&+ \left(\frac{1}{2}(F_{t-1} \cdot \Delta L_t \cdot \vec{f}^{gov}_{t-1}) + \frac{1}{2}(F_t \cdot \Delta L_t \cdot \vec{f}^{gov}_t)\right) \\
&+ \left(\frac{1}{2}(F_{t-1} \cdot L_{t-1} \cdot \Delta\vec{f}^{gov}_t) + \frac{1}{2}(F_t \cdot L_t \cdot \Delta\vec{f}^{gov}_t)\right) \\
\\
&+ \left(\frac{1}{2}(\Delta F_t \cdot L_{t-1} \cdot \vec{f}^{gfcf}_{t-1}) + \frac{1}{2}(\Delta F_t \cdot L_t \cdot \vec{f}^{gfcf}_t)\right) \\
&+ \left(\frac{1}{2}(F_{t-1} \cdot \Delta L_t \cdot \vec{f}^{gfcf}_{t-1}) + \frac{1}{2}(F_t \cdot \Delta L_t \cdot \vec{f}^{gfcf}_t)\right) \\
&+ \left(\frac{1}{2}(F_{t-1} \cdot L_{t-1} \cdot \Delta\vec{f}^{gfcf}_t) + \frac{1}{2}(F_t \cdot L_t \cdot \Delta\vec{f}^{gfcf}_t)\right) \\
\\
&+ \left(\frac{1}{2}(\Delta F_t \cdot L_{t-1} \cdot \vec{f}^{invent}_{t-1}) + \frac{1}{2}(\Delta F_t \cdot L_t \cdot \vec{f}^{invent}_t)\right) \\
&+ \left(\frac{1}{2}(F_{t-1} \cdot \Delta L_t \cdot \vec{f}^{invent}_{t-1}) + \frac{1}{2}(F_t \cdot \Delta L_t \cdot \vec{f}^{invent}_t)\right) \\
&+ \left(\frac{1}{2}(F_{t-1} \cdot L_{t-1} \cdot \Delta\vec{f}^{invent}_t) + \frac{1}{2}(F_t \cdot L_t \cdot \Delta\vec{f}^{invent}_t)\right) \\
&+ \epsilon_t,
\end{aligned}
\tag{2.11}
$$

where final demand $\vec{f}$ is further decomposed for $\vec{f}^{hh}$ private households, non-profits $\vec{f}^{np}$, government $\vec{f}^{gov}$, investment $\vec{f}^{gfcf}$ and changes in inventories $\vec{f}^{invent}$.

Part II

Essay on Network Dynamics in Economics

Chapter 3

An Empirical Model of Buyer-Supplier Matching in Dynamic Trade Networks

Co-written by Philipp Mundt

Abstract

Employing a stochastic actor-oriented model for network dynamics, we study buyer-supplier matching within and across countries of the OECD. Our model approaches these dynamics from the perspective of individual nodes and thus enables us to identify the driving forces behind the formation of trade relationships. Compared to more standard econometric techniques, one major advantage of the actor-oriented model is that it resolves the endogeneity problem in dynamic trade networks that arises when structural network characteristics determine the formation of ties. Extending this framework to the multiplex dynamics of co-evolving one-mode and two-mode networks, we do not only model the formation of trade relationships between OECD countries in the one-mode network but also their affinity to import from and export to non-OECD economies via the consideration of two additional two-mode networks. The empirical value of our approach is demonstrated by fitting the model to data from the OECD Inter-Country Input-Output Tables. Building on this approach, we find that geography, supplier heterogeneity in terms of productivity, labor costs, output share, and economic complexity, trade frictions and costs of trade, as well as technological similarities and complementarities determine the formation of trade relationships. At the same time, the analysis also shows that trade relationships across OECD countries are strongly related to the countries' import and export activities in the two bipartite networks.

3.1 Introduction

The focus of research in the field of international trade has largely shifted in recent years from the study of aggregate exchange between countries towards a more fine-grained analysis of the microeconomic origins of trade relationships. Fueled by the growing availability of detailed trade and production data, current research provides a detailed account of trade patterns on the level of disaggregated sectors or even individual firms. At the same time, the pertinent literature has started to approach international trade from a network perspective because it is well-known that intermediate inputs such as raw materials, parts, and supporting services account for the majority of today's international trade, especially for highly industrialized economies.

Consistent with the view that modern production is organized in global value chains, this perspective acknowledges that producers are embedded in a complex network of upstream suppliers and downstream customers and has contributed to a better understanding of cross-country co-movement in growth and inter-industry reallocation (Shea, 2002; Melitz, 2003; Foerster et al., 2011), properties of input-output architecture (Acemoğlu et al., 2012; Carvalho et al., 2016; Bernard et al., 2019b), propagation of supply- and demand-side shocks (Acemoglu et al., 2016; Barrot and Sauvagnat, 2016; Magerman et al., 2016; Baqaee and Farhi, 2019), relevance of inefficient or imperfect markets (Jones, 2011; Kikkawa et al., 2019; Liu, 2019), buyer-supplier matching (Atalay et al., 2011; Bernard and Moxnes, 2018), and firm efficiency and performance (Tintelnot et al., 2018; Bernard et al., 2019b). Yet, although the structure of these firm or industry-level relationships is defined, to some extent, by exogenously given technological constraints, trade networks are an inherently dynamic phenomenon that change under the influence of producers' decisions to form and dissolve network ties. Nonetheless, there has been a surprising lack of attention to the endogenous dynamics of trade networks until recently, particularly in empirical research.[1]

To fill this void, the purpose of the present study is to provide a careful statistical analysis of the formation of input-output interlinkages in the global production network. In this respect, the paper makes three main contributions. First, we propose a stochastic actor-oriented model (SAOM) to identify the main driving forces behind the evolution of trade relationships within and across countries of the OECD, controlling for a rich set of variables reflecting different dimensions of supplier heterogeneity and country-level specificities (for an introduction to the SAOM, see Snijders et al., 2010). The stochastic actor-oriented approach, which was originally introduced in the literature on social networks by Snijders (1996, 2001) and later refined by Snijders et al. (2007, 2013), is an alternative to the prominent exponential random graph model (ERGM). Relative to the ERGM, which uses a global optimization problem to draw insights into the network formation process, the SAOM is a behavioral model that approaches network formation from the perspective of individual agents (or actors). These actors make sequential decisions on the formation or dissolution of net-

[1]Yet, progress has been made recently to understand the evolution and endogenous dynamics of production and information networks in economics (Carvalho and Voigtländer, 2014; Acemoglu et al., 2014; Zou, 2019; Acemoglu and Azar, 2020).

work ties based on their position in the network and their individual characteristics, implying that the SAOM provides insights into the microfoundations of network formation, which renders this framework as a promising tool to study endogenous trade networks.[2] Building on this approach, we find that geography and policies aiming at a reduction of barriers to trade, intersectoral differences in productivity, labor costs, output share, and economic complexity, technological similarities and complementarities, as well as several structural network properties such as reciprocity and triadic closure determine the formation of network linkages, which is largely consistent with recent theorizing on endogenous production networks (Atalay et al., 2011; Carvalho et al., 2016; Tintelnot et al., 2018).

Second, we extend the baseline framework to the modeling of two additional bipartite networks reflecting the import and export activities of economies in the OECD with non-OECD countries. This approach enables us to analyze patterns of structural change and dependencies in global value chains between industrialized and developing countries that would remain hidden in the study of a single (global) trade network where intra-OECD trade and import and export decisions to non-OECD economies would be integrated into a one-mode network. One of our most interesting results emerging from this extension is that free trade agreements (FTA) have an asymmetric effect on OECD and non-OECD countries. While FTA evidently increase export relationships from the OECD into non-OECD economies, they even reduce OECD import relationships from non-OECD countries, which casts doubts on the economic benefits of FTAs at least from developing countries' perspective. We also find strong evidence for common choices in the OECD with respect to the selection of trading partners in non-OECD economies and that the economic complexity of the economies determines their import and export activities. In particular, our results indicate that higher complexity of OECD economies leads to more export ties but reduces import ties with non-OECD countries.

Third, we study the multiplex dynamics of the co-evolving one-mode OECD network and the two bipartite trade networks reflecting import and export preferences from and to non-OECD economies to see to what extent these trade patterns may be

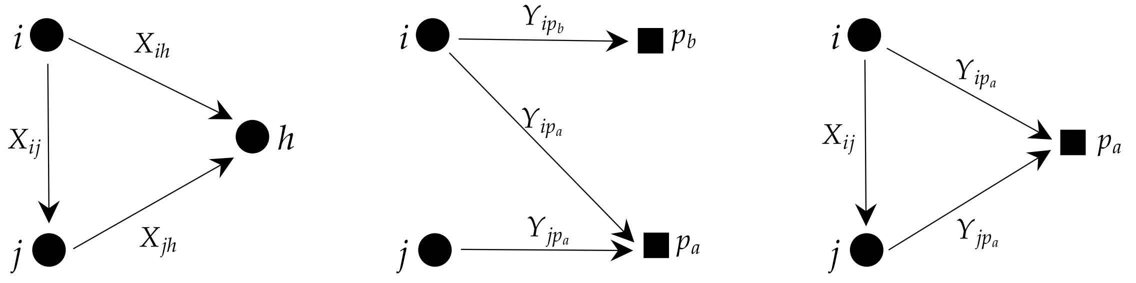
[2]Additionally, there are several technical advantages of the SAOM. Since the SAOM incorporates the information from network panel data which leads to additional statistical variation that can be exploited for parameter estimation, the model is less susceptible to convergence problems. As such SAOMs seem to be the natural choice for studying network dynamics (Broekel et al., 2014). Moreover, ERGMs are extremely computationally intensive which impairs their application in the context of large networks. ERGMs are also prone to identification problems because standard Markov Chain Monte Carlo (MCMC) or fixed density estimation methods with improved convergence properties cannot guarantee the uniqueness of the estimated parameters, even for large samples, as argued by Chandrasekhar and Jackson (2014).

mutually dependent. One interesting finding arising from this joint analysis is that OECD import activities from specific industries outside the OECD promote the creation of export activities to the same non-OECD industries, which is not true for exports into non-OECD countries. This finding would be consistent with the dependency theory predicting that developing countries tend to be locked into the unproductive pattern of producing and selling mostly primary goods to highly industrialized economies and importing finished goods (Ahiakpor, 1985; Blum et al., 2010).

The remainder of this article is organized as follows: Section 3.2 introduces the SAOM and sets out our empirical framework. Section 3.3 describes the data and discusses the main properties of the OECD production network and its non-OECD trade preferences. Section 3.4 presents the estimation results for the network formation and covariates effects using the SAOM, while Section 3.5 summarizes and concludes.

3.2 Model

This section lays out the empirical framework employed for the dynamic modeling of trade relationships. To estimate the driving forces behind the formation of these relationships, we consider a stochastic actor-oriented model (SAOM) in the spirit of Snijders (2001) that is extended to the modeling of co-evolving one-mode and two-



(a) One-mode network (b) Two-mode network (c) One-mode tie, two-mode agreement

Figure 3.1: Example of an one-mode and a two-mode network and their joint consideration in a multiplex network. Circles and squares represent nodes in two disjunct sets, representing sectors in OECD and non-OECD economies, respectively. In panel (a) links X represent trade relationships between OECD country-sectors, while ties Y in panel (b) illustrate import activities of OECD (non-OECD) countries in node set 1 from non-OECD (OECD) countries in node set 2. The graph in panel (c) represents a mixed triplet of both network types where OECD country-sector i has a trade relation with OECD country-sector j and both agree on the activity p_a with the same non-OECD country-sector.

mode (or bipartite) networks as in Koskinen and Edling (2012) and Snijders et al. (2013). In the field of social network analysis, one-mode networks are used to model relationships between entities, while bipartite networks are employed to model the association between these entities and certain activities (or preferences) such as, e.g., the individuals' membership in sport clubs (Fujimoto et al., 2018) or the choice of investment portfolios (Fang et al., 2019). In our application, graphically illustrated in Fig. 3.1, the one-mode network represents sector-level input-output relationships within and across OECD economies, while we consider two additional bipartite networks consisting of two types of nodes: industries in OECD countries and non-OECD economies. One of the two bipartite networks represents the export activity of OECD countries into non-OECD economies, while the second two-mode network describes the preferences of OECD economies to import from non-OECD countries. One major advantage of the SAOM approach is the possibility to analyze the multiplex dynamics of the co-evolving networks, enabling us to study potential interdependencies between the different types of trade networks. For example, we may ask to what extent trade relationships between OECD-countries determine their preference to export to or import from non-OECD countries and vice versa.

3.2.1 Rate and Evaluation Functions

The one-mode network consists of N OECD country-sectors that are connected by directed ties X_{ij} for $i, j \in N(i \neq j)$, where $X_{ij} = 1$ if country-industry $i = 1, ..., N$ buys intermediates from country-industry $j = 1, ..., N$, and $X_{ij} = 0$ otherwise.[3] The two node sets of the bipartite networks consist of the same N OECD country-sectors and a set of P trade preferences p for non-OECD import or export partners. Thus, the bipartite network has tie variables Y_{ip} for $i \in N$ and $p \in P$, with $Y_{ip} = 1$ if OECD country-sector i participates in the (import or export) trading activity with the preferred non-OECD partner p, and $Y_{ip} = 0$ otherwise.

The one and two-mode networks are observed at times $t_1, \ldots, t_M$, yielding $M \geq 2$ panel waves of observations $X(t_1), \ldots, X(t_M)$, and $Y(t_1), \ldots Y(t_M)$. To estimate the mechanism governing network formation, the SAOM splits the period between two consecutive network observations into a sequence of micro steps that can become arbitrarily small, implying that time in the model is continuous. At each of these micro time steps, one node (or actor) is randomly selected and has the opportunity to change one tie $X_{ij}(t)$ or $Y_{ip}(t)$ based on her individual preferences that depend on the current

[3]This specific choice of binary coding reflects the underlying assumption of the SAOM that network ties are controlled by the sender of the tie. When links represent monetary flows as in the present setting, the sender is the demanding sector which has a choice between different suppliers.

network structure and nodes' individual (exogenous) characteristics, implying that the process governing tie formation is a continuous-time Markov chain. Thus, the SAOM approaches the process of network formation from the viewpoint of individual actors whose decisions and dynamic interactions give rise to incremental changes of the network over time.[4]

In the model, the change of network ties is described by two components: a rate function for the modeling of change frequencies (determining the expected number of micro steps), and an evaluation function for the modeling of the change determination process. Here we simplify the analysis by assuming a constant rate of change across all actors. The specification of the evaluation function, on the other hand, reflects alternative hypotheses on the mechanism of network formation. For the one-mode network the evaluation function for actor i reads

$$f_i^X(x,y) = \sum_k \beta_k^X e_{ki}^X(x,y), \tag{3.1}$$

which represents a linear combination of effects $e_{ki}^X(x,y)$ that depend on the current network structure and node idiosyncrasies captured by appropriate covariates. These effects are weighted by the parameter β_k^X reflecting the strength of the k-th effect.[5] For example, network-related effects may represent tendencies towards reciprocity or triadic closure, while covariate-related effects may depend either on the individual characteristics of the actors (e.g., their productivity, size, economic complexity of the output) or variables depending on pairs of actors (so-called dyadic covariates such as geographic distance or the existence of free trade agreements between two countries). Similarly, the evaluation function for changes of the two-mode network is given by[6]

$$f_i^Y(x,y) = \sum_k \beta_k^Y e_{ki}^Y(x,y). \tag{3.2}$$

3.2.2 Estimation

The parameters of the model are estimated with the method of simulated moments, implemented in the RSiena package for R (Ripley et al., 2011), using Markov Chain Monte Carlo (MCMC) techniques (see, e.g., Snijders et al., 2007, p.13-19 for technical details). This routine employs a simulation algorithm that consists of three phases. The first phase determines the sensitivity of the statistics used to evaluate convergence of

[4]Thus, the SAOM provides a complementary view to the more time-honored exponential random graph model (ERGM) that estimates the network formation mechanism from an aggregate optimization problem in which the network is created at once rather than link by link as in the SAOM.

[5]More specifically, this weight quantifies the contribution to the log-probability of increasing the network variable by one unit when the respective effect is increased by one unit.

[6]See Appendices 3.A-3.A for further details.

the simulated networks to the empirical data, e.g. the total number of changed ties or the number of reciprocated relationships, with respect to the parameters. The second phase updates these provisional parameters in an iterative process by calculating the deviation of these moments from the observed values at the end of the observation period and adjusting the parameters using the stochastic approximation technique. The last phase tests whether the average statistics of a large number of network simulations are close enough to the target values. The criteria to evaluate the performance of the model is the overall maximum convergence ratio and the t-statistic for deviations from targets.

3.2.3 Model Specification

To test the influence of alternative hypotheses on the formation of trade relationships, we specify the canonical statistics for consideration in the evaluation functions in Eqs. (3.1)-(3.2) based on theoretical considerations. For the one-mode network, the different effects are summarized in Table 3.1.[7] By the type of the explanatory variable, we distinguish elementary (or structural network) effects and different forms of covariate effects.

Elementary effects capture the influence of the current network structure on the formation of ties. They are standard in the field of social network analysis, but the pertinent literature has also stressed the impact of the topology of input-output networks on the formation of trade relationships (e.g., Atalay et al., 2011). While the density effect captures the general tendency of producers to establish relationships, the reciprocity effect reflects the tendency to form bi-directional ties between two trading partners. Consistent with theoretical work arguing that producers preferably select indirect suppliers at close geodesic distance, we also consider the transitive triplets effect. For the focal country-industry i this effect is defined as the number of triadic relations in which i has a trading relationship to country-sectors j and k and additionally j has a trade relation to k. In economic terms, such tendencies for triadic closure have been attributed to innovation processes (Carvalho and Voigtländer, 2014) and the diffusion of information according to which firms use existing relationships to search for new business partners to reduce search and informational frictions (Chaney, 2014).[8] Moreover, to assess the empirical relevance of local hierarchical structures, we also consider the so-called 3-cycles effect, defined by the sum of paths of length three completing a cycle between three industries, in addition to the transitive triplets effect.

[7] Details on the measurement of the different effects are provided in Section 3.3.

[8] In the literature on social networks, this effect has been characterized as a “friends of friends become my friends” mechanism (Jackson and Rogers, 2007).

Level	Type	Definition	Hypothesis
Elementary effects			
Industry	Outdegree density	$\sum_j x_{ij}$	Tendency to have trade relationships.
Industry	Reciprocity	$\sum_j x_{ij}x_{ji}$	Tendency to have mutual relationships.
Industry	Transitive triplets	$\sum_{j,h} x_{ij}x_{ih}x_{hj}$	Tendency to select suppliers at network distance two.
Industry	3-cycles	$\sum_{j,h} x_{ij}x_{jh}x_{hi}$	Tendency for a circular relationship within a group of three.
Constant covariate effects			
Industry	Technological similarity	$\sum_j x_{ij}(sim^v_{ij} - \hat{sim^v})$	Producers select suppliers operating in the same industry.
Country	Regional similarity	$\sum_j x_{ij}(sim^v_{ij} - \hat{sim^v})$	Producers select suppliers operating in the same region.
Variable covariate effects			
Industry	Output share activity	$v_i x_{i+}$	Large producers have many suppliers.
Industry	Output share popularity	$\sum_j x_{ij}v_j$	Large producers are more frequently selected as suppliers.
Industry	Productivity activity	$v_i x_{i+}$	Productive producers have many suppliers.
Industry	Productivity popularity	$\sum_j x_{ij}v_j$	Productive producers are more frequently selected as suppliers.
Industry	Labor cost activity	$v_i x_{i+}$	Labor costs determine the tendency to form ties.
Industry	Labor cost popularity	$\sum_j x_{ij}v_j$	Producers select suppliers with competitive labor costs.
Country	Economic complexity activity	$v_i x_{i+}$	Producers selling complex products have many suppliers.
Country	Economic complexity popularity	$\sum_j x_{ij}v_j$	Producers selling complex products are more frequently selected.
Dyadic covariate effects			
Industry	Technological complementarity	$\sum_j x_{ij}(w_{ij} - \overline{w})$	Producers select suppliers selling complementary inputs.
Country	Geographical distance	$\sum_j x_{ij}(w_{ij} - \overline{w})$	Producers select suppliers at close geographical distance.
Country	Free trade agreements	$\sum_j x_{ij}(w_{ij} - \overline{w})$	Producers select suppliers from economies with trade agreements.

Table 3.1: Specification of the effects for the one-mode network. Network ties are denoted by x_{ij}, and v_i are actor-specific covariates. Dyadic covariates refer to a pair of nodes and are denoted by w_{ij}. A + sign denotes summation over the pertinent index. If not explicitly stated otherwise, we consider the deviation of covariates from their sectoral average to account for sectoral differences.

To investigate the influence of sector or country-level characteristics on the formation of trade relationships, we additionally consider individual and dyadic covariate effects. Using time-invariant covariates we implement effects for technological (based on the sectoral classification) and regional (based on geographical zones) similarity. The former accounts for the presence of intra-industrial trade, while the latter captures the well-known finding that global input-output interlinkages are clustered within countries or more broadly defined regions (e.g., Cerina et al., 2015).[9] Moreover, we consider a set of time-dependent covariates to investigate the influence of different dimensions of supplier heterogeneity on both the activity to form new relationships and the popularity to be selected as a counterparty using so-called activity (or ego) and popularity (or alter) effects. The activity effect measures whether actors with high values of the covariate v are more active in creating new ties and thus have a higher outdegree, while the popularity effect measures whether actors with higher v are more frequently selected.

As covariates for these activity and popularity effects, we include size, productivity, labor costs, and an index for the economic complexity of the produced output because these variables have been identified as crucial variables determining buyer-supplier matching (e.g., Bernard and Moxnes, 2018; Tintelnot et al., 2018; Bernard et al., 2019b). We also consider geographical distance between the countries to model the effect of transportation costs on trade that is not captured by the regional similarity statistic (Disdier and Head, 2008). Moreover, we control for free-trade agreements between the different countries that aim to reduce trade barriers such as tariff and non-tariff measures (Baier and Bergstrand, 2007) and input complementarities in production processes (input-output relationships between industries arising from technological constraints).

Next, we turn to the effects for the two bipartite networks in Table 3.2. In addition to the different covariate activity effects that are largely identical to those in one-mode network, here we consider the elementary 4-cycles effect that captures the tendency of OECD countries to have overlapping export / import activities with non-OECD economies.[10] More specifically, the 4-cycles effect tests the hypothesis that OECD countries importing from or exporting to the same non-OECD countries tend to have more partners in common, which could be attributed to concentration or specialization ten-

[9]The similarity effect measures whether ties tend to occur more often between actors with similar values of the covariate v.

[10]A crucial difference to the one-mode network is the absence of popularity effects in the two-mode network. Here the latter are nonsensical because nodes in the second mode are interpreted as “activities” instead of focal actors.

dencies of trade as argued by Krugman (1981); Hummels et al. (1999); Frankel and Romer (1999). Moreover, the consideration of free trade agreement effects in the two bipartite networks enables us to study potential asymmetries in the benefits and trade stimulating effects of trade agreements between developing and developed countries as considered by McQueen (2002); Freund (2003); Baier and Bergstrand (2007); Eicher and Henn (2011).

Level	Type	Definition	Hypothesis
Elementary effects			
Industry	Outdegree density	$\sum_p y_{ip}$	Tendency to export to or import from non-OECD economies.
Industry	4-cycles	$\frac{1}{4}\sum_{p,k,h} y_{ip}y_{ik}y_{hp}y_{hk}$	Tendency to make common choices regarding the selection of non-OECD trade partners.
Variable covariate effects			
Industry	Output share activity	$v_i y_{i+}$	Large producers have many non-OECD relationships.
Country	Economic complexity activity	$v_i y_{i+}$	Producers of complex output have many non-OECD relationships.
Industry	Productivity activity	$v_i y_{i+}$	Productive producers have many non-OECD relationships.
Industry	Labor cost activity	$v_i y_{i+}$	Producers with competitive labor costs have many non-OECD relationships.
Dyadic covariate effects			
Industry	Technological complementarity	$\sum_j x_{ij}(w_{ij} - \overline{w})$	Trade activities depend on technological complementarity.
Country	Geographic distance	$\sum_j x_{ij}(w_{ij} - \overline{w})$	Trade activities depend on geographic distance.
Country	Free trade agreements	$\sum_j x_{ij}(w_{ij} - \overline{w})$	Trade activities depend on free trade agreements.

Table 3.2: Specification of the effects for the two-mode network. Network ties are denoted by y_{ip}, and v_i are actor-specific covariates. Dyadic covariates refer to a pair of nodes and are denoted by w_{ij} A $+$ sign denotes summation over the pertinent index.

We are also interested in the question whether trade relationships between OECD countries in the one-mode network determine their propensity to export to or import from certain non-OECD economies and vice versa. To this end, we consider the interaction between the one-mode and the two bipartite networks based on the effects in Table 3.3.

Level	Type	Formulation	Hypothesis
Multiplex effects			
Industry	Agreement from preferences	$\sum_{j \neq h} x_{ij} y_{ih} y_{jh}$	Industry agreement with respect to their trade activities.
Industry	Agreement to preferences	$\sum_{j \neq h} x_{ij} y_{ih} x_{hj}$	Trade activities leading to agreement along industries.
Industry	Agreement from with same prerequisites	$\sum_{j \neq h} x_{ij} y_{ih} y_{jh} I\{v_i = v_j\}$	Trade activities due to same economic conditions.
Industry	Entrainment effect of trade preferences	$\sum_j y_{ij}^{impnet} y_{ij}^{expnet}$	Trade activities are mutually dependent.

Table 3.3: Specification of the effects for the interaction of one-mode and two-mode networks. One-mode network relations are denoted by x_{ij}, while two-mode network relations are given by y_{ip}. As before, v_i are sector-specific covariates, and w_{ij} refer to dyadic covariates.

In the stochastic actor-oriented model, important mechanisms to model interactions between the different types of networks are agreement and entrainment effects. Agreement effects, graphically illustrated in panel (c) of Figure 3.1, are considered to test whether the presence of input-output relationships between OECD country-sectors lead to the same export or import activities in the bipartite network, or if sourcing strategies from / to certain economies outside the OECD lead to trade dependencies between OECD economies. In the supply chain literature, such triadic dependencies among trading partners have been identified as relevant mechanisms for exploiting sourcing differentials between different suppliers, allowing buyers to benefit from cooperation and competition when suppliers exhibit partially overlapping capabilities (Dubois and Fredriksson, 2008). Finally, the entrainment effect tests if imports from non-OECD economies also imply exports into these countries and vice versa, which could be rationalized with the hypothesis that raw materials, preliminary or semi-finished products are first imported from non-OECD countries and later exported to non-OECD countries as processed goods, consistent with the Prebisch-Singer hypothesis (Prebisch, 1950; Singer, 1950), or the dependency theory put forth by Ahiakpor (1985).

3.3 Data and Descriptive Statistics

Our analysis builds on several data sources. Network data on production input interlinkages are obtained from the OECD Inter-Country Input-Output (ICIO) Tables (OECD, 2018a). These include transactions in intermediate goods and services across 36 ISIC Rev. 4 industries for all 36 OECD and 28 non-OECD countries over the period 2005-2015 as well as additional information on each sector's value added, gross output, and final demand.[11] From these annual input-output tables we construct 11 realizations of the one-mode network (so-called network panel waves), describing trade relationships within and across OECD countries, and two bipartite networks reflecting the import and export preferences of OECD countries vis-à-vis non-OECD economies, represented by binary adjacency matrices. Following Carvalho (2014), we aim to reduce noise in the fluctuations of input-output relationships by deleting links that account for less than 1% of the sector's total input purchases used for total output.

Several additional explanatory variables are obtained from the OECD Structural Analysis (STAN) Database (OECD, 2018b,c). Our measure of size are output shares, measured in terms of the share of sector-specific output to global output in that industry in a given year. Production efficiency is measured by total factor productivity

$$TFP_{i,t} = VA_{i,t}/(EMPN_{i,t}^{\alpha_{i,t}} \cdot CFC_{i,t}^{1-\alpha_{i,t}}), \tag{3.3}$$

where $VA_{i,t}$ represents value added, $EMPN_{i,t}$ denotes the number of employees, $CFC_{i,t}$ is the consumption of fixed capital, and $\alpha_{i,t}$ refers to the labour share.[12] We measure labour costs in terms of the ratio of total compensation of employees to employment in a given country-sector.

Our measure of geographic similarity is derived from UN geoscheme codes (UN, 2006) that assign each country into 1 of 10 global regions such as Northern America, Western Europe, or Eastern Asia. Building on this classification, two countries are considered similar when they are part of the same geozone. In a similar way we proceed with technological similarity that is operationalized based on common ISIC Rev. 4 industry classification codes of two production units. Moreover, to capture the potential impact of tariffs and non-tariff barriers to trade on the formation of input-output architecture, we rely on data from Dür et al. (2014) and count the number of existing free trade agreements between countries. We also control for transportation costs between countries by measuring the geographic distance between the capitals of the different OECD and non-OECD countries. Further, since we expect structural differences in the type of products produced in developed and developing countries, we account for the economic complexity of the output. The pertinent data come from the Atlas of Economic Complexity by Hidalgo and Hausmann (2009); Hausmann et al. (2014), providing a comprehensive measure of economic complexity on the country-level. Finally, we

[11] A list of the available OECD and non-OECD countries and the different sectors is provided in Table 3.8 in Appendix 3.A.

[12] Based on an comparison of alternative measures for different countries, Blades and Meyer-zu Schlochtern (1998) have argued convincingly that consumption of fixed capital is the preferred measure of the capital stock.

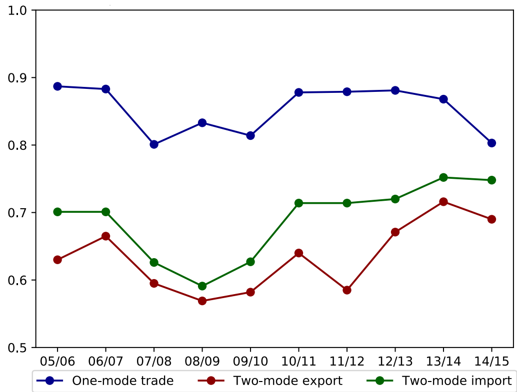
account for technological complementarities in production functions by constructing a dyadic indicator variable that turns one if the output of a specific sector belongs to the technology set of another industry, and zero otherwise. This dummy is estimated from the input-output data at the beginning of the sample period. More specifically, for each sector i, we use input-output data for 2005 and compile of list of other sectors $j \neq i$ (the technology set) supplying their intermediates to i. Sectoral specificities are eliminated by subtracting the sectoral mean from the different covariates. Moreover, to improve the performance of the simulation algorithm, we transform continuous data into percentiles, ranging from one for the smallest observations to ten for the largest ones.

We begin our empirical analysis by discussing some key characteristics of the global production network in Table 3.4.

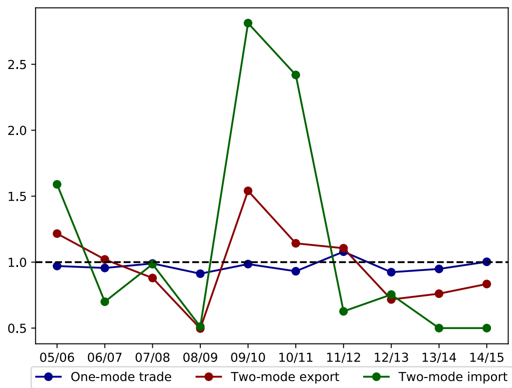
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
One-mode network											
Density (in %)	0.641	0.640	0.638	0.637	0.632	0.631	0.628	0.631	0.628	0.626	0.626
Average degree	8.299	8.285	8.262	8.252	8.184	8.171	8.133	8.174	8.133	8.103	8.105
Number of ties	10756	10737	10707	10695	10606	10590	10541	10593	10540	10501	10504
Two-mode export network											
Density (in %)	0.071	0.074	0.074	0.072	0.060	0.067	0.069	0.071	0.066	0.063	0.061
Average degree	0.714	0.746	0.749	0.725	0.603	0.674	0.694	0.713	0.668	0.639	0.618
Number of ties	925	967	971	940	781	874	900	924	866	828	801
Two-mode import network											
Density (in %)	0.021	0.023	0.022	0.021	0.018	0.023	0.026	0.024	0.023	0.021	0.019
Average degree	0.213	0.231	0.217	0.216	0.183	0.228	0.262	0.242	0.231	0.211	0.191
Number of ties	276	299	281	280	237	295	339	314	300	273	248

Table 3.4: Descriptive network statistics for the one-mode and the bipartite export and import networks for the period 2005-2015.

The density of the one-mode network, defined as the number of existing ties relative to the maximum number of relationships, ranges between 0.63-0.64%. Thus, consistent with prior studies (e.g., Carvalho, 2014; Carvalho and Tahbaz-Salehi, 2019, for US data), we find that the OECD production network is extremely sparse, implying that most production units have few upstream and downstream relationships to other industries. The average (total-) degree of the one-mode network confirms this impression. Our estimate suggests that the average industry in a OECD country has 8 intra-OECD relationships. We obtain qualitatively similar results for the two bipartite networks. Yet density, average degree, and the total number of connections for the export network exceed the pertinent statistics for the import network by a factor of 3, implying that trade relationships between members of the OECD and non-OECD countries are mainly export-driven from a OECD country's viewpoint.[13]



(a) Jaccard indices



(b) Construction-to-destruction ratio

Figure 3.2: Panel (a) shows the time evolution of the Jaccard index for all three types of networks. The Jaccard index is defined as $N_{11}/(N_{11} + N_{01} + N_{10})$, where N_{11} is the number of stable ties, N_{01} the number of newly created ties, and N_{10} denotes the number of dissolved relationships. Panel (b) illustrates the construction-to-destruction ratio reflecting the number of created ties relative to the number of dissolved relationships. Further information on the network persistence can be found in Appendix 3.A.

How much fluctuating are individual trade relationships over time? To answer this question, we consider two measures of network persistence in Figure 3.2. While the Jaccard index in panel (a) quantifies the percentage of stable relationships from one network panel wave to the next, the construction-to-destruction measure in panel (b) quantifies the ratio of created to dissolved network ties over two consecutive periods. Both measures indicate that trade relationships within and across OECD economies are more stable than ties representing import or export activities between OECD and non-OECD economies. The Jaccard index for the one-mode network ranges between 0.8 and 0.9, implying that 10-20% of trade relationships between OECD countries change from one year to the next. At the same time, the observed rate of change for the import and export relationships with non-OECD economies is higher, ranging between 25 and 45%.

Interestingly, import ties turn out to be more stable than export linkages into non-OECD economies, which might be explained with the relatively inelastic demand for primary goods and raw materials imported from these economies. Moreover, the descriptives also testify to more intense fluctuations in the global trade network during the 2007/08 financial and banking crisis and the subsequent recovery period, which suggests that the trade collapse after the crises is not merely reflected in adjustments

[13]Of course, the degree neglects information on the intensive margin of trade, i.e. the volume of trade conditional trade relationships.

at the intensive margin (i.e. changes in the value of trade conditional on existing relationships), but also led to significant changes of the network structure. We will turn to the estimation of the driving forces behind these fluctuations in the next section.

3.4 Results

We estimate our dynamic network model for the whole period 2005-2015 using 11 annual waves of network panel data and assuming constant parameters across years. To ease the exposition, we divide the results into three parts. Section 3.4.1 discusses our findings pertaining to intra-OECD trade in the one-mode network, while Section 3.4.2 presents the results for the two bipartite networks for trade relationships between OECD and non-OECD countries. Section 3.4.3 provides the estimation results for the multiplex dynamics of both network types, shedding light on the mutual dependence between intra-OECD trade and import and export relationships with developing economies.

3.4.1 One-Mode Network Dynamics

Good convergence was achieved in all simulation runs, with all individual t-ratios less than 0.1 as recommended by Ripley et al. (2011). Considering the results in Table 3.5, we find a negative density and a positive reciprocity effect (both statistically significant at the 1% level).

While the negative outdegree density testifies to capacity constraints in the accumulation of linkages, consistent with the sparseness of the production network described in Section 3.3, reciprocated ties may originate in bargaining strategies as reciprocity is an effective means to elicit cooperation from trading partners as noted by Rhodes (1989).[14] Moreover, the positive coefficient of the transitive triplets effects confirms recent theorizing on endogenous production networks stressing the importance of geodesic distance on the formation of trade relationships. This mechanism indicates locally hierarchical structures, especially when considered jointly with the negative 3-cycles effect capturing generalized reciprocity with clusters of mutual customers and suppliers, which would not be consistent with local hierarchy. We may attribute this tendency for local hierarchy to the well established finding that trade between industries within a country is more likely than trade across countries (Storper, 2004).

[14]Reciprocity is also a very relevant mechanism in many social networks (Wasserman et al., 1994).

Level	Type	Coefficient	Standard error	Significance
Elementary effects				
Industry	Outdegree density	-5.3638	0.0792	***
Industry	Reciprocity	0.3082	0.031	***
Industry	Transitive triplets	0.4641	0.0058	***
Industry	3-cycles	-0.0633	0.0066	***
Constant covariate effects				
Industry	Technological similarity	0.6409	0.1176	***
Country	Regional similarity	1.2488	0.0422	***
Variable covarate effects				
Industry	Output share activity	-0.1018	0.0046	***
Industry	Output share popularity	0.1969	0.0049	***
Industry	Productivity activity	0.0386	0.0054	***
Industry	Productivity popularity	0.0114	0.0043	***
Industry	Labor cost activity	-0.0063	0.0043	
Industry	Labor cost popularity	0.0261	0.0044	***
Country	Economic complexity activity	-0.0561	0.0046	***
Country	Economic complexity popularity	0.0292	0.0046	***
Dyadic covariate effects				
Industry	Technological complementarity	0.4615	0.0249	***
Country	Geographic distance	-0.2042	0.0084	***
Country	Free trade agreements	0.0334	0.007	***

Table 3.5: Estimated coefficients for the OECD one-mode network. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

We also obtain strong evidence for the importance of geography on trade. The negative estimate of geographic distance suggests that the propensity to establish trade relationships declines with geographical distance and transportation costs, while the positive estimate of regional similarity emphasizes the relevance of regional cooperations, supporting theories of industrial agglomeration that date back at least to Marshall (1920) and the literature of regional integration relating to the work of Sapir (1992); Vamvakidis (1998); Mattli et al. (1999).

The intended effect of free trade agreements (FTA) to foster economic cooperation between countries is reflected in our results as we obtain a positive coefficient for the pertinent FTA effect, consistent with results in Baier and Bergstrand (2007). Yet, at least for trade within the OECD, we find that the strength of this mechanism is relatively small. A potential explanation of this result is that OECD member countries already exhibit relatively low barriers to trade, rendering the influence of such agreements on the formation of additional or the dissolution of existing ties relatively unimportant.

Next, we turn to the activity and popularity effects. The positive output-share popularity effect with a coefficient of 0.1969 ± 0.0049 suggests that industries with a large output-share are more frequently selected as new suppliers. Since the output-share is essentially an alternative measure of size, this finding provides an economic rationale for the relevance of the preferential attachment mechanism, initially proposed by Barabsi and Albert (1999); Albert and Barabási (2000) applied to the formation of production networks by Atalay et al. (2011). Interestingly, the output share activity effect is negative, contradicting the intuition that large production units should have more suppliers. The explanation for this might be twofold: First, as we consider the industry instead of the firm-level, firm-specific features could have averaged out in the process of aggregation. Second, the negative effect could originate in tendencies towards higher specialization as in Krugman (1981); Hummels et al. (1999); Kramarz et al. (2019), implying that large producers tend to be less diversified and, therefore, also have fewer import ties.

TFP-related effects are positive and statistically significant for both activity and popularity, implying that very productive industries are actively demanding more inputs from different sources and are also more popular to be selected as suppliers by other industries. Theoretical explanations for the latter mechanism have been provided by Gualdi and Mandel (2016); Oberfield (2018) on a firm and Acemoglu and Azar (2020) at a sectoral level. One economic justification for the relevance of the productivity effect is that high productivity can lead to cost advantages that allow industries to offer lower prices compared to less efficient suppliers and, therefore, to attract more customers in the competitive process.

The remaining effects pertaining to labour costs and economic complexity yield somewhat ambiguous results. While the labour costs activity effect is not significant, we find that industries with labour costs above the industry-wide average are more popular in being selected as suppliers. At this point we can merely speculate that the positive coefficient of labour costs popularity reflects superior quality, consistent with empirical evidence in the extant literature reporting that rich countries tend to import relatively more from countries that produce high-quality goods (Hallak, 2006). Regarding the influence of the economic complexity the model predicts that industries with higher economic complexity indices (ECIs) are less active in sourcing inputs. At the same time, the positive economic complexity popularity effect means that countries producing more complex output are more frequently selected as trading partners. Speculating about the nature of the negative activity effect sourcing inputs, we may assume that complex sectors seek to optimize and consolidate their supplier structure. At the same time, industries that produce complex goods are more sophisticated in their economic abilities and tend to be more frequently selected as suppliers as more downstream sectors depend on their complex inputs. Thus, economic complexity is not merely a crucial determinant for growth and prosperity, but also a relevant factor shaping the formation of input-output topology and trade networks.

3.4.2 Two-Mode Network Dynamics

Next, we turn to the estimation results for the two-mode networks on export and import preferences with non-OECD economies in Table 3.6. Among the elementary effects, exports and imports exhibit the same sign on outdegree density. Like in the one-mode network, these densities are negative due to the sparseness of the trade networks (a positive effect would imply a tendency to converge to a fully connected network). The positive and highly significant 4-cycles effects suggests a strong tendency for trade activity homophily (or overlapping trade preferences), i.e. OECD country-sectors that export to or import from the same non-OECD country-sectors also tend to have more trading partners in common. We also find a significantly positive effect for technological complementarity and a negative effect for geographical distance. The relevance of these mechanisms is evident in the context of intra-OECD trade, so it should not come as a surprise that they are also decisive factors determining trade relationships with non-OECD economies.

Strikingly, the coefficients for the two technological complementarity effects in imports and exports are off the charts, compared to the other estimated coefficients, which lends additional empirical support to the view that these effects are utterly important for the formation of trade relationships. Moreover, the negative geographic distance effect supports suggestions by Chaney (2016) and Allen (2014) about search and information frictions, as transport costs may not fully account for the negative impact of distance on trade and geography tends to matter for trade in a specific way.

Interestingly, our results indicate asymmetric effects of free trade agreements on the import and export activities with non-OECD economies. While the existence of FTA evidently enhances the exports of OECD to non-OECD countries, the same is not true for the imports from non-OECD economies, suggesting the existence of unequally distributed benefits from FTAs between developed and developing countries as described by (McQueen, 2002). One economic rationale for these dependencies is that recent FTAs include labour and environmental standards and the enforcement of these standards may impose constraints on the access to international markets, thereby reducing the exports of non-OECD into OECD countries. At the same time, the same regulations may reduce cost advantages of non-OECD countries leading to a substitution of non-OECD imports with intra-OECD trade (McGee and Yoon, 2003).[15] Thus, our results provide new evidence on the asymmetric gains of trade from a dynamic network perspective.

Level	Activity	Type	Coefficient	Standard error	Significance
Elementary effects					
Industry	Export	Outdegree density	-19.473	1.552	***
Industry	Export	4-cycles	0.2277	0.0113	***
Industry	Import	Outdegree density	-9.1232	0.5696	***
Industry	Import	4-cycles	0.1683	0.0106	***
Variable covariate effects					
Industry	Export	Output share activity	0.6559	0.0263	***
Industry	Export	Productivity activity	-0.0533	0.025	**
Industry	Export	Labour cost activity	0.0552	0.0211	***
Country	Export	Economic complexity activity	0.0716	0.018	***
Industry	Import	Output share activity	-0.01	0.0325	
Industry	Import	Productivity activity	-0.0357	0.0483	
Industry	Import	Labour cost activity	-0.1212	0.036	***
Country	Import	Economic complexity activity	-0.0725	0.0309	**
Dyadic covariate effects					
Industry	Export	Technological complementarity	6.9301	1.5181	***
Country	Export	Geographic distance	-0.2201	0.0151	***
Country	Export	Free trade agreements	0.3044	0.0237	***
Industry	Import	Technological complementarity	2.7543	0.3885	***
Country	Import	Geographic distance	-0.1806	0.0293	***
Country	Import	Free trade agreements	-0.2718	0.0917	***

Table 3.6: Estimated coefficients for the two-mode networks. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Considering the different covariates, our estimates indicate a positive complexity effect for export activity and a negative influence of complexity on import activity. This

[15]It is noteworthy that similar results have been reported by Bernard and Dhingra (2016) who analyzes the consequences of the US-Colombia Free Trade Agreement on firm-to-firm matching, finding that US exporters increased their average prices, reduced their export volume and the number of their import partners in the Colombian market.

implies that especially complex goods produced in the OECD are exported to non-OECD countries depending on these imports and complex economies in the OECD exhibit a lower propensity to import from non-OECD countries (and trade more with other countries within the OECD). The latter is consistent with the view that developed countries tend to have comparative advantages in producing complex goods while developing countries specialize more on primary and less complex products. The labour cost effects support this interpretation in the sense that the production of complex goods in OECD countries requires a higher proportion of high-skilled labour that earns wages above the average, leading to the positive labour cost activity effect for exports. Moreover, high labour cost industries in the OECD are less dependent on imports from non-OECD because the coefficient of labour cost import activity is evidently negative. We also find that OECD countries with high market shares in a given industry tend to have more export linkages to non-OECD economies. Moreover, it is interesting to see that disproportionately productive industries within the OECD do not have more export links to non-OECD countries because the coefficient is significantly negative.

3.4.3 Multiplex Dynamics

Estimation of the joint evolution of the three networks in Table 3.7 testifies to strong interaction effects between intra-OECD and OECD-non-OECD trade as we find that the agreement between OECD country-sectors i and j to export to or import from the same non-OECD country-sector promotes the creation or maintenance of a trade relationship between i and j in the one-mode network. Moreover, the positive coefficients of the two agreement to export / import activity effects means that the existence of a trade relationship between two OECD country-sectors in the one-mode network leads to agreement with respect to export or import activities with non-OECD partners. All four coefficients are significant and positive. A comparison of the pertinent coefficients suggests that the agreement to import activity effect is quantitatively most important. Among various ways to rationalize this effects, we may argue that existing connections lower the costs of finding new partners, implying that the network of OECD trade relationships provides information about appropriate third party suppliers in non-OECD economies.

To ease the economic interpretation of the different effects, we study the interaction of the agreement effects with covariates. In particular, we try to address the question if similarity of the country-sectors with respect to their individual characteristics can explain the agreement effects mentioned before. Yet, at least with respect to agreement from export activity this is not the case because three out of four covariates (output shares, labour costs, economic complexity) are insignificant and the fourth has the opposite sign, suggesting that a common level of productivity between two OECD country-sectors reduces the probability to form trade relationships in the one-mode network although they share the same preferences to export to the same industries in non-OECD countries.

Relationship	Type	Coefficients	Standard error	Significance
Multiplex network effects				
One-mode formation from export activity	Agr. from trade act.	0.3175	0.0271	***
Export activity to one-mode formation	Agr. to trade act.	1.3567	0.0367	***
One-mode formation from export activity, same output shares	Agr. from trade act. with same prereq.	-0.0293	0.0371	
One-mode formation from export activity, same productivity	Agr. from trade act. with same prereq.	-0.1255	0.0536	**
One-mode formation from export activity, same labour costs	Agr. from trade act. with same prereq.	-0.0225	0.0399	
One-mode formation from export activity, same economic complexity	Agr. from trade act. with same prereq.	-0.0154	0.0511	
Export activities promote the creation of import activities	Entrainment effect of trade act.	1.2608	0.8648	
One-mode formation from import activity	Agr. from trade act.	0.2874	0.1242	**
Import activity to one-mode formation	Agr. to trade act.	2.974	0.1124	***
One-mode formation from import activity, same output shares	Agr. from trade act. with same prereq.	-0.0543	0.1515	
One-mode formation from import activity, same productivity	Agr. from trade act. with same prereq.	0.3511	0.1743	**
One-mode formation from import activity, same labour costs	Agr. from trade act. with same prereq.	-0.0851	0.1188	
One-mode formation from import activity, same economic complexity	Agr. from trade act. with same prereq.	0.2715	0.1243	**
Import activities promote the creation of export activities	Entrainment effect of trade act.	3.3899	0.533	***

Table 3.7: Estimated coefficients for the multiplex dynamics of the one-mode and two-mode network. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively. Agr. stands for agreement, act. for activities and prereq. for prerequisites.

Regarding the one-mode formation from import activity effects, it appears that at least some variation in the data can be explained by a common productivity and economic complexity. Thus, OECD country-industries with a similar productivity and economic complexity have a positive tendency to form import ties to the same non-OECD countries.

Last but not least, another interesting finding refers to the entrainment effect of trade activities. It implies that OECD import activities from specific industries outside the OECD promote the creation of export activities to the same non-OECD industry. This leads us to conclude that OECD industries use non-OECD sectors especially in their upstream activities to source inputs such as raw materials, and then sell (pre)finished goods back to these non-OECD countries. Even though this result does not directly relate to the Prebisch-Singer hypothesis, it provides empirical support to the dependency theory, predicting that resources flow from the economic "periphery" of developing countries to the "core" of wealthy and highly industrialized countries. The fact that the same is not true for the entrainment effect from exports on imports reflects favorably on this interpretation.

3.5 Concluding Remarks

This paper examines how a global input-output network, describing production linkages within and across OECD industries as well as their relationship to non-OECD countries, is subject to significant changes in the formation of industrial interlinkages due to several driving forces at the mesoeconomic and macroeconomic level. To achieve this, we use a comprehensive data set on supplier-customer links among OECD and non-OECD countries, to develop a set of facts about the underlying evolutionary process guiding developments in one-mode and two-mode dynamic trade networks. With this, we do not only model the formation of trade relationships between OECD countries in a one-mode network and their affinity to import from and export to non-OECD economies but also extent the framework to the multiplex dynamics of co-evolving one-mode and two-mode networks.

To estimate the process of tie formation, we proposed an empirical model of network formation that enables us to assess the quantitative importance of several alternative mechanisms that have been brought forward in the theoretical debate in recent years. The key elements of contributions are: First, in context of one-mode network formation process we identify several important features for international trade, such as network proximity, geographical distance or technological complement, being decisive drivers for decision processes in trade relations. Moreover, we confirm established facts about size, productivity and free trade agreements. Second, referring to the findings of the export and import preference networks, we can confirm results on geodesic distance and industry-related size in trade activities while providing new insights on asymmetric gains from free trade agreements comparing developing and developed countries. Third, we provide insights in the multiplexity of network tie formation of both network types. An interesting finding that emerges from the joint analysis is that OECD import activities from certain non-OECD industries encourage the creation of export activities to the same non-OECD industries. However, this does not apply to exports to non-OECD countries.

Since surprisingly little attention has been paid to the endogenous dynamics of trade networks until recently, especially in empirical research, the present paper aims to fill this gap by providing a careful statistical analysis of the formation of input-output linkages in the strand of endogenous production networks. Some of the characteristics mentioned above have been identified as crucial determinants of shock propagation and macroeconomic outcomes. Thus, the remarkable stability of these aggregated characteristics, even at the sectoral level, testifies to their general importance even or especially in a dynamically developing environment. Therefore, we believe that the results reported in this paper provide valuable insights into the dynamics of macroeconomic input-output relationships at the global level of economic activity. However, by limiting the level of analysis to industrial aggregates, we make the study vulnerable to biases in sample aggregation, thus hindering the generalization of the results to a more fine-grained context.

With our paper we merge the established and already identified empirical regularities on economic trade networks with dynamic network formation processes providing a sound intuition that hopefully inspires the development of new theories relating to dynamic trade networks. In this context, the heterogeneity of industries, both in terms of buyers and suppliers in production depending on their different economic prerequisites in a global context of import and export activities has provided fruitful insights. This in mind, allows for several avenues of future research: Our explorations indicate that the main mechanisms of network formation, while highly relevant, are merely discussed at the global economic level and therefore provide a selective picture of the actual mechanism of tie formation processes from a theoretical perspective. Nonetheless, as the SAOM modeling approach proves to be an effective tool to estimate channels of network formation from an economic perspective, further efforts need to be put into the understanding of socio-economic (mis)behavior leading to the outcomes we observe. Understanding these phenomena, we might elaborate more just and sustained growth strategies while redressing inequality and exploitation. The SAOM methodology is flexible enough to accommodate for the multiplexity of economic dimensions. There are no limits to imaginable economic research interests. Unfortunately, the additional computational complexity limits the application of this modeling approach in very large networks today. However, believing in Moore's law this should not be a problem in the near future.

3.A Appendix

Details of the Model

Change Frequencies

At the current states of x and y, the rate functions $\lambda_i^X(x,y)$ and $\lambda_i^Y(x,y)$ represent the expected number of occasions per time unit where actor i has the opportunity to make a change in one single outgoing ties in $X_{ij}(t)$ or $Y_{ip}(t)$, respectively.

Using properties of Markov processes, the waiting time until the arrival of the next change opportunity for an actor in the one-mode and two-mode network can be modeled by a exponential distribution with expected value

$$\frac{1}{\lambda_+^X(x,y) + \lambda_+^Y(x,y)}, \tag{3.4}$$

$$\text{where } \lambda_+^X(x,y) = \sum_{i \in \mathcal{N}} \lambda_i^X(x,y) \text{ and } \lambda_+^Y(x,y) = \sum_{i \in \mathcal{N}} \lambda_i^Y(x,y).^{16} \tag{3.5}$$

This allows to compute the probability of a change opportunity for actor i in the one-mode network as

$$\frac{\lambda_i^X(x,y)}{\lambda_+^X(x,y) + \lambda_+^Y(x,y)}. \tag{3.6}$$

In a similar vein, the opportunity for change probability in the two-mode network can be formalized as

$$\frac{\lambda_i^Y(x,y)}{\lambda_+^X(x,y) + \lambda_+^Y(x,y)}. \tag{3.7}$$

Conditional Choice Probability

The evaluation functions $f_i^X(x,y)$ and $f_i^Y(x,y)$ represent the relative tendency of actor i to make a change towards state x (y) in the one-mode or two-mode network, respectively, when the other network has state y (x) and vice versa. These propensities are measured on a log-probability scale, and their value may only be compared between changes allowed from a given current state of a single bond to another. Since the link variables X_{ij} and Y_{ip} are dichotomous, a change of a link variable can be regarded as a switching (toggle) of the link variable from X_{ij} (Y_{ij}) to $1 - X_{ij}$ ($1 - Y_{ip}$). For one of the given network states x and y we denote a switch or toggle from X_{ij} (Y_{ij}) to $1 - X_{ij}$ ($1 - Y_{ip}$) as $x^{(\pm ij)}$, or $y^{(\pm ip)}$ respectively, while all other variables remain the same.

[16] Replacing an index by a $+$ sign denotes the summation over this index.

Based on this formalism, the conditional probability of actor i changing a tie to j in the one-mode network is given by

$$P\{X(t) \text{ changes to } x^{ij} | X(t) = x, Y(t) = y\} = \frac{\exp(f_i^X(x^{(\pm ij)}, y))}{\exp(f_i^X(x, y)) + \sum_{h \in \mathcal{N}(i)} \exp(f_i^X(x^{(\pm ih)}, y))}. \tag{3.8}$$

Thus, when actor i has the opportunity to change a one-mode tie, the one-mode networks that can be obtained as a result of this change are all $x^{(\pm ij)}$ for $j \neq i$ as well as the current x. The resulting evaluation function for the changed network has the values $f_i^X(x^{(\pm ij)}, y)$. Furthermore, $\mathcal{N}(i)$ denotes the set of all actors except i, thus $\mathcal{N}(i) = \{j \in \mathcal{N} | j \neq i\}$.

When actor i has the opportunity to change a two-mode tie, the two-mode networks that can be obtained are $y^{(\pm ip)}$ for any $p \in \mathcal{P}$, given that the current state of the system is (x, y). Therefore, the conditional probability of changing the preference tie to p is

$$P\{Y(t) \text{ changes to } y^{ip} | X(t) = x, Y(t) = y\} = \frac{\exp(f_i^Y(x, y^{(\pm ip)}))}{\exp(f_i^Y(x, y)) + \sum_{b \in \mathcal{P}} \exp(f_i^Y(x, y^{(\pm ib)}))}, \tag{3.9}$$

where for potentially emerging networks the evaluation function takes the values $f_i^Y(x, y^{(\pm ip)})$.

Data

OECD countries	Code	Non-OECD economies	Code	Industry	Code	ISIC Rev.4
Australia	AUS	Argentina	ARG	Agriculture, forestry and fishing	D01T03	01, 02, 03
Austria	AUT	Brazil	BRA	Mining and extraction of energy producing products	D05T06	05, 06
Belgium	BEL	Brunei Darussalam	BRN	Mining and quarrying of non-energy producing products	D07T08	07, 08
Canada	CAN	Bulgaria	BGR	Mining support service activities	D09	09
Chile	CHL	Cambodia	KHM	Food products, beverages and tobacco	D10T12	10, 11, 12
Czech Republic	CZE	China (People's Republic of)	CHN	Textiles, wearing apparel, leather and related products	D13T15	13, 14, 15
Denmark	DNK	Colombia	COL	Wood and products of wood and cork	D16	16
Estonia	EST	Costa Rica	CRI	Paper products and printing	D17T18	17, 18
Finland	FIN	Croatia	HRV	Coke and refined petroleum products	D19	19
France	FRA	Cyprus 2	CYP	Chemicals and pharmaceutical products	D20T21	20, 21
Germany	DEU	India	IND	Rubber and plastic products	D22	22
Greece	GRC	Indonesia	IDN	Other non-metallic mineral products	D23	23
Hungary	HUN	Hong Kong, China	HKG	Basic metals	D24	24
Iceland	ISL	Kazakhstan	KAZ	Fabricated metal products	D25	25
Ireland	IRL	Malaysia	MYS	Computer, electronic and optical products	D26	26
Israel 1	ISR	Malta	MLT	Electrical equipment	D27	27
Italy	ITA	Morocco	MAR	Machinery and equipment, nec	D28	28
Japan	JPN	Peru	PER	Motor vehicles, trailers and semi-trailers	D29	29
Korea	KOR	Philippines	PHL	Other transport equipment	D30	30
Latvia	LVA	Romania	ROU	Other manufacturing; repair and installation of machinery and equipment	D31T33	31, 32, 33
Lithuania	LTU	Russian Federation	RUS	Electricity, gas, water supply, sewerage, waste and remediation services	D35T39	35,36, 37, 38, 39
Luxembourg	LUX	Saudi Arabia	SAU	Construction	D41T43	41, 42, 43
Mexico	MEX	Singapore	SGP	Wholesale and retail trade; repair of motor vehicles	D45T47	45, 46, 47
Netherlands	NLD	South Africa	ZAF	Transportation and storage	D49T53	49, 50, 51, 52, 53
New Zealand	NZL	Chinese Taipei	TWN	Accomodation and food services	D55T56	55, 56
Norway	NOR	Thailand	THA	Publishing, audiovisual and broadcasting activities	D58T60	58, 59, 60
Poland	POL	Tunisia	TUN	Telecommunications	D61	61
Portugal	PRT	Viet Nam	VNM	IT and other information services	D62T63	62, 63
Slovak Republic	SVK			Financial and insurance activities	D64T66	64, 65, 66
Slovenia	SVN			Real estate activities	D68	68
Spain	ESP			Other business sector services	D69T82	69, 70, 71, 72, 73, 74, 75, 77, 78, 79, 80, 81, 82
Sweden	SWE			Public admin. and defence; compulsory social security	D84	84
Switzerland	CHE			Education	D85	85
Turkey	TUR			Human health and social work	D86T88	86, 87, 88
United Kingdom	GBR			Arts, entertainment, recreation and other service activities	D90T96	90, 91, 92, 93,94,95, 96
United States	USA			Private households with employed persons	D97T98	97, 98

Table 3.8: Country and industry overview for all elements within the OECD Inter-Country Input-Output (ICIO) Tables.

Network Persistence

Periods	05/06	06/07	07/08	08/09	09/10	10/11	11/12	12/13	13/14	14/15
One-mode network										
$0 \Rightarrow 0$	1666932	1666934	1666435	1666696	1666633	1667069	1667072	1667084	1667059	1666670
$0 \Rightarrow 1$	632	649	1178	929	1081	661	707	643	721	1149
$1 \Rightarrow 0$	651	679	1190	1018	1097	710	655	696	760	1146
$1 \Rightarrow 1$	10105	10058	9517	9677	9509	9880	9886	9897	9780	9355
Hamming distance	1283	1328	2368	1947	2178	1371	1362	1339	1481	2295
Two-mode export network										
$0 \Rightarrow 0$	1305207	1305204	1305170	1305271	1305322	1305286	1305217	1305297	1305381	1305404
$0 \Rightarrow 1$	236	197	227	157	265	208	251	147	121	136
$1 \Rightarrow 0$	194	193	258	316	172	182	227	205	159	163
$1 \Rightarrow 1$	731	774	713	624	609	692	673	719	707	665
Hamming distance	430	390	485	473	437	390	478	352	280	299
Two-mode import network										
$0 \Rightarrow 0$	1306030	1306027	1306023	1306043	1306041	1305998	1305987	1306011	1306041	1306070
$0 \Rightarrow 1$	62	42	64	45	90	75	42	43	27	25
$1 \Rightarrow 0$	39	60	65	88	32	31	67	57	54	50
$1 \Rightarrow 1$	237	239	216	192	205	264	272	257	246	223
Hamming distance	101	102	129	133	122	106	109	100	81	75

Table 3.9: Change and persistence of network ties. $0 \Rightarrow 0$ refers to links which are absent in both waves, while $0 \Rightarrow 1$ describes newly created ties. $1 \Rightarrow 0$ refers to dissolved trade relationships. $1 \Rightarrow 1$ indicates trade relationships that persist over two consecutive years. Hamming distance measuring the number of tie variables that differ between two consecutive waves. For example, in the year 2015 27 ties were newly created and 54 existing ties were dissolved in the two-mode import network. Their sum (81) reflects the Hamming distance.

Part III

Essay on Statistical Regularities in Economics

Chapter 4

What's Left After the Hype? An Empirical Approach Comparing the Distributional Properties of Traditional and Virtual Currency Exchange Rates

Abstract

This paper provides an empirical analysis of the distributional properties and statistical regularities of virtual, intra-virtual and traditional currency exchange rates. To perform the analysis, the most relevant virtual, intra-virtual and foreign currency exchange rates between October 2015 and December 2018 are examined. The analysis shows that, in spite of their differing mode of formation, daily log-returns of all currency types share tent-shaped empirical densities, one of the characteristics of a Laplace distribution at semi-log scale. This peculiar property has also been examined thoroughly in other fields of economic literature. Moreover, the empirical results show that virtual and traditional currencies hold the same functional form, even after the 2018 hype. However, in spite of these similarities virtual and intra-virtual currencies display fatter tails and steeper towering peaks than regular foreign currencies which underscores the rather speculative nature of this asset class.

4.1 Introduction

Money and innovation are central to capitalist economies. While money can be described as a medium of exchange, being fully liquid and used to make or receive payments for goods and services (Groth, 2015), innovation is more difficult to define. In its Oslo Manual, the Organisation for Economic Co-operation and Development (OECD) describes innovation as "the implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organizational method, in-business practices, work-place organization or external relations" (OECD, 2005).

However, whereas the concept and technology of (fiat) money remained constant for decades, the concept of innovation has evolved tremendously; see, for example, Schumpeter (1934) or Boer and During (2001). Innovation reaches into all aspects of our lives. It was therefore only a matter of time before innovation began to affect our concept of money, and how we transfer and receive means of payment.

In 2009, Satoshi Nakamoto, an anonymous person or group, published a paper entitled "Bitcoin: A Peer-to-Peer Electronic Cash System". The paper describes a concept that enables online payments to be made between two parties without the need for an intermediate entity, such as a financial institution. Instead, they propose a protocol, arranged as a decentralized network using a chain of hash-based proof-of-work, forming a record that cannot be forged. Nakamoto (2009) demonstrates the system's unforgeability using the example of a binomial random walk. Records cannot be forged because changing the information in the chain requires that the majority of nodes in the network approve this change and write it into a general ledger. Such a ledger is also referred to as a "Blockchain" (Nakamoto, 2009).

Nevertheless, it is important to explore how virtual currencies such as Bitcoin, Etherium and Litecoin are formed, and how they compare to traditional money around the world. In traditional fiat money systems, the theory is that central banks cover for aspects such as price stability by controlling the interbank interest rate via open market operations; in contrast, most virtual currencies are "mined". In the case of Bitcoin, for example, involving a competitive and decentralized process, specific hardware and software are used to solve cryptographic hash functions. In the process, a "hash" describes a hexadecimal number with a particular target difficulty that needs to be explored by the nodes in the system to create or solve for a new block (Bitcoin Project, 2016). Based on a predetermined schedule Miners who successfully participate in solving the block are rewarded a specific amount of Bitcoins, bound to their computational power (CoinDesk, 2016).

Before exploring the key question of this paper, which is how virtual currencies are distributed and how they behave compared to traditional currencies, other matters of interest must first be addressed. First, why do virtual currencies, such as Bitcoin, have a value? What determines their price? How are they traded? Are virtual currencies a safe means of payment, and what are their disadvantages compared to traditional currencies? In addition, light must also be shed on the growing body of research into this phenomenon in recent years, which will be performed in the next section.

The main motivation for this paper is to investigate several virtual currencies, including Bitcoin, and to compare their statistical properties and regularities using a parsimonious continuous probability distribution. Another objective is to compare, how the stylized facts of real and virtual currency exchange rates differ, and whether they exhibit similarities. In an attempt to provide an extensive picture, an analysis is therefore undertaken of the four most significant exchange rates of actual to virtual currencies, three intra-virtual currency exchange rates and the four foreign currency exchange rates, including three major and one minor fiat currency.

The remainder of this paper is organized as follows: Section 4.2 introduces the academic and public debate, creating a reference to the current literature and answering the aforementioned question. Section 4.3 describes the data and descriptive statistics. Section 4.4 introduces the empirical set-up, and outlines the method used to construct the empirical analysis. Afterwards section 4.5 contains the main results, setting their implications into relation to current findings. The paper concludes with section 4.6, also highlighting a number of limitations and ideas for future research.

Outlining the results briefly, it was possible to show that virtual, intra-virtual and foreign currency log-returns are tent-shaped one of the characteristics of a Laplace distribution at semi-log scale and share the same functional form. This peculiar property has also been well examined thoroughly in other fields of economic literature (Bottazzi and Secchi, 2003; Fagiolo et al., 2007; Alfarano and Milaković, 2008). Furthermore, virtual and intra-virtual currencies exhibit higher volatility, fatter tails and steeper towering peaks than regular foreign currencies, while following the stylized facts of asset returns more extremely to some extent. It may therefore be better to consider virtual and intra-virtual currencies as a speculative instrument rather than an alternative to traditional currencies.

4.2 Public Perception and Rising Academic Interest in Recent Years

Following the functional definitions of Mishkin (2004), any type of money is a medium of exchange, unit of account and store of value. Kiyotaki and Wright (1989) examine the defining characteristic of money, also including velocity, acceptability, and liquidity. In contrast to commodity money, which holds a value due to its physical properties, or fiat money, which is used as legal tender and mainly bases its value on trust in central authorities, virtual money is created by means of computational processes using a commonly predictable rate. Nonetheless, virtual money also bases its value on trust and the conviction that its concept is superior to other forms of currency.

Like other currencies, the price of virtual money is determined by supply and demand. To facilitate this process, it is traded through currency exchanges. However, this results in a number of limitations. Regulatory barriers are not only high for currency exchanges (they operate as "money transmitters" and must therefore be registered with the Financial Crimes Enforcement Network in the US, for example), they must also provide an online infrastructure that is sufficiently strong to withstand hacking attacks. The number of relevant high-volume exchanges is therefore small (Böhme et al., 2015). In December 2016, the three most significant exchanges (OKCoin, Huobi and BTC China) accounted for more than 97% of all (Bitcoin) trades over a six-month period (Bitcoinity, 2016).

However, the exchange provider market has experienced extreme fluctuations, especially over the last two years. By January 2018, the three largest exchanges had utterly changed (then bitfinex, coinbase, bitflyer) and accounted for a total market share of only 53% of all (Bitcoin) transactions over a six-month period (Bitcoinity, 2018). By January 2019, after the hype, Bit-x and GDAX superseded coinbase and bitflyer. Still, the total market share of the top three providers remained unchanged (Bitcoinity, 2019). Hence, the description provided in Böhme et al. (2015) is unlikely to remain valid, since the market for virtual currencies is growing tremendously, with more exchange providers competing for a piece of the cake.

Trading platforms such as Bitfinex are no different than typical trading platforms used to trade foreign currencies, stocks or futures. At Bitfinex, for example, it is possible to trade Bitcoin (BTC), Ethereum (ETH) and Litecoin (LTC) on a spot price, that describes a current price in the marketplace at which the currencies can be bought or sold for immediate delivery. In the process, the platform offers different display options, like candlesticks, ranging from one minute to one day. Also, conventional indicators such as the Moving Average Convergence/Divergence (MACD) or Relative Strength Index (RSI) can be used to identify or ease trading opportunities and decisions (Bitfinex, 2017). However, the authors of Böhme et al. (2015) point out that virtual currencies mirror a payment platform rather than what economists consider a currency, because Bitcoin exchange refers to a fixed amount of a conventional currency. In another economic appraisal, Yermack (2015) concludes that Bitcoin might function somewhat like a speculative investment rather than a currency. In contrast, Kristoufek (2015) find from a wavelet coherence analysis that standard fundamental factors — usage in trade, money supply and price level — play a role in Bitcoin price in the long term, which is generally in line with monetary economics theory.

Nonetheless, safety plays a vigilant role in trading and dealing with virtual currencies. When Nakamoto (2009) proposed his electronic cash system, he argued that, due to the expansion of online commerce, financial institutions deprave to trusted third parties handling electronic payment processes, while continuing to inherit the weaknesses of being based on the trust model. It is not possible for financial institutions to reverse a transaction, such as when an individual is not liquid or when an operation is canceled without additional costs. These additional costs even increase in the case of non-reversible payments of non-reversible services. Therefore, Nakamoto (2009) recommends, "an electronic payment system based on cryptographic proof instead of trust, allowing any two willing parties to transact directly with each other without the need for a trusted third party". Following this line of argumentation, this system would protect sellers from fraud because it would be computationally impractical to reverse transactions, and buyers would be safeguarded through routine escrow mechanisms. Thus, in theory, as a means of payment, the system would be a suitable alternative to traditional methods of payment.

However, the design of virtual currencies, such as Bitcoin, exhibits characteristic risks that vary compared to other currencies or methods of payment. In a review paper, Böhme et al. (2015) focus on aspects such as market, counterparty, transaction, operational, legal and regulatory risks. They argue that the sharp movements within the exchange rate (USD/BTC) between 2013 and 2015 could serve as a source of concern when being used for transactions or store of value. Moore and Christin (2013) finds that counterparty risk also has a virtue influence, when contemplating virtual currency exchanges. In their study, they conclude that, out of 40 surveyed exchanges, 18 closed after a median lifetime of 381 days, five of which failed to reimburse customers who held currencies on their accounts. Hence, loss of funds is a considerable risk. The most well-know case was Mt. Gox, which went bankrupt in 2014.

Moreover, due to the growing interest in virtual currencies, the costs of transacting virtual currencies among wallets or the general trade fee on trading platforms can also become a significant bottleneck A wallet is a software program where virtual coins can be stored. Wallets facilitate the sending and receiving of virtual coins, and give ownership of the balance to the user. With transaction, deposit/withdrawal and trading fees of up to 5% of the order value per operation, one could talk of predatory pricing behavior. In January 2018, for example, the average Bitcoin transaction fee was 28 USD; in mid-December 2017 it was 55 USD per transaction. By January 2019, the transaction fee dropped to 20 USD cents per execution (Bitinfocharts, 2019). Finally, another risk related to the transaction of virtual currencies is time. Although the average seven-day transaction time (January 2019) is 15 minutes to execute an order, this can vary widely. In early 2018, for example, transactions took an average of 2,521 minutes (almost two days) to settle (Blockchain.info, 2019a).

Furthermore, besides transaction risk, operational risk must also be considered. In an article, Barber et al. (2012) describe the risk of a history-revision attack. The authors describe the case where, if any party in mining were able to gain more than 50% of the computational power in the decentralized network mining the coins, the entire coin base could be replaced by a figment of its forgery. This, however, seems to be rather difficult to achieve. In December 2016, the largest mining pool "AntPool" had a mining market share of 19% (Blockchain.info, 2016). Throughout the turbulent year of 2017, this did not change (Blockchain.info, 2018). By 2019 however, AntPool was ousted by BTC.com, with a mining capacity of the total mining market of 15.4% (Blockchain.info, 2019b). It is important to note a concentration among the country of origin of the different mining pools: 81% of these are based in China (Buybitcoinworldwide.com, 2018). Nevertheless, the scientific community is aware of these security issues and challenges (Lin and Liao, 2017) and is searching for solutions, for example, by means of quantum computation which would improve the level of security by the laws of physics, a state not achievable from a non-quantum information theoretic viewpoint (Ikeda, 2018).

Finally, Bitcoin and other virtual currencies are subject to legal and regulatory risks. Böhme et al. (2015) therefore raise concerns in the field of (financial) crime (for example, money laundering) and consumer protection. However, there is a more vigilant concern: politics and monetary policy regulation. In an opinion of the European Central Bank (ECB) as early as in 2016, the authors state that "[...]the reliance of economic actors on virtual currency units, if substantially increased in the future, could in principle affect the central banks' control over the supply of money with potential risks to price stability[...]" (European Central Bank, 2016). Also, in 2017, the ECB continued to warn about the dangers of investing in digital currencies. In September, Vice President Victor Constancio compared the Bitcoin hype with the 17th century tulip mania (CNBC, 2017). Benoît Cœuré, Member of the Executive Board of the ECB, argued: "Bitcoin is not a currency; it is a financial instrument which creates major risks for investors because its value is highly unstable, [...]" (European Central Bank, 2017). Benoît Cœuré was proven correct after the Bitcoin price dropped from around 20,000 USD in early 2018 to under 5,500 USD by the end of the year, dubbing it in an interview with the Financial Times "an evil spawn of the financial crisis" (Financial Times, 2018). Similar views were shared by Augustin Carstens, General Manager of the Bank for International Settlements (BIS) (CNBC, 2018).

Yanis Varoufakis, Greek economist and Greece's former Minister of Finance, warned in an article about decentralized and "apoliticized" money such as Bitcoin, arguing that the only way to steer a course between Ponzi growth and stagnation is to exercise a degree of rational, collective control over the supply of money. Since such control is bound to be political, because different monetary policy decisions affect various groups of people, the only way to guarantee money for the people by the people is through a democratically controlled and collective agency (Varoufakis, 2013). Another strong argument is the deliberate choice of countries or states to devalue or depreciate their national currencies against foreign currencies. By devaluating their currency, countries can increase their market competitiveness in an open economy, while inheriting the currency purchasing power over some time in their own country. This approach would not be possible for countries that participate in a decentralized network, that renounce national control over their own monetary policy. This does not apply for countries within the Euro area, the monetary union of 19 of the 28 European Union Member States. Hence, it is understandable that regulatory institutions and the public are becoming increasingly sensitive to the subject of virtual currencies, their properties and effects on the economy. Not surprisingly, academic research in this field has gathered pace in recent years, too.

Discussing the economics of Bitcoin, Buchholz et al. (2012) analyze the effect of volatility on prices using autoregressive conditional heteroscedasticity (ARCH) and generalized autoregressive conditional heteroscedasticity (GARCH) models. The authors find that the effect is significant. Briere et al. (2015) investigate whether Bitcoin could serve as a financial asset to diversify a investor's portfolio. Weighing up the risks they found it to be an attractive opportunity for the future. In Kondor et al. (2014) the authors analyze Bitcoin transactions, and find that sublinear preferential attachment governs the evolution of wealth distribution within the transaction network. Examining the relationship between several virtual currencies, looking at how network effects influence the cryptocurrency market, Gandal and co-authors conclude that there may be statistical arbitrage opportunities due to comovement among several of these currencies (Gandal and Halaburda, 2016).

However, little attention has been paid to the statistical properties of virtual currencies; and how they compare to traditional currencies. While Osterrieder and Lorenz (2016) conduct a risk assessment for Bitcoin and its extreme tail behavior, Osterrieder (2016) looks beyond Bitcoin and examines the statistics of other virtual currencies, too, finding that they all exhibit heavy tails. Similar methods are applied by Chu et al.

(2015), fitting fifteen of the most popular parametric distributions in finance to the log-returns of Bitcoin. They find that the Generalized Hyperbolic Distribution (GH) fits the empirical distribution of Bitcoin best.

Still, there are some constraints to these results. The five parameters required to estimate the GH are not parsimonious in an economic sense. Also, as a general form, it builds a superclass of several distributions, including the Student's t-Distribution, the Laplace, and Hyperbolic distribution, which makes it harder to distinguish the origins of its distributional properties. Moreover, the variety of parameters makes it difficult to interpret the underlying basis from an economic perspective. Due to its semi-heavy tails and the variety of parameters, the GH is often used to model financial markets and such like. Nevertheless, narrower distributions with fewer parameters may be of greater interest. Thus, it is interesting to read that the Laplace distribution and the Exponential Power or Subbotin distribution were the best fitting distributions, with two and three parameters, respectively (Chu et al., 2015).

Moreover, applying (asymmetric) Laplace distributions to financial data, especially to exchange rate returns, has proven to be an effective solution. Kozubowski and Podgórski (2001) fit (asymmetric) Laplace laws to daily log exchange rate returns, finding that they reflect properties of empirical data much better than other two-parametric distributions. Additionally, due to their one-dimensional and multivariate densities, which have convenient computational forms, estimation procedures are practical and more comfortable to implement. Kotz et al. (2001) also provide examples of this being true for other fields in finance such as stock market returns, option pricing and value-at-risk models. Moreover, the peculiar properties of the Laplace distribution are also examined in detail in other fields of economic literature, see Bottazzi and Secchi (2003); Fagiolo et al. (2007); Alfarano and Milaković (2008); Castaldi and Dosi (2009) for example.

Furthermore, to compare the properties of exchange rate returns of traditional, foreign currencies and of virtual and intra-virtual currencies, it may be insightful to compare the stylized facts of logarithmic returns. Across a wide range of different speculative markets, certain universal properties have been found in recent years. These include uncorrelated raw returns, and an alternation of periods of low volatility with periods of high volatility. The former can be recognized by the fact that the autocorrelation function of raw returns tends to be significantly different to zero for all time lags, the latter by the fact that the autocorrelation function of absolute returns is positive and decays slowly. Hence one could talk of a long-range dependence or a long memory effect. These stylized facts, and others, are discussed in detail in Guillaume et al. (1997);

Stanley and Mantegna (2000); Lux and Ausloos (2002); Johnson et al. (2003); Sornette (2003); Westerhoff et al. (2009).

4.3 Data and Descriptive Statistics

The data can be clustered into four groups. In the first group, we find the four largest virtual currencies in recent years, valued against the US Dollar (USD). These are Bitcoin (BTC), Etherium (ETH), Ripple (XRP) and Litecoin (LTC). The hard forks Bitcoin Cash (BCH) and Etherium Classic (ETC) are disregarded. A hard fork is a radical change to the protocol that makes previous transactions valid (or vice versa); it constitutes a permanent divergence from the previous version of the blockchain. The second group consists of virtual currencies valued against the largest virtual currency on the market, BTC. In this group, we find ETH, XRP and LTC again. The third group includes the four foreign exchange rates valued against the Euro (EUR). These are the US Dollar USD, the British Pound Sterling (GBP) and the Japanese Yen (JPY). Moreover, the Turkish Lira (TRY) is examined as an example for a more volatile currency in recent years. Finally, the fourth group contains the above currency groups, but pooled into one sample. However, it is important to note that these sets have been standardized by subtracting the mean and dividing by the standard deviation, which leads to the mean being equal to zero and sigma to unity. This corrects for multi-modality in the pooled datasets.

Since virtual currencies and their trading on exchanges are relatively new, it is a cumbersome task to find the right data and frequencies, and compare them against traditional currencies over a longer period of time. To establish frequency comparability between all three currency groups, daily log-returns have been used, as traditional currency exchange rates are not free available on an intra-day level. Moreover, statistically it is an important factor to use same frequencies when trying to determine regularities between datasets. As show in Ruppert (2004), e.g. regressing two time series and considering two different frequencies (e.g. daily and weekly log-returns), it is shown than the variance of coefficients is approximately five times smaller using daily than weekly-log returns. Aggregational gaussianity is another factor. Increasing the time scale of Δt over which returns are calculated, the distribution tends to look like a normal distribution. This has also been proven for foreign currencies by Boothe and Glassman (1987). Therefore, comparing the shape of distributions on different time scales needs to be viewed with caution (Cont, 2001).

However, as intra-day data are available for virtual currencies the statistical and distributional analysis in this paper is also conducted for the four major virtual currencies with an exchange rate price frequency of four hours. The results can be found in the appendix 4.A, as the focus of this paper lies on daily returns. Nevertheless, the results will be referenced in the according upcoming sections. Both data for virtual currencies valued against USD and the intra-virtual exchange rates valued against BTC were retrieved from poloniex.com for daily exchange rates between October 1, 2015 and December 31, 2018. Poloniex was chosen because it has a complete API database for several virtual currencies (StackExchange, 2016). The same time period applies to intra-day virtual currencies. Since the ECB does not release reference rates on weekends or public holidays the time period for traditional currency pairs was extended to April 1, 2014 and December 31, 2018, so as to gain a comparatively large sample set. Euro foreign exchange reference rates were retrieved from the European Central Bank website in an XML hyperlink format.

Providing an overview of the virtual currency market, Table 4.1 contains information about the four exchange rates under consideration. As can be seen, Bitcoin achieves by far the largest market capitalization and trade volume, followed by Ethereum. It is intriguing to see how the virtual currency market has changed tremendously within the space of just one year. Table 4.3 and Table 4.4 provide the same data for January 2017 and January 2018, where the market capitalization of Bitcoin increased 15 times and the capitalization of Etherium even 136 times, before dropping to the level displayed below.

Currency	Market Cap	Volume (24h)	Available Supply	Maximum Supply
Bitcoin	$62,826,332,646	$5,149,012,954	17,495,200 BTC	21,000,000 BTC
Ethereum	$12,367,840,710	$2,508,432,990	104,509,749 ETH	—
Ripple	$13,064,861,972	$415,643,519	41,040,405,095 XRP	100,000,000,000 XRP
Litecoin	$1,896,529,537	$572,027,104	60,131,350 LTC	84,000,000 LTC

Table 4.1: Market capitalization and trade volume information for selected virtual currencies. Table notes: The information was retrieved from CoinMarketCap (2019).

The development of virtual currency market shares is also extraordinary. In January 2017, Bitcoin accounted for 85.3% of the Top 100 virtual currencies with the four currencies accounting for 93.1% of the market (CoinMarketCap, 2017). By January 2018, the market share of Bitcoin had shrunk to 34%, and the four aforementioned currencies accounted for a market share of only 66%. By January 2019, Bitcoin had recovered to 54%, with the four currencies accounting for 78%.

Appendix Figure 4.5 shows the price time series of the virtual currencies in group one. The hype around virtual currencies for all four currency pairs, which started in early 2017 and ended abruptly as the bubble burst in the second quarter of 2018, is clearly visible. The price time series for the second group of virtual currencies valued against BTC are shown in appendix Figure 4.6. It is apparent that the price is a fraction of the Bitcoin, and that the hype is less distinct than in the first group.

To enable a comparison of virtual and actual foreign currency distributions, the four currency pairs valued against the Euro (USD, GBP, JYP, TRY) were evaluated in the third group. Appendix Figure 4.7 shows the development of the foreign exchange rate price times series accordingly. In contrast to the first two groups, the price development in group three is less erratic; even though a trend behavior is visible, it does not constitute an exponential-like explosion, as is the case with virtual currency prices. An exception is the Turkish Lira, due to economical and political up- and downswings, which thus is an interesting candidate to compare.

In general, one limitation should be noted. The worldwide foreign exchange market is the largest financial market in the world. According to the Triennial Central Bank Survey of the Bank for International Settlements (2016), the "turnover in global foreign exchange (FX) markets averaged $5.1 trillion per day in 2016." Thus, this should not be neglected when comparing both markets. Even though the virtual currency exchange market grew considerably last year, it is by far incomparable in size. When dividing the total market capitalization of the whole virtual currency market today by turnover in global foreign exchange markets, it corresponds to less than 8% of traditional foreign exchange turnovers, for one day in April.

Following the stylized facts in Cont (2001), a brief look must also be taken of the descriptive statistics of different exchange rates. As can be seen in Table 4.2, on average, the USD declined against all virtual currencies over the sample period (the mean change is positive). This is also true for intra-day virtual currencies, as can be seen in appendix Table 4.9. A mixed picture is revealed for the intra-virtual and foreign exchange rates. Virtual currencies may therefore prevail in this picture, although this might be due to the vast number of market entries in previous years, which was not experienced in the traditional foreign exchange market.

Looking at the standard deviation, it is striking that both virtual currency groups experience a much higher standard deviation than foreign currencies. Such volatility, which, by implication, is higher is also apparent when volatility clustering is considered in the different datasets. This is visualized in Figure 4.8, Figure 4.9 and Figure 4.10 in the appendix. It can be seen from these figures that virtual and intra-virtual log-returns exhibit much higher distortions than traditional currencies, while the clustering itself is more diffuse, too. Moreover, it is interesting to note that both volatility clustering and its intensity gathered pace over time. For example, the exchange rate returns of USD/LTC and BTC/LTC increased immediately after the hype around virtual currencies in early 2017. Nevertheless, actual currencies also display volatility clustering over time, albeit in a more sequential fashion.

Currency	N	Mean	SD	Median	Skewness	Kurtosis
Virtual Currencies						
USD/BTC	1209	0.00222	0.04115	0.00316	-0.23246	6.75165
USD/LTC	1209	0.00195	0.05974	-0.00061	1.47794	16.0588
USD/ETH	1209	0.00427	0.06987	0.00003	0.39261	7.4049
USD/XRP	1209	0.00335	0.08162	-0.00134	2.42715	34.8900
Intra-Virtual Currencies						
BTC/LTC	1209	-0.00029	0.04655	-0.00321	3.08048	33.3706
BTC/ETH	1209	0.00202	0.06042	-0.00287	0.68773	7.62495
BTC/XRP	1209	0.00113	0.07697	-0.00390	2.86402	42.6591
Foreign Currencies						
EUR/USD	1215	-0.00015	0.00534	-0.00009	-0.22685	6.91474
EUR/GBP	1215	0.00006	0.00539	-0.00007	0.91837	12.3740
EUR/JPY	1215	-0.00010	0.00587	0.	-1.04531	14.2434
EUR/TRY	1215	0.0005	0.01014	0.00010	2.20288	34.0457
Pooled Currency Groups						
Pooled Virtual Currencies	4836	$-1.26 \cdot 10^{-17}$	0.99969	-0.04111	1.01631	16.2763
Pooled Intra-Virtual Currencies	3627	$6.36 \cdot 10^{-18}$	0.99972	-0.06758	2.21075	27.8849
Pooled Foreign Currencies	4860	$1.46 \cdot 10^{-17}$	0.99969	-0.01220	0.46227	16.8945

Table 4.2: Descriptive statistics of virtual, intra-virtual and foreign exchange rates. Table notes: Variables are log-returns of the respective currencies. Pooled Exchange Rates are standardized for mean zero and standard deviation one.

The skewness of the exchange rate log-returns yields an inconsistent picture across classes. The log-returns for USD/BTC have a slightly negative (left) skew, which tends to be more pronounced than for other currencies; in contrast, BTC/XRP exchange rate log-returns experience an extreme positive (right) skew, outweighing all other pairs. All currencies are leptokurtic, i.e. they have positive excess kurtosis, which is more peaked and fat-tailed than the Gaussian distribution. Similar results were provided by Boothe and Glassman (1987) back in the late 1980s for foreign exchange rate returns and by Osterrieder and Lorenz (2016) for virtual currencies in recent years. When increasing the frequency, the descriptive statistics in appendix Table 4.9 reveal that the kurtosis increases significantly, which indicates high intra-day dynamics. Comparing daily virtual, intra-virtual and foreign currencies, the kurtosis of the first two types of currency is more extreme than that of the latter type. This becomes increasingly apparent when the currencies are pooled.

4.4 Empirical Framework

To receive logarithmic exchange rate returns, let P_i be the closing price of a virtual, intra-virtual or foreign currency at time i. Hence, a return in one period can be defined as the relative change of P between j and i, where $j = i - 1$. Thus the simple net return is given by Equation (4.1)

$$R_i = \frac{P_i - P_{i-1}}{P_{i-1}} = \%\Delta P_i. \tag{4.1}$$

Following this approach, we can define continuously compound daily returns, r_i as Equation (4.2)

$$r_i = ln(\frac{P_i}{P_j}) = ln(P_i) - ln(P_j), \tag{4.2}$$

where r_i can be called the log return. There are several advantages of using log returns, namely log-normality, approximate raw-log equality and time-additivity. These advantages result in the simplicity of multi-period returns, amongst others Ruppert (2004).

Building on the theoretical findings described above, the two distributions adduced to compare the distributional properties of traditional and virtual currency exchange rates are the Laplace distribution and the Exponential Power, or Subbotin, distribution. These distributions which are fitted against the virtual, intra-virtual and foreign currencies in this paper, are frequently used in finance; see Satchell and Knight (2000) for the Subbotin distribution, for example. To the Laplace distribution an important economist already drew attention in the early 20th century. In his article entitled "The principal averages and the laws of error which lead to them", Keynes (1911) focused on the Laplace distribution, emphasizing the importance it gives to the median of sample errors. This was also supported by Crum (1923) in his survey of interest rates. It therefore, comes as no surprise that the Laplace distribution also plays a vigilant role in the attempt to find the best fitting distribution for data. Kotz et al. (2001) states that "an area where the Laplace and related distributions can find most interesting and successful applications is modeling of financial data", also noting that they can be successfully applied to changes in currency exchange rate (Kotz et al., 2001).

Briefly introducing the relevant distributions, let the probability density function (PDF) of r be represented as $f(r)$. Consequently, the aforementioned two distributions are specified as follows:

- the Exponential Power Distribution (Subbotin, 1923)

$$f(r) = \frac{\kappa}{2\sigma\Gamma(\frac{1}{\kappa})} \exp\left\{-\left(\frac{|r-\mu|}{\sigma}\right)^{\kappa}\right\} \tag{4.3}$$

  for $-\infty < r < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$ and $\kappa > 0$, where $\Gamma(.)$ defines the gamma function, defined as $\Gamma(a) = \int_0^\infty t^{a-1} \exp(-t)\mathrm{d}t$.

- the Laplace Distribution (Laplace, 1774)

$$f(r) = \frac{1}{2\sigma} \exp\left(-\frac{|r-\mu|}{\sigma}\right) \tag{4.4}$$

  for $-\infty < r < \infty$, $-\infty < \mu < \infty$ and $\sigma > 0$.

The distributions were fitted by the method of Maximum Likelihood, using Mathematica to estimate the relevant parameters. Since estimating the standard errors for the Laplace and Subbotion distribution using Fisher information is not trivial (the Laplace distribution is unimodal (single-peaked) and is thus not continuously differentiable), standard errors of the parameters for the Laplace distribution were estimated using a bootstrap method. The bootstrap estimates the standard errors of parameters by drawing data points out of the dataset at random, replicating the same length of the dataset; it also estimates parameters using the maximum likelihood method. This procedure is repeated $10,000$ times. The standard deviation is then mapped on the transposed list of bootstrap estimates. The same method was used to solve for standard errors for the Subbotin distribution. The results for all estimates are shown in appendix Table 4.5, with standard errors in parentheses. The results for the virtual intra-day currency exchange rate returns are displayed in Table 4.10 in the appendix.

It is interesting to briefly interpret the special case for $\kappa = 1$ of the Subbotin distribution, where the Subbotion distribution equals a Laplace distribution. Hence, both distributions are nested (Cox and Hinkley, 1979). As can be seen in appendix Figure 4.11, the first three virtual exchange rate shape parameters are slightly below unity, when considering the top of the double standard error bands shown in gray. However, USD/XRP and the shape parameter for intra-virtual currencies differ to those of first three somehow. The red dots describe the estimated parameter κ; the gray dots indicate positive and negative double standard errors. The results are remarkable for foreign exchange rates. The shape parameter is basically unity. As such, the shape may be best described by the Laplace distribution for traditional currencies, as well as for the most important virtual currencies. Intriguing are also the shape parameter values shown in Table 4.10 for the intra-day virtual currency data; except for USD/BTC, the κ values tend to be around unity.

To validate the visual impressions of appendix Figure 4.11, supporting goodness of fit tests were also conducted. These tests were the Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) statistics, which are well-known and widely used methods for discriminating between fitted distributions. The results of the test statistics are given in Table 4.6 in the appendix. In general, the smaller the value of the test statistics of the distribution, the better the distribution fits the data. The tests support the Laplace distribution best, considering the constraints of the Subbotin distribution (number of parameters / degrees of freedom), which must also be considered. The Laplace distribution is significant especially for almost all virtual and traditional currencies (daily log returns), and also achieves lower statistics values than the Subbotin distribution in some cases. The test statistics for the virtual currency intra-day data can be found in appendix Table 4.11.

More importantly, a likelihood ratio test was used to distinguish nested distributions, which is true for the Laplace and Subbotin distribution(Cox and Hinkley, 1979). It can be introduced as follows: Let L_1 be the maximum value of the likelihood without an additional assumption, and let L_0 be the maximum value of the likelihood where the parameters are restricted and reduced in number, based on an assumption. In our case, we want to distinguish whether the Subbotin or Laplace distribution yield significantly different results. L_1 displays the likelihood of the Laplace distribution and L_0 the likelihood of the Subbotin distribution, where the shape parameter κ is fixed to unity. Then the ratio forms as $\lambda = L_0/L_1$ and χ^2 can be calculated by $\chi^2 = -2ln\lambda$. If the calculated value is significantly higher than the counter value to the $100(1-\alpha)$ percentile point of a Chi-Square distribution with k degrees of freedom, we can reject the hypothesis that, in our case, the Subbotin distribution would yield significantly better results than the Laplace distribution. The Subbotin has three parameters and the Laplace two parameters, thus the degrees of freedom $k = 1$. As can be seen in appendix Table 4.7, the Laplace distribution yields better, but more parsimonious, test results than the Subbotin distribution. Therefore, the likelihood ratio test fails to confirm that the Subbotin distribution yields significantly better results than the nested Laplace distribution, when exploring virtual and intra-virtual currencies with daily returns. Thus the likelihood ratio test also backs the previously conducted goodness of fit tests. This robustness will be heightened when visually inspecting the outcome in the section 4.5. The ratio test results for the intra-day virtual currencies are displayed and briefly discussed in Table 4.12 in the appendix.

Finally, to examine the comparability of daily virtual and intra-virtual to traditional foreign currencies, the two stylized facts shortly introduced in section 4.2 are reviewed. An examination is therefore undertaken to determine whether exchange rate increments are uncorrelated, and whether changing volatility regimes results in long-range dependence. In statistics, the autocorrelation, or serial correlation, of a random process is the Pearson correlation between values of the process at different times, as a function of the two times or the time lag. This function is applied versatilely in financial and economic time series analysis (Dunn, 2014; Yaffee and McGee, 2000). If these phenomena can be confirmed, it would be another indication that virtual, intra-virtual and actual foreign exchange rates share some statistical and distributional properties.

4.5 Results

Figure 4.1 shows the binned empirical densities of the virtual exchange rate log-returns, displaying the characteristic tent-shape of a Laplace distribution on a semi-log scale. Dispersions of the log-return in the lower log-scale are more frequent and extreme, similar to the intra-virtual group, but less so than the traditional currency group described below. As such, USD/BTC provides a more consistent picture, whereas the log-returns of USD/LTC and USD/XRP, for example, exhibit extreme events on the right-hand side of the distribution. The results obtained for the intra-day virtual exchange rate returns are shown in Figure 4.15.

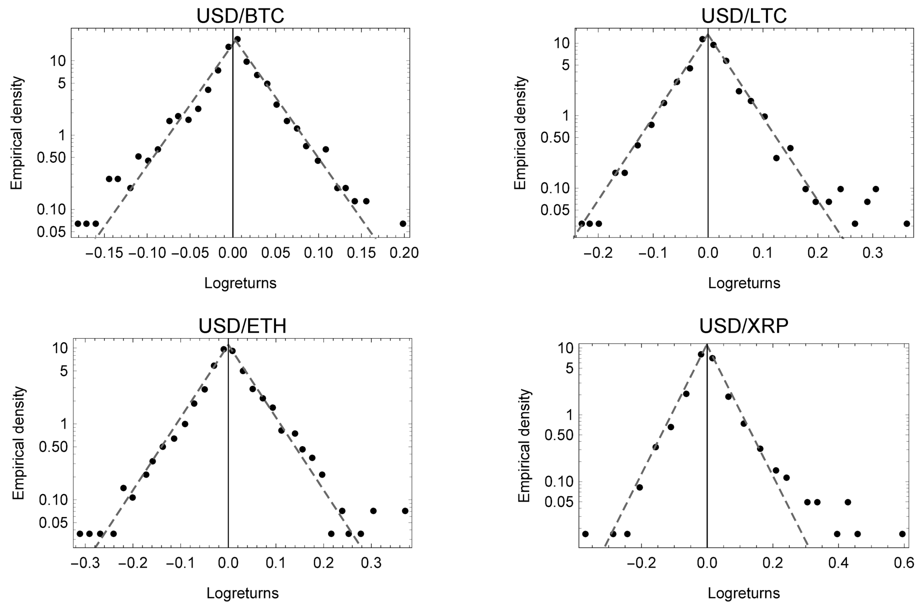


Figure 4.1: Empirical densities of virtual currencies. Empirical densities of currency exchange rate log-returns for the virtual currencies.

Figure 4.2 shows the binned empirical densities of the intra-virtual exchange rate log-returns, also displaying the characteristic tent-shape of a Laplace distribution on a semi-log scale. As with virtual log-returns, intra-virtual log-returns are sharper (more leptokurtic) than the Laplace itself, when considering the shape parameter of the Subbotin distribution. Moreover, the values of parameter κ are further from unity than virtual currencies. However, intra-virtual daily log-returns seem to fit a superimposed Laplace distribution quite well. Comparing the log-return on the x-axis to that of virtual and foreign currencies, we see that they are more similar to the virtual currencies, but more extreme than foreign currencies, while still inheriting the same functional form.

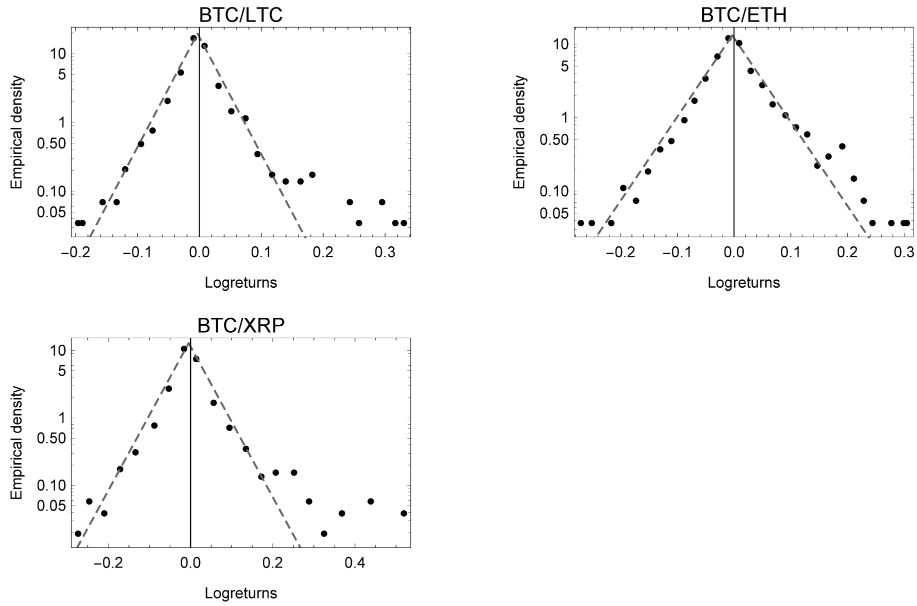


Figure 4.2: Empirical densities of intra-virtual currencies. Empirical densities of currency exchange rate log-returns for the intra-virtual currencies.

The binned empirical densities of foreign exchange rates on a semi-log scale are shown in Figure 4.3. Strikingly, as already described, the shape parameter κ perfectly matches the Laplace nested in the Subbotin distribution. It therefore, comes as no surprise that the results support the findings of Boothe and Glassman (1987) from the late 1980s for foreign exchange rate returns. Moreover, comparing the three plots, the analysis shows that daily log-returns of all currency types, share tent-shaped empirical densities, in spite of their differing mode of formation. This peculiar property also applies, for example, to firm growth or profit rates, as explained above. Also, when comparing the daily empirical densities of virtual currencies in Figure 4.1 to those of foreign currencies in Figure 4.3, we can see that extreme values of log-returns occur much more often and in a more extreme fashion, and therefore generate heavier tails. Moreover, even the hype of 2017 and 2018 leads to significant differences (as can be seen in the lower part of the semi-log plots), yet the same functional form prevails.

Before pooling the different groups, they were standardized by subtracting the mean and dividing by the standard deviation to avoid multi-modality in the shown

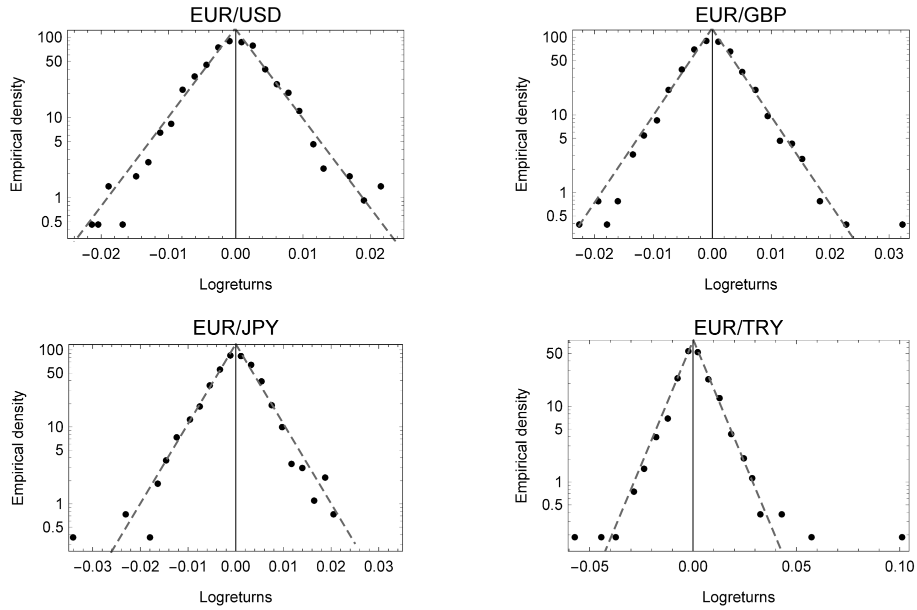


Figure 4.3: Empirical densities of foreign currencies. Empirical densities of foreign currency exchange rate log-returns for the currencies valued against the Euro.

pooled empirical densities. After pooling the different exchange rate log-returns contained in one group, it is vital to note that the general tent-shape and their distribution regularities prevail, while the shape parameter κ remains almost constant for virtual and intra-virtual currencies, slightly increasing for foreign currencies; this still points to a Laplace distribution.

Figure 4.4 shows the results obtained. Both virtual, and intra-virtual currencies have many more outliers in the lower scale of the empirical density plot, especially to the right indicating the high returns during the hype. Moreover, this may be a clear indication of the volatility and upheavals faced by these currencies in recent years.

But what is implied by the fact that all currency pairs of log-returns point, some more strongly than others, to a Laplace distribution? One explanation could be that traders of virtual currencies follow positive feedback strategies (buy when prices rise, sell when prices fall). This means that, "if rational speculators early buying triggers positive feedback trading, then an increase in the number of forward-looking speculators can increase volatility about fundamentals (De Long et al., 1990)." This explanation also holds for firm growth rates, for example, as explained in Bottazzi and Secchi (2003). This may be a suitable explanation, as many agents followed the trend of rising virtual currency prices, especially over the last two years, when the hype around virtual currencies gathered pace. It can also result from extrapolative expectations about the future price of an asset or trend shaping, driving prices up exponentially. A similar phenomenon was shown by Garcia et al. (2014) for Bitcoin, where two positive feedback loops were identified, implying a constant increase in price. However,

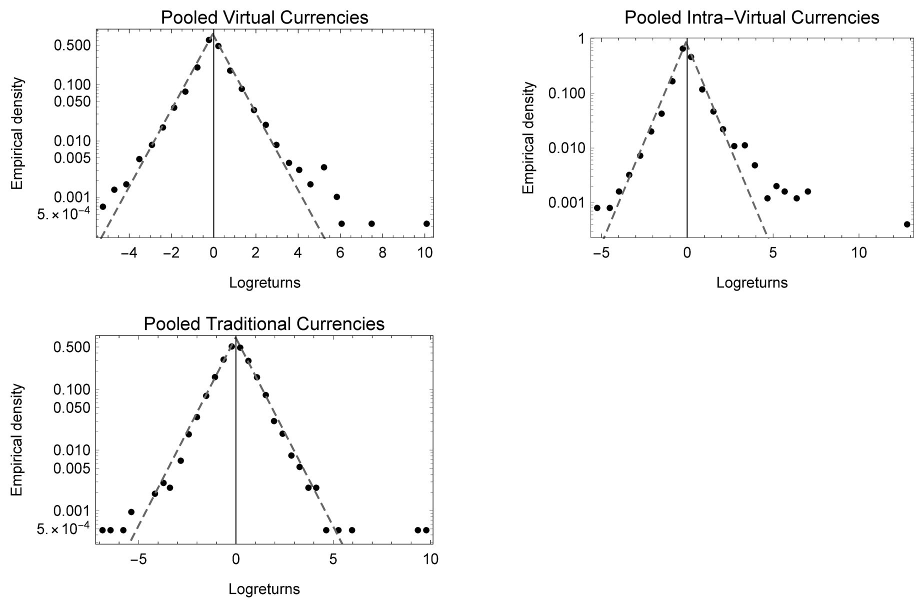


Figure 4.4: Empirical densities of standardized and pooled currencies. Empirical densities of standardized and pooled currency log-returns for all three groups of virtual, intra-virtual and foreign currency exchange rates.

these explanations fail to explain the sudden negative shocks that drove down Bitcoin prices. Now, that Bitcoin and other virtual currencies have lost a substantial share of their value over a few months, the aspect of sudden negative shocks may also be investigated in the future.

But why are log returns rather tent-shaped? The Laplace distribution is also called double exponential distribution, because it can be viewed as two exponential distributions brought together, back-to-back. This makes the Laplace distribution more leptokurtic than normal distribution and results in a single towering peak. Following Kotz et al. (2001), the Laplace random variable can be represented as the difference between two i.i.d. exponential random variables. According to McGill (1962), this is important because "the difference, and hence the Laplace distribution, provides a characterization of the error in a timing device that is under periodic excitation."

To compare the properties of exchange rate returns of traditional, foreign currencies with virtual and intra-virtual currencies, further stylized facts of logarithmic returns were analyzed, as explained in section 4.2. The plots on the left side of Figure 4.12, Figure 4.13 and Figure 4.14 in the appendix show the autocorrelation function of virtual, intra-virtual and foreign exchange rate returns. Analogously the results for the intra-day virtual currencies can be found in appendix Figure 4.16. As in a wide range of different speculative markets, the universal property of uncorrelated raw returns can also be found in virtual and intra-virtual currencies. This means that the evolution of all three types of exchange rate resemble a random walk. Interpreting this factor economically, one could also argue that, in the sense of Cont (2001), the (weak) condition of the efficient market hypothesis (EMH) is fulfilled, as arbitrage does not seem possible, and new information is immediately factored in the price-forming process.

Interestingly, this is not the case for the USD/LTC and USD/XRP exchange rate returns, all intra-day virtual some daily intra-virtual currencies and the EUR/TRY, at least on visual inspection. Also, the Ljung-Box and Box-Pierce tests reject the null hypothesis that the data are uncorrelated to lag-eight at the 5 percent level. The same holds for all intra-virtual exchange rate returns. The results of the Box-Ljung and Box-Pierce tests are given in Table 4.8 for the daily data and in Table 4.13 for the virtual intra-day data in the appendix. Hence, the autocorrelation encountered in all five daily currency pairs and the intra-day data would open the door for statistical arbitrage, a simple strategy with positive expected earnings (Cont, 2001). With the high transaction costs mentioned in section 4.2, this arbitrage strategy may not be a suitable explanation for this effect. Nonetheless, the unequal possession of several individuals in this exclusive market may be an explanation. As is well known, in several markets a disproportionately small number of individuals accumulate large sums of virtual currencies (Bloomberg, 2017). Hence, if their market capital becomes overpowering, they may be able to capitalize on their position by transferring outstanding amounts of capital, while still leveraging the transaction investment.

Another stylized fact that helps to distinguish whether virtual and intra-virtual exchange rates resemble actual foreign currencies is long-range dependence or long memory. This means that periods of low volatility are followed by periods of high volatility; they can be recognized by the fact that the autocorrelation function of absolute returns is positive, and decays slowly. Here, it is true for virtual and intra-virtual currencies, as can be seen on the right side of Figure 4.12, Figure 4.13 and Figure 4.14 and also for intra-day virtual currencies of Figure 4.16 in the appendix. Intriguingly, the long-range dependence of virtual and intra-virtual currencies decays much more quickly, and the correlation is far higher than in traditional currencies, rather resembling other, more speculative asset categories. There may be two reasons for this: First of all, the virtual currency market is far more speculative, and is driven by trend-following agents who rely on increasing prices, while avoiding the market or not knowing about it. Examples are the volatility increase in USD/LTC or BTC/LTC trade in early 2017, as can be seen in 4.8 and 4.9 in the appendix. Another explanation may be that, in contrast to the foreign exchange market, the virtual currency market is more exclusive and thus contains fewer participants by several orders of magnitude, in contrast to the large number of agents in the foreign exchange market.

4.6 Discussion and Conclusion

The virtual and intra-virtual currency market share the rather turbulent dynamics of a traditional foreign exchange market, and even more extremely to some extent. It has been empirically shown that virtual, intra-virtual and foreign currency log-returns are tent-shaped, have the characteristics of a Laplace distribution at the semi-log scale, and share the same functional form. However, this non-Gaussianity questions conditions of the central limit theorem, more specifically the independence assumption. Furthermore, virtual and intra-virtual currencies exhibit greater volatility, fatter tails and steeper towering peaks than regular foreign currencies, while following the stylized facts of asset returns in a more extreme fashion to some extent. It may therefore be better to view virtual and intra-virtual currencies as a speculative investment, as noted by Yermack (2015). The results also contribute to a recent strand of literature comparing and classifying virtual currencies to other asset categories. Since autocorrelation has also been identified in daily log-returns for one virtual currency and all intra-virtual currencies, statistical arbitrage may be possible. This supports the findings of Gandal and Halaburda (2016), identifying arbitrage strategies due to co-movement.

To answer the question of what is left after the hype, it can be said that daily log-returns of all currency types share, tent-shaped empirical densities, one of the characteristics of a Laplace distribution at the semi-log scale, in spite of the differing mode of formation, technology and speculative intensity in recent years. This peculiar property is shared with several asset types and applies to firm growth and profit rates, for example Bottazzi and Secchi (2003); Fagiolo et al. (2007); Alfarano and Milaković (2008). Moreover, the results highlight the fact that virtual and traditional currencies hold the same functional form, even after the hype of 2018.

Although a direct comparison of the results is not possible due to the difference in frequencies it is important to note that when shifting the frequency to intra-day price observations, the distributional properties of virtual currencies seem to change, too. Even though the Laplace and Subbotin distribution seem to fit the data well in the tent-shaped middle, they fail to describe the numerous extreme events happening within a day being pictured by the tails. Therefore, finding a distribution which fits higher frequency data of virtual and traditional data as good as the Laplace does for daily log-returns would be a valuable investigation in the future. Yet, due to the lag of free available intra-day data of traditional currencies, this might be a cumbersome task. Furthermore, as evidence for autocorrelation has been presented, it would be interesting to investigate whether statistical arbitrage strategies can be or have been realized for virtual currencies, or whether this is merely an artifact in the perception or representation of any information used and induced by the technology involved. It would also be interesting to explore whether there are any interdependencies between the currency groups. To this end, it would be interesting to see how, and if, idiosyncratic external shocks (Brexit, the presidential election in the US) influenced the price of foreign and virtual currencies, for instance. A first guess would be that there is no such effect for virtual and intra-virtual currencies, while there is for foreign currencies.

Summing up, what is left after the hype is a peculiar statistical regularity shared with many asset classes. But, in the future, virtual currencies may be of interest as an asset, decoupled from the conventional financial system. However the risk, and therefore uncertainty, in the virtual exchange rate market seems to be higher due to its speculative nature, and must be valued with caution, especially following the extreme hype in 2017 and the massive crash of 2018.

4.A Appendix

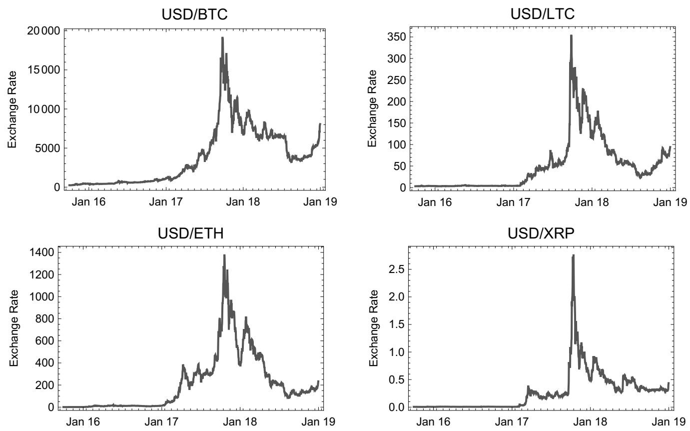


Figure 4.5: Price time series of virtual currencies. Closing price development of virtual currency exchange rates.

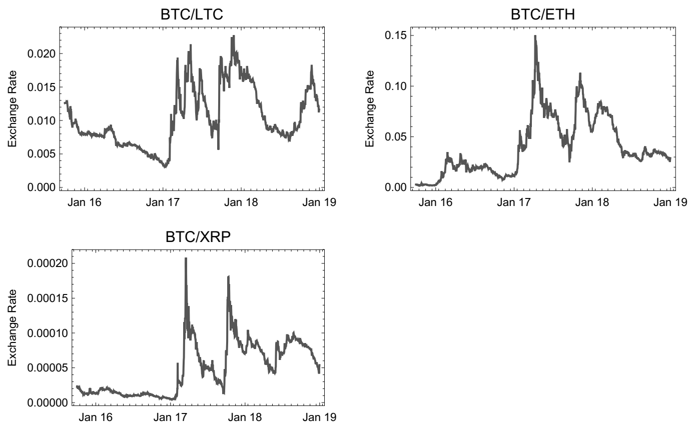


Figure 4.6: Price time series of intra-virtual currencies. Closing price development of intra-virtual currency exchange rates.

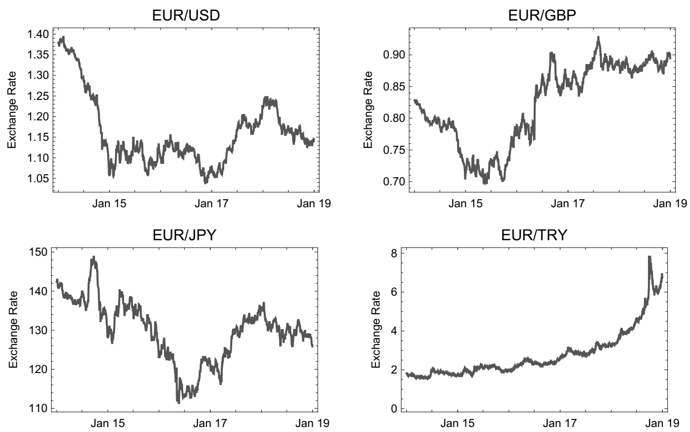


Figure 4.7: Price time series of foreign currencies. Closing price development of foreign currency exchange rates.

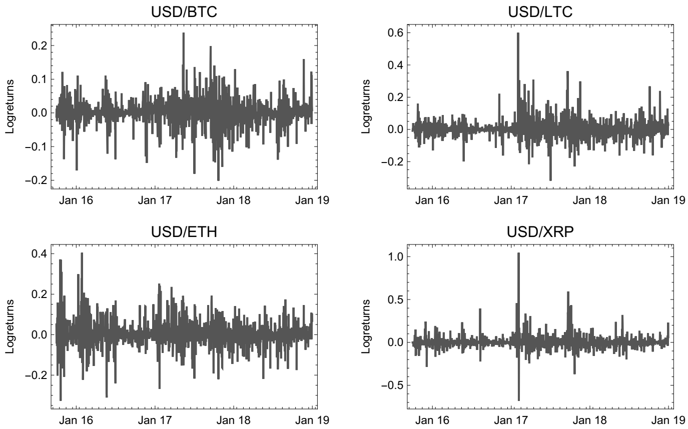


Figure 4.8: Volatility clustering in virtual currency log-returns. Volatility clustering in log-returns of the virtual to real currency exchange rate group.

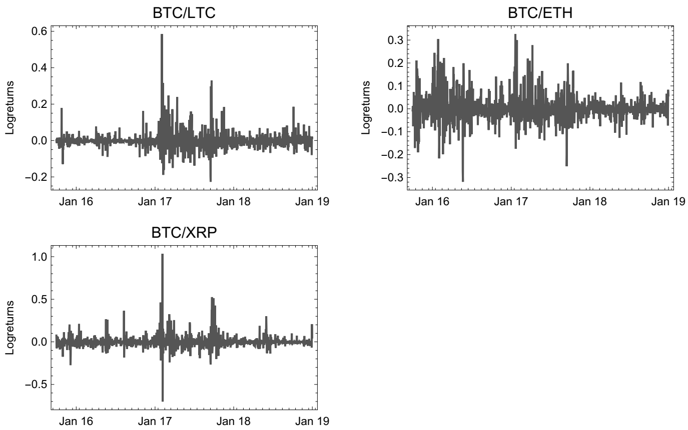


Figure 4.9: Volatility clustering in intra-virtual currency log-returns. Volatility clustering in log-returns of the intra-virtual to virtual currency exchange rate group.

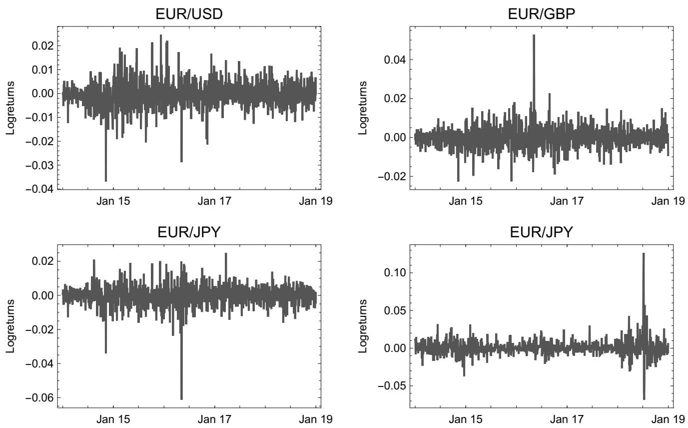


Figure 4.10: Volatility clustering in foreign currency log-returns. Volatility clustering in log-returns of foreign currency exchange rate group.

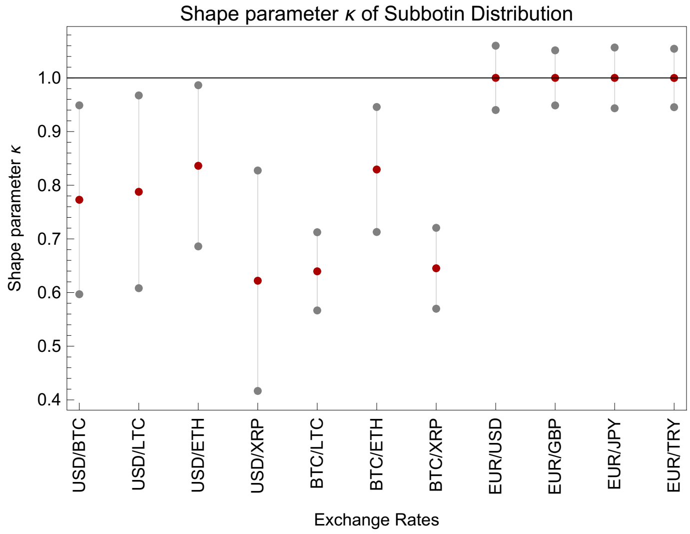


Figure 4.11: Shape parameter κ of Subbotin distribution. Shape parameter κ of Subbotin distribution for virtual, intra-virtual and foreign currency exchange rate log-returns (daily).

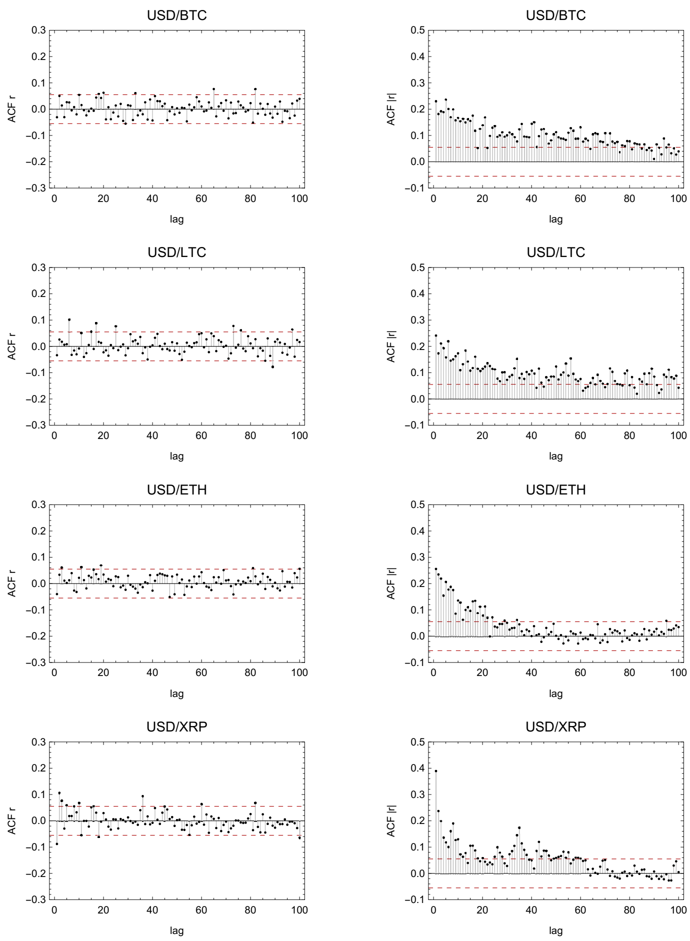


Figure 4.12: Autocorrelation functions of virtual currencies. Autocorrelation function of raw and absolute returns of virtual currency exchange rates.

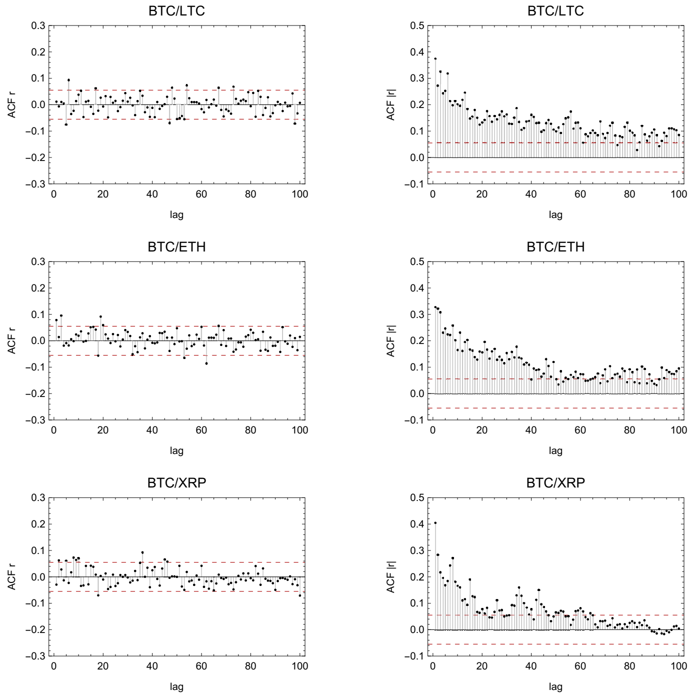


Figure 4.13: Autocorrelation functions of intra-virtual currencies. Autocorrelation function of raw and absolute returns of intra-virtual currency exchange rates.

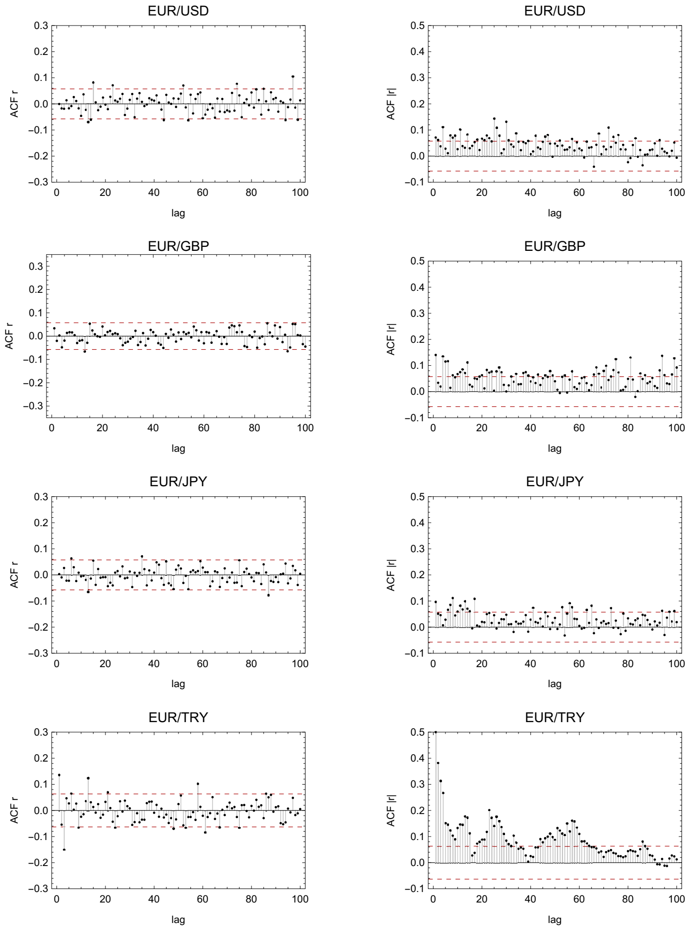


Figure 4.14: Autocorrelation functions of foreign currencies. Autocorrelation function of raw and absolute returns of foreign currency exchange rates.

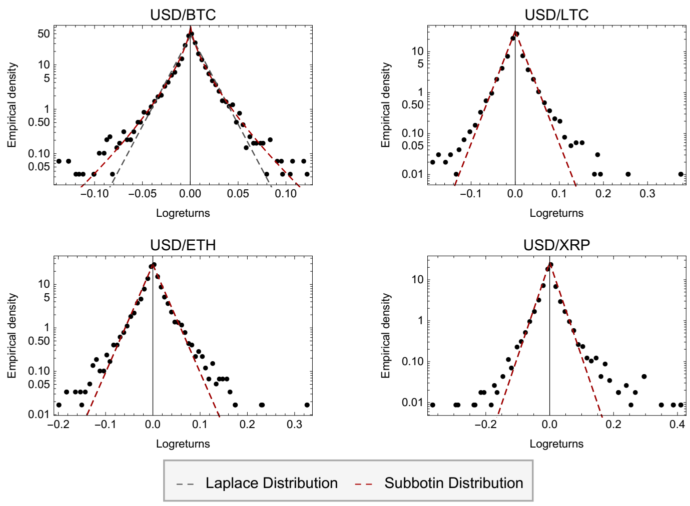


Figure 4.15: Empirical densities of intra-day virtual currencies. Empirical densities of currency exchange rate log-returns for the intra-day virtual currencies.

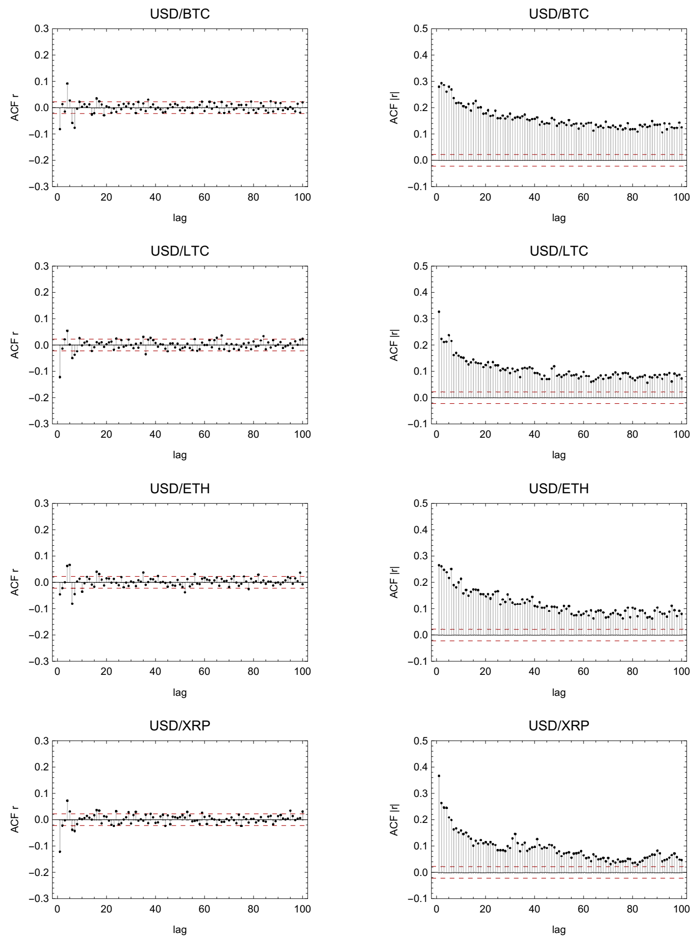


Figure 4.16: Autocorrelation functions of intra-day virtual currencies. Autocorrelation function of raw and absolute returns of intra-day virtual currency exchange rates.

Currency	Market Cap	Volume (24h)	Available Supply	Maximum Supply
Bitcoin	$227,578,237,659	$11,534,500,000	16,802,262 BTC	21,000,000 BTC
Ethereum	$127,953,096,946	$5,216,660,000	96,973,835 ETH	—
Ripple	$70,987,542,244	$1,846,320,000	38,739,142,811 XRP	100,000,000,000 XRP
Litecoin	$13,051,269,334	$1,041,820,000	54,751,458 LTC	84,000,000 LTC

Table 4.3: Market capitalization and trade volume information for selected virtual currencies - January 2018. The information was retrieved from CoinMarketCap (2017).

Currency	Market Cap	Volume (24h)	Available Supply	Maximum Supply
Bitcoin	$15,078,898,496	$95,527,900	16,135,862 BTC	21,000,000 BTC
Ethereum	$936,697,745	$8,675,030	88,424,437 ETH	—
Ripple	$235,014,889	$409,345	36,856,524,148 XRP	100,000,000,000 XRP
Litecoin	$200,552,099	$18,074,500	49,582,331 LTC	84,000,000 LTC

Table 4.4: Market capitalization and trade volume information for selected virtual currencies - January 2017. The information was retrieved from CoinMarketCap (2017).

	Laplace Distribution		Subbotin Distribution		
	$\hat{\mu}$ (SE)	$\hat{\sigma}$ (SE)	$\hat{\kappa}$ (SE)	$\hat{\mu}$ (SE)	$\hat{\sigma}$ (SE)
Virtual Currencies					
USD/BTC	0.00316	0.02747	0.77291	0.00309	0.02433
	(0.00059)	(0.00088)	(0.08801)	(0.00057)	(0.00147)
USD/LTC	0.00061	0.03822	0.787747	$-1.14 \cdot 10^{-9}$	0.03411
	(0.00089)	(0.00132)	(0.08980)	(0.00105)	(0.00192)
USD/ETH	0.00003	0.04709	0.83623	$-7.73 \cdot 10^{-9}$	0.04339
	(0.00112)	(0.00149)	(0.07508)	(0.00105)	(0.00218)
USD/XRP	-0.00134	0.04649	0.62206	$-1.03 \cdot 10^{-9}$	0.036448
	(0.00122)	(0.00192)	(0.10271)	(0.00143)	(0.00304)
Intra-Virtual Currencies					
BTC/LTC	-0.00321	0.02557	0.63956	-0.00316	0.02005
	(0.00063)	(0.00111)	(0.03643)	(0.00089)	(0.00074)
BTC/ETH	-0.00287	0.04009	0.82933	-0.00315	0.03676
	(0.00130)	(0.00130)	(0.05824)	(0.00149)	(0.00150)
BTC/XRP	-0.00390	0.04106	0.64525	-0.00345	0.03232
	(0.00117)	(0.00189)	(0.03769)	(0.00176)	(0.00119)
Foreign Currencies					
EUR/USD	-0.00009	0.00390	1.00009	-0.00009	0.0.00390
	(0.00016)	(0.00010)	(0.02997)	(0.00015)	(0.00010)
EUR/GBP	-0.00007	0.00385	1.00007	-0.00007	0.00385
	(0.00017)	(0.00010)	(0.02566)	(0.00017)	(0.00011)
EUR/JPY	0.	0.00419	1.00004	0.00014	0.00419
	(0.00017)	(0.00011)	(0.02831)	(0.00016)	(0.00011)
EUR/TRY	0.00010	0.00629	0.99999	0.00016	0.00664
	(0.00018)	(0.00022)	(0.02721)	(0.00021)	(0.00027)
Pooled Currency Groups					
Virtual Currencies	-0.04111	0.63885	0.75841	-0.04111	0.55943
	(0.00583)	(0.01109)	(0.01801)	(0.00266)	(0.01580)
Intra-Virtual Currencies	-0.06758	0.58214	0.68894	-0.06589	0.48008
	(0.00965)	(0.01335)	(0.01615)	(0.01677)	(0.00877)
Foreign Currencies	0.00473	-0.72007	1.17660	0.00413	0.76720
	(0.01715)	(0.01144)	(0.04494)	(0.01522)	(0.01210)

Table 4.5: Fitted distributions, parameter estimates and standard errors. Table notes: Variables are log-returns of the respective currencies. The Pooled Exchange Rates are standardized for mean zero and standard deviation one.

	Kolmogorow-Smirnow		Anderson-Darling	
	Laplace	Subbotin	Laplace	Subbotin
Virtual Currencies				
USD/BTC	0.04139	0.02105*	2.33460*	0.58989*
USD/LTC	0.03362*	0.09708	1.44721*	19.2695
USD/ETH	0.03601*	0.02703*	2.42895*	1.65442*
USD/XRP	0.04145	0.04801	5.00501	2.82935
Intra-Virtual Currencies				
BTC/LTC	0.05466	0.02262*	7.62011	1.44359*
BTC/ETH	0.03585*	0.03016*	3.61771	.583180
BTC/XRP	0.05478	0.04081	8.52987	2.97168
Foreign Currencies				
EUR/USD	0.03188*	0.03188*	1.59082*	1.59082*
EUR/GBP	0.02998*	0.02998*	1.45344*	1.45344*
EUR/JPY	0.03294*	0.03294*	1.65519*	1.65519*
EUR/TRY	0.02894*	0.03668*	1.47207*	2.62343*

Table 4.6: Kolmogorow-Smirnow and Anderson-Darling goodness of fit test statistics for fitted distributions. *: Significant for a significance level of $\alpha = 0.05$.

Currency	$-2\ln\lambda$-Value	χ^2-Benchmark	Statement
Virtual Currencies			
USD/BTC	18.553	3.841	Reject
USD/LTC	-730.876	3.841	Accept
USD/ETH	8.311	3.841	Reject
USD/XRP	79.182	3.841	Reject
Intra-Virtual Currencies			
BTC/LTC	77.626	3.841	Reject
BTC/ETH	9.392	3.841	Reject
BTC/XRP	75.266	3.841	Reject
Foreign Currencies			
EUR/USD	$1.09 \cdot 10^{-}12$	3.841	Accept
EUR/GBP	1.085	3.841	Accept
EUR/JPY	$0.11 \cdot 10^{-}12$	3.841	Accept
EUR/TRY	3.595	3.841	Accept

Table 4.7: Likelihood ratio test results. Degree of freedom $k = 1$. Since the Subbotion distribution equals a Laplace distribution for $\kappa = 1$, it comes as no surprise that the Likelihood ratio test yields better results because the κ parameters reported in Table 4.5 are close to unity.

	Ljung-Box		Box-Pierce	
	Statistic	P-Value	Statistic	P-Value
Virtual Currencies				
USD/BTC	7.07728	0.52831	7.04688	0.53158
USD/LTC	21.5188	0.00588	21.3853	0.00619
USD/ETH	11.5488	0.17249	11.4915	0.17537
USD/XRP	41.1377	$1.96 \cdot 10^{-6}$	40.9697	$2.11 \cdot 10^{-6}$
Intra-Virtual Currencies				
BTC/LTC	21.3024	0.00638	21.1649	0.00672
BTC/ETH	20.3187	0.00919	20.2461	0.00944
BTC/XRP	19.4751	0.01251	19.3628	0.01303
Foreign Currencies				
EUR/USD	2.45668	0.96372	2.44254	0.96435
EUR/GBP	6.03101	0.64375	6.00286	0.64691
EUR/JPY	8.82376	0.35738	8.76778	0.36226
EUR/TRY	56.2427	$2.59 \cdot 10^{-9}$	56.4588	$2.29 \cdot 10^{-9}$

Table 4.8: Ljung-Box and Box-Pierce test results. Table notes: The null hypothesis that data is uncorrelated to lag 8 is or is not rejected at the 5 percent level. *: These tests are rejected at the 5 percent level.

Currency	N	Mean	SD	Median	Skewness	Kurtosis
USD/BTC	8021	0.00044	0.01748	0.00047	-0.19794	14.0661
USD/LTC	8021	0.00045	0.02765	0.	0.58395	69.3499
USD/ETH	8021	0.00073	0.02867	0.	0.49880	18.0121
USD/XRP	8021	0.00054	0.03520	0.	1.05859	57.2758

Table 4.9: Descriptive statistics of intra-day virtual exchange rates. Table notes: Variables are log-returns of the respective currencies.

	Laplace Distribution		Subbotin Distribution		
	$\hat{\mu}$ (SE)	$\hat{\sigma}$ (SE)	$\hat{\kappa}$ (SE)	$\hat{\mu}$ (SE)	$\hat{\sigma}$ (SE)
Intra-Day Virtual Currencies					
USD/BTC	0.00048	0.01071	0.067661	0.00034	0.00877
	(0.00011)	(0.00015)	(0.01741)	(0.00021)	(0.00015)
USD/LTC	0.	0.01579	1.	0.	0.01579
	$(6.31 \cdot 10^{-11})$	(0.00025)	$(2.34 \cdot 10^{-10})$	$(4.91 \cdot 10^{-12})$	(0.00025)
USD/ETH	0.	0.01747	1.	0.	0.01747
	(0.00003)	(0.00025)	(0.14374)	$(1.57 \cdot 10^{-9})$	(0.00145)
USD/XRP	0.	0.01887	1.	0.	0.01887
	(0.)	(0.00032)	(0.)	(0.)	(0.00032)

Table 4.10: Fitted distributions, parameter estimates and standard errors for intra-day exchange rates. Table notes: Variables are log-returns of the respective currencies.

	Kolmogorow-Smirnow		Anderson-Darling	
	Laplace	Subbotin	Laplace	Subbotin
Intra-Day Virtual Currencies				
USD/BTC	0.07611	0.02196	66.1883	4.46575
USD/LTC	0.04523	0.04523	38.7224	38.7224
USD/ETH	0.04893	0.04893	29.3384	29.3384
USD/XRP	0.06409	0.06409	62.2389	62.2389

Table 4.11: Kolmogorow-Smirnow and Anderson-Darling goodness of fit test statistics for fitted distributions on virtual currency intra-day data. *: Significant for a significance level of $\alpha = 0.05$.

Currency	$-2 \ln \lambda$-Value	χ^2-Benchmark	Statement
Intra-Day Virtual Currencies			
USD/BTC	342.780	3.841	Reject
USD/LTC	$0.05 \cot 10^{-12}$	3.841	Accept
USD/ETH	$0.07 \cot 10^{-12}$	3.841	Accept
USD/XRP	$0.09 \cot 10^{-12}$	3.841	Accept

Table 4.12: Likelihood ratio test results for virtual intra-day data. Table notes: Degree of freedom k = 1. Since the Subbotion distribution equals a Laplace distribution for $k = 1$, it comes as no surprise that the Likelihood ratio test yields better results because the κ parameters for the intra-day data reported in Table 4.10 are close to unity. This leads to the impression that the Subbotin distribution fits the intra-day data better, as the test statistics displayed in Table 4.10 are also lower, however not significant.

	Ljung-Box		Box-Pierce	
	Statistic	P-Value	Statistic	P-Value
Intra-Day Virtual Currencies				
USD/BTC	208.251	$6.15 \cdot 10^-40$*	208.089	$6.66 \cdot 10^-40$*
USD/LTC	188.155	$1.00 \cdot 10^-35$*	188.045	$1.05 \cdot 10^-35$*
USD/ETH	158.326	$1.65 \cdot 10^-29$*	158.190	$1.76 \cdot 10^-29$*
USD/XRP	201.597	$1.53 \cdot 10^-38$*	201.481	$1.62 \cdot 10^-38$*

Table 4.13: Ljung-Box and Box-Pierce test results for virtual intra-day data. Table notes: The null hypothesis that data are uncorrelated to lag 8 is or is not rejected at the 5 percent level. *: These tests are rejected at the 5 percent level.

Summary and Outlook

"All theories tend to shape the facts they try to explain; any theory may thus turn into a procrustean bed. Our proposed theoretical formulation is designed to protect the investigator from this danger: it does not permit him [her] to draw any special or general conclusions before he [she] or someone else completes the always difficult and seldom glamorous task of ascertaining the necessary facts."

- Wassily Leontief -

Lecture to the memory of Alfred Nobel, December 11, 1973

Complex economic dynamics at different stages of the economic process can have a lasting effect on the underlying economic structure, its participants and the instruments they use. The aspects responsible for these dynamics include the position and size of such economic entities within their network of interaction, individual prerequisites such as productivity and complementarity , as well as their collective behavior. This dynamic environment can also create regularities in the form of statistical distributions in various market settings, which are neither intended nor planned by those interacting. This thesis addresses some of these issues from a variety of perspectives.

Chapter 1 introduces a parsimonious model that allows for the adjustment of sales shares in response to productivity shocks, and operationalizes it to quantify the impact of idiosyncratic shocks on fluctuations in the aggregate. In the process, we demonstrate that previous approaches substantially underestimate the importance of microeconomic shocks, leading to higher-order macroeconomic effects, because they fail to account for what we call the mesoeconomic interaction of granularity and network effects. Mesoeconomic objects are still subject to microeconomic forces, in our case the microeconomic elasticities of substitution, but the properties resulting from the interaction between industries cannot be reduced to their individual micro behavior. The mesoeconomic interaction of granularity and network effects is therefore less intuitive when the two conditions work in opposite directions. To estimate the relative importance of these two effects, we also derive the polar cases of an economy that is homogeneous in its network structure but heterogeneous in its size distribution (the purely granular case), and an economy that is heterogeneous in its network structure but homogeneous in its size distribution (the purely networked case). Building on these three cases, the main findings in this chapter suggest that a precise and stable perception of business cycle fluctuations can only be achieved if account is taken of the significant asymmetries in the characteristics of the entities and their interactions. With this in mind, our mesoeconomic aggregation rule suggests a number of directions for future research. First, since homogeneity is assumed within industries, which is a limiting assumption in any input-output table at the industry level, the (technological) heterogeneity that exists within industries cannot be fully captured. Therefore, applying our rule to firm data would allow a more systematic analysis and behavioral assessment at a fine-grained level. Second, since it is a non-parametric model, we are unable to study the impact of specific structural characteristics such as geographical or market frictions and non-technological shocks in general. This issue is tackled in part in chapter 3, albeit from a different perspective. Third, another important area for future research relating to the findings of Chapter 1 could be a systematic analysis of the relationships between financial networks or monetary policy and the extent of contagion and cascading failure of production networks.

Chapter 2 introduced a joint method based on structural decomposition analysis (SDA) and principal component analysis (PCA), which was applied to input-output data in order to decompose cyclical and apparently synchronized European GDP time series into their constituent parts, namely into a markup, inter-industry and final goods market component. We show further that the application of a PCA allows us to disentangle the aggregate and idiosyncratic part of these constituent elements, investigating whether pairwise linkages or higher-order linkages lead to the observed comovement of European business cycles. Empirical evidence suggests that macroeconomic volatility in Europe is mainly explained by final goods markets rather by than changes in productivity or market power. We further provide evidence that the large degree of comovement of GDP time series is mainly attributable to aggregate developments through higher-order connections rather than direct pairwise linkages. Taken together, Part I aims to provide new insights into how sizable higher-order macroeconomic effects and regularities, such as cyclicities, emerge from their underlying economic structure.

Despite the importance of heterogeneous interactions in the research of macroeconomic fluctuations, depending on international trade and production networks, there has been a surprising lack of attention to the fact that economic interactions are embedded in an evolutionary process of production and exchange across borders globally, which evolves over time. This issue is tackled in Part III. In the process, Chapter 3 elaborates this matter by providing a stochastic actor-oriented model (SAOM) for network dynamics. This framework allows us to model a variety of network graph types, and infer a wide range of influences on network changes over time. This allows us to estimate the effects of specific prerequisites, and to test corresponding hypotheses. For this, it is assumed that the network develops as a stochastic, actor-oriented process, which is partly determined endogenously depending on the network structure, and partly exogenously depending on specific node characteristics. Since the SAOM incorporates the information from at least two observation waves, creating an additional heterogeneity of the network statistics that can be exploited for parameter estimation, it makes the model less susceptible to convergence problems. As such SAOMs seem to be the natural choice for studying network dynamics. Using them, we identify various important features for production networks in international trade, such as network proximity, geographical distance and technological complements, which are crucial drivers for decision processes in trade link formation. We are also able to confirm results on the size of entities in trade activities, while providing new insights on (asymmetric) gains from free trade agreements, comparing developing and developed countries. In addition, we contribute to the discussion that examines the direction of, and reasons for, trade flows in global value chains. We find that OECD import activities from specific industries outside the OECD promote the creation of export activities to these very non-OECD industries. However, this so-called entrainment effect is not conditional from exports on imports. This adds to discussions in the international trade literature on dependencies and intermediate relations in production. As such, the present work represents an important step in filling the empirical gap in the relatively young field of endogenous production networks. Some of the characteristics mentioned above have been identified as crucial determinants of shock propagation and macroeconomic outcomes. Thus, the remarkable stability of these aggregated characteristics, even at the sectoral level, testifies to their general importance even, or especially, in a dynamically developing environment. However, since the SAOM modeling approach proves to be an effective tool to estimate channels of network formation from an economic perspective, further efforts need to be made in the understanding of socioeconomic (mis)behavior leading to the outcomes we observe. Understanding these phenomena may enable us to elaborate more just and sustained growth strategies, while redressing inequality and exploitation. The SAOM methodology is flexible enough to accommodate for the multiplexity of economic dimensions.

Finally, Chapter 4 provides an empirical analysis of the distributional properties and statistical regularities of virtual, intra-virtual and traditional currency exchange rates. Despite differing from the topics presented above, this analysis also aims to provide insights into interactions in markets which lead to aggregate regularities. In this case, it can be observed in the form of statistical distributions. The analysis shows that, in spite of their differing modes of formation, daily log-returns of all currency types share tent-shaped empirical densities, one of the characteristics of a Laplace distribution at semi-log scale. This peculiar property has also been examined thoroughly in other fields of economic literature. To address the question preceding the title about what is left after the hype, it can be said that virtual currencies also exhibit to a peculiar statistical regularity that is shared by many other asset classes, but is not foreseen or intended by the interacting market participants. More interesting aspects for future research may involve questions about the implications of these peculiar properties for the competitive environment that private and institutional investors face when competing in these currency markets. It would also be worthwhile elaborating questions related to whether statistical arbitrage strategies can be, or have been, realized for virtual currencies, or whether this is merely an artifact in the perception or representation of any information used and induced by the technology involved. Despite being decoupled from the conventional financial system until now, and being equipped with new technological possibilities, virtual currencies offer opportunities and risks for real and financial economic systems, a development that should also be monitored closely.

In summary, the theoretical and empirical analysis in this dissertation shows the importance of several aspects that need to be taken into account when considering the economic interaction of heterogeneous actors within economic systems. As such, this thesis may serve as a starting point for a broader research agenda that examines the necessary facts leading to new knowledge about, theories of, and a better understanding of complex economic processes.

Bibliography

Acemoglu, D., Akcigit, U., and Kerr, W. (2016). Networks and the Macroeconomy: An Empirical Exploration. *NBER Macroeconomics Annual*, 30(1):273–335.

Acemoglu, D. and Azar, P. D. (2020). Endogenous production networks. *Econometrica*, 88(1):33–82.

Acemoglu, D., Bimpikis, K., and Ozdaglar, A. (2014). Dynamics of information exchange in endogenous social networks. *Theoretical Economics*, 9(1):41–97.

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

Ahiakpor, J. C. (1985). The success and failure of dependency theory: The experience of Ghana. *International Organization*, pages 535–552.

Akaike, H. (1973). *Information Theory and an Extension of the Maximum Likelihood Principle*, pages 199–213. Springer New York, New York, NY.

Albert, R. and Barabási, A.-L. (2000). Topology of evolving networks: Local events and universality. *Physical Review Letters*, 85(24):5234.

Alfarano, S. and Milaković, M. (2008). Does classical competition explain the statistical features of firm growth? *Economics Letters*, 101(3):272–274.

Allen, T. (2014). Information frictions in trade. *Econometrica*, 82(6):2041–2083.

Ascari, G., Fagiolo, G., and Roventini, A. (2015). Fat-tail distributions and business cycle models. *Macroeconomic Dynamics*, 19(2):465–476.

Atalay, E. (2017). How important are sectoral shocks? *American Economic Journal: Macroeconomics*, 9(4):254–80.

Atalay, E., Hortaçsu, A., Roberts, J., and Syverson, C. (2011). Network structure of production. *Proceedings of the National Academy of Sciences of the United States of America*, 108(13):5199–5202.

Baier, S. L. and Bergstrand, J. H. (2007). Do free trade agreements actually increase members' international trade? *Journal of International Economics*, 71(1):72–95.

Bank for International Settlements (2016). Triennial central bank survey. Technical report, Monetary and Economic Department.

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

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

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

Barber, S., Boyen, X., Shi, E., and Uzun, E. (2012). Bitter to better - how to make bitcoin a better currency. *Lecture Notes in Computer Science*, 7397:399–414.

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

Baxter, M. and Kouparitsas, M. A. (2005). Determinants of business cycle comovement: A robust analysis. *Journal of Monetary Economics*, 52(1):113–157.

Bergholt, D., Larsen, V. H., and Seneca, M. (2019). Business cycles in an oil economy. *Journal of International Money and Finance*, 96:283–303.

Bernard, A. B. and Dhingra, S. (2016). Importers, exporters and the division of the gains from trade.

Bernard, A. B., Dhyne, E., Magerman, G., Manova, K., and Moxnes, A. (2019a). The Origins of Firm Heterogeneity: A Production Network Approach. *NBER Working Paper Series*, 1(25441).

Bernard, A. B., Dhyne, E., Magerman, G., Manova, K., and Moxnes, A. (2019b). The Origins of Firm Heterogeneity: A Production Network Approach. *NBER Working Paper Series*, 1(25441).

Bernard, A. B. and Moxnes, A. (2018). Networks and Trade. *Annual Review of Economics*, 10(1):65–85.

Böhme, R., Christin, N., Edelman, B., and Moore, T. (2015). Bitcoin: Economics, technology, and governance. *Journal of Economic Perspectives*, 29(2):213–38.

Bitcoin Project (2016). Frequently asked questions about bitcoins. https://bitcoin.org/en/faq. [Online; accessed 2016-12-27].

Bitcoinity (2016). Exchanges. https://bitcoinity.org/markets/list?currency=ALL&span=6m. [Online; accessed 2016-12-27].

Bitcoinity (2018). Exchanges. https://bitcoinity.org/markets/list?currency=ALL&span=6m. [Online; accessed 2018-01-14].

Bitcoinity (2019). Exchanges. https://bitcoinity.org/markets/list?currency=ALL&span=6m. [Online; accessed 2019-01-22].

Bitfinex (2017). The world's leading cryptocurrency trading platform. https://www.bitfinex.com/. [Online; accessed 2019-01-22].

Bitinfocharts (2019). Transaction fees. https://bitinfocharts.com/comparison/bitcoin-transactionfees.html#1y. [Online; accessed 2019-01-22].

Blackorby, C. and Russell, R. R. (1989). Will the real elasticity of substitution please stand up? (A comparison of the Allen/Uzawa and Morishima elasticities). *American Economic Review*, 79(4):882–888.

Blades, D. and Meyer-zu Schlochtern, J. (1998). How to represent capital in international comparisons of total factor productivity. *Second Meeting of the Canberra Group on Capital Stock Statistics*.

Blanco-Arroyo, O., Ruiz-Buforn, A., Vidal-Tomás, D., and Alfarano, S. (2018). On the determination of the granular size of the economy. *Economics Letters*, 173:35–38.

Blockchain.info (2016). Hashrate distribution. https://blockchain.info/pools. [Online; accessed 2016-12-28].

Blockchain.info (2018). Pool share. https://blockchain.info/pools. [Online; accessed 2018-01-14].

Blockchain.info (2019a). Average confirmation time. https://blockchain.info/de/charts/avg-confirmation-time?timespan=30days. [Online; accessed 2019-01-22].

Blockchain.info (2019b). Pool share. https://blockchain.info/pools. [Online; accessed 2019-01-22].

Bloomberg (2017). The Bitcoin whales: 1,000 people who own 40 percent of the market. https://www.bloomberg.com/news/articles/2017-12-08/the-bitcoin-whales-1-000-people-who-own-40-percent-of-the-market. [Online; accessed 2018-02-08].

Blum, B. S., Claro, S., and Horstmann, I. (2010). Facts and figures on intermediated trade. *American Economic Review*, 100(2):419–23.

Boehm, C., Flaaen, A., and Pandalai-Nayar, N. (2018). Input linkages and the transmission of shocks: Firm-level evidence from the 2011 Tōhoku earthquake. *The Review of Economics and Statistics*, forthcoming.

Boer, H. and During, W. E. (2001). Innovation, what innovation? A comparison between product, process and organisational innovation. *International Journal of Technology Management*, 22(1-3):83–107.

Bonacich, P. (1991). Simultaneous group and individual centralities. *Social Networks*, 13(2):155–168.

Boothe, P. and Glassman, D. (1987). The statistical distribution of exchange rates: Empirical evidence and economic implications. *Journal of International Economics*, 22(3):297–319.

Bottazzi, G., Li, L., and Secchi, A. (2019). Aggregate fluctuations and the distribution of firm growth rates. *Industrial and Corporate Change*, 28(3):635–656.

Bottazzi, G. and Secchi, A. (2003). Why are distributions of firm growth rates tent-shaped? *Economics Letters*, 80(3):415–420.

Briere, M., Oosterlinck, K., and Szafarz, A. (2015). Virtual currency, tangible return: Portfolio diversification with Bitcoin. *Journal of Asset Management*, 16(6):365–373.

Broekel, T., Balland, P.-A., Burger, M., and van Oort, F. (2014). Modeling knowledge networks in economic geography: A discussion of four methods. *The Annals of Regional Science*, 53(2):423–452.

Buchholz, M., Delaney, J., Warren, J., and Parker, J. (2012). Bits and bets, information, price volatility, and demand for bitcoin. *Economics*, 312:2–48.

Buybitcoinworldwide.com (2018). Origin of mining pools. https://www.buybitcoinworldwide.com/mining/pools/. [Online; accessed 2018-01-14].

Carvalho, V. M. (2014). From micro to macro via production networks. *The Journal of Economic Perspectives*, 28(4):23–47.

Carvalho, V. M., Nirei, M., Saito, Y. U., and Tahbaz-Salehi, A. (2016). Supply Chain Disruptions: Evidence from the Great East Japan Earthquake. *Columbia Business School Research Papers*, 17(5).

Carvalho, V. M. and Tahbaz-Salehi, A. (2019). Production networks: A primer. *Annual Review of Economics*, 11:635–663.

Carvalho, V. M. and Voigtländer, N. (2014). Input Diffusion and the Evolution of Production Networks. *NBER Working Paper Series*, 1(20025):54.

Castaldi, C. and Dosi, G. (2009). The patterns of output growth of firms and countries: Scale invariances and scale specificities. *Empirical Economics*, 37(3):475–495.

Cerina, F., Zhu, Z., Chessa, A., and Riccaboni, M. (2015). World Input-Output Network. *PLoS ONE*, 10(7).

Chandrasekhar, A. G. and Jackson, M. O. (2014). Tractable and consistent random graph models. Technical report, National Bureau of Economic Research.

Chaney, T. (2014). The Network Structure of International Trade. *American Economic Review*, 104(11):3600–3634.

Chaney, T. (2016). Networks in International Trade. *The Oxford Handbook of the Economics of Networks*, pages 753–775.

Chu, J., Nadarajah, S., and Chan, S. (2015). Statistical analysis of the exchange rate of Bitcoin. *PLOS ONE*.

Clauset, A., Shalizi, C. R., and Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review*, 51(4):661–703.

CNBC (2017). Interview with Vitor Constancio. https://www.cnbc.com/2017/11/29/investors-are-taking-risk-by-buying-expensive-cryptocurrencies.html. [Online; accessed 2018-01-14].

CNBC (2018). Bitcoin savaged by BIS as 'a bubble, ponzi scheme and an environmental disaster'. https://www.cnbc.com/2018/02/06/bitcoin-and-cryptocurrencies-are-ponzi-bubbles-says-bis.html. [Online; accessed 2018-12-13].

CoinDesk (2016). How bitcoin mining works. http://www.coindesk.com/information/how-bitcoin-mining-works/. [Online; accessed 2016-12-27].

CoinMarketCap (2017). Crypto-currency market capitalizations. http://coinmarketcap.com/currencies/. [Online; accessed 2017-01-31].

CoinMarketCap (2019). Crypto-currency market capitalizations. `http://coinmarketcap.com/currencies/`. [Online; accessed 2019-01-22].

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1:223–236.

Cox, D. R. and Hinkley, D. V. (1979). *Theoretical statistics*. CRC Press.

Crum, W. L. (1923). The use of the median in determining seasonal variation. *Journal of the American Statistical Association*, 18(141):607–614.

De Grauwe, P. (2020). *Economics of the Monetary Union*. Oxford University Press, USA, 13th edition.

De Haan, J., Inklaar, R., and Jong-A-Pin, R. (2008). Will business cycles in the Euro area converge? A critical survey of empirical research. *Journal of Economic Surveys*, 22(2):234–273.

De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance*, 45(2):379–395.

Di Giovanni, J. and Levchenko, A. A. (2010). Putting the parts together: Trade, vertical linkages, and business cycle comovement. *American Economic Journal: Macroeconomics*, 2(2):95–124.

Di Giovanni, J., Levchenko, A. A., and Méjean, I. (2014). Firms, destinations, and aggregate fluctuations. *Econometrica*, 82(4):1303–1340.

Di Giovanni, J., Levchenko, A. A., and Mejean, I. (2018). The micro origins of international business-cycle comovement. *American Economic Review*, 108(1):82–108.

Dietzenbacher, E. and Los, B. (1998). Structural Decomposition Techniques : Sense and Sensitivity. *Economic Systems Research*, 5314.

Disdier, A.-C. and Head, K. (2008). The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and statistics*, 90(1):37–48.

Domar, E. D. (1961). On the measurement of technological change. *The Economic Journal*, 71(284):709–729.

Dosi, G., Grazzi, M., Tomasi, C., and Zeli, A. (2012). Turbulence uderneath the big calm? The micro-evidence behind Italian productivity dynamics. *Small Business Economics*, 39(4):1043–1067.

Dosi, G., Napoletano, M., Roventini, A., and Treibich, T. (2019). Debunking the granular origins of aggregate fluctuations: From real business cycles back to Keynes. *Journal of Evolutionary Economics*, 29(1):67–90.

Drozd, L. A., Kolbin, S., and Nosal, J. B. (2017). Long-run trade elasticity and the trade-comovement puzzle. *FRB of Philadelphia Working Paper*, 17-42.

Dubois, A. and Fredriksson, P. (2008). Cooperating and competing in supply networks: Making sense of a triadic sourcing strategy. *Journal of Purchasing and Supply Management*, 14(3):170–179.

Dunn, P. F. (2014). *Measurement and data analysis for engineering and science*. CRC press.

Dür, A., Baccini, L., and Elsig, M. (2014). The design of international trade agreements: Introducing a new dataset. *The Review of International Organizations*, 9(3):353–375.

Ebeke, M. C. H. and Eklou, K. M. (2017). The granular origins of macroeconomic fluctuations in Europe. *IMF Working Paper Series*, 17(229).

Eicher, T. S. and Henn, C. (2011). In search of WTO trade effects: Preferential trade agreements promote trade strongly, but unevenly. *Journal of International Economics*, 83(2):137–153.

European Central Bank (2016). Opinion of the European Central Bank of 12 october 2016 on a proposal for a directive of the European Parliament and of the Council amending Directive (EU) 2015/849. https://www.ecb.europa.eu/ecb/legal/pdf/en_con_2016_49_f_sign.pdf. [Online; accessed 2016-12-24].

European Central Bank (2017). Interview with Benoît Cœuré. https://www.ecb.europa.eu/press/inter/date/2017/html/ecb.in171121.en.html. [Online; accessed 2018-01-14].

Fagiolo, G., Napoletano, M., and Roventini, A. (2007). How do output growth-rate distributions look like? Some cross-country, time-series evidence. *The European Physical Journal B*, 57(2):205–211.

Fang, V., Hovehlmann, L., and Lux, T. (2019). Peer effects in professional analysts' choice of their portfolio of companies. *CEU Working Paper Series*.

Feiwel, G. R. (1975). *The Intellectual Capital of Michal Kalecki: A Study in Economic Theory and Policy*. University of Tennessee Press.

Financial Times (2018). ECB official dubs bitcoin 'evil spawn of the financial crisis'. https://www.ft.com/content/92c4737e-e8ed-11e8-885c-e64da4c0f981. [Online; accessed 2018-12-13].

Foerster, A. T., Sarte, P.-D. G., and Watson, M. W. (2011). Sectoral versus aggregate shocks: A structural factor analysis of industrial production. *Journal of Political Economy*, 119(1):1–38.

Frankel, J. A. and Romer, D. (1999). Does trade cause growth? *American Economic Review*, 89(3):379–399.

Frankel, J. A. and Rose, A. K. (1998). The endogenity of the optimum currency area criteria. *The Economic Journal*, 108(449):1009–1025.

Freund, C. (2003). *Reciprocity in free trade agreements*. The World Bank.

Fujimoto, K., Snijders, T. A., and Valente, T. W. (2018). Multivariate dynamics of one-mode and two-mode networks: Explaining similarity in sports participation among friends. *Network Science*, 6(3):370–395.

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

Gabaix, X. (2016). Power laws in economics: An introduction. *Journal of Economic Perspectives*, 30(1):185–206.

Galí, J. (2008). Monetary policy, inflation, and the business cycle. *Princeton University Press*.

Gandal, N. and Halaburda, H. (2016). Can We Predict the Winner in a Market with Network Effects? Competition in Cryptocurrency Market. *Games*, 7(3):16.

Garcia, D., Tessone, C. J., Mavrodiev, P., and Perony, N. (2014). The digital traces of bubbles: Feedback cycles between socio-economic signals in the Bitcoin economy. *Journal of the Royal Society Interface*, 11(99):20140623.

García-Algarra, J., Mouronte-López, M. L., and Galeano, J. (2019). A stochastic generative model of the world trade network. *Scientific Reports*, 9(1):1–10.

Garin, J., Pries, M. J., and Sims, E. R. (2018). The relative importance of aggregate and sectoral shocks and the changing nature of economic fluctuations. *American Economic Journal: Macroeconomics*, 10(1):119–48.

Giammetti, R. (2020). Tariffs, domestic import substitution and trade diversion in input-output production networks: An exercise on Brexit. *Economic Systems Research*, 3:318–350.

Giammetti, R., Russo, A., and Gallegati, M. (2020). Key sectors in input–output production networks: An application to Brexit. *The World Economy*, 43(4):840–870.

Gorman, W. M. (1961). On a class of preference fields. *Metroeconomica*, 13(2):53–56.

Gräbner, C., Heimberger, P., Kapeller, J., and Schütz, B. (2020). Is the Eurozone disintegrating? Macroeconomic divergence, structural polarisation, trade and fragility. *Cambridge Journal of Economics*, 44(3):647–669.

Groth, C. (2015). *Lecture Notes in Macroeconomics*. MIMEO, Copenhagen.

Gualdi, S. and Mandel, A. (2016). On the emergence of scale-free production networks. *Journal of Economic Dynamics and Control*, 73:61–77.

Guillaume, D. M., Dacorogna, M. M., Davé, R. R., Müller, U. A., Olsen, R. B., and Pictet, O. V. (1997). From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics*, 1(2):95–129.

Hallak, J. C. (2006). Product quality and the direction of trade. *Journal of International Economics*, 68(1):238–265.

Hausmann, R., Hidalgo, C., and Bustos, S. (2014). The atlas of economic complexity: Mapping paths to prosperity. *Center of International Development at Harvard University*.

Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, 106(26):10570–10575.

Hohnisch, M. and Westerhoff, F. (2008). Business cycle synchronization in a simple Keynesian macro-model with socially transmitted economic sentiment and international sentiment spill-over. *Structural Change and Economic Dynamics*, 19(3):249–259.

Hulten, C. R. (1978). Growth accounting with intermediate inputs. *The Review of Economic Studies*, 45(3):511–518.

Hummels, D., Ishii, J., and Yi, K. M. (1999). The nature and growth of vertical specialization in world trade. *Journal of International Economics*, 54(1):75–96.

Ikeda, K. (2018). Security and privacy of blockchain and quantum computation. *Advances in Computers*, 111:199–228.

Imry, Y. (2008). *Introduction to Mesoscopic Physics*. Oxford University Press.

Jackson, M. O. and Rogers, B. W. (2007). Meeting strangers and friends of friends: How random are social networks? *American Economic Review*, 97(3):890–915.

Johnson, N. F., Jefferies, P., Hui, P. M., et al. (2003). Financial market complexity. *OUP Catalogue*.

Jones, C. I. (2011). Intermediate goods and weak links in the theory of economic development. *American Economic Journal: Macroeconomics*, 3(2):1–28.

Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika*, 18(1):39–43.

Keynes, J. (1911). The principal averages and the laws of error which lead to them. *Journal of the Royal Statistical Society*, 74(3):322–331.

Kikkawa, K., Magerman, G., and Dhyne, E. (2019). Imperfect Competition in Firm-to-Firm Trade. *SSRN Electronic Journal*.

Kirman, A. P. (1992). Whom or what does the representative individual represent? *Journal of Economic Perspectives*, 6(2):117–136.

Kiyotaki, N. and Wright, R. (1989). On money as a medium of exchange. *Journal of Political Economy*, 97(4):927–954.

Kleinert, J., Martin, J., and Toubal, F. (2015). The few leading the many: Foreign affiliates and business cycle comovement. *American Economic Journal: Macroeconomics*, 7(4):134–59.

Kondor, D., Pósfai, M., Csabai, I., and Vattay, G. (2014). Do the rich get richer? An empirical analysis of the Bitcoin transaction network. *PLOS ONE*, 9(2):1–10.

Kose, M. A. and Yi, K.-M. (2006). Can the standard international business cycle model explain the relation between trade and comovement? *Journal of International Economics*, 68(2):267–295.

Koskinen, J. and Edling, C. (2012). Modelling the evolution of a bipartite network—peer referral in interlocking directorates. *Social Networks*, 34(3):309–322.

Kotz, S., Kozubowski, T., and Podgorski, K. (2001). *The Laplace distribution and generalizations: A revisit with applications to communications, economics, engineering, and finance*. Springer Science & Business Media.

Kozubowski, T. J. and Podgórski, K. (2001). Asymmetric Laplace laws and modeling financial data. *Mathematical and Computer Modelling*, 34(9):1003–1021.

Kramarz, F., Martin, J., and Mejean, I. (2019). Volatility in the small and in the large: The lack of diversification in international trade. *Journal of International Economics*, 122:103276.

Kristoufek, L. (2015). What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. *PLOS ONE*, 10(4):e0123923.

Krugman, P. (1981). Intraindustry Specialization and the Gains from Trade. *Rethinking International Trade*, 89(5):959–973.

Kuha, J. (2004). AIC and BIC: Comparisons of assumptions and performance. *Sociological Methods & Research*, 33(2):188–229.

Laplace, P.-S. (1774). Mémoire sur la probabilité des causes par les évènements. *Mémoires présentés par divers savants [à lcadémie royale des sciences], Paris, Imprimerie royale*, 6:621–656.

Larch, M., Wanner, J., and Yotov, Y. V. (2018). Bi-and unilateral trade effects of joining the Euro. *Economics Letters*, 171:230–234.

Leontief, W. (1986). *Input-Output Economics*. Oxford University Press, New York, 2nd edition.

Leontief, W. W. (1936). Quantitative input and output relations in the economic systems of the United States. *The Review of Economic Statistics*, pages 105–125.

Lin, I.-C. and Liao, T.-C. (2017). A survey of blockchain security issues and challenges. *IJ Network Security*, 19(5):653–659.

Liu, E. (2019). Industrial Policies in Production Networks. *The Quarterly Journal of Economics*, 134(3):1883–1948.

Livan, G., Alfarano, S., Milaković, M., and Scalas, E. (2015). A spectral perspective on excess volatility. *Applied Economics Letters*, 22(9):745–750.

Lucas, R. E. (1977). Understanding business cycles. In *Carnegie-Rochester Conference Series on Public Policy*, volume 5, pages 7–29. Elsevier.

Lux, T. and Ausloos, M. (2002). *Market fluctuations I: Scaling, multiscaling, and their possible origins*. Springer.

Magerman, G., De Bruyne, K., Dhyne, E., and Van Hove, J. (2016). Heterogeneous firms and the micro origins of aggregate fluctuations. *National Bank of Belgium Working Paper Series*, 1(312).

Marshall, A. (1920). *Industry and trade*. MacMillan London, 3rd edition.

Mattli, W. et al. (1999). *The logic of regional integration: Europe and beyond*. Cambridge University Press.

McGee, R. W. and Yoon, Y. (2003). Labor standards and trade agreements: Raising standards or raising the walls of protectionism? *The International Trade and Finance Association, Forthcoming*.

McGill, W. J. (1962). Random fluctuations of response rate. *Psychometrika*, 27(1):3–17.

McQueen, M. (2002). The EU's free trade agreements with developing countries: A case of wishful thinking? *World Economy*, 25(9):1369–1385.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Mishkin, F. (2004). *The Economics of Money, Banking, and Financial Markets*. The Addison-Wesley series in economics. Pearson.

Moore, T. and Christin, N. (2013). Beware the middleman: Empirical analysis of Bitcoin-exchange risk. In *Proceedings of Financial Cryptography 2013*.

Nakamoto, S. (2009). Bitcoin: A peer-to-peer electronic cash system.

Oberfield, E. (2018). A theory of input–output architecture. *Econometrica*, 86(2):559–589.

OECD (2005). Oslo manual. `http://dx.doi.org/10.1787/9789264013100-en`. [Online; accessed 2016-12-25].

OECD (2018a). OECD Inter-Country Input-Output Tables. `https://www.oecd.org/sti/ind/inter-country-input-output-tables.htm`. [Online; accessed 2020-06-18].

OECD (2018b). STAN Industrial Analysis. `https://stats.oecd.org/Index.aspx?DataSetCode=STANI4_2016#`. [Online; accessed 2020-06-18].

OECD (2018c). Trade in Employment - TiM indicators. `https://stats.oecd.org/Index.aspx?DataSetCode=TIM_2019_MAIN`. [Online; accessed 2020-06-18].

Osterrieder, J. (2016). The statistics of Bitcoin and cryptocurrencies. *Advanced Risk and Portfolio Management Paper*.

Osterrieder, J. and Lorenz, J. (2016). A statistical risk assessment of Bitcoin and its extreme tail behaviour. *Big Data and Innovative Financial Technologies Research Paper Series*, pages 1–13.

Ozdagli, A. and Weber, M. (2017). Monetary policy through production networks: Evidence from the stock market. Technical report, National Bureau of Economic Research.

Prebisch, R. (1950). The economic development of Latin America and its principal problems. *Economic Bulletin for Latin America*.

Puty, C. A. C. B. (2018). Sectoral mark-ups in US manufacturing. *Structural Change and Economic Dynamics*, 46:107–125.

Quesnay, F. (1758). *Tableau économique des physiocrates*, volume 1. Calmann-Lévy.

Rhodes, C. (1989). Reciprocity in trade: The utility of a bargaining strategy. *International Organization*, pages 273–299.

Ripley, R. M., Snijders, T. A., Boda, Z., Vörös, A., and Preciado, P. (2011). Manual for RSIENA. *University of Oxford, Department of Statistics, Nuffield College*, 1:2011.

Ruppert, D. (2004). *Statistics and Finance - An Introduction*. Springer-Verlag New York, New York.

Saint-Paul, G. (1993). Productivity growth and the structure of the business cycle. *European Economic Review*, 37(4):861–883.

Sapir, A. (1992). Regional integration in Europe. *The Economic Journal*, 102(415):1491–1506.

Satchell, S. and Knight, J. (2000). *Return Distributions in Finance*. Quantitative Finance. Elsevier Science.

Schumpeter, J. A. (1934). *The theory of economic development*. Harvard University Press.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

Shea, J. (2002). Complementarities and comovements. *Journal of Money, Credit, and Banking*, 34(2):412–433.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48(1):1–48.

Singer, H. W. (1950). The distribution of gains between borrowing and investing countries. *American Economic Review*, 40(2):473–485.

Snijders, T. A. (1996). Stochastic actor-oriented models for network change. *Journal of Mathematical Sociology*, 21(1-2):149–172.

Snijders, T. A. (2001). The statistical evaluation of social network dynamics. *Sociological Methodology*, 31(1):361–395.

Snijders, T. A., Lomi, A., and Torló, V. J. (2013). A model for the multiplex dynamics of two-mode and one-mode networks, with an application to employment preference, friendship, and advice. *Social Networks*, 35(2):265–276.

Snijders, T. A., Steglich, C. E., and Schweinberger, M. (2007). Modeling the co-evolution of networks and behavior. *Longitudinal Models in the Behavioral and Related Sciences*, 31(4):41–71.

Snijders, T. A., van de Bunt, G. G., and Steglich, C. E. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, 32(1):44–60.

Snowdon, B., Vane, H. R., and Wynarczyk, P. (1994). *A modern guide to macroeconomics: An introduction to competing schools of thought*. Edgar Elgar Publishing.

Sornette, D. (2003). Why stock markets crash.

StackExchange (2016). Ether historical prices. http://ethereum.stackexchange.com/questions/635/ether-historical-prices/638#comment11082_638. [Online; accessed 2016-12-20].

Stanley, H. and Mantegna, R. (2000). *An introduction to econophysics*. Cambridge University Press, Cambridge.

Stella, A. (2015). Firm dynamics and the origins of aggregate fluctuations. *Journal of Economic Dynamics and Control*, 55:71–88.

Storper, M. (2004). Territories, flows, and hierarchies in the global economy. *Reading Economic Geography*, page 271.

Subbotin, M. T. (1923). On the law of frequency of error. *Matematicheskii Sbornik*, 31(2):296–301.

Summers, L. H. (1986). Some skeptical observations on real business cycle theory. *Federal Reserve Bank of Minneapolis Quarterly Review*, Fall:23–27.

Syverson, C. (2004). Market structure and productivity: A concrete example. *Journal of Political Economy*, 112(6):1181–1222.

Syverson, C. (2019). Macroeconomics and market power: Context, implications, and open questions. *Journal of Economic Perspectives*, 33(3):23–43.

Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R., and Vries, G. J. (2015). An illustrated user guide to the world input-output database: The case of global automotive production. *Review of International Economics*, 23(3):575–605.

Tintelnot, F., Kikkawa, A. K., Mogstad, M., and Dhyne, E. (2018). Trade and Domestic Production Networks. *NBER Working Paper Series*, 18(25120):98.

Tweedle, J. (2018). Correlated shocks within firms. *Economics Letters*, 163:95–97.

UN (2006). Methodology: Standard country or area codes for statistical use. `https://unstats.un.org/unsd/methodology/m49/`. [Online; accessed 2020-06-18].

Vamvakidis, A. (1998). Regional integration and economic growth. *The World Bank Economic Review*, 12(2):251–270.

Varoufakis, Y. (2013). Bitcoin and the dangerous fantasy of 'apolitical' money. `https://yanisvaroufakis.eu/2013/04/22/bitcoin-and-the-dangerous-fantasy-of-apolitical-money/`. [Online; accessed 2017-01-31].

Wasserman, S., Faust, K., et al. (1994). *Social network analysis: Methods and applications*, volume 8. Cambridge University Press.

Westerhoff, F. et al. (2009). Exchange rate dynamics: A nonlinear survey. *Handbook of Research on Complexity*, pages 287–325.

Wolff, E. N. (1985). Industrial Composition, Interindustry Effects, and the U.S. Productivity Slowdown. *The Review of Economics and Statistics*, 67(2):268–277.

World Bank (2018). The World Bank Data. `https://data.worldbank.org/indicator/NY.GDP.PCAP.KD.ZG?locations=SE-FI-EU`. [Online; accessed 2018-07-04].

World Input-Output Database (2017). WIOD home. `http://www.wiod.org/home`. [Online; accessed 2017-04-13].

Yaffee, R. A. and McGee, M. (2000). *An introduction to time series analysis and forecasting: with applications of SAS® and SPSS®*. Elsevier.

Yermack, D. (2015). Is Bitcoin a real currency? An economic appraisal. In *Handbook of Digital Currency*, pages 31–43. Elsevier.

Yu, X., Dosi, G., Lei, J., and Nuvolari, A. (2015). Institutional change and productivity growth in China's manufacturing: The microeconomics of knowledge accumulation and "creative restructuring". *Industrial and Corporate Change*, 24(3):565–602.

Zou, Y. (2019). Endogenous Production Networks and Gains from Trade. *Universitat Pompeu Fabra Working Papers*, pages 1–62.