Production networks and industrial growth: interconnectedness as the cornerstone of the economy

by

Eszter Molnár

School of Engineering, Faculty of Science and Technology

A thesis submitted for the degree of

Doctor of Philosophy

December, 2023

Declaration

I declare that the work presented in this thesis is, to the best of my knowledge and belief, original and my own work. The material has not been submitted, either in whole or in part, for a degree at this, or any other university.

Eszter Molnár

Production networks and industrial growth:

interconnectedness as the cornerstone of the economy

Eszter Molnár, BSc, MSc.

Abstract

This thesis delves into the application of network science in economic sciences, focusing on inter-industry production networks, offering a new perspective on how economic interdependencies can be analysed. The first empirical chapter starts with the study's theoretical background, discusses the application of network science to understand complex questions within production networks, and introduces the threshold problem. After constructing the 2007 and 2012 US national production network using data from the Bureau of Economic Analysis Input-Output Accounts, it looks at the sensitivity of the interactions in the network topology and central industries to different threshold values, highlighting the importance of these thresholds in shaping the network. Secondly, the thesis examines the relationship between national production network characteristics and industrial growth. It introduces a network-based growth model based on the national production network and provides results that reveal that these characteristics can account for up to 29% of growth variance. This outcome highlights the potential of national production network metrics as explanatory variables in understanding industrial growth. After, the focus shifts to industry-specific analysis, specifically the storage battery industry and its supply chain. It introduces the concept of the ego network, discusses the challenges and relationships within this industry and explores the storage battery ego network. It analyses separately the simple upstream and downstream ego network and also the advanced ego network at the industry and network level. It identifies the indispensable industries (referring to them as integrator, allocator, and mediator industries) within the ego network, shedding light on supply chain dynamics and their role in ensuring resilience. The last part continues with the industry-specific perspective and examines the role of the storage battery ego network in industrial growth. It finds that 51% of an industry's success is linked to its position within the ego network. In summary, this thesis advances the understanding of the intricate relationship inside production networks and emphasises the critical role of threshold values, the impact of network characteristics on growth, and the significance of industry-specific ego production networks.

Acknowledgements

This thesis would not have been possible without the support of several individuals who, in one way or another, contributed their valuable help in completing this study.

I am deeply grateful to Dr. Dénes Csala, my advisor, whose guidance not only steered me through this study but also profoundly influenced my academic journey and worldview. His mentorship provided invaluable lessons that have shaped me as a researcher and as a person, too. Köszönöm, hogy megmutattad ezt az egyedülálló hozzáállást az élethez.

Balázsnak, aki a több ezer kilométer ellenére a világ minden pontjáról támogatott és hazavárt. Nincsenek szavak, amelyekkel kifejezhetem a hálámat a végtelen türelemért, amellyel végighallgattad az összes mélypontot és bizonytalanságot. Te voltál az inspiráció és a motiváció ennek az útnak minden egyes lépése mögött.

Anyának, apának és Borókának az otthont, ahol nyugodtan lehetett morcosnak lenni és az összes bátorító szót, eszmecserét, amellyel kihúztak a gödrökből.

I am incredibly thankful for the advice and openness of my examiners, Dr. Silvia Fierăscu and Dr. Robert Read.

I would also like to thank Dr. Richard Harper, who gave me constructive feedback on critical thinking and the importance of communicating research in an accessible way. I would like to acknowledge the financial and professional support from the Material Social Futures Doctoral Training Centre in Lancaster.

Thanks to my cohort, Jonathan, George, Dom, and William, for asking challenging questions and to other members of the training programme, especially Petter and Christina. Thank you, Anita, for the discussions and gracias, Lucía, por los largos paseos.

Contents

1 Introduction **1**

1.1 Background and summary 1

1.2 Research objective and questions 6

1.3 Methodological overview 9

1.4 Thesis structure 13

2 From data to map: the production network and the threshold sensitivity of the topology **17**

2.1 Network science as a tool to map complex questions 17

2.2 The production network and the threshold-dependent behaviour 21

2.3 Strategy for analysing the threshold dependence of the topology 26

2.3.1 Data source 26

2.3.2 Production network generation from data 27

2.3.3 Monetary transaction threshold 28

2.3.4 Algorithms for random and scale-free models 29

2.3.5 Measure for degree distribution comparison (Kolmogorov-Smirnov test) 30

2.3.6 Production network centrality measures 31

2.4 Findings for threshold sensitivity of production network topology 32

2.5 Discussion 37

3 **Explaining growth: the production network-based prediction of industrial growth** **41**
3.1 Industrial growth in the context of the production network 41
3.2 Empirical strategy to analyse industrial growth using networks 44
3.2.1 Data source 44
3.2.2 Constructing the manufacturing part of the production network . . 45
3.2.3 Monetary transaction threshold 46
3.2.4 Description and analysis of network metrics 47
3.2.5 Measuring industrial growth 52
3.2.6 The network-based growth model 53
3.3 Results 57
3.3.1 The best-performing network growth model 57
3.3.2 The production network 60
3.3.3 The production network metrics explaining growth 62
3.4 Discussion 65

4 **Ego production network perspective: the case study of the storage battery industry** **67**
4.1 The ego production network 67
4.2 Energy storage, the storage battery industry and its supply chain 73
4.3 Methodology 81
4.3.1 Data source and the presence of the storage battery industry in it . 81
4.3.2 Constructing the ego production network 83
4.4 Simple ego network of storage battery industry 86
4.4.1 Upstream ego network 86
4.4.2 Downstream ego network 88
4.4.3 Discussion and comparison of upstream and downstream network . 93
4.5 Advanced ego network of storage battery industry 94

4.5.1 Node-level measures in the advanced ego network 95
4.5.2 Network-level measures in the advanced ego network 115
4.6 Conclusions from the ego network analysis 130

5 Ego network growth model: the case study of the storage battery industry 134

5.1 Introduction 134
5.2 Strategy to analyse industry growth in the ego production network 136
5.3 Results of the storage battery ego network growth model 138
5.3.1 Factors influencing growth 142
5.3.2 Storage battery manufacturing industry in the national production network growth model 151
5.4 Discussion 154

6 Conclusions 156

6.1 Overview of principal findings 156
6.2 Implications 159
6.3 Postface and the Material Social Futures Doctoral Training Centre 163

Appendix A Explaining growth: the production network-based prediction of industrial growth 165

A.1 Variable distribution (production network) 165
A.2 Correlation matrix (production network) 166

Appendix B Ego network growth model: the case study of the storage battery industry 167

B.1 Correlation matrix (storage battery ego network) 168
B.2 Variable distributions (storage battery ego network) 169

References 170

List of Figures

1.1 Methodological structure for Chapters 2 and 3 10

1.2 Methodological structure for Chapters 4 and 5 12

2.1 Example of a two-industry production network 28

2.2 Global edge weight distribution (log-lin scale) 33

2.3 Kolmogorov-Smirnov statistic value of the distributions compared (lower value means more similar) 34

2.4 Node degree distribution of the production network 36

2.5 Top industries according to threshold ζ 37

3.1 Model accuracy scores based on threshold and rate type. 57

3.2 Weight and modularity in the production network. 58

4.1 Storage battery supply chain 78

4.2 Storage battery industry upstream and downstream simple production network (2007) 91

4.3 Storage battery industry upstream and downstream simple production network (2012) 92

4.4 Changes in the integrator (by number) industries in the storage battery ego network between 2007 and 2012, by number of industry contacts (in-degree) 102

4.5 Changes in the integrator (by amount) industries in the storage battery ego network between 2007 and 2012, by resource volume (in-degree weighted) . 105

4.6 Changes in the allocator (by number) industries in the storage battery ego network between 2007 and 2012, by number of industry contacts (out-degree) 107
4.7 Changes in the allocator (by amount) industries in the storage battery ego network between 2007 and 2012, by resource volume (out-degree weighted) 110
4.8 Changes in the mediator (by number) industries in the storage battery ego network between 2007 and 2012, by number of industry contacts (betweenness) . . . 112
4.9 Changes in the mediator (by amount) industries in the storage battery ego network between 2007 and 2012, by resource volume (betweenness weighted) 114
4.10 Sparse and dense networks . . . 116
4.11 Storage battery industry advanced ego network (2007) . . . 124
4.12 Storage battery industry advanced ego network (2012) . . . 125
4.13 Storage battery industry advanced ego network (2007 & 2012) — Medium environment (0.05 dollars threshold) . . . 128
4.14 Storage battery industry advanced ego network (2007 & 2012) — Micro environment (0.1 dollars threshold) . . . 129

5.1 Storage battery industry advanced manufacturing-industry ego network (2007 & 2012) . . . 140

A.1 Distribution of every measure calculated in the production network . . . 165

B.1 Distribution of every measure calculated in the storage battery ego network (2007) . . . 169
B.2 Distribution of every measure calculated in the storage battery ego network (2012) . . . 169

List of Tables

3.1 The remaining number and percentage of the original edge number in the production network . . . 46
3.2 Variable definitons . . . 56
3.3 Sample of the most connected industries based on all analysed network measures (2012) . . . 61
3.4 Descriptive statistic . . . 62
3.5 Regression for industrial growth using Eq. 3.5 . . . 63

4.1 NAICS index entries for 335911 - Storage battery manufacturing industry . 82
4.2 NAICS major industry classifications . . . 83
4.3 Upstream and downstream ego network general statistics . . . 93
4.4 Node-level centrality metrics and their implications for ego production network . . . 99
4.5 Colours of the industry classifications on the advanced ego network and node-level figures . . . 100
4.6 Possible directed relationships for three nodes without self-loops . . . 115
4.7 Storage battery advanced ego network measures descriptive statistics (2007) 117
4.8 Storage battery advanced ego network measures descriptive statistics (2012) 119
4.9 Storage battery industry advanced ego network complexity change between 2007 and 2012 . . . 121
4.10 Number and ratio of nodes and edges in the medium-level and micro-level advanced storage battery ego network (2007 & 2012) . . . 126

5.1 Number and ratio of nodes and edges in the storage battery ego network compared to the national network next to the 0.005 dollars threshold value (2007 & 2012) . . . 137
5.2 Storage battery manufacturing ego network for the growth model (2007 & 2012) . . . 139
5.3 Ego network growth model accuracy scores based on growth rate type (0.0005 dollars threshold) . . . 141
5.4 Ego network growth model accuracy scores compared to national production network growth model scores . . . 142
5.5 Storage battery industry (manufacturing) ego network growth model descriptive statistics . . . 143
5.6 Sample of the most connected industries in the storage battery ego network based on all analysed network measures (2007) . . . 145
5.7 Sample of the most connected industries in the storage battery ego network based on all analysed network measures (2012) . . . 146
5.8 Regression for industrial growth results using the storage battery ego network framework (Eq. 3.5) . . . 148
5.9 Storage battery industry growth rates in the national production network growth model (2007 & 2012) . . . 151
5.10 Storage battery industry centrality metrics in the national production network growth model (2007 & 2012) . . . 152

A.1 Correlation matrix . . . 166

B.1 Correlation matrix . . . 168

Chapter 1

Introduction

1.1 Background and summary

A systematic view of complex problems and questions remains incredibly relevant in our fast-paced world. In an age where everything is interconnected, and information is constantly pouring in, having a systematic approach is crucial for understanding intricate issues. This approach encourages us to break big problems into smaller, more manageable parts, which makes it easier to analyse and find underlying patterns. It also reminds us to think about different angles and potential outcomes when looking for solutions. Whether in science, business, politics, or any other field, the systematic view offers a valuable way to tackle complex challenges and make well-informed decisions that can shape a better future.

The application of complex system tools in economic sciences has brought about a profound transformation in how to analyse and model economic processes (Foster, 2005). Traditional economic models often simplify real-world complexities, assuming rational behaviour and/or linear relationships. However, complex system tools enable economists to capture the intricate, linear, nonlinear, and interconnected nature of actual economies. Techniques like agent-based modelling, network science, and chaos theory allow researchers to explore emergent phenomena, feedback mechanisms, and the influence of diverse agents in economic decision-making. Embracing these tools empowers economists to develop more realistic and nuanced models, providing fresh perspectives on economic crises, market dynamics, and policy interventions. This shift towards complex systems thinking represents a significant development in economic sciences, promising deeper insights into the workings of economies.

Network science offers a compelling solution to the complexities inherent in economic sciences (Barabási, 2013; Kenett and Havlin, 2015). This view argues that in our modern world, where intricate relationships and information flow are the norm, it makes sense to view economic systems as networks. Networks are a collection of nodes and their relationships to each other stored in a mathematical format. In this perspective, nodes represent entities like individuals, businesses, or financial institutions, and the links between them signify various interactions and connections. Network science provides a robust framework for unravelling the intricate web of these economic connections. It helps to understand how financial shocks spread, information disseminates, and systemic risks arise. By using tools like network topology analysis and centrality measures, literature can pinpoint critical nodes and vulnerabilities within economic networks, providing valuable insights for both policymakers and economists. Harnessing the analytical power of network science enables us to navigate the intricacies of economic systems more effectively, leading to more profound comprehension and more informed strategies for addressing economic challenges.

The importance of national production networks in this context cannot be emphasised enough. As highlighted, industries rely on each other's products and services, forming a complex web of interdependence. To understand this intricate network and its impact on our economy, researchers use tools from network science and data analysis. They create what's called a "production network," which is like a map showing how different industries are connected (Vasco M. Carvalho and Tahbaz-Salehi, 2019). This network helps in the exploration of questions about how industries relate to each other and how changes in one sector can affect the entire economy. Imagine industries as groups of businesses that make similar things. For example, the tech industry includes companies that produce computers, smartphones, and other electronic devices. Now, think of all these industries as pieces of a puzzle that fit together to create our economy. Economic researchers who take the network science point of view like to study this puzzle to see how it works; using math and data, they try to understand how money flows between these industries. It's like figuring out which puzzle pieces are connected and how strong those connections are.

Therefore, a production network is a concept that blends principles from graph and network theory with the economic idea of input-output relationships. This fusion creates a visual representation of industrial connections, allowing us to extract specific metrics that describe the network's structure. In the interrelated global landscape, understanding how a country's industries and businesses are interconnected within its borders is vital for assessing its overall economic well-being and adaptability. These networks reveal the intricate links between different sectors and help to see how disruptions or policy changes can have far-reaching consequences. By examining national production networks, policymakers can make more informed choices about economic diversification, supply chain resilience, and targeted measures to stimulate growth. Furthermore, during crises like the recent pandemic, the insights gained from these networks prove invaluable in developing effective strategies to protect and revive the domestic economy. In short, national production networks are an essential tool for navigating the complexities of today's economic environment, offering a deeper understanding and more informed approaches to economic policy and planning.

Given these possibilities and these methodological innovations, this thesis focuses on production network research, exactly on a national production network where all inter-industry monetary transactions are represented. The data taken from the United States Bureau of Economic Analysis has enabled the production of the national production network, which can be built in the most detailed version across countries with sectors represented in plenty of subcategories. Collecting this type of data is very time-consuming and expensive, and the US has the most detailed and accessible data source for industry interactions.

In addition to exploring this framework, the thesis also dives into a particular aspect of production network research. These networks are like a big tangle of connections between industries, where almost every industry is linked to others. To make this complexity more manageable, researchers must be selective about which connections they keep and which ones they remove (Radicchi, Ramasco, and Fortunato, 2011). It is like trimming a tree to keep it healthy. The thesis takes a close look at how this trimming process affects the overall shape of the network, like whether it makes it look more like a bush or a tree with fewer branches. This helps in understanding how the network topology changes when we remove some of these connections. This technique, focusing on the most important connections and ignoring the less important ones, is called "thresholding". The thesis examines precisely the network topology distortion as a function of link removal (threshold value).

Next to industry relationships, understanding how industries grow and thrive within an economy is a complex question that economists have been trying to solve (Klenow and Rodríguez-Clare, 1997). Traditionally, research has often looked at individual industries in isolation, overlooking the intricate web of connections that tie them together. However, this approach can be limiting because, as presented above, industries don't exist independently – they depend on and interact with each other. To address this gap, recent studies have started merging traditional economic theories with concepts from complex systems science, aiming to gain a fresh perspective on economic complexity and growth (César A. Hidalgo and Hausmann, 2009).

To repeat, one way to make sense of this complexity is by using network models, which are like maps of how different parts of the economy connect and interact. These networks can help us understand not only international trade relationships but also how industries are linked within a single nation's economy. By studying these production networks, we can uncover valuable insights about how industries' positions within these networks influence their growth and, in turn, affect the nation's economic health. This shift towards a network-based approach offers a promising new perspective to contribute to our understanding of industrial dynamics.

Hence, after this closer look at the United States national production network, the thesis uses this concept as the foundation of a network-based industrial growth model. The growth model is inspired by Kali and Reyes, 2007. They use an international perspective and focus on the economic growth of countries, adding their trade network features as another dimension to already used growth models. However, the focus of the thesis is the industries in a national context. What is less clear in the literature is how the network framework works in a national environment and on its own as a separate model. This dissertation offers an examination of the relationship between the economic growth of industries in association with their role in the national economy's production cycle. The industrial growth model presented in this thesis is built entirely on national production network topological features.

Having examined carefully and thoroughly the national landscape, the thesis moves on to bring this complex system perspective into assessing a particular industry. It analysed the storage battery industry as a case study from the perspective of the production network framework and the network-based growth model. This serves as a case study for the viability of this method when zooming in on a specific sector. Also, it highlights the critical challenges faced in the production network of the indispensable storage battery industry that is a crucial component of energy storage and, thus, of the transition to renewable energy. The framework can be applied to other sectors, too.

The core production network of the storage battery industry is built from the national production network, called the ego network. This is a well-defined subset of the whole national network, where the main organising industry is the storage battery industry. It can be interpreted as the industry production network of storage battery manufacturing.

In the first layer, the interdependencies and the key industries in the storage battery ego network are examined. In the second layer, the thesis uses the ego network for the network-based growth model that will give the model built on a closely related subset of the national production cycle, in other words, on a tighter environment in the sense of transactions.

Taken together, the recent pandemic and even crises before that have created challenges not only for policymakers but also for scholars, particularly in the field of economics. Traditional tools and models were criticised for their inability to assist decision-makers in real time as the situation unfolded. Consequently, more and more researchers have been exploring alternative tools and frameworks to analyse economic problems, seeking potentially improved approaches. The research presented in this thesis aligns with these recent endeavours documented in the literature.

1.2 Research objective and questions

The overall objective of this research is to discover the added value that systemic perspective within that network science can bring to the field of economics, especially to the inter-industry production networks research area.

This thesis aims to explore inter-industry dependencies on two scales: a national one and a specific one, using the storage battery industry segment as a case study. On both scales, there are two primary goals:

Firstly, the purpose is to build, investigate and prove the utility of the production network framework. On the national scale, the network includes all inter-industry monetary transactions in a national economy. While on the one-industry scale, it is the specific core production network of one industry in a broader sense. The storage battery industry is explored at the one-industry scale. This core production network is more than just a simple supply chain but less than a whole national production network. In network science terms, it is a storage battery ego network, a particular well-defined fragment of the national production network. There is a growing consensus that to truly understand the complex dynamics of modern industries, we need to move beyond just looking at geographic clusters and embrace the idea of industrial networks. Studies like those conducted by Vasco M Carvalho, 2008, Reyes, Schiavo, and Fagiolo, 2010, Kali and Reyes, 2010, Acemoglu et al., 2012 supports this perspective. So, the goal of this section is to use network science tools to examine all the complex input-output national trade relationships. After, the intent is to zoom in on the storage battery industry in this context.

Secondly, the idea is to ascertain if this framework is viable as the basis of an economic growth model. As trade connections have become more intricate, the dynamics of industrial growth are no longer limited solely to geographic boundaries. This raises the fundamental question of what factors drive the growth of industries integrated into the national trade network. The question has been asked on an international scale, but so far, not at a national level (Fagiolo, Reyes, and Schiavo, 2010). To address this research gap, the thesis employs a quantitative approach to assess how industrial network characteristics influence growth at the industry level. The objective here is to identify in what proportion the inter-industry relations are responsible for the individual industries' growth. The model takes the roles a sector holds in the national or the storage battery production cycle that also represents its inter-industry linkages and provides an answer to the extent to which this contributes to the industry's growth. The inter-industry relations and the specific roles of an industry in the production network are represented by the production network metrics. These network centrality measures are quantitative metrics used to assess the importance or prominence of industries within the national production network and the storage battery ego network. These measures help identify which nodes are more central or influential based on their connectivity and relationships within the network. In simpler terms, once again, the analysis aims to determine whether an industry's network characteristics can influence its growth. Another layer that comes above this is the comparison of the national and the ego network context.

The thesis is structured into four empirical chapters, each dedicated to distinct research questions:

1. **How does the United States national inter-industry production network look like?**

    - Do different inter-industry relation thresholds highlight different features?
    - Does the threshold change affect the topology, including identifying the central industries?

2. **To what extent can the growth of an industry be determined purely based on its role in the national production cycle that has been defined**

by network characteristics?

- Which network and centrality measures explain the best economic growth?

3. **How can a specific industry and its supply chain be explored and analysed from the network science perspective on the basis of data? Using the storage battery industry as a case study.**

   - How can the national production network framework be customised to explore a specific industry's production chain dependencies?
   - What can we learn from the simple and advanced storage battery ego network?
   - Which are the indispensable industries, in the form of integrator, allocator and mediator industries, in the storage battery ego network?

4. **To what extent can we determine the growth of the industries present in the storage battery supply chain purely from their role in the storage battery production cycle defined by the ego network characteristics?**

   - Which network and centrality measures explain the best economic growth in the context of the storage battery ego network case?
   - How does the network growth model based on the national entity work on a specific industry ego network? The case of the storage battery industry.
   - Has the closely related production environment more explanatory power than the national production cycle? (Does the specific industry ego network have more explanatory power than the national production network? The case of the storage battery industry.)

In the case of the threshold sensitivity analysis, the expectation is that the threshold influences can suitably affect the topology of the production network, even with small threshold values. For the growth models, it is anticipated the network characteristics to affect economic growth and the ego network characteristics even more. The expectation is that the small, well-defined conditions have more explanatory power than the national ones. There is no prior anticipation regarding which network variables will carry greater significance or substantially influence the analysed dependent variables. This uncertainty arises from the limited literature examining the relationship between trade networks and growth.

1.3 Methodological overview

The work builds on the emergent field of economic networks (Schweitzer, Fagiolo, Sornette, Vega-Redondo, Vespignani, et al., 2009; Schweitzer, Fagiolo, Sornette, Vega-Redondo, and White, 2009). This is a progressive multidisciplinary approach that explores the complex net of connections and relationships among economic agents, with the aim of understanding how these networks influence and shape economic phenomena. Researchers in this field employ a variety of methodologies from economics, mathematics, network science, sociology, and other disciplines to uncover fundamental principles governing network formation, evolution, and behaviour, as well as to elucidate the consequences of network structures on economic outcomes.

In this broad area of study, this analysis focuses on economic networks that encompass interactions in the form of financial transactions between different sectors. Thus, the thesis utilises industry-level data to examine the research questions.

The first two empirical chapters explore the national context. Figure 1.1 shows the methodological structure for these two chapters.

The research data to conduct the quantitative analysis is drawn from two main sources. Firstly, for every chapter, the starting point is the national production network. This presents a fascinating challenge when it comes to data, as obtaining a sufficiently extensive dataset that includes industry-level information for numerous disaggregated sectors, potentially on a national scale, is quite challenging. To capture the intricate dynamics of national trade connections, the dataset used includes the Input-Output Accounts for the United States of America, as the US has the most detailed data on separate inter-industry financial transactions. A comprehensive description of the dataset

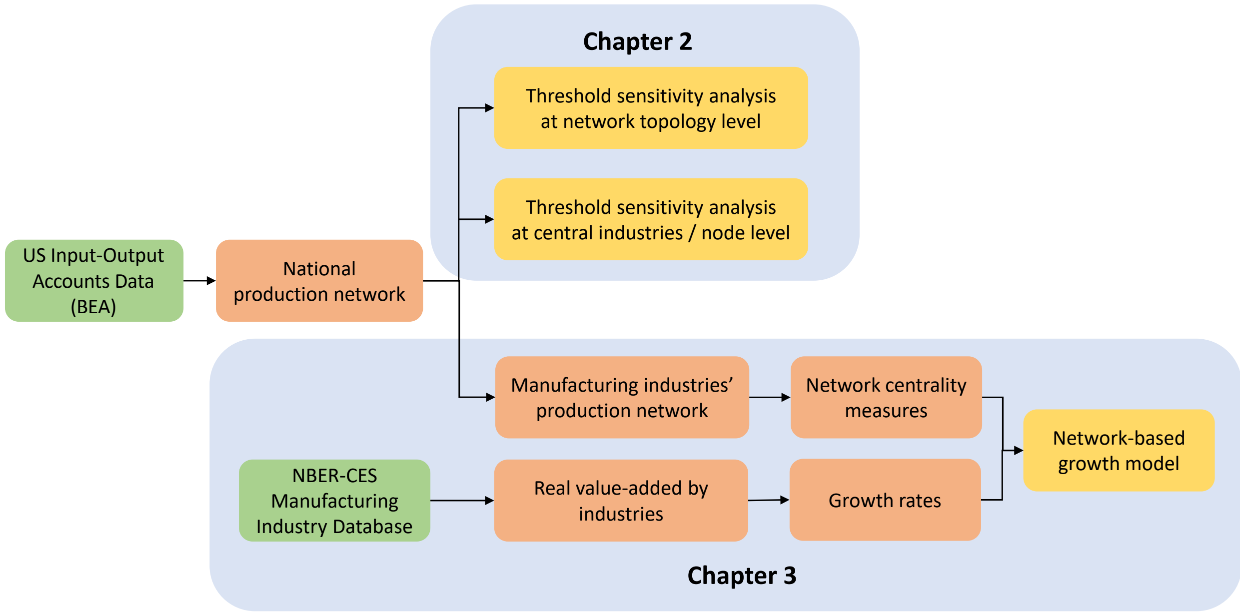


Figure 1.1: Methodological structure for Chapters 2 and 3

can be found at *Concepts and Methods of the U.S. Input-Output Accounts — U.S. Bureau of Economic Analysis (BEA)* 2023.

The Bureau of Economic Analysis publishes the Industry Economic Accounts generally at three levels of detail: sector (21 industry groups), summary (71 industry groups), and detail (405 industry groups). Estimates at the detail-level are produced roughly every five years, with the last two releases being from 2007 and 2012 at the time when the thesis was written. It uses the last releases (2007, 2012) of the most disaggregated version with more than four hundred industries (detail level), specifically the Total Requirements matrix (*Total Requirements Data — U.S. BEA Input-Output Accounts* 2023). These are industry-by-industry matrices showing the normalised monetary transaction values inside. Economic values are normalised because they show the inputs by industry required (directly and indirectly) from another industry in monetary terms in order to deliver one dollar of industry output to final users. From this matrix, the thesis built the 2007 and 2012 US national production networks. The outcome is a directed network where individual nodes correspond to industries, and each edge (or connection) signifies a trade association between two industries. The details of the transformation process are elaborated upon in their respective chapters.

The goal of the second chapter, after constructing the network from raw data, is to investigate the 2007 and 2012 production networks from a threshold sensitivity viewpoint. The 2007 and 2012 Input-Output Accounts are the last releases at the time of the project. For this, the whole national production network is used, and the thesis examines if the network topology and the central industries change according to the monetary transaction threshold used in their definition. The different threshold values are used to cut out generally the smallest inter-industry transactions in value, and this chapter examines how much this distorts the structure of the national production network. In other words, whether different thresholds highlight different features of the network. Thirty different threshold values are explored, and at every threshold, the topology of the national production network is compared to two well-known network structures: random and scale-free network structures. Specific algorithms are used with the original national production network's parameters to generate the random and scale-free graphs used for the comparison. Afterwards, a statistical similarity test is employed to compare the US national production network's observed node degree distribution to what we would expect from a random and scale-free graph with the same parameters. Also, the central industries are examined at different threshold levels. The procedure and the reasoning behind it are explained in detail in Chapter 2.

For the third chapter, the research needs to put together this national production network framework with the economic growth variables. The objective is to analyse how the attributes of the national industrial networks influence the growth of industries. Thus, the thesis introduces the second main data source: the National Bureau of Economic Research (NBER) and the US Census Bureau's Center for Economic Studies (CES) Manufacturing Industry Database (Becker, Gray, and Marvakov, 2021). The NBER-CES issues annual industry-level data from 1958-2018 on output, employment, payroll, investment, capital stocks, and various industry-specific price indexes. From all this, the industry-level (real) value-added metric is used to calculate growth rates. Because these were growth rates for the manufacturing industries, the thesis used the manufacturing industry part of the national production network. Even if narrowing it down to manufacturing industries, this left plenty of detail with a network of more than one hundred fifty industries. The dataset used has the structure of a balanced panel consisting of 156 observations (N) and two years (T). From this network, twenty different network centrality metrics are computed. The network metrics are used as the independent variables in the linear regression, and the industry growth rate is the dependent variable. The chapter also tried several different growth rates, all in the interval between 2002-2017. The two networks are from 2007 and 2012. The regression was inspired by Kali and Reyes, 2007, but this network-based growth model followed that only partially. They also used industry-specific metrics in their model; however, this thesis only uses network topology metrics as independent variables. A comprehensive explanation of this procedure, along with the underlying rationale, is provided in Chapter 3.

The next two empirical chapters investigate the storage battery industry context and are built based on the first two empirical chapters. Hence, their methodological structure shown in Figure 1.2 is very similar to the one presented in Figure 1.1.

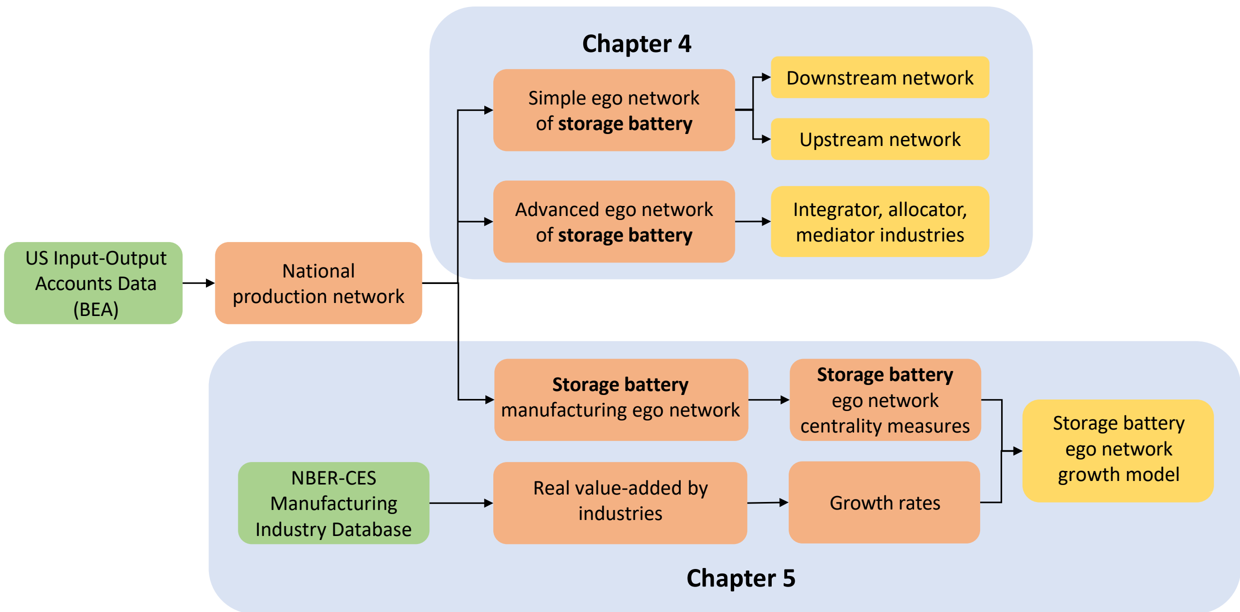


Figure 1.2: Methodological structure for Chapters 4 and 5

Nonetheless, there are some important differences, too, as the goals and some methods differ when analysing a specific industry in this context or an entire national trade system.

The goal of Chapter 4 is to examine the storage battery ego network and to map the indispensable relations and industries in it. A secondary aim is to show the viability of the production framework examined in Chapter 2 when analysing one specific industry and its production network. This chapter also starts with the 2007 and 2012 national production network derived from the US Input-Output Accounts, and from that, it constructs the simple and advanced storage battery ego network. These are all well-defined subsets of the entire national network. When focusing on the simple ego network, it is divided even more into upstream and downstream suppliers part. Afterwards, adding more complexity to the advanced ego network part, with the help of this framework, the thesis defines the integrator, allocator and mediator industries based on network centrality measures in the storage battery production cycle.

For Chapter 5, the same approach is used as in Chapter 3. Instead of narrowing the national production network to the manufacturing industries, it is narrowed down to manufacturing industries present in storage battery production. Thus, that is the 2007 and 2012 storage battery ego network. From this, the same twenty network centrality metrics are computed. Because the network is different, this is a closer environment in the sense of financial transactions; the measures are also different. Afterwards, the chapter employed the network growth model built for the national production network to this ego network. In this case, the independent variables are the storage battery ego network centrality measures, and the dependent variable is the growth rate of every industry.

1.4 Thesis structure

The thesis is divided into six chapters.

In **Chapter 1**, the thesis initiates by establishing the research's background and aim and also formulates the research questions intended to be tackled in this dissertation. This involves exploring the interdependencies between industries through the framework of national and industry-specific production networks and quantifying their contribution to industrial growth.

Chapter 2 begins by laying out a part of the theoretical dimensions of the research. The context for network science as a tool to map complex questions and dive into the production network field of study is provided. Also, the introduction of the question of threshold dependency is elaborated on as the main focus of this chapter. Afterwards, the empirical strategy designed to investigate the threshold dependency of the US national production network topology is described. It includes the data source: the inter-industry monetary transaction from the US Input-Output Accounts, the transformation of this data to the 2007 and 2012 national production networks and the methods used for analysing its structural distortion according to the monetary threshold change. The chapter finds that the industry inter-dependence network is highly exposed to the threshold value. Thus, different thresholds highlight different features of the network. The topology of the production network with no threshold or minimal threshold value cannot be classified into a well-defined network structure. However, as monetary transactions increasing in number and size are disregarded, the topology takes a different shape. Not just the overall topology changes but some core central industries that served as hubs dropped to the bottom of the top lists as the number of transactions included decreased. The chapter explains that the threshold value chosen makes a significant difference in the production network's structure; therefore, we must carefully consider and keep in mind the limitations of the research made on production networks using a particular threshold.

Chapter 3 builds on this approach and delves into the topic of economic growth. The aim is to define to what extent industrial growth can be determined purely from production network characteristics. It starts by examining the limited literature on industrial growth in the context of production networks. It will then go on to the development of the network-based growth model. This is a linear regression model with the network centrality measures being the independent variables and the industrial growth the dependent variable. The same US national trade relationships are used for the production network, and the network centrality metrics are derived from these.

The network centrality measures represent the roles that industries hold in the national production cycle, such as their central or influential power, based on their connectivity and relationships within the network. The growth of the industries is measured by real value-added changes from the NBER-CES Manufacturing Industry Database. The results are pioneering in the sense that this is the first model in the rapidly expanding field of industrial growth, focusing solely on inter-industry relationships and their effectiveness in explaining growth. The outcome of the network-based growth model on the US national trade suggests that network characteristics can define industrial growth up to a certain level. Almost a third of the change in the industry is explained by its location in the production chain topology alone. Even though this chapter is based only on relational metrics (excluding industry-specific metrics), the percentage reported here contributes to our understanding of industry dynamics. In the results section, the question of which network and centrality metrics explain the best economic growth is also covered.

In **Chapter 4**, the thesis shifts the perspective from the national dimension to the industry-specific analysis. The aim is to explore the storage battery industry and its supply chain through the lens of the production network framework. It starts by describing the theoretical background behind ego network science in the inter-industry relations context. Followed by this is the contextualisation of the storage battery industry from the renewable energy transition side and the formalisation of battery supply chain challenges. Afterwards, the methodology is described. The chapter constructed three different storage battery ego networks using the same US national trade data: simple upstream network, simple downstream network and advanced network. The simple ones are constrained only to the direct monetary transactions between the storage battery industry and any other supplier or receiver industry, while the advanced network also includes the transactions between the suppliers in which the storage battery industry is not present. These networks are analysed separately as each holds different kinds of information about the storage battery's supply environment. At the advanced ego network part, the network-level and the node-level metrics are explored separately. The thesis maps through the lenses of production networks the indispensable relations and industries according to the data in the production cycle of the storage battery industry.

Chapter 5 remains at the industry-specific perspective and builds on the network-based growth model designed in Chapter 3. The concept behind the model is the same; however, in this chapter, the storage battery ego network characteristics are the explanatory variables. As this is a closer and narrower environment in the sense of inter-industry relations, the expectation is to have a higher influential power on growth than the national context. The outcome is remarkable, as it suggests that a significant portion, precisely half, of an industry's success can be attributed to its position within the immediate production network and its proximity to key interfaces in the industry.

In the concluding **Chapter 6**, the thesis circles back to the research objectives and presents the overview of the principal findings through the thesis. It also delves into the implications both from the academic and policy sides. As of the former, bridging the gap in production network research, rethinking traditional economic metrics and models and also the industry-specific insights. For the latter, examples include resilience and risk management, industrial growth strategies, and supporting key industries. Next to conclusions, it provides guidance for future endeavours.

Chapter 2

From data to map: the production network and the threshold sensitivity of the topology

2.1 Network science as a tool to map complex questions

Network science is an interdisciplinary field that focuses on the analysis, modelling, and understanding of complex systems composed of interconnected elements, often referred to as nodes, and the relationships or interactions between these elements, represented as edges or links. It draws upon concepts and methodologies from various disciplines, including mathematics, physics, computer science, sociology, biology, and more, to explore the structure, behaviour, and dynamics of networks in diverse domains (Newman, 2003).

The foundation of modern network science can be traced back to the 18th century when Swiss mathematician Leonhard Euler, in his paper, "Solutio problematis ad geometriam situs pertinentis," introduced a revolutionary concept - the graph (Euler, 1741). He employed nodes (vertices) and edges (lines) to represent interconnected objects and relationships, respectively. Euler's solution to the "Seven Bridges of Königsberg" problem marked the first formal proof in the field of graph theory, setting the stage for the study of networks in mathematics.

The development of graph theory in the 20th century by eminent Hungarian mathematicians like Dénes Kőnig, Pál Erdős and Alfréd Rényi further expanded the theoretical underpinnings of networks (Konig, 1936; Erdős and Rényi, 1960).

The turning point for the discipline came in the mid-20th century with the recognition of the practical applications of network modelling. Stanley Milgram's "six degrees of separation" experiment in the 1960s brought attention to the small-world phenomenon in social networks, where surprisingly short paths connect individuals (Milgram, 1967). Duncan Watts and Steven Strogatz formalized this concept, igniting interest in studying small-world networks (Watts and Strogatz, 1998). These developments piqued interest in examining networks beyond mere mathematical abstraction, sparking a multidisciplinary approach to understanding the real-world structures and dynamics of complex systems.

In the early 2000s, the research of Albert-László Barabási and Réka Albert introduced the concept of scale-free networks, a characteristic that reflects the coexistence of hubs alongside numerous vertices with low degrees in real networks (Albert, H. Jeong, and Barabási, 2000; Albert and Barabási, 2002). This research explained the emergence of power-law degree distributions in various complex systems, including the World Wide Web. Network science, once largely confined to mathematics, became increasingly interdisciplinary.

As the 21st century progressed, network science's popularity soared, driven by its applicability across domains. It found applications in social networks, transportation systems, biological networks, information networks, epidemiology, and more (Pastor-Satorras and Vespignani, 2001; Barrat et al., 2004). Researchers from physics, sociology, biology, computer science, and other fields converged to explore the intricate structures and dynamics of networks in various contexts (Girvan and Newman, 2002).

Hence, network science is not only present in one big discipline. For example, in computer science, it is used for analysing websites, internet traffic and information dissemination. It can even be used for studying food chains and living systems in natural sciences or for modelling in mathematics and physics. Social network analysis can be useful in business when organizing communication leadership or exploring the customer base and the supply chain, and it can also help discover criminal and terrorist networks in legal theory. On another level, Barabási, 2003 says that we can notice that the networks of genes working together are key to understanding how cancer works and how it can be treated. The famous specialist in network analysis emphasizes that we have only examined some parts of the whole so far. We have tried to correct some problematic components and blamed a few defective genes for causing different diseases. But what if this is not precisely the case? What if there is an error somewhere in the relations, and something goes wrong with the system? Barabási, 2003 says that we know nearly everything that should be known about smaller parts, but we are far from understanding nature, society, and the world as a whole. Our models to describe reality are mostly simplified, and putting these simplified elements together is harder than we thought. This is because by forcing simplification, we were faced with complexity. Elements could be put together in many different ways in complex systems, and it would take billions of years to try out every possibility. Therefore, network analysis opens a new window in every discipline through which we can examine different problems: the window of complexity and embeddedness.

As in the words of Jorge Luis Borges, "Everything is connected to everything." (Borges, 1998) it becomes more and more obvious that nothing can happen in isolation. We live in a world in which everything is connected to everything else. We don't know everyone, but surely, there is a possible connection between any two persons in the human network. In the same way, there is a connection between any two companies in the world or any two chemical elements in the human body. Nothing and no one is excluded from the closely related network of life (Barabási, 2003).

Networks are especially important in economics-related modelling. Our economy is embedded in structures of social relations (Granovetter, 1973), which consist of networks that initiate commercial interaction respectively, allowing these interactions to come into being. Granovetter thinks that the relationship between clients and salesmen is at least that important – if not even more important – than the commercial transactions themselves. The transactions do not only take place between strangers but often between individuals who already know each other for a long time.

The embeddedness of markets is also important because relations are costly. If relations had no expenses, society would behave like an entirely interconnected network. But relationships do have a cost; thus, the embeddedness becomes extreme, and markets are limited by the structures of the social networks that surround them.

An example of this is the study carried out by Borondo et al., 2014 in which they analysed the boundaries between the meritocratic and topocratic embedded markets. A system is meritocratic if the individuals' compensation "power" is defined by their abilities and merits, but it is topocratic if this is defined mainly by their position in the network. The model set up two channels of compensation: a meritocratic one, where the individuals get their compensation for the content they have created, and a topocratic one, where the compensation takes place based on the numbers of the shortest paths which passed through the individual within the network, namely on the number of relations in which the individual played the role of a mediator. The two separate channels allowed them to investigate the payments and also to examine if the individual got these payments because of the content or because of his role as a mediator. The research has proved that the entirely connected system is entirely meritocratic and becomes increasingly topocratic as the density of relationships gradually decreases. Entirely connected systems do not exist in real social systems. However, the model suggests that a system which heads from a sparse network to a more closely connected one becomes more and more meritocratic. Taking today's technological changes into consideration, this is one of the most important results of the research. The recent changes in communication technologies have increased the connectivity of our society, decreasing the expenses of social and economic interactions simultaneously.

The history of network science is a testament to the enduring human quest for understanding interconnected systems. From Euler's elegant graph theory to the modern study of complex networks, this journey reflects not only the evolution of our comprehension of networks but also the profound impact of network science on our contemporary world.

Modern systems are inherently complex and interconnected. Traditional reductionist approaches often fall short in capturing the intricate relationships and emergent phenomena within these systems. Network science provides a holistic perspective that helps unravel this complexity (Lewis, 2011). Networks foster a systems thinking approach. By viewing systems as interconnected networks, researchers gain insight into how changes in one part of the system can ripple through the entire network, with cascading effects on its structure and behaviour (Börner, Sanyal, and Vespignani, 2007).

With network-like (societal) structures serving as the backbone of information flow, transportation logistics, and social interactions, network science offers indispensable tools to optimise infrastructure, predict disease outbreaks, mitigate financial risks, and even unravel the mysteries of our interconnected brain. In a world facing grand challenges such as climate change and information security, network science empowers us with the means to address these complex issues systematically, making it an essential and ever-evolving discipline in our interconnected era.

2.2 The production network and the threshold-dependent behaviour

Inter-industry relations are profoundly interconnected; therefore, we must use techniques that view industries from a systematic perspective and understand the complexity of relations. No industry can exist and develop individually. All sectors are interdependent through the exchange of products and services (Xu and Liang, 2019). Given the increasing complexity of our society, we cannot examine a problem without the help of multidisciplinary approaches. In this study, I use network and data science tools to explore traditional economic science issues, specifically the question of industry inter-relatedness through the lens of complex production networks.

From the economics perspective, an industry is a branch of an economy that produces closely related raw materials, goods, or services, while from the business point of view, it is a group of companies that are related based on their primary business activities. There are dozens of industry classifications, and these classifications are typically grouped into larger categories called sectors.

Leontief's (W. Leontief, 1970; Dietzenbacher and Lahr, 2008) economic input-output models represent, in mathematical form, the monetary transactions between industry sectors. They specify what goods and services (output) are consumed by other industries (input).

The network science approach to input-output models is not a new concept. There are mainly two basic approaches currently being adopted in this research area. One is the analysis from a supply chain perspective, using company-level data (Wu, 2015; Brintrup and Ledwoch, 2018; Perera, Bell, and Bliemer, 2017), and the other is the industry perspective, using country or global-level input-output accounts.

In the industry perspective approach, a considerable amount of literature has been published using the World Input-Output Database (Xu and Liang, 2019; Soyyiğit and Boz, 2018; Grazzini and Spelta, 2015; Baldwin and Lopez-Gonzalez, 2015; Soyyiğit and Çırpıcı, 2017), covering 40 countries in the 2013 release and 43 countries in the 2016 release, all with 35 (2013) and 56 (2016) sectors (Timmer et al., 2015). Although this data source can cover most countries, it only contains information on very aggregated sectors. In recent years, researchers have also investigated various approaches to the input-output transaction data of the US economy as systematised by the Bureau of Economic Analysis (BEA) (*Input-Output Accounts Data — U.S. Bureau of Economic Analysis* 2021). Most of the studies focused on the sector and summary level of the input-output accounts containing 21 (sector-level) and 71 (summary-level) aggregated industries (Foerster and J. Choi, 2017; Duan, 2012).

Sectoral inter-dependencies are pivotal in connecting microeconomic shocks with business fluctuations and cycles, especially during the supply-chain fluctuations of today. Researchers claim that production networks derived from input-output models provide a valuable account in opening the black box of co-movement and propagation mechanisms that shape aggregate outcomes. Carvalho (Vasco M. Carvalho, 2014) points out that production network research can advise scholars on the origins of aggregate fluctuations and policymakers on how to be ready and recuperate from disadvantageous shocks - long studied in the system dynamics modelling community - that disturb production chains (Keith, J. D. Sterman, and Struben, 2017; Dykes and J. Sterman, 2017).

Horvath (Horvath, 1998; Horvath, 2000) and Acemoglu et al. (Acemoglu et al., 2012) argue that the structure of the production network plays a crucial role in determining the aggregate behaviour of the system. Network structure is the set of nodes and edges of a network. Nodes representing industries in the production network carry certain quantitative properties, which are represented as weights. The edges representing the monetary transactions between industries also carry weights, but in the case of a production network they are also directed - and the weights vary significantly in terms of directions.

When the production network is outstandingly asymmetric, for example, when few sectors are in a dominant role as suppliers, idiosyncratic shocks lead to aggregate fluctuations. If the production organisation is dominated by a small number of hubs supplying inputs to many different sectors, disturbances at these crucial nodes will affect the global production system, determining losses in production and welfare (Acemoglu et al., 2012). Bigio and La'O (Bigio and La'O, 2016) also demonstrate that overall network topology defines the strength of each channel. They compare two production networks: a horizontal and a vertical economy, and show that the network centrality of sectors matters for how they affect aggregate output. Carvalho (Vasco M. Carvalho, 2014) also advocates for the significance of network topology by comparing the amplification of micro-level volatility and the network multiplier a horizontal economy with no input trade, a vertical economy with a source and a sink, and a star/hub-and-spoke economy with a central node/s.

The US Bureau of Economic Analysis publishes the Industry Economic Accounts generally at three levels of detail: sector (21 industry groups), summary (71 industry groups), and detail (405 industry groups). For example, at the sector-level, Durable goods are included among the 21 industry groups. This sector at the summary-level contains Primary metals, Machinery, Computer and electronic products and 8 other sub-sectors. While broken down even further to detail-level, the Primary metals include 10 industries (Iron and steel mills and ferroalloy manufacturing, Ferrous metal foundries, etc.), the Machinery includes 28 industries (Farm machinery and equipment manufacturing, Semiconductor machinery manufacturing, etc.), and Computer and electronic products include 20 industries (Electronic computer manufacturing, Telephone apparatus manufacturing, etc.). These industry classifications are all grouped hierarchically into three levels.

Although extensive research has been carried out from an industry perspective, just a few studies exist that develop a network of at least 400 detail-level industries. This can be a key problem because the networks built on a highly aggregated level with just a few nodes and connections might not represent the industry interdependencies accurately.

On the one hand, the topology of a detail-level industry network can differ considerably from an aggregate-level network. The first one tends to be way denser with more lower-weighted connections, thereby behaving differently. On the other hand, some essential links could be hidden in an aggregate-level network. For example, embedded in a highly-weighted connection, several detail-level links could have been hidden that might be more important than the other present aggregated ones. The whole map of industrial interdependence could change when analysing these separately. The detail-level input-output account data might allow us to discover a more representative picture not just with more separable sectors but with way more supporting connections between industries in number and validity. This increased granularity of understanding allows for better identification of key industries undergoing change and more efficient tracking and understanding of innovation.

The general tendency in the research community is to use a threshold value and consider only a percentage of the monetary transactions to be present in the production network. As the monetary transactions will become the edges between industry-industry pairs when constructing the production network, they do this mainly to make the data more manageable. Initially, in the input-output models, there are almost i^2 supporting transactions, with i being the number of industries. Every industry is connected to almost every other, and almost every industry is affecting revenue generation in every other. The production networks obtained from such detail-level input-output models are exceptionally dense and challenging to analyse without offering much insight.

This problem can be mitigated by carefully choosing the level of intersection to minimise the distortion in the network structure caused by the removal of data (Radicchi, Ramasco, and Fortunato, 2011; De Benedictis et al., 2014). The goal is to prune a tiny fraction of the total weight so that as much information as possible is preserved (García-Algarra, Mouronte-López, and Galeano, 2019). Hence, most scholars impose a threshold and exclude a percentage of the smaller monetary transactions - thus "severing" these links in the production network. While not explicitly referred to thus far, in this thesis I refer to this cut-off point as threshold ζ.

Collectively, these studies outline a critical need to examine how the threshold ζ influences the topology of the production network and, thus, the propagation mechanisms and aggregate fluctuations. While we might agree that production networks can provide a bridge between micro and macro (Vasco M. Carvalho, 2014), very little is currently known about how much the threshold value ζ chosen by scholars distorts this bridge.

In this chapter of the thesis, I discover the considerable change of the production network topology to the threshold value ζ by analysing a detail-level input-output model with a high number of sub-sectors to understand the most accurate structure of the network of industries. The expectation is that different thresholds highlight significantly different features of the network.

The *Strategy for analysing the threshold dependence of the topology* section presents the chosen dataset, the production network model, and the topological and node-level analysis. I present the outcome in the *Findings for threshold sensitivity of production network topology* and summarise everything in the *Discussion* section.

2.3 Strategy for analysing the threshold dependence of the topology

2.3.1 Data source

I used the last releases of the detail-level United States Input-Output Accounts Data (*Input-Output Accounts Data — U.S. Bureau of Economic Analysis* 2021) to conduct this research. The Bureau of Economic Analysis generally publishes the Industry Economic Accounts at three levels of detail: sector (21 industry groups), summary (71 industry groups), and detail (405 industry groups). Estimates at the detail level are produced roughly every five years, with the last two releases being from 2007 and 2012. While there are Historical Benchmark Input-Output Tables from 1947 until 2002, these reflect industry definitions that vary across years, and BEA advises that they should not be used as a time series. BEA uses its Industry Codes for the three levels of detail, but it also defines how these relate to the 2012 North American Industry Classification System (NAICS) code structure (*North American Industry Classification System* 2021).

I used the last two releases, the 2007 and 2012 Total Requirements table - Industry by Industry, including 405 industries and inter-industry purchases (*Total Requirements Data — U.S. BEA Input-Output Accounts* 2023). Additional explanations regarding the Total Requirements derivation can be found on BEA's website (*Total Requirements Definition — U.S. Bureau of Economic Analysis* 2021).

A noteworthy aspect where the dataset used in the research falls short is its lack of transaction costs within supply relations. This omission obscures crucial economic relationships that play a significant role in understanding the dynamics of supply chains. While the dataset captures transactions, it does not delve into their economic implications, particularly regarding transaction costs. Transaction costs encompass various factors, such as the time required to secure links in the supply chain, risks associated with transactions, levels of trust between parties, considerations of quality, and pricing dynamics. These factors are essential in comprehensively analysing supply chain cost structures and efficiencies, yet they are not accounted for in the dataset. Moreover, the dataset's nature as a post hoc compilation implies that it reflects a specific moment in time, potentially overlooking the ongoing costs incurred by firms in their pursuit of optimal suppliers or customers. Despite this limitation, the dataset remains valuable due to its richness in disaggregated economic sectors. Collecting transaction cost data for each transaction within such a vast and diverse dataset would be an immensely time-consuming task. The dataset's granularity in capturing various economic sectors allows for detailed analysis and insights into sector-specific dynamics, contributing to a comprehensive understanding of economic activities. Therefore, while the absence of transaction cost data poses a challenge to fully assessing supply chain dynamics, the dataset's breadth and depth offer valuable opportunities for exploring other facets of economic interactions and behaviours within and across sectors.

Another important limitation of the dataset used in the research is its lack of consideration for market power, ownership, and control dynamics within supply relations. While the dataset captures various transactional elements, it overlooks the influence of market power on the volume, direction, and values of transactions. Regardless of a firm's or industry's position within the network, market power significantly shapes transactional dynamics through its impact on competition and bargaining power. This oversight can lead to underestimation or overestimation of values, especially concerning international transactions and international data; the distortion is way less in national data. While acknowledging these limitations, it's important to recognise the inherent complexity of capturing such dynamics within a dataset, particularly at the national level encompassing numerous industries.

2.3.2 Production network generation from data

The BEA Total Requirements table has a network representation in which an element w_{ij} represented the nominal amount of goods i used as input by sector j, with $i,\ j = 1, ..., N$, where N was the number of sectors.

For a two-industry example of a network from the Total Requirements table, see

Figure 2.1. The network built from this database is a directed weighted graph. The vertices are the industries (i and j), and the directed connections are the monetary transactions, the nominal flow of goods between sectors. The weight of each link is the economic value, representing how much an industry supports the development of the other (w_{ij} and w_{ji}). For example, industry i requires w_{ij} dollar input from industry j to produce one dollar output to final users (typically $w_{ij} \neq w_{ji}$).

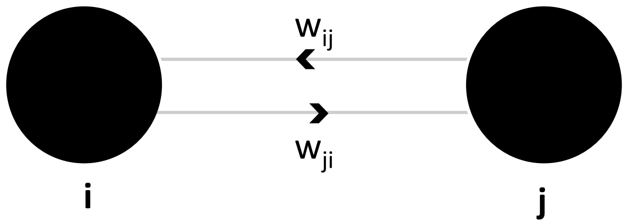


Figure 2.1: Example of a two-industry production network

The input-output models also indicate if the output of an industry is required as input to the same industry (w_{ii}). For example, the Oil and gas extraction industry could produce the oil and gas to power its own equipment, or the computer design and manufacturing industry can produce computers that are used to plan the next generation of computers (*Economic Input-Output Life Cycle Assessment. Carnegie Mellon University Green Design Institute* 2008). I didn't use these self-sector transactions in the analysis.

2.3.3 Monetary transaction threshold ζ

After building the production network from the Input-Output Accounts Data, I defined thirty different edge cut-off thresholds ζ from 0.00001 to 0.15 with 0.005 equal intervals. The threshold ζ is associated with the value of the links, representing inter-industry trade in monetary expression. For example, at the end of the analysis, when the threshold value ζ is 0.15, I only consider those connections that weigh at least 0.15. In other words, only those monetary transactions become edges in the resulting production network where at least 0.15 dollars is needed from one industry to produce one dollar output in another. At this threshold ζ, in the 2012 production network, only 142 industries (nodes) from the starting 405 and 156 monetary transactions (edges) are present. I chose this long interval and these small steps to represent the best topology change according to the cut-off ζ.

2.3.4 Algorithms for random and scale-free models

Then, I *pruned* the graph according to the considered cut-off thresholds ζ. In other words, I removed the edges that have a lower weight. I generated a random and scale-free graph with the same parameters as the production network at each cut-off point ζ. For both, I used the NetworkX graph generator algorithms (*NetworkX Python package* 2014).

The goal is to find out whether the network follows one of the canonical topologies - because then it can draw conclusions about its behaviour. For this, I generate several networks and use a distance metric to measure similarity.

For the random graph, I used the directed *G(n,m)* algorithm, with the parameters n being the node number and m the edge number. I gave the same node and edge number for the random graph generator algorithm as it is in the original production network.

For the scale-free graph, I used the algorithm implemented after Bollobás et al. (Bollobás et al., 2003) with the same node number and by calculating α, β and γ parameters to fit the production network. The probability of adding a new node in this algorithm is distributed between the parameters α, β and γ, hence the sum of these should be 1. α is the probability of adding a new node connected to an existing node chosen randomly according to the in-degree distribution, while γ is the opposite of this: adding a new node according to the out-degree distribution. β is the likelihood of adding an edge between two existing nodes. One existing node is chosen randomly according to the in-degree distribution and the other is chosen randomly according to the out-degree distribution. In simple terms, α and γ define whether the directed network is scale-free from the in-degree or the out-degree perspective. I represented these parameters by checking the top in-degree and out-degree values in the original production network. In this context, this means comparing the top buyer to the top supplier in terms of the number of different transactions.

2.3.5 Measure for degree distribution comparison (Kolmogorov-Smirnov test)

However, there are some other methods too (Barrat et al., 2004; Newman, 2005); I chose the Kolmogorov-Smirnov (KS) statistical test to compare the observed node degree distribution to what we would expect from a random and a scale-free graph with the same parameters because it is widely used when comparing degree distributions for networks (Deng et al., 2011; Muchnik et al., 2013; Gómez, Kaltenbrunner, and López, 2008; Gjermëni, 2017).

As the production network is a directed network, I distinguished between in-degree and out-degree:

$$k_{in,i} = \sum_{j=1}^{N} a_{ij} \qquad k_{out,i} = \sum_{j=1}^{N} a_{ji} \tag{2.1}$$

where $k_{in,i}$ is the incoming degree of node i, representing the number of incoming edges onto the node, a_{ij} is 1 if there is a directed transaction between industry i (buyer) and industry j (seller). In other words, if industry i requires the output of industry j, otherwise it is 0. The outgoing degree is $k_{out,i}$, representing the number of links that point from node i to other nodes. In this case, a_{ji} is 1 if the industry i is a supplier to industry j. In these measures, self-sector transactions, indicating if the output of an industry is required as input to the same industry, are not included (a_{ii}).

I analysed and compared in-degree and out-degree distribution separately at each threshold ζ by observing when the network behaves like a random or scale-free graph. I used the SciPy two-sample KS statistical test based on (Hodges, 1958). The KS statistic quantified the distance between the empirical distribution functions of the two samples (the observed degree distribution and the random/scale-free graph's degree distribution), hence creating a comparable metric for similarity. The smaller this value was, the more likely the two samples were drawn from the same distribution.

2.3.6 Production network centrality measures

I calculated two network centrality measures for every node at six cut-off points ζ from 0.0 to 0.1 with 0.02 equal intervals on the directed weighted graph. In this interval, monetary transactions are exponentially disregarded, and at 0.1, there are still more connections than nodes in the production network. With 6 cut-off points ζ in this area, I can illustrate the change in the central industries well.

The degree is defined as the total number of links of a node. The production network is weighted and directed; thereby, I computed the weighted in-degree (in-strength) and the weighted out-degree (out-strength) of each node:

$$k^{w}_{in,i} = \sum_{j=1}^{N} w_{ij} \qquad k^{w}_{out,i} = \sum_{j=1}^{N} w_{ji} \tag{2.2}$$

where $k^{w}_{in,i}$ is the in-strength of industry i, representing the sum of all incoming flows of goods, the nominal inputs used by the sector i, w_{ij} is the transaction value between industry i (buyer) and industry j (seller). In other words, the required amount from industry j to produce one dollar of output from industry i. The outgoing degree for industry i is $k^{w}_{out,i}$, representing the sum of the outflow of goods from node i to other nodes. In this case, w_{ji} is the nominal input that industry i supplies to industry j. I didn't use self-sector transaction weights representing in what amount the output of an industry is required as input to the same industry (w_{ii}).

I chose the weighted out-degree and the weighted *PageRank* out-degree centrality (Page et al., 1999; Langville and Meyer, 2004) as key metrics. The first represents how big and indispensable an industry is in monetary terms, *globally* in the entire production network, and the second represents how key an industry is in terms of *location* in the network.

The reason behind the PageRank algorithm is that a node is systemically important if its neighbours are important and/or the neighbours of the neighbours are important. However, the production network is a directed network, and in directed networks, we distinguish the links based on their directions. Defining the centrally located nodes in undirected networks is relatively straightforward, while in directed networks is much more complicated. Most centrality metrics in network research, such as the PageRank, are based on in-degree (the number of edges pointing to the node). Hence, in this case, based on the number of suppliers. Although, the production network is quite a particular network in this sense. One might ask if those industries are the central ones that need the most resources or the ones that pump the most supply into the network. The critical sectors that dominate economic activity are all hidden if we only consider the in-degree-based centrality metrics. That being the case, we calculate the weighted PageRank out-degree centrality on the reversed production network. The inverted network contains the same nodes and edges, but the direction of the edges is reversed.

At each threshold ζ, I show the top 20 central industries and compare how they changed according to the cut-off.

Besides this, I also compare the weighted in-degree and out-degree distribution without a cut-off ζ to dig deeper into the asymmetry analysis.

2.4 Findings for threshold sensitivity of production network topology

The results provide a *first* insight into the production networks' topology change according to the threshold value ζ.

I first analysed the topological features and found that the industry inter-dependence network topology is highly sensitive to the chosen threshold ζ. Starting with the smaller monetary value transactions, as I disregarded the higher and higher ones, the production network's topology transformed; from the out-degree perspective, it leaned towards a scale-free structure and from the in-degree perspective towards a random structure.

Figure 2.2 shows the edge weight distribution throughout the whole production network. In other words, the normalised monetary transaction values distribution in the national economy as it is present in the US input-output accounts. It is quite apparent from this figure that the vast majority of monetary transactions are minimal in their value. Therefore, the very small threshold ζ values cut out the most transactions.

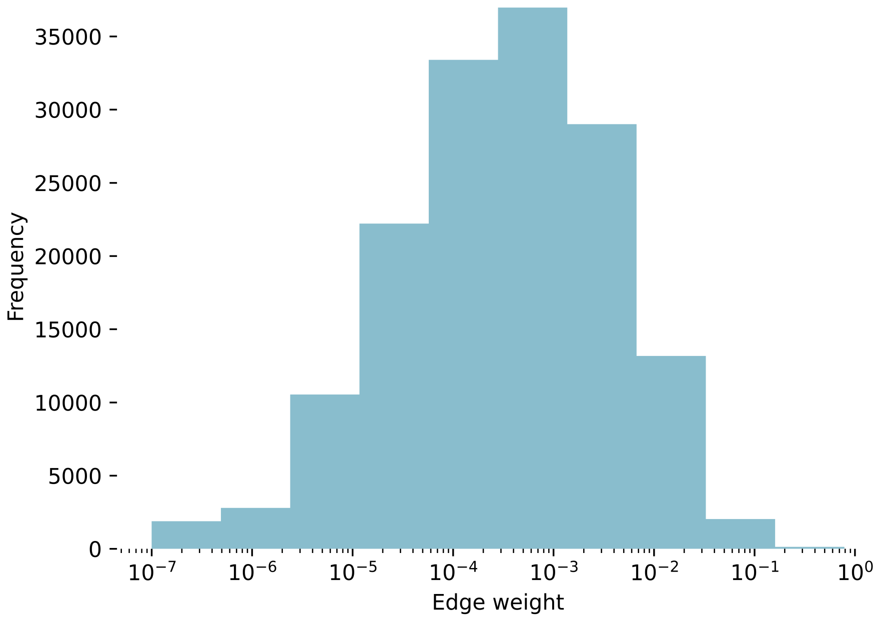


Figure 2.2: Global edge weight distribution (log-lin scale)

Figure 2.3 shows the Kolmogorov-Smirnov statistical test's value at various edge cut-off thresholds ζ. I compared the production networks' observed in-degree and out-degree distribution at each threshold ζ to what we would expect from a random and scale-free network with the same parameters. The closer the KS statistic value is to 0, the more likely the two samples are drawn from the same distribution.

With the threshold ζ increasing, the edge numbers decline at a very fast pace in the beginning. From the starting value of more than 150 000 edges, only around a third, 50 000 edges, remain at the very small threshold of 0.001 dollars. The same goes for the threshold of 0.005 dollars with 20 000 edges and the threshold of 0.01 dollars with 10 000. The node number during these thresholds stays the same. After that, the decline is a bit slower in the edge number, and the nodes start to decrease slowly, too. I end the analysis at a threshold of 0.15 dollars with 142 nodes and 156 edges.

Neither the in-degree nor the out-degree perspective is similar to the random or scale-free distribution with no threshold or very low threshold value ζ. After a while, the in-degree distribution inclines towards a random structure and the out-degree towards a

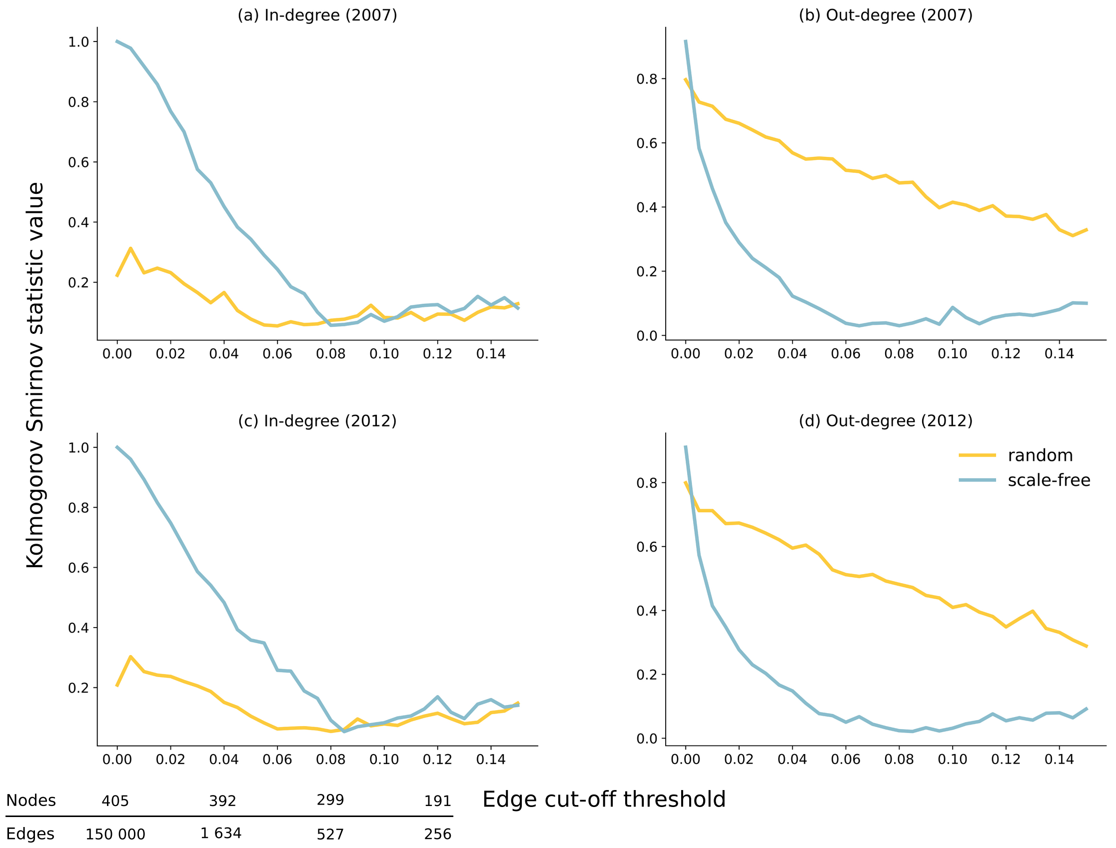


Figure 2.3: Kolmogorov-Smirnov statistic value of the distributions compared (lower value means more similar)

scale-free.

According to out-degrees, the scale-free shift appears around the 0.05 threshold value ($\zeta \approx 0.05$) when the KS statistic falls below 0.1 (in the 2007 network 0.08, and in the 2012 network 0.07) and the p-value exceeds the critical 0.05 (in the 2007 network 0.15, and in the 2012 network 0.23). It stays in the range until the final ζ threshold value of 0.15, analysed in the 2007 network with a KS statistic of 0.1 and a p-value of 0.48 and in the 2012 network with a KS statistic of 0.09 and a p-value of 0.59. The production network is the most scale-free at the 0.065 ζ threshold value in 2007 (KS = 0.029, p-value = 0.99) and at the 0.085 ζ threshold value in 2012 (KS = 0.021, p-value = 0.99).

The in-degree network perspective starts to have a topology more similar to a random network after the 0.05 threshold value (in the 2007 network, KS = 0.07, p-value = 0.21, and in the 2012 network, KS = 0.08, p-value = 0.18). Just as with the out-degree perspective, the in-degree stays in the random range too, having at the end of the comparison (at the 0.15 threshold) in the 2007 network a KS statistic of 0.12 and a p-value of 0.19 and in the 2012 network a KS statistic of 0.17 and a p-value of 0.09. It is most similar to a random network at the 0.06 ζ threshold value in 2007 (KS = 0.054, p-value = 0.67) and at the 0.08 ζ threshold value in 2012 (KS = 0.053, p-value = 0.78).

The results showed that the threshold changes the structure of the production network. I might actually say that even with these very small step thresholds, the production network becomes a different entity at every cut-off, especially during the first phase of imposing the threshold. However, that is not entirely true, as the same industries and the highest transactions in value remain in the network. While the topology slowly shifts. From both dimensions, the most notable change is between the 0.0 and 0.05 thresholds. At the 0.05 threshold while still having the vast majority of the industries and the most important monetary transactions in value (in the 2007 network: 373 nodes, 1154 edges, and in the 2012 network: 363 nodes, 1232 edges) the topology from the in-degree perspective takes the shape of a random network and from the out-degree side the shape of a scale-free structure. At this point, 0.8% of all original transactions are present in the network.

Carvalho (Vasco M. Carvalho, 2014) in his production network research accounts for about 80% of the total value of input trade, which would leave us with 5 000 edges out of 150 000, 3.5% of all transactions in the network, around the 0.02 cut-off point ζ ($\zeta \approx 0.02$).

In-degree and out-degree distributions of the production network behave very differently, as is expected. It is more similar to a random network from the in-degree perspective, whereas, from the out-degree perspective, it leans towards a scale-free topology.

The scale-free nature of trade networks is a broadly analysed subject in production network research, too (Gualdi and Mandel, 2016; Liu, Shen, and Tan, 2021).

To discover more about this asymmetry, I show the weighted in-degree and out-degree distribution of the entire production network without cut-off ζ in Figure 2.4.

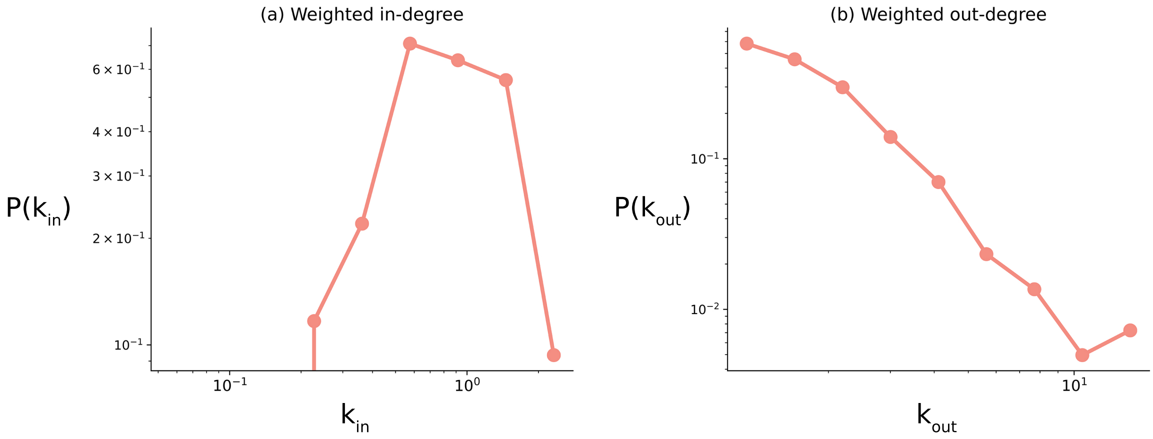


Figure 2.4: Node degree distribution of the production network

Some scholars have already approached the question of asymmetry in directed networks (Grazzini and Spelta, 2015; Horvath, 1998; Horvath, 2000; Acemoglu et al., 2012; Luo and Whitney, 2015; L. Wang and X. Wang, 2017). In the case of the production network, it is quite transparent the underlying explanation.

The topological difference comes from the presence of critical sectors that dominate economic activity, the so-called "commanding heights". Vladimir Lenin used this phrase in the early 1920s, referring to the control of key segments of a national economy. The difference between raw and processed materials is evident in the degree distribution comparison. The core industries that drive the economy formed hubs and push the topology to the scale-free range from the out-degree perspective. While from the in-degree perspective, most sectors need the same amount of resources; therefore, the topology leans towards a random network.

According to size and centrality measures, the results show that some core industries tend to remain on top at each threshold ζ. Figure 2.5 reveals the node-level threshold ζ sensitivity of the production network. I calculated the top 20 sectors according to weighted out-degree and weighted PageRank out-degree centrality at six cut-off points ζ. I found that the top three industries stay the same: 1. Iron, still mills, ferroalloy manufacturing, 2. Oil and gas extraction, 3. Petroleum refineries.

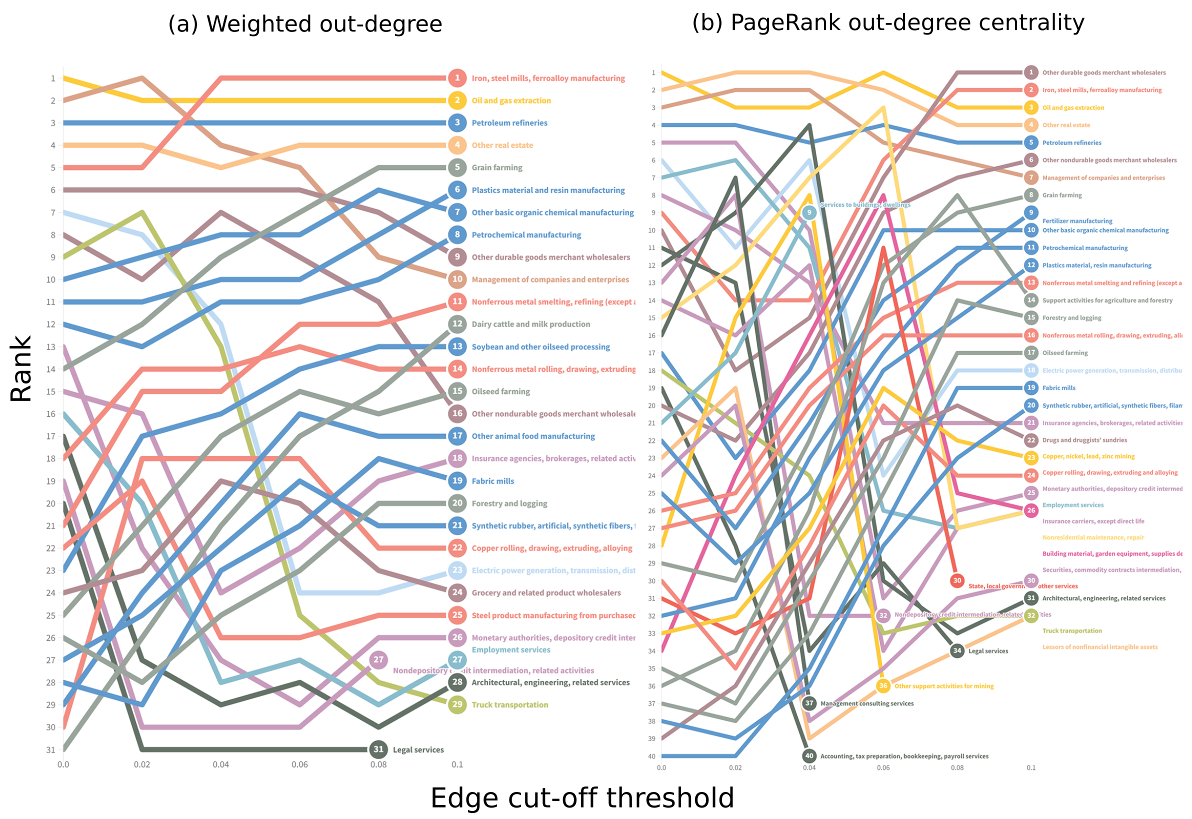


Figure 2.5: Top industries according to threshold ζ

Because the threshold ζ change cuts off the minor value transactions, some industries leading the top lists drop to the bottom of the rankings. These are usually the industries that support almost all the other ones but with smaller transactions, such as Truck transportation and Electric power generation, transmission, and distribution. Therefore, we can conclude that the threshold ζ cuts out some significant hubs that amplify aggregate fluctuations.

2.5 Discussion

In this chapter, I aimed to analyse the widely used threshold's influence on the production network topology.

There is a growing body of literature that recognises the importance of production network research, but no attention has been paid yet to the thresholding methods in this framework. My hypothesis was that changing the percentage of monetary transactions that researchers include in the network (edge cut-off threshold ζ) changes the structure of the network and the core industries to a large extent. In other words, different thresholds highlight different features of the network. Hence, if the topology is highly sensitive to the threshold, that can have further implications for studies based on this.

I proved the hypothesis by examining the network topology and centrality metrics under different thresholds ζ on a network derived from the US input-output accounts data for 2007 and 2012, by showing that the topology is more similar to random at one cut-off and scale-free at the other,

I discovered that the industry inter-dependence network is highly exposed to the edge cut-off threshold ζ. The topology of a production network with no threshold or very low threshold ζ cannot be classified as a random or scale-free network structure. However, as I disregard monetary transactions increasing in number and size, the topology from the in-degree perspective leans towards a random network structure and, from the out-degree side, takes the shape of a scale-free network.

The node-level threshold ζ susceptibility analysis shows us that not just the overall topology changes but some core industries that served as hubs dropped to the bottom of the top lists as I decreased the number of transactions included. These industries mainly support numerous other industries with small monetary transactions.

In his study, Carvalho (Vasco M. Carvalho, 2014) mentions that hubs shorten distances. While cutting out these transactions, we lose these hubs and increase distances. These monetary transactions contribute directly to the production networks' connectedness and complexity, which explains the propagation (and sometimes circularity) of local shocks and disturbances. For example, a reasonable question to ask would be, if we do not consider them all, in what proportion will the propagation modelling be distorted?

Even small-value supporting transactions could make a difference. For example, while we might cut out most of the Truck transportation transactions, a disturbance in the industry could impact all the others dependent on it, even with just a small amount. If there is an interruption in Truck transportation, manufacturing industries shake too, because, among other effects, probably, the spare parts don't arrive in time. Therefore, if we do not consider these transactions, we lose some essential propagation mechanisms.

These small monetary transactions not only directly contribute to the propagation, as Truck transportation directly supports other industries, but they also function as intermediaries. If we cut them, we ignore all the propagation mechanisms along the chain. The Truck transportation disturbance might impact the manufacturing industries, but this shock might propagate even further in smaller proportion to the third-forth industries connected to the manufacturing ones.

Another aspect that is quite important but I didn't capture through this analysis is that these interdependencies are in no way linear. For example, the Truck transportation industry might contribute only 5% to Car manufacturing. Still, if the Truck transportation industry stops, the other sector falls to zero without access to that crucial part. This limitation of this chapter can be overcome by examining exactly this resilience aspect in future work.

This chapter has shown that the threshold value ζ chosen makes a significant difference in the production network's topology; therefore, we must carefully consider and keep in mind the limitations of the research made on production networks using a particular threshold ζ.

The determination of the threshold value in this context is inherently heuristic, as emphasised throughout the discussion. It is recognised that the final and optimal threshold value is not a straightforward mathematical solution but rather contingent upon the specific case at hand. The choice of this threshold value isn't solely based on a fundamental criterion but involves empirical testing to gauge its impact and relevance within the context of the network being analysed. This process acknowledges that different goals may necessitate different threshold values; therefore, a singular universally applicable threshold doesn't exist. The empirical testing conducted involved observing the network's topology shift, particularly noting a significant change occurring at a threshold of 0.05

within the monetary transaction dashboard. The theoretical implications of this approach highlight the necessity of considering context-specific factors and allowing thresholds to be endogenously generated from the data rather than imposing arbitrary values. This perspective aligns with the methodological and philosophical stance within economics, advocating for theoretically and empirically justifiable thresholds rather than universally standardised ones. Thus, the framework developed through this empirical testing process offers insights into how similar methodologies could be generalised across various sectors and industries, emphasising the importance of context-driven threshold determination rather than adopting preset values without justification.

As a further step of the research, it would be reasonable to explore the distribution fitting with other methods too (Barrat et al., 2004; Newman, 2005; Barabási and Pósfai, 2016) and to analyse the change of some concrete propagation mechanisms of particular industries in the context of threshold modification.

Also, another further development of the study could be to examine the clustering structure of the industries in the production network (Fagiolo, Reyes, and Schiavo, 2008; Fagiolo, Squartini, and Garlaschelli, 2013; Bartesaghi, Clemente, and Grassi, 2022; Bartesaghi, Clemente, and Grassi, 2023), compare it to the sectoral categories, and study how this is distorted through the threshold change.

Chapter 3

Explaining growth: the production network-based prediction of industrial growth

3.1 Industrial growth in the context of the production network

The question of industrial growth has been widely studied in the economics literature. Still, recent research has primarily focused on the characteristics of individual industries rather than on the complexity of the economy in which they are embedded. However, industries do not diffuse and develop in isolation. Those conventional econometric tools that do not consider the evolving and complex nature of the economy's interconnections might be insufficient to investigate growth.

This is an encouragement for the research community to understand better how economic sectors interconnect. The latest studies try to address this issue by merging standard economic theory with new concepts from complex systems science in order to obtain a different point of view. This novel perspective aims to extend beyond standard economics (Arthur, 2014).

Networks provide a good starting point for mapping complex systems (Schweitzer, Fagiolo, Sornette, Vega-Redondo, Vespignani, et al., 2009; Schweitzer, Fagiolo, Sornette, Vega-Redondo, and White, 2009). The economy can be thought of as a collection of highly interacting networks (Farmer et al., 2012).

Some of the critical networks needed to understand the economy in these terms map international trade (Fagiolo, Reyes, and Schiavo, 2009; Fagiolo, Reyes, and Schiavo, 2010; Reyes, Schiavo, and Fagiolo, 2010), aim to recognise clusters, dependencies (Snyder and Kick, 1979; Nemeth and Smith, 1985; Guo-hong et al., 2010) and structural autonomy (Sacks, Ventresca, and Uzzi, 2001) in global trade networks, assess the complexity of an economy or product in export-import networks (C. A. Hidalgo et al., 2007; César A. Hidalgo and Hausmann, 2009; Hausmann et al., 2013), investigate the increasing effect of global networks for technological improvements (McNerney et al., 2022), examine the economic crises through the lenses of global networks (Kali and Reyes, 2010; Lee et al., 2011) and even discover the co-authorship of economic literature (Goyal, Leij, and Moraga-González, 2006). From the business point of view, there is the rise of supply network science with firms as agents (T. Y. Choi, Dooley, and Rungtusanatham, 2001; Pathak et al., 2007; Brintrup and Ledwoch, 2018), determining knowledge networks in firms (Pinch et al., 2003; Giuliani, 2007) and credit chains to analyse bankruptcy propagation (Battiston et al., 2007).

Most of this research aims to understand the economy on a global scale using international trade data, mainly in the light/context of globalisation. Very little attention is paid to national economies. Although, since Leontief (W. W. Leontief, 1936; W. Leontief, 1986), more and more diverse data has become available to integrate the production network framework in the study of industry growth at the national level too. The national economy is a series of sectors (rather than a set of countries) interconnected through exchanging goods or economic transactions. Thus, the economy can be viewed as a network consisting of nodes, represented by sectors, and directed links, represented by monetary transactions between sectors, and this framework can be built from existing economic input–output models (R. E. Miller and Blair, 2009).

Fisher, 2006 was the first to show the feasibility of the national economic input-output matrix as a directed weighted network, focusing on identifying the "central sectors" in the US economy. Xu, Allenby, and Crittenden, 2011 investigated the US production network from the clustering side and wanted to understand how diversified interconnectedness affects the economy's resilience. Vasco M Carvalho, 2008 and Acemoglu et al., 2012 also argue that the structure of the production network plays a crucial role in determining the aggregate behaviour of the system, while also Vasco M. Carvalho, 2014 examines whether this network perspective can shed new light on the comovement of industries and business cycle fluctuations.

Even though some research has been carried out on production networks in national economies, no studies have been found which investigate specifically whether the location of an industry in this national production network topology is able to explain the macroeconomic dynamics of its growth and development. This research question has been asked on a global scale firstly by Kali and Reyes, 2007, and researchers tended to examine the world input-output data rather than a national one. Therefore, this study is pioneering in this sense.

Besides this, it significantly contributes to research on industrial growth as I move further from the traditional approach. While studies to date realised that complexity and interconnections are a must in growth research, they still used the network framework only as an additional ingredient/factor to existing economic models. However, such methods are unsatisfactory as I agree with Farmer et al., 2012 that *the only way to make a major advance in economic modelling is to explore entirely new approaches rather than make incremental modifications to existing models.*

This chapter fits within the objectives of the current research direction, and it goes even further by aiming to discover to what extent we can determine/predict the growth of an industry purely from its place in the production cycle and how effective the production network model alone is in explaining industrial growth. To date, production network characteristics leading to one industry's growth and the other's decline remain speculative. This study in this chapter set out to investigate these issues.

I use network-based metrics that consider not only the magnitude of trade but also the industry's influence on the national economy. These measures encompass the structure and functioning of the production network and can provide a relevant approach to industrial growth alongside current measures based on trade volumes.

Using these network characteristics, I find that the position of a sector in the production network considerably impacts economic growth. Therefore, the suggestion is that the network approach to national economic trade has the potential for beneficial applications in national finance, policy and development.

3.2 Empirical strategy to analyse industrial growth using networks

3.2.1 Data source

I continue to use the detail level United States Input-Output Accounts Data (*Input-Output Accounts Data — U.S. Bureau of Economic Analysis* 2021), but now I slice the network to the manufacturing industries.

The Bureau of Economic Analysis publishes the Industry Economic Accounts generally at three levels of detail: sector (21 industry groups), summary (71 industry groups), and detail (405 industry groups). Estimates at the detail level are produced roughly every five years, with the last two releases being from 2007 and 2012. While there are Historical Benchmark Input-Output Tables from 1947 until 2002, these reflect industry definitions that vary across years, and BEA advises that they should not be used as a time series. BEA uses its Industry Codes for the three levels of detail, but it also defines how these relate to the 2012 North American Industry Classification System (NAICS) code structure (*North American Industry Classification System* 2021).

For this chapter, I use the data on manufacturing industries from the last two releases, the 2007 and 2012 Total Requirements table - Industry by Industry, including 156 industries from the Durable and Nondurable goods sector and their inter-industry purchases.

To quantify industrial growth, I use the data from the National Bureau of Economic Research (NBER) and the US Census Bureau's Center for Economic Studies (CES) Manufacturing Industry Database (Becker, Gray, and Marvakov, 2021). The NBER-CES issues annual industry-level data from 1958-2018 on output, employment, payroll, investment, capital stocks, and various industry-specific price indexes.

3.2.2 Constructing the manufacturing part of the production network

The BEA Total Requirements table has a network representation in which an element w_{ij} represents the nominal amount of goods i used as input by sector j, with i, $j = 1, ..., N$, where N is the number of sectors.

For a two-industry example of the network built from the Total Requirements table, see Figure 2.1. The network built from this database is a directed weighted graph. The vertices are the industries (i and j), and the directed connections are the monetary transactions, the nominal flow of goods between sectors. The weight of each link is the economic value, representing how much an industry supports the development of the other (w_{ij} and w_{ji}). For example, industry i requires w_{ij} dollar input from industry j to produce one dollar output to final users (typically $w_{ij} \neq w_{ji}$).

The input-output models also indicate if the output of an industry is required as input to the same industry (w_{ii}). For example, the Oil and gas extraction industry could produce the oil and gas to power its own equipment, or the computer design and manufacturing industry can produce computers that are used to plan the next generation of computers (*Economic Input-Output Life Cycle Assessment. Carnegie Mellon University Green Design Institute* 2008). I excluded these self-sector transactions that indicate self-loops in the production network, as these would make the network more complex and might be misleading, too.

3.2.3 Monetary transaction threshold ζ

After constructing the production network from the Input-Output Accounts Data, I define three different edge cut-off thresholds ζ: 0.0001, 0.001 and 0.01.

The threshold ζ is associated with the value of the links, representing inter-industry trade in monetary expression. For example, when the threshold value ζ is 0.01, I only consider those connections that weigh at least 0.01. In other words, only those monetary transactions become edges in the resulting production network where at least 0.01 dollars is needed from one industry to produce one dollar output in another. I also analysed the no-cut-off production network, which is almost a fully connected graph. At this point, I incorporate all monetary transactions, even the tiny ones.

At all thresholds ζ, the production network includes all manufacturing industries initially present in the input-output accounts. For the last cut-off, 0.01 is chosen because that is the last value where all industries are included as nodes (all 156 manufacturing industries). I would have lost nodes from the production network at a higher threshold value. Table 3.1 shows the percentage of edges present in the production network according to threshold ζ.

Table 3.1: The remaining number and percentage of the original edge number in the production network

	Production network (2007)		Production network (2012)	
Threshold ζ	Edge number	%	Edge number	%
0.0	24 145	100%	24 118	100%
0.0001	14 320	59%	13 799	57%
0.001	4 955	21%	4 651	19%
0.01	1 312	5%	1 237	5%

I work with several thresholds in order to cover the threshold sensitivity of the topology to differing trade magnitudes Chapter 2). Much of the current production network literature uses only one threshold, typically less than 5% of the original transaction (Vasco M. Carvalho, 2014).Others, while still using a few cut-off values, disregard the weights of the remaining links after imposing the threshold. For example, Kali and Reyes, 2007 using a global input-output database accounts for a link to be present if the value is higher than x% a country's total exports. In other words, he assigns the value 1 to the link if it exceeds the threshold and 0 if it does not. In this chapter, after constructing the production network with the threshold value, I still consider the values of the remaining monetary transactions in the analysis.

3.2.4 Description and analysis of network metrics

One of network science's most significant current discussions is how to approach and analyse networks with specific characteristics, such as multilayer networks, dynamic networks, or special cases of weighted directed networks, like this production network. Most research is centralised around simple undirected networks. However, some networks are more complex than this by default. These networks cannot be simplified and downgraded to simple networks because we need to pay attention to the unique properties to maintain some fundamental factors that determine the basis of the map provided by the network.

Such a complex network is the production network too. Two main features characterise the production networks. Firstly, monetary transactions are one-way transactions. It matters pretty much who buys from whom, and the relationship might be different the other way around. Consequently, the connections between industries in the production network are directed. Secondly, monetary transactions have a value, and one transaction differs from another in the magnitude of the values. While there might be a reciprocal transaction between two industries, it might differ in terms of values. Therefore, industry connections also carry the property of value called edge weights.

Considering this, I list and calculate the metrics commonly considered to be the most representative of a network's inherent topological structure and information flow along its edges. The reason behind this is to cover as accurately as possible every dimension of the industry's position in the production network.

As discussed above, in most networks, there is no direction of the connections. A classic example of this is social connections. If someone is a friend of another person on Facebook, the relationship is the same the other way around. However, in directed networks, we distinguish the links based on their directions, such as in a Twitter network, one might follow the other one, but that does not mean that the other one follows back. Network science suggests that centrally located nodes in networks have essential roles. Defining the centrally located nodes in undirected networks is relatively straightforward, while in directed networks is much more complicated. In social networks, opinion leaders are followed by many others; hence they are the central nodes. Most networks work like this. Therefore, most centrality metrics in network research are based on in-degree (the number of edges pointing to the node).

The production network is quite a particular network in this sense. One might ask if those industries are the central ones that need the most resources or the ones that pump the most supply into the network. The critical sectors that dominate the economic activity, the so-called "commanding heights" (Yergin and Stanislaw, 2002), are all hidden if we only consider the in-degree-based centrality metrics. That being the case, I also calculate the outward centrality metrics on the reversed production network.

On the other hand, in simple measures that are not weighted (in-degree, out-degree, degree, etc.), I do not consider the magnitude of the value of the monetary transactions (the weights of the links). These metrics only consider that a transaction (link) is present. In other words, all transactions have equal value, signalling the presence of a link equal to one. Because the original production network is almost a fully connected graph, these location measures tend to be less relevant and have almost the same value for every industry/node. The only distinguishing factor, in this case, the magnitude of the monetary transaction (the weights of the links), is not considered. This scenario happens only when I do not impose a threshold. When using a cut-off (for 0.0001, 0.001 and 0.01), I also consider the weight indirectly as I cut out the small transaction from the production network.

Weight means distance in some measures, while in others means connection strength. In the production network, the weight is the monetary transaction amount. Hence, the weight is directly proportional to the connection's importance, "strength". I assume that the more considerable the amount of the transaction, the more essential the supplying link is. Therefore, at those metrics where weight means distance (betweenness and closeness centralities), I change the weight to 1/weight.

The starting level of network analysis is the node degree (Jennings, 1937; Barabási and Pósfai, 2016). We distinguish three types of degrees in a directed network. *In-degree* measures the number of direct suppliers an industry has, the number of inward connections. In other words, the number of sectors an industry needs directly as a resource. Its minimum value is 0 if the industry does not need resources, and its maximum value is *n-1*, where n means the number of industries in the production network. At the maximum value, it needs the resource of all other industries directly. For this network, n is 156 industries, so the maximum value is 155. *Out-degree* is the opposite of in-degree. It is the number of sectors an industry directly supplies, the number of outward connections. Its minimum value is also 0, which happens if the industry does not supply any other industry. Its maximum value is also 155. In this case, it supplies all the other industries directly. *Degree* is the aggregate of these two, the number of sectors an industry needs as a resource and supplies directly. It is the number of inward and outward connections. Its minimum value is 0 if the industry does not need any and does not supply any resources; ergo, it is isolated. Its maximum value is *2(n-1)*, where n is the number of industries in the production network. An industry with the maximum value supplies and needs the resources of all other industries directly. For this production network, the maximum value is 310.

The other dimension of degrees is when we also consider the monetary transactions values. In this case, *in-degree weighted* is the amount of resource one industry needs directly from other industries to produce one dollar output for the final consumer. It is the sum of inward connection weights. In contrast, *out-degree weighted* is the amount of resource one industry supplies directly to other industries, the sum of outward connection weights. *Degree weighted* is still the aggregate of these two, the amount of resource an industry needs directly from other industries to produce one dollar output for the final consumer and the amount it supplies directly to other industries, the sum of inward and outward connection weights.

Moving on to more complex measures, *betweenness centrality* is one of them. It presumes that the key to a node's success in a network lies in its role as a mediator between two groups; in other words, it is located among many other nodes (Freeman, 1977; Easley and Kleinberg, 2010). If the shortest path from one node to the other leads through a third, the node from the middle can be determinative. Precisely for this production network, it measures how often an industry plays the role of a mediator in a supply chain and how often the route through this industry is the shortest way to access other industries. Its minimum value is 0 if the industry does not occur on any shortest path, and its maximum value is 1 if it is present on all. *Betweenness weighted centrality* presumes the same assumption but also considers the magnitude of monetary transactions. As previously stated, I assume that the higher the transaction value, the more essential the connection is; hence the distance is smaller, and the two industries tend to be "closer". Its minimum value is 0 if the industry does not occur on any shortest path, and its maximum value is 1 if it is present on all.

Another complex measure, *closeness centrality*, assumes that a node is in a central position when others can easily and quickly reach it (Bavelas, 1950; Sabidussi, 1966). It quantifies the number of steps by which all other network nodes can reach the chosen node, compared to the scenario where all nodes are just one step away from the node in question. The maximum closeness value is 1, meaning all the other nodes can reach the chosen node in just one step. Hence denotes for this production network that all industries are direct suppliers to the industry in inquiry. All smaller closeness values calculate the distance by which all industries could reach the chosen one, compared to directly supplying it. Analysing the out-going aspect of this measure is the *outward closeness centrality*. It assumes the outward dimension of closeness centrality that a node is in a central position when it can easily and quickly reach others. Outward closeness is the number of steps by which the node can get to all other network nodes, compared to the scenario in which all nodes are reachable in one step. Here the maximum value is also 1, meaning the analysed industry supplies all other sectors directly. Smaller outward closeness values compute the distance the chosen industry could reach the others, compared to directly supplying them.

When considering the weights, *closeness weighted centrality* assumes that a node is in a central position when others can easily and quickly reach it. However, it also takes into account the magnitude of monetary transactions. I make the same assumption as with the betweenness weighted measure about the essentiality of the connections. From the outgoing perspective, *outward closeness weighted centrality* assumes the outward dimension of closeness weighted that a node is in a central position when it can easily and quickly reach others. The measure also considers the magnitude of monetary transactions. It is the sum of the edge weights the node can get to all other network nodes, compared to the scenario in which all nodes are reachable in one step with a weight equal to the maximum, 1. Outward closeness weighted values calculate the distance the chosen industry could reach the others, compared to directly supplying them with the maximum amount.

Eigenvector centrality takes the origin of a node's connections into account: where they come from, how popular the nodes are and where they are located in the network (Newman, 2018). It is based on the assumption that not only the number of neighbours matters but also how influential they are. In the production network, the specific traits of trading industries might increase or moderate an industry's influence on others and the whole production network. Consequently, industries depend on one another indirectly as well as directly. The simple eigenvector centrality is also an in-degree-based metric. For the production network, that means that centrally located industries need the most diverse types of resources and are supplied by those sectors that still need diverse resources. This does not give us the most accurate picture, as intuitively, we would think centrally located industries are not those that are the most "dependent". Therefore I compute out-edges perspective too, named *reversed eigenvector centrality*, which gives us the centrally located industries that supply the most resource to the production network and mainly supply to those still essential suppliers in the production circle. *Eigenvector weighted centrality* is a very similar indicator to the simple eigenvector. However, it also considers the size of the monetary transactions. Centrally located industries are those that need the most resources in terms of value (not diversity) and are also supplied by those in need of a significant amount of resources. We are also interested in central position industries from the other interpretation, so I compute *reversed eigenvector weighted centrality* that evaluates the centrally located industries that supply to those still essential suppliers in the production network with the highest amount of monetary transactions.

PageRank centrality is perhaps one of the more famous centrality measures, also a subclass of eigenvalue-based measures initially designed to rank websites in the search engine through an algorithm (Page et al., 1999; Langville and Meyer, 2004). The reason behind the PageRank is very similar in that a node is systemically crucial if its neighbours are essential and/or the neighbours of the neighbours are important. We then compute *reversed PageRank centrality*, *PageRank weighted centrality*, and *reversed PageRank weighted centrality* with the same logic used at eigenvector centralities.

3.2.5 Measuring industrial growth

I computed four different industry-level (real) value-added growth rates from the NBER-CES Manufacturing Industry Database for every industry in 2007 and 2012. I use the change in the VADD (total value added in one million dollars) variable as a proxy for industrial growth and expansion.

Growth rate R_1 is the industry's value-added of the year in question directly compared to the value 5 years before (Eq. 3.1a). Specifically presented for the analysed two years in Eq. 3.1b.

$$R_{1_{(t)}} = \frac{V_t}{V_{t-5}} - 1 \tag{3.1a}$$

$$R_{1_{(2007)}} = \frac{V_{2007}}{V_{2002}} - 1 \qquad R_{1_{(2012)}} = \frac{V_{2012}}{V_{2007}} - 1 \tag{3.1b}$$

The other three rates are due to reduce the effect of outlier years. Growth rate R_2 answers the question of how different the year is and whether it follows the trend. It compares the year in the inquiry industry's value-added to its average value for the last

5 years (Eq. 3.2a), presented for the specific years analysed in Eq. 3.2b.

$$R_{2_{(t)}} = \frac{V_t}{\overline{X}} - 1 \qquad \overline{X} = \frac{1}{5} \sum_{i=t-5}^{t-1} V_i \tag{3.2a}$$

$$R_{2_{(2007)}} = \frac{V_{2007}}{\frac{1}{5}\sum_{i=2002}^{2006} V_i} - 1 \qquad R_{2_{(2012)}} = \frac{V_{2012}}{\frac{1}{5}\sum_{i=2007}^{2011} V_i} - 1 \tag{3.2b}$$

Growth rate R_3 is the industry's average annual growth rate in the last 5 years, including the particular year (Eq. 3.3a). For the specific years 2007 and 2012 in Eq. 3.3b.

$$R_{3_{(t)}} = \frac{1}{5} \sum_{i=t-4}^{t} \frac{V_i}{V_{i-1}} - 1 \tag{3.3a}$$

$$R_{3_{(2007)}} = \frac{1}{5} \sum_{i=2003}^{2007} \frac{V_i}{V_{i-1}} - 1 \qquad R_{3_{(2012)}} = \frac{1}{5} \sum_{i=2008}^{2012} \frac{V_i}{V_{i-1}} - 1 \tag{3.3b}$$

Growth rate R_4 is the industry's average annual growth rate in 10 years when the year at issue is in the middle of that 10 years (Eq. 3.4a). For the specific years 2007 and 2012 in Eq. 3.4b.

$$R_{4_{(t)}} = \frac{1}{10} \sum_{i=t-4}^{t+5} \frac{V_i}{V_{i-1}} - 1 \tag{3.4a}$$

$$R_{4_{(2007)}} = \frac{1}{10} \sum_{i=2003}^{2012} \frac{V_i}{V_{i-1}} - 1 \qquad R_{4_{(2012)}} = \frac{1}{10} \sum_{i=2008}^{2017} \frac{V_i}{V_{i-1}} - 1 \tag{3.4b}$$

3.2.6 The network-based growth model

This chapter's analysis's main interest is to investigate to what extent can fundamental network metrics be useful in predicting industry growth. If a solid statistical relationship exists, this could be used to signal early warning about declining or failing industries.

The dataset used has the structure of a balanced panel consisting of 156 observations (N) and two years (T). In order to identify the best model, I produce 16 separate linear regression models with the four different edge cut-off thresholds ζ and the four distinct growth rates. After I chose the best-performing model with the specific threshold ζ and growth rate, I zoomed in and examined the parameter coefficients.

The regression described in this chapter is inspired by the economic growth model initially studied in Kali and Reyes, 2007, but it only follows that model partially. Unlike other researchers, I only use network topology metrics in this model as independent variables.

The industrial growth regression is specified as follows:

$$\begin{aligned} Growth_{it} = \alpha_0 &+ \beta_1 InDeg_{it} + \beta_2 OutDeg_{it} + \beta_3 Deg_{it} + \beta_4 InDegW_{it} + \beta_5 OutDegW_{it} \\ &+ \beta_6 DegW_{it} + \beta_7 Bet_{it} + \beta_8 BetW_{it} + \beta_9 Clos_{it} + \beta_{10} OutClos_{it} + \beta_{11} ClosW_{it} \\ &+ \beta_{12} OutClosW_{it} + \beta_{13} Eig_{it} + \beta_{14} OutEig_{it} + \beta_{15} EigW_{it} + \beta_{16} OutEigW_{it} \\ &+ \beta_{17} Pg_{it} + \beta_{18} OutPg_{it} + \beta_{19} PgW_{it} + \beta_{20} OutPgW_{it} + \beta_{21} Year + \varepsilon_{it} \end{aligned} \tag{3.5}$$

Where the variables are described in Table 3.2. The dependent variable, $Growth_{it}$, takes the values of R_1, R_2, R_3 and R_4 one by one, and I compute different models with each growth rate to discover the best rate type. All *Rate* values are calculated based on the NBER-CES Manufacturing Industry Database.

The independent network variables enter the model on the right-hand side of the regression, also explained in Table 3.2. These input variables are all synthetic variables constructed by calculating the fundamental network metrics of the underlying production network based on the BEA US Input-Output Accounts.

We do this sequence four times, as I have four differently defined production networks at four threshold ζ values: 0, 0.0001, 0.001 and 0.01. Consequently, the independent network variables change at every threshold ζ. In the end, I choose the best-performing linear regression model.

The choice of linear regression as the primary statistical approach in the analysis was driven by the intention to effectively highlight the argument and perspective of the new production framework. While acknowledging the theoretical emphasis on nonlinear relationships within the network and the network metrics highlighting these, linear regression was chosen as a pragmatic method to quantify the relationships and their impact on growth. Consideration was given to more advanced statistical approaches, such as exponential random graph models, which could potentially offer a more nuanced understanding by incorporating economic variables alongside network topologies. However, the decision to stick with linear regression was influenced by the desire to initially explore and understand the different coefficients for various variables and assess their individual contributions to growth. Despite recognising the potential for more sophisticated models to further enhance the analysis, linear regression was deemed sufficient for making a compelling argument regarding the significance of the production network framework. This choice was also pragmatic, considering the complexity of integrating network variables into the model and the overarching goal of presenting a clear and coherent argument within the given context. While alternative approaches were considered, including Lasso, Ridge regression and others, the decision to focus on linear regression was justified by its ability to effectively support the main argument of the research while acknowledging the potential for future refinement and sophistication in subsequent studies.

Table 3.2: Variable definitons

Name	Definition	Variable
Growth rate R_1	The industry i value-added of the year t compared to the value 5 years before.	$Growth_{it}$
Growth rate R_2	The industry i value-added of the year t compared its average value for the last 5 years.	$Growth_{it}$
Growth rate R_3	The industry i average annual growth rate in the last 5 years, including year t.	$Growth_{it}$
Growth rate R_4	The industry i average annual growth rate in 10 years, when the year t is in the middle of that 10 years.	$Growth_{it}$
In-degree	The number of direct suppliers industry i has at time t. (The number of inward connections.)	$InDeg_{it}$
Out-degree	The number of direct suppliers industry i has at time t. (The number of outward connections.)	$OutDeg_{it}$
Degree	The number of industry i's direct suppliers and to whom it directly supplies at time t. (The number of inward + outward connections.)	Deg_{it}
In-degree (w)	The amount of resource industry i directly needs at time t to produce one dollar output. (Sum of inward link weights.)	$InDegW_{it}$
Out-degree (w)	The amount of resource industry i supplies to the production network at time t. (Sum of outward connection weights.)	$OutDegW_{it}$
Degree (w)	The amount of resource industry i needs and supplies to the production network at time t. (Sum of inward + outward link weights.)	$DegW_{it}$
Betweenness	The centrality measure of how often industry i is on the shortest path between two other industries at time t.	Bet_{it}
Betweenness (w)	The centrality measure of how often industry i is on the shortest path between two other sectors at time t, considering the weights of the links. The bigger the weight, the closer the two industries are.	$BetW_{it}$
Closeness	The centrality measure of how quickly all the other industries can reach industry i at time t in the production network, compared to reaching it in one step.	$Clos_{it}$
Outward closeness	The centrality measure of how quickly industry i can reach all the other industries at the time t in the production network, compared to reaching it in one step.	$OutClos_{it}$
Closeness (w)	The centrality measure of how quickly all the other industries can reach industry i at time t in the production network, compared to reaching it in one step and considering the weights of the links. The bigger the weight, the closer the two industries are.	$ClosW_{it}$
Outward closeness (w)	The centrality measure of how quickly industry i can reach all the other industries at the time t in the production network, compared to reaching it in one step and considering the weights of the links. The bigger the weight, the closer the two industries are.	$OutClosW_{it}$
Eigenvector	A centrality measure of an industry i's popularity at time t in terms of the centrality of the industries that supply it.	Eig_{it}
Reversed eigenvector	A centrality measure of an industry i's popularity at time t in terms of the centrality of the industries to whom it supplies.	$OutEig_{it}$
Eigenvector (w)	A centrality measure of an industry i's popularity at time t in terms of the centrality of the industries that supply it, considering the magnitude of the supply.	$EigW_{it}$
Reversed eigenvector (w)	A centrality measure of an industry i's popularity at time t in terms of the centrality of the industries to whom it supplies, considering the magnitude of the supply.	$OutEigW_{it}$
PageRank	A centrality measure of an industry i's popularity at time t in terms of the centrality of the industries that supply it.	Pg_{it}
Reversed PageRank	A centrality measure of an industry i's popularity at time t in terms of the centrality of the industries to whom it supplies.	$OutPg_{it}$
PageRank (w)	A centrality measure of an industry i's popularity at time t in terms of the centrality of the industries that supply it, considering the magnitude of the supply.	PgW_{it}
Reversed PageRank (w)	A centrality measure of an industry i's popularity at time t in terms of the centrality of the industries to whom it supplies, considering the magnitude of the supply.	$OutPgW_{it}$

3.3 Results

3.3.1 The best-performing network growth model

The first question aims to choose the best-performing linear regression model. Figure 3.1 compares the 16 model accuracy scores. By the model score, I mean the R-square of the linear regression.

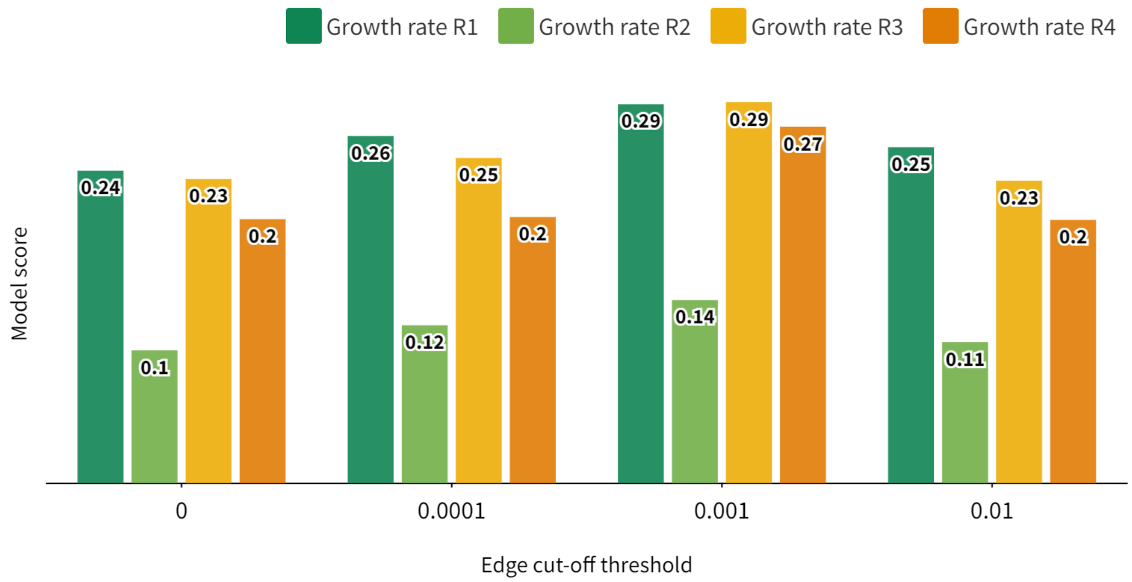


Figure 3.1: Model accuracy scores based on threshold and rate type.

From the chart, it can be seen that all values are above zero. This indicates that even if I choose a different combination than the best performance, only topological metrics at all rates and cut-offs can explain a percentage of industrial growth.

It is apparent that the best result is at the 0.001 edge cut-off threshold value. Both growth rates R_1 and R_3 give us almost 30%. While growth rate R_1 is the industry value-added of the year in question compared directly to the value 5 years before, R_3 is the average annual growth rate for the last 5 years. I choose to stick with the growth rate R_3 because R_1 does not filter out outlier years.

This result suggests that considering only the network metrics, we can explain 30% of industry growth measured by (real) value-added change.

Consequently, location in the production network matters, and it is crucial. It provides essential insights into the success of industries. If we create a new industry, 30% of its success or even failure would be explained by the other's success, to which it links, on the key interfaces of the industry. Another definition of the result can be that if we look at all industries in the production network, we can then choose to enter an industry as an investor where the interfaces/points of contact are successful because this is a significant part of whether the industry will succeed.

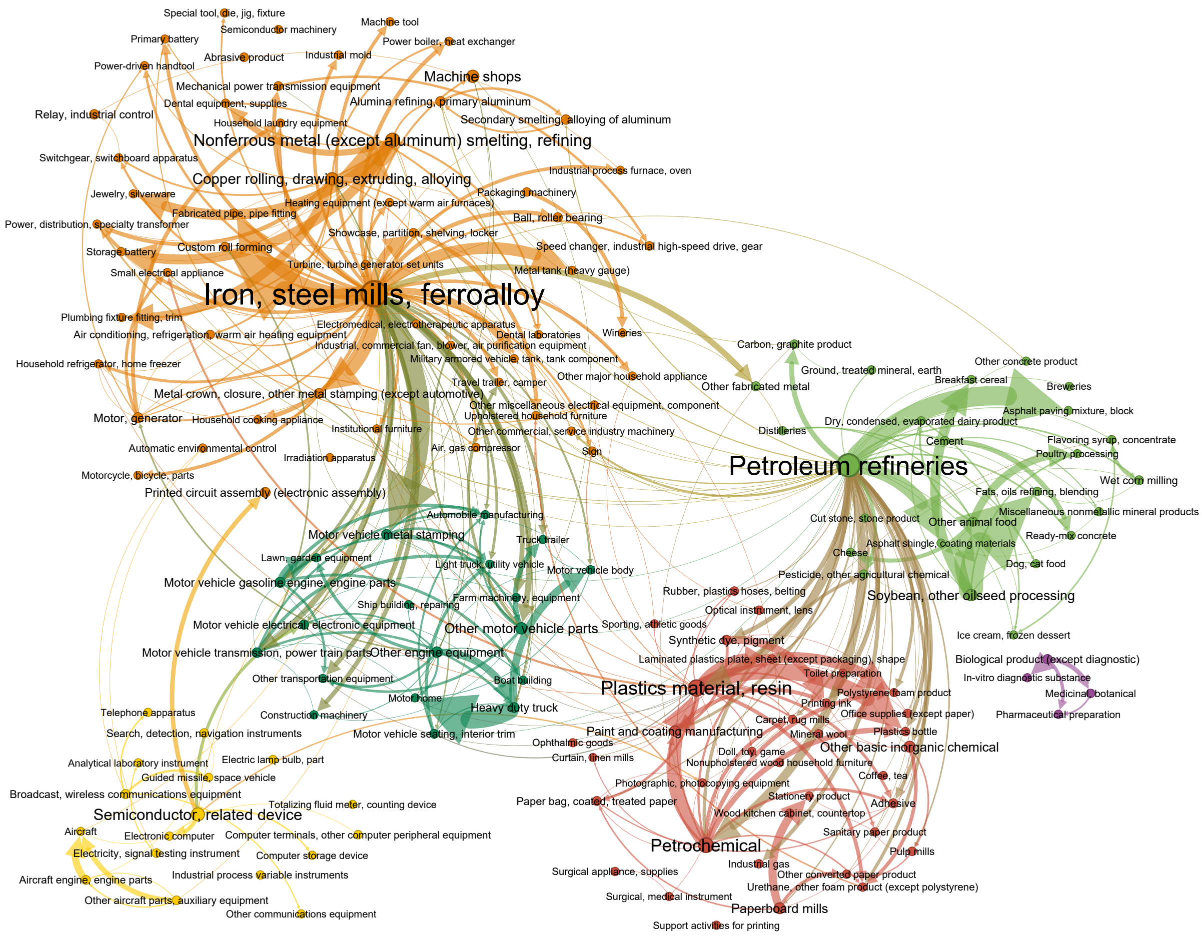


Figure 3.2: Weight and modularity in the production network.

This investigation's model performances are higher than those of other studies. While Kali and Reyes, 2007, studying the 1990s global network, also got 0.296 at one of the models, he included other economic variables too, such as GDP, human and physical capital, which intuitively are better explanatory variables than the topological features. Prado, 2017, analysing the world input-output data between 2006 and 2009, got an R-squared of 0.766 only with the economic variables, such as GDP, industry openness, labour productivity, country openness and inflation. When the study included a few network explanatory variables (out-degree, eigenvalue, betweenness, network density), it could only improve the result with 0.002, to 0.768 on the total sample. He found a difference between the effects depending upon the analysis period, but still reaching only 0.09 explanatory power with the network metrics besides the economic variables. Therefore, the results represent a step-change in formulating the fundamental network's explanatory power for predicting industry growth.

The decision to exclusively utilise network variables as explanatory variables in the regression analysis stemmed from a deliberate intention to quantify the contribution of the production network framework to understanding industry dynamics. The primary goal of the project was to assess how the framework and its associated relationships could provide valuable insights into economic growth and dynamics. By focusing solely on network metrics, the aim was to emphasise the significance of these relationships and their role in shaping industrial dynamics. However, it's acknowledged that this approach may have limitations, particularly in terms of explanatory power, when compared to models incorporating both economic and network variables. Reflecting critically on this choice, it's recognised that including economic variables alongside network metrics would likely enhance the model's explanatory power by capturing additional dimensions of industry dynamics. Economic variables are inherently powerful in explaining various aspects of industrial behaviour and performance. Therefore, while the decision to focus solely on network variables aligns with the project's specific objectives, it's acknowledged that incorporating economic variables would provide a more comprehensive understanding of industry dynamics. However, some variables that Kali and Reyes, 2007 and Prado, 2017 used would not make sense in a national context, only on international trade data. Ultimately, the project's emphasis on quantifying the contribution of the production network framework serves as a valuable tool for highlighting its significance and offering a new perspective on industrial dynamics.

3.3.2 The production network

The US industrial network drawn in Figure 3.2 is dominated by fully connected nodes from the out-degree perspective, defined as industries that supply 100 per cent of other manufacturing industries directly: Petroleum refineries, Petrochemical, Plastics material and resin, Iron and steel mills and ferroalloy, Machine shops and Semiconductor and related device. These support all the other 155 industries in the production network.

The top industries based on other network centrality metrics are presented in Table 3.3.

Table 3.3: Sample of the most connected industries based on all analysed network measures (2012)

Measure name	Industry with the highest value	Highest measure value
In-degree	Heavy duty truck manufacturing	52
Out-degree	Petroleum refineries	155
Degree	Other motor vehicle parts manufacturing	197
In-degree weighted	Motor home manufacturing	1.036
Out-degree weighted	Iron, steel mills and ferroalloy manufacturing	8.542
Degree weighted	Iron, steel mills and ferroalloy manufacturing	8.799
Betweenness	Construction machinery manufacturing	0.094
Betweenness weighted	Iron, steel mills and ferroalloy manufacturing	0.166
Closeness	Motor home manufacturing	0.442
Outward closeness	Petrochemical manufacturing	1.0
Closeness weighted	Motor home manufacturing	0.003
Outward closeness weighted	Petroleum refineries	0.032
Eigenvector	Motor home manufacturing	0.142
Reversed eigenvector	Plastics material and resin manufacturing	0.229
Eigenvector weighted	Motor home manufacturing	0.541
Reversed eigenvector weighted	Petroleum refineries	0.753
PageRank	Dog and cat food manufacturing	0.085
Reversed PageRank	Machine shops	0.034
PageRank weighted	Dog and cat food manufacturing	0.065
Reversed PageRank weighted	Petroleum refineries	0.186

Table 3.4: Descriptive statistic

Variable	Obs	Mean	Std. Dev.	Min	Max
Growth rate R_3	312	0.034	0.081	-0.241	0.667
In-degree	312	30.788	9.020	6	57
Out-degree	312	30.788	45.812	0	155
Degree	312	61.577	45.538	11	197
In-degree weighted	312	0.388	0.190	0.024	1.036
Out-degree weighted	312	0.388	1.053	0	8.646
Degree weighted	312	0.777	1.048	0.106	8.881
Betweenness	312	0.008	0.014	0	0.094
Betweenness weighted	312	0.018	0.033	0	0.168
Closeness	312	0.360	0.026	0.281	0.442
Outward closeness	312	0.436	0.264	0	1
Closeness weighted	312	0.002	0.000	0.001	0.003
Outward closeness weighted	312	0.004	0.004	0	0.035
Eigenvector	312	0.076	0.025	0.008	0.152
Reversed eigenvector	312	0.043	0.067	0.000	0.228
Eigenvector weighted	312	0.055	0.058	0.001	0.541
Reversed eigenvector weighted	312	0.019	0.078	0.000	0.794
PageRank	312	0.006	0.008	0.002	0.085
Reversed PageRank	312	0.006	0.009	0.001	0.034
PageRank weighted	312	0.006	0.008	0.002	0.065
Reversed PageRank weighted	312	0.006	0.021	0.001	0.199

3.3.3 The production network metrics explaining growth

The descriptive statistics for the independent variables used in Equation 3.5 are presented in Table 3.4. The network variable distributions are shown in Appendix A.1, and the correlation matrix of the explanatory variables is also presented in Appendix A.2.

Table 3.5: Regression for industrial growth using Eq. 3.5

Name	Coefficient
In-degree	-0.010***
Out-degree	0.005***
Degree	-0.005***
In-degree weighted	0.050**
Out-degree weighted	-0.034**
Degree weighted	0.015
Betweenness	-0.054
Betweenness weighted	0.262***
Closeness	-0.545*
Outward closeness	0.059*
Closeness weighted	-29.067
Outward closeness weighted	-6.063**
Eigenvector	6.109***
Reversed eigenvector	-0.257
Eigenvector weighted	-0.327**
Reversed eigenvector weighted	-0.472*
PageRank	0.230
Reversed PageRank	-2.187
PageRank weighted	0.656
Reversed PageRank weighted	4.588***
Constant	33.419***
R-sq	0.287
N	312

Significance level: *p<0.2 **p<0.05 ***p<0.01

The linear regression results are summarised in Table 3.5. The explanatory variable *in-degree* is significantly negative, meaning if someone creates a new industry that needs many resources and therefore depends on many things, that is not the best road to success.

The *degree* parameter is also significantly negative. If an industry has a lot of connections (whether supplying or supplier ones) with industries in general, that also has adverse effects.

Betweenness weighted being significantly positive suggests that it contributes to an industry's success if it is an essential mediator in most supply chains.

Based on *eigenvector*, that is also significantly positive; it is essential how much resource an industry needs and, more importantly, if this resource comes from fewer, more prominent players or smaller ones. It is more valuable if it comes from fewer significant actors. In this case, by significant actors, I mean industries in central positions requiring many resources. Even intuitively, if fewer industries support an industry, the change in its diffusion can be better and more stably predicted than if it depends on plenty of industries as suppliers.

On the other hand, based on *reversed PageRank weighted*, that is also significantly positive in the model; it matters what other industries an industry supports and by how much. Considering this metric, it is more beneficial for industrial growth to support other important/central industries with high value. So it is easier to predict the industry's "fate" if it is essential to the cycle, supporting other industries with high-value resources that also support others with high-value resources. In this case, its change/growth can be more accurately estimated by its position in the network. Not the density of connections that matters; having a lot of small ones, but it is more stable to supply a few crucial industries.

Closeness being negative and slightly less significant than those mentioned above suggests that creating an industry in the middle of the net, dependent on many others, wouldn't be the best strategy. In return, *outward closeness* having a positive sign indicates that it is good if the industry supports and is close to most other industries.

3.4 Discussion

This chapter discusses whether purely the network characteristics of an industry influence its growth. Before this study, it was difficult to make predictions about how much one industry's place alone in the production cycle influences its growth. Most research, until now, if used the newly developed production network dimension, added this framework to already existing traditional models and research.

Our results make a pioneering contribution to the rapidly expanding field of industrial growth modelling by focusing on analysing only the production network model and its effectiveness in explaining growth.

The outcome suggests that network characteristics can define industrial growth up to a certain level. Almost a third of the change in the industry is explained by its location in the supply chain topology alone. Of course, I can improve the model by considering the obvious economic metrics, for example, industry GDP, but the goal is limited to topology metrics only. Even though the current study is based on these metrics, the 29% reported here sheds new light on traditional models as this is a significant amount that cannot be overseen and essentially contributes to industrial growth investigation.

Another perhaps expected finding to emerge from this study is that it is more beneficial for an industry's growth to supply other central and essential industries with high-value monetary transactions. The amount of supplying relationships is less notable for the industry than to supply a few crucial industries.

It should be noted that it is reasonable to question whether the largest financial transaction might be the most essential one. The value is not necessarily directly proportional to essentiality. Despite this, I do not have a better economic metric to symbolise supplying necessity on the sectoral level. The same happens with real value-added change, which might not fully explain an industry shrink or diffusion. Also, this study, limited to two years, needs a broader perspective. Notwithstanding these shortcomings, the study's suggestions are valid in the current research environment.

While I conclude how accurate the production model is in general terms on a national economy, it would be a fruitful area for further work to analyse one or a few sectors based on this network model. This model and the information gained through it can be used to develop targeted interventions and policies aimed at different industries during a crisis or recession.

Chapter 4

Ego production network perspective: the case study of the storage battery industry

4.1 The ego production network

Industries, and thus companies in those industries, rely on the capabilities of other companies in other industries to provide goods and services. When companies choose to collaborate, they establish networks of interdependence between them, and hence inter-industry reliance networks, which have an impact on various aspects, such as the generation and dissemination of resources (Gulati, 1995; Schilling and Phelps, 2007).

Supply chain refers to the interrelated chain of organisations, individuals, activities, information, and resources involved in producing, distributing, and delivering goods or services from the point of origin to the final consumer. It encompasses various stages, including procurement of raw materials, manufacturing, transportation, storage, and retailing.

Recent studies have contended that the supply chain possesses characteristics of a complex adaptive system, emerging with multiple resource flows. A complex adaptive system refers to a dynamic and interconnected network composed of diverse components that interact and adapt in response to their environment (Holland, 1992; Gell-Mann, 1994; Lansing, 2003; Holland, 2006). These systems exhibit emergent behaviour, meaning that the collective interactions of the components lead to new patterns and properties that are not directly predictable from the behaviour of individual parts. Complex adaptive systems are characterised by self-organisation, non-linear relationships, feedback loops, and the ability to learn and evolve over time. Examples include ecosystems (Levin, 1998), the internet (Phister, 2010), infrastructure (Oughton et al., 2018), economies (Tesfatsion, 2003; Gintis, 2006), social networks (Benham-Hutchins and Clancy, 2010; Hasgall, 2013), and even the language (Ellis and Larsen-Freeman, 2009) and the human brain (Morowitz, 2018).

Therefore, it is recommended to employ analytical tools from network science and complex system analysis to examine and understand it (T. Y. Choi, Dooley, and Rungtusanatham, 2001; Pathak et al., 2007). While companies have the autonomy to select their own customers and primary suppliers, they generally lack control over the suppliers' choices regarding their own purchases. Researchers have proposed that due to this emergent behaviour, companies may find themselves sharing common suppliers, forming interconnected network structures rather than simple linear chains. This observation has led to the introduction of the term "supply network" within the field of supply chain literature (Brintrup and Ledwoch, 2018; Demirel, 2022).

Since then, researchers tended to focus more and more on examining supply networks with different aims. Some studied the idea that decisions made by companies regarding partnerships are impacted by the social context that emerges from prior collaborations and the concerns of strategic interdependence by analysing interactions between firms (Gulati, 1995). Others focused explicitly on using network analysis to examine the structural aspects of supply networks from both material flow and contractual ties of automotive industry companies (Kim et al., 2011).

Several studies have attempted to discover the issue of risks in these complex systems of supply networks, such as one analysis's outcome of a considerable association between global supply network structure and both risk diffusion and supply network health (Basole and Bellamy, 2014). Another introduced the term nexus supplier to the pool of already critical suppliers in terms of potential risk factors, too, as a key actor due to its network position and the resultant portfolio of inter-organisational ties (T. Yan et al., 2015).

Also, the topological analysis of supply networks has gained some attention. There have been investigations on the interfirm collaboration networks in the global electronics industry, proving that high-performing companies have a specific supply network shape and degree distribution (Basole, 2016). Other researchers related the supply network structure of the US public multiple industry companies to productive efficiency by resulting in some network metrics most closely associated with supply chain efficiency (Kao et al., 2017).

All this research trend towards expanding supply network science literature shows a critical need to approach and analyse supply chains with complex system perspectives and tools. The difference between examining supply relationships in chains or networks is clearly visible. A review paper on supply network science (Brintrup and Ledwoch, 2018) distinguishes between two types of research direction: macroscopic and microscopic. Macroscopic views investigate traits of the supply networks that are observable at large scales, such as resilience, robustness or holistic network indicators. At the same time, the microscopic point of view chooses to examine measures that characterise an individual node's position in the network. Brintrup and Ledwoch emphasise that there is a need to unify these various types of networks and views. This chapter addresses this gap and examines the storage battery network from both standpoints.

Research until now mainly focused on interfirm relations when examining supply networks. Hence, the term supply network refers to intercompany relationships. This makes perfect sense as supplier-buyer relationships are the most concrete and tangible between companies. However, if we want to discover the supply network of a specific industry, this presumes that we choose a company or several companies in the industry and map the links between them. This limits the study to the level of those firms and their specific supplier choices. This chapter aims to examine all sectoral relationships surrounding the storage battery manufacturing sector and have a general, comprehensive view of all industrial dependencies in the production of storage batteries. Hence, not to limit the project to a few companies in the industry. Consequently, the dimension of this research is a bit different, focusing on inter-industry relations from the perspective of production networks, taking into account the lessons presented in the first two chapters. The difference lies in the actors present in the network and the defining relations between them. Supply networks are constructed mostly between firms, and the relations might be very diverse, from conceptual to material flows. However, in this thesis, I focus on production networks where actors are the different economic sectors and inter-industry monetary transactions as relations between them. I use the term production network in this definition.

Once again, if we want to build the production network of an industry using inter-industry relations, the starting point to discover these relationships is the actual monetary transactions between industries. As explained in Chapter 2, these monetary transactions are quantified and normalised in national and international input-output accounts. Also, in Chapter 2, I describe thoroughly how to get from these raw data sources, specifically from the very detailed US accounts, the US national production network framework. The national production network framework is a map of all transactions occurring between all sectors in a national economy. It includes everything from the tiniest monetary transaction to the most significant one. To emphasise this even more, specifically, the US national production network contains 405 industries and 150 000 transactions between them. The basis of the analysis will be this national network; however, to zoom in only on a specific industry production network, I will introduce the term and concept of the ego network in this chapter.

An **ego network** consists of a central node, known as the ego, along with its immediate connections or ties to other nodes (first-degree connections) (Burt, 1980; Freeman, 1982; Borgatti and Halgin, 2011). Additionally, it can encompass the connections among the nodes directly linked to the ego.

Nowadays, the ego network perspective is most used in online social networks research focusing on information diffusion (Arnaboldi, Conti, La Gala, et al., 2016; Arnaboldi, Conti, Passarella, and Dunbar, 2017) and on investigating community structure and social circles (Leskovec and Mcauley, 2012; Arnaboldi, Conti, Passarella, and Pezzoni, 2012; A. Biswas and B. Biswas, 2015). However, the range of this field is extensive, even from scientific collaboration networks (Lu et al., 2021) to methodological papers (Everett and Borgatti, 2005; Hampton and W. Chen, 2021) on ego network study.

Ego network research has been helpful in various economic contexts, too, such as innovation diffusion and adaption according to network structure in firms (Carnovale and Yeniyurt, 2015; Kumar and Zaheer, 2019), stability of alliances, structural holes and power in market-level collaboration networks (Burt, 1995; Ahuja, 2000), and the role that ego networks have on new joint venture formations (Carnovale and Yeniyurt, 2014).

In this chapter, the ego network is an inter-industry transaction network, with the ego being the storage battery manufacturing industry. Once again, the phrase ego network is used with the meaning of a specific subset of the whole national inter-industry transaction network. It is a subset of connections and relationships that a specific industry, referred to as the "ego industry," maintains with other industries within a larger network. In this network structure, the ego industry is the focal point, and it is primarily connected to industries with which it directly interacts and exchanges resources. In other words, it is that industry's core production network.

Storage batteries play a crucial role as vital components for storing energy, making them an essential solution for addressing the challenges associated with energy transition. Consequently, the industry itself warrants particular attention, as elaborated upon in the next section of this chapter. Furthermore, due to various risks within the supply chain of storage batteries that could potentially lead to disruptions in production, it is crucial to acknowledge the growing body of literature highlighting the significance of studying the supply chain of storage batteries. This chapter aims to contribute to the advancement of research in this specific field from the inter-industry perspective.

Despite the extensive coverage of supply chain analysis in the field of economics, no existing studies have been discovered that approach the subject from a systematic inter-industry data-oriented perspective. Therefore, the main objective of this chapter is to fill this gap by examining and analysing the storage battery industry and its inter-industry supply chain through the lens of data, specifically using complex systems analysis tools like network science.

Another purpose of this chapter is to show the viability of the model developed above when focusing on a particular industry, such as the production network framework (Chapter 2). Consequently, this chapter adopts a quantitative approach to analysing the production network of the storage battery industry.

It explores the ego production network built from inter-industry monetary transactions, using the production network framework and highlighting the key integrator, allocator and mediator industries in storage battery production. The aim of this part is to see how data can provide new perspectives on how the storage battery industry connects with other industries in its production network and to examine what this new perspective (the ego production network framework) tells us about the storage battery industry's supply chain dependencies and the indispensable industries in it.

The contribution of this study can be summarized in two ways. Firstly, this chapter shows the viability of the production framework developed in Chapter 2 when analysing one specific industry and its production network. It can be read as a case study of how to use the system and network perspective on the economy and industry expansion explained in the above chapter only on one subset of the national production cycle, specifically on the surrounding supporting environment of one industry. Thus also highlighting the differences between analysing a general national production network and a specific ego network.

This chapter is organised into three main sections, as follows:

Firstly, I present in the *Energy storage, the storage battery industry and its supply chain* chapter the critical role of energy storage in the transition to renewables by providing the storage battery industry as a solution. After that, I summarise an overview of the storage battery industry with benefits in several other industries and the challenges faced in their supply chain, concluding the need for a detailed systematic data-based production network investigation.

Secondly and thirdly, I map the storage battery ego network and analyse the network independently. After presenting the data source and how the storage battery is included in the data, I indicate the methodology of constructing the ego production network from this data source. I analyse the storage battery ego network from two dimensions: first, I start with a so-called simple ego network, then move on to a more advanced one. The simple network is constrained only to the direct monetary transactions between the storage battery industry and any other supplier or receiver industry, while the advanced network also includes the transactions between the suppliers in which the storage battery industry is not present. I analyse these networks separately as each holds different kinds of information about the storage battery's production environment. Hence, the next section is the study of the simple ego network from the upstream and downstream view and comparing these two. At the advanced ego network part, I explore the network-level and the node-level metrics separately. I map through the lenses of production networks the indispensable relations and industries according to the data in the production cycle of the storage battery industry. The simple and advanced perspectives carry different approaches and conclusion types that I explain in detail in the section.

4.2 Energy storage, the storage battery industry and its supply chain

During the last decades, there have been growing concerns about climate change voiced by many leading energy researchers (MacKay, 2009; Armaroli, Balzani, and Serpone, 2013). To respond to skyrocketing global demand, energy production will have to double by 2050 (*IRENA — Energy transition outlook* 2023), and in order to fulfil this demand, every 15 seconds, we drain an Olympic-sized swimming pool full of oil. Knowing these facts, it's quite reasonable to ask how we can protect the planet yet keep the machine running.

Scientists repeatedly urge for putting the solution in motion: energy transition. According to the International Renewable Energy Agency (*IRENA — International Renewable Energy Agency* 2023), the sustainable energy transition is a pathway towards transforming the global energy sector from fossil-based to zero-carbon by the second half of this century. There is a great need for this transition because the largest source of greenhouse gas emissions from human activities is from burning fossil fuels for electricity, heat, and transportation. Besides technological changes switching from an economy based on energy stocks (fossil fuel deposits) to one based on energy flows (renewable energy production rates) requires a social paradigm shift, too (Sgouridis and Csala, 2014). Reconfiguring energy systems poses profound technological, social and economic questions and is one of the central policy challenges facing industrial countries (C. A. Miller, Iles, and Jones, 2013). Neither private markets nor governments seem likely to stimulate a transition on their own (Fri and Savitz, 2014). According to one view, energy transitions take an incredibly long time to occur. Another view argues that many transitions - at varying scales involving different things, including fuels, services, and end-user devices - have occurred quite quickly. Sovacool (Sovacool, 2016) analysed both sides and concluded that energy transitions are path dependent and cumulative. Studying path dependence isn't easy; it requires analysis of context. An approach that is systemic is needed. A reasonable one would be network science. Therefore, to address these questions, I use precisely these methods in my research.

The transition to renewables will demand more flexibility from the energy systems. These energy sources are driven by the availability of their corresponding natural resources rates, such as wind speed or insolation. These are intermittent – and their peak production times do not coincide with the peak consumption times of societal needs. Therefore, we need a buffer to get a two-match. Among several options for increasing this flexibility, energy storage is a promising one. But no such technology stands out simultaneously in all technical characteristics; therefore, selection should be driven on a case base analysis (Gallo et al., 2016). Vaclav Smil (Smil, 2010; Smil, 2018; Smil, 2019) claims that energy storage is the main issue in this field, and there is limited attention to this problem: "Give me mass-scale storage, and I don't worry at all. With my wind and photovoltaics, I can take care of everything. [But] we are nowhere close to it."

Consequently, energy storage is also frequently employed to mitigate the slight variations in energy production from both small and large power generation sources. Additionally, storage enhances dependability and reinforces the resilience of systems at both major and minor substation levels. However, energy storage is also widely utilized in transportation (Thackeray, Wolverton, and Isaacs, 2012), such as electric cars, trains, and bicycles. In the past, energy storage systems have been prohibitively costly and not financially feasible on a business-oriented magnitude. Nevertheless, significant advancements in energy storage technologies have resulted in cost reductions and enhanced technological implementations.

One of the most widely used and convenient forms of energy storage is batteries. Batteries can be classified based on various criteria, including their electrodes' chemistry, construction, and applications. They are primarily organised into primary and secondary batteries (Barak, 1980; Linden, 1984; Crompton, 2000). Primary batteries, also known as disposable batteries, are non-rechargeable and intended for single use. Once their energy is depleted, they cannot be recharged. Examples include alkaline batteries, zinc-carbon batteries, and lithium primary batteries. Secondary batteries, also called rechargeable batteries, can be recharged and used multiple times. They are commonly used in various applications and can be recharged by applying an external electrical current. Examples include lead-acid batteries, nickel-cadmium batteries, nickel-metal hydride batteries, and lithium-ion batteries. I call the latter, the secondary battery type, as the storage battery in this chapter.

Therefore, a **storage battery**, also known as an accumulator or rechargeable battery, is a type of electrical device designed to store and release electrical energy through reversible chemical reactions. It consists of one or more electrochemical cells that convert chemical energy into electrical energy during charging and vice versa during discharging (Borah et al., 2020). Unlike disposable primary batteries, which cannot be recharged, storage batteries can be repeatedly charged and discharged, making them suitable for a wide range of applications.

Storage batteries typically consist of two electrodes—a positive electrode (cathode) and a negative electrode (anode)—immersed in an electrolyte solution. An external power source applies a voltage across the battery terminals during charging, causing a chemical reaction at each electrode (Kiehne, 2003). This reaction results in the accumulation of electrical energy within the battery, which is stored in the form of chemical potential energy. When the battery is connected to an electrical load, the stored energy is released as a flow of electrons, creating an electric current that can power various devices (Kiehne, 2003).

Common types of storage batteries include lead-acid batteries, nickel-cadmium (NiCd) batteries, nickel-metal hydride (NiMH) batteries, and lithium-ion (Li-ion) batteries (Nishi, 2001; Franco, 2015). Each type has its own unique characteristics, such as energy density, self-discharge rate, cycle life, and environmental impact, which make them suitable for specific applications ranging from small portable electronics to electric vehicles and renewable energy storage systems (J.-K. Park, 2012).

The storage battery industry has become increasingly important in recent years due to its crucial role in the energy transition and the need for decarbonisation. The development and deployment of storage batteries have the potential to transform the way we generate, distribute, and consume energy and can offer a range of **benefits** and opportunities for various stakeholders (Wicki and Hansen, 2017).

One of the primary benefits of storage batteries is their ability to provide *grid stability and flexibility* (Popovich et al., 2021). As the share of variable renewable energy sources, such as wind and solar, increases in the power system, storage batteries can help balance supply and demand and maintain grid stability (Vartanian, 2010; Subburaj, Pushpakaran, and Bayne, 2015). Storage batteries can also help avoid the need for expensive upgrades to transmission and distribution infrastructure by providing localized energy storage and load management and also by providing backup power during emergencies and maintaining the reliability of the power supply (T. Chen et al., 2020; D. Choi et al., 2021).

Another essential benefit of storage batteries is their potential to *reduce energy costs and increase efficiency* (Duffner et al., 2020). Storage batteries can be used for peak shaving, load levelling, and frequency regulation, which can help avoid expensive peak demand charges and reduce the need for additional generation capacity (Asif and Singh, 2017; S. Ziegler, Juhyun Song, and E. Trancik, 2021). Storage batteries can also improve the efficiency of renewable energy systems by smoothing out fluctuations in energy output (Ruan, Walker, and Zhang, 2016; Turcheniuk et al., 2021).

They can also contribute to *environmental benefits* by reducing greenhouse gas emissions and promoting sustainable energy practices (Ruetschi, 1993; Pistoia, 2010). Storage batteries can help reduce the need for fossil fuel-based peaker plants and provide an alternative to diesel generators for backup power and can also help support the deployment of renewable energy sources by providing backup power and reducing curtailment (Shiau et al., 2009; Cusenza et al., 2019; B. Jeong et al., 2020).

Moreover, the storage battery industry can drive *innovation and economic growth* by creating new business opportunities and supporting technological advancements (Litjens, Worrell, and Sark, 2018; Popovich et al., 2021). The industry can also help create jobs and promote local economic development by attracting investment and supporting research and development (X. Li, Chalvatzis, and Stephanides, 2018; Ajanovic and Haas, 2019; Aqidawati et al., 2022).

There is also a growing importance of storage batteries in *transportation and mobility*, including the electrification of vehicles and the development of charging infrastructure (Shiau et al., 2009; Ajanovic and Haas, 2019; Popovich et al., 2021). Another potential for storage batteries is to play a key role in *emerging technologies*, such as energy storage for space exploration (Mercer et al., 2012; Bugga and Brandon, 2020) or as a component of the Internet of Things (Golpîra and Bahramara, 2020).

Finally, the storage battery industry can *empower consumers* to manage their energy consumption, reduce their carbon footprint, and increase their energy independence (Agnew and Dargusch, 2017). Storage batteries can help households and businesses reduce their reliance on the grid and take control of their energy usage (Boulaire et al., 2019).

In conclusion, the storage battery industry is of significant importance and significance due to its potential to transform the energy system and deliver a range of benefits for various stakeholders. The sector has the potential to provide grid stability and flexibility, energy security and reliability, reduce energy costs and increase efficiency, contribute to environmental benefits, drive innovation and economic growth, contribute to transportation and emerging technology efficiency and empower consumers. However, addressing the industry's challenges will be crucial for successfully delivering these benefits. Because despite the benefits, there are also significant **challenges** facing the storage battery industry. These include supply chain volatility, technological limitations, regulatory barriers, and economic uncertainties. Addressing these is key for the industry to realise its full potential.

One of these critical challenges is the volatility of the industry's **supply chain** (Jaffe, 2017; Sun et al., 2019). In the following paragraphs, I present every step of the storage battery supply chain (Figure 4.1), and I try to summarise the key risks at these stages.

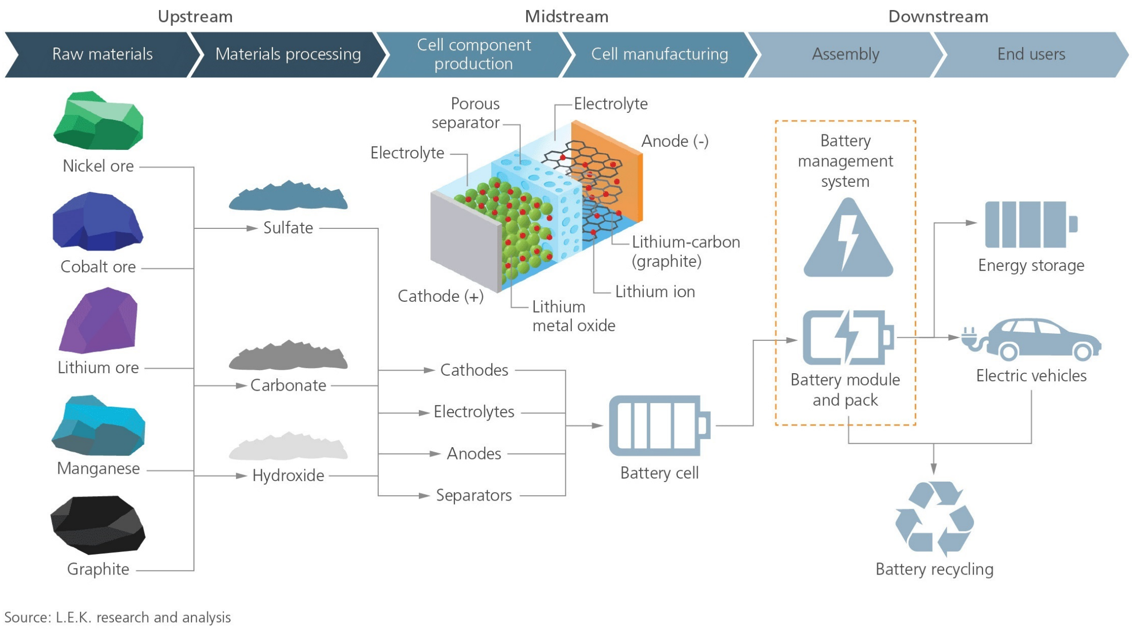


Figure 4.1: Storage battery supply chain

The first stage of the supply chain of storage batteries is *raw material extraction* (Weimer, Braun, and Hemdt, 2019). Lithium, cobalt, nickel, and manganese are the primary raw materials used in battery production (Backhaus, 2021). These minerals are mainly found in countries such as China, the Democratic Republic of Congo, Australia, and Chile (Jiali Song et al., 2019). The extraction of these minerals can have significant environmental and social impacts, including deforestation, water pollution, and child labour (Silva Lima et al., 2022). Therefore, ensuring that the raw materials used in battery production are ethically sourced is important. One approach to promoting ethical raw material extraction is through certification schemes such as the Responsible Minerals Initiative (RMI) (*Responsible Minerals Initiative* 2023). The RMI is a multi-stakeholder initiative that aims to promote responsible mineral supply chains. The initiative provides a framework for companies to assess and manage mineral extraction's social and environmental risks. Many battery manufacturers have joined the RMI, and they are required to demonstrate that their supply chains are free from human rights abuses, child labour, and environmental harm.

Once the raw materials have been extracted, they are transported to *battery production* facilities (Väyrynen and Salminen, 2012). Battery production involves several stages, including material processing (J. Li, Daniel, and Wood, 2011), electrode fabrication (Gonçalves, Lanceros-Méndez, and Costa, 2022), cell assembly, and module and pack assembly (Maiser, 2014; Saw, Ye, and Tay, 2016). Battery production is a highly technical process that requires specialized equipment and skilled labour (Kwade et al., 2018). The major battery manufacturers include Tesla, LG Chem, Samsung SDI, and Panasonic (E. M. Sarkar, T. Sarkar, and Bharadwaj, 2018; Cooke, 2020; Moores, 2021; Y. Park, Nakaoka, and Y. Chen, 2022). The location of battery production facilities is often influenced by factors such as access to raw materials, labour costs, and government policies. For example, many battery production facilities are located in China, which has a significant supply of raw materials and low labour costs (Y. Wang et al., 2019). However, there are concerns about the environmental and social impacts of battery production in China, including air pollution and poor working conditions (Kuijp, Huang, and Cherry, 2013).

After battery construction, the batteries need to be transported to their final destination, such as an electric vehicle manufacturer or a renewable energy project. *Transportation and logistics* play a critical role in the supply chain of storage batteries, as they can affect the cost, efficiency, and sustainability of the supply chain. Transportation and logistics are particularly important for the global supply chain of storage batteries, which involves shipment across continents (J. Yan et al., 2019). The transportation of batteries involves significant energy use and emissions, which can contribute to climate change. Therefore, optimising the transport and logistics of storage batteries to reduce their environmental impact is essential. One approach to reducing the environmental impact of transportation and logistics is through the use of renewable energy. For example, Tesla has developed a fleet of electric trucks powered by renewable energy. Using electric trucks can reduce the emissions associated with transportation and logistics while also promoting the use of renewable energy (Cooke, 2020).

The final stage of the supply chain of storage batteries is *end-of-life management* (Dunn et al., 2014). Storage batteries have a limited lifespan, after which they need to be disposed of or recycled. The disposal of batteries can have significant environmental impacts, as batteries contain hazardous materials such as lead and cadmium (Z. Chen et al., 2022). Therefore, it is crucial to ensure that storage batteries are disposed of or recycled safely and sustainably. The recycling of storage batteries can provide significant environmental and economic benefits. Recycling can help to reduce the demand for raw materials, as the materials contained in batteries can be recovered and reused (Beaudet et al., 2020).

Therefore, we can conclude that storage batteries as critical energy storage tools are a key solution for the problem of energy transition. Hence, the industry deserves special attention. Also, because of the reason that many risks in its supply chain can cause disruptions in storage battery production. Consequently, a growing body of literature recognises the importance of storage battery supply chain research. This chapter contributes to this research area from the perspective of network science and inter-industry production networks.

Whilst supply chain analysis is a central topic in the economics literature, no studies have been found that approach the subject from the systematic data perspective. This chapter aims to address the question of how can the storage battery industry and its supply chain be explored and analysed from the network side perspective on the basis of data. By using complex system analysis tools, such as network science, I will map a national storage battery ego network and examine the critical inter-industry interdependencies in it.

4.3 Methodology

The goal of this chapter is to map the ego production network of the storage battery industry based only on data and only on the production network framework. This framework was already developed and described for the entire US national production network in Chapter 2. While it is inspired by the approach designed there, this section focuses specifically on the storage battery industry and its surroundings. It contains an extensive analysis of the 2007 and 2012 storage battery ego network from several dimensions.

4.3.1 Data source and the presence of the storage battery industry in it

As a starting point, I use the same data source as in Chapter 2, the detail level United States Input-Output Accounts Data for years 2007 and 2012 (*Input-Output Accounts Data — U.S. Bureau of Economic Analysis* 2021).

BEA uses its Industry Codes for the three levels of detail, but it also defines how these relate to the 2012 North American Industry Classification System (NAICS) code structure (*North American Industry Classification System* 2021). NAICS represents a universally accepted categorisation framework that facilitates grouping an establishment's field of operation. Government statistical bodies rely on the NAICS classification system to examine and disseminate statistical information about the United States economy. The NAICS code structure employs a hierarchical arrangement of six-digit codes, effectively organising all economic undertakings into twenty distinct sectors. Each six-digit industry category within the NAICS classification system contains more specific and detailed sub-categories that provide a finer level of granularity in describing economic activities. Their most detailed classification is represented by "Index Entry" descriptions.

The storage battery manufacturing industry has the 335911 NAICS code, and the companies listed in this US industry are mainly engaged in manufacturing storage batteries. I show its Index Entries according to the NAICS Association in Table 4.1.

Table 4.1: NAICS index entries for 335911 - Storage battery manufacturing industry

2007 NAICS	2012 NAICS	Index Entries for 335911
335911	335911	Alkaline cell storage batteries (i.e., nickel-cadmium, nickel-iron, silver oxide-zinc) manufacturing
335911	335911	Automobile storage batteries manufacturing
335911	335911	Batteries, rechargeable, manufacturing
335911	335911	Batteries, storage, manufacturing
335911	335911	Lead acid storage batteries manufacturing
335911	335911	Lithium batteries, storage, manufacturing
335911	335911	Marine storage batteries manufacturing
335911	335911	Nickel-cadmium storage batteries manufacturing
335911	335911	Rechargeable battery packs made from purchased battery cells and housings
335911	335911	Rechargeable nickel-cadmium (NICAD) batteries manufacturing
335911	335911	Storage batteries manufacturing

The first digit of the NAICS code refers to the big industry group classification. I show all industry groups in Table 4.2. The storage battery manufacturing industry with the 335911 NAICS code falls into the Manufacturing industries, specifically durable goods.

Table 4.2: NAICS major industry classifications

NAICS code first digit	Industry
1	Agriculture, forestry, fishing and hunting
2	Mining, quarrying, extraction, utilities and construction
3	Manufacturing (durable and nondurable goods)
4	Wholesale, retail trade, transportation and warehousing
5	Information, finance, real estate, rental, management of companies, professional, technical and administrative services
6	Educational and health services
7	Arts, entertainment, recreation, accommodation and food services
8	Other services
9	Public administration

4.3.2 Constructing the ego production network

The network representation used in Chapter 2 is valid here, too. An element w_{ij} represented the nominal amount of goods industry i used as input by sector j, with i, $j = 1, ..., N$, where N was the number of sectors.

For a two-industry network example of the Input-Output Accounts, see Figure 2.1. The network built from this database was a directed weighted graph. The vertices were the industries (i and j), and the directed connections were the monetary transactions, the nominal flow of goods between sectors. The weight of each link was the economic value, representing how much an industry supports the development of the other (w_{ij} and w_{ji}). For example, industry i requires w_{ij} dollar input from industry j to produce one dollar output to final users (typically $w_{ij} \neq w_{ji}$).

The input-output models also indicate if the output of an industry is required as input to the same industry (w_{ii}). For example, the Oil and gas extraction industry could produce the oil and gas to power its own equipment, or the computer design and manufacturing industry can produce computers that are used to plan the next generation of computers *Economic Input-Output Life Cycle Assessment. Carnegie Mellon University Green Design Institute* 2008. We didn't use these self-sector transactions in our analysis.

After building the production network for the entire national economy, I focused on the storage battery ego network.

I used two approaches with different goals when selecting the ego network. Both ego networks zoomed in on the storage battery manufacturing industry; thus, both networks had the storage battery industry at their centre. The two approaches gave me two different perspectives and, hence, different interpretations of the subset of this national production network while still using the same data source. Consequently, I could study the storage battery ego network as detailed as possible with this data source. The first approach was simple, focusing specifically on the direct production relationships of the storage battery manufacturing industry. The second, a more advanced one, was built on the first one and took a closer look at the surrounding inter-industry relationships of the suppliers.

The idea behind the first approach was to map out all direct suppliers of the storage battery industry and limit the analysis at this step. Direct industries are directly linked to the industry in question (here: storage battery manufacturing industry) and execute direct exchange/business with them. In our ego network, the direct exchange is the direct monetary transaction, thereby creating a direct connection between two industries.

The benefit of this network is that I can clearly see the direct connections between the storage battery manufacturing industry and any other industry in need or need of the storage battery industry. With this approach, I also see the distribution of direct suppliers based on the amount of the monetary transaction made. Therefore, I can make assumptions about which inter-industry relationships are essential for the storage battery industry.

In order to have an even more transparent picture, I separated all upstream and downstream direct suppliers into two different ego networks. For the storage battery manufacturing industry, upstream refers to the material and service inputs needed for the production of the storage battery, while downstream is the opposite end, where the storage battery gets distributed. Upstream direct monetary transactions are the ones when the storage battery manufacturing industry is the buyer, and the other industry is the seller. At the level of the ego network, the storage battery industry, as a node in the network, is the target node, while the other sector is the source node. By contrast, downstream direct monetary transactions are the ones when the storage battery manufacturing industry is the seller, and the other industry is the buyer. The direction of the connection is also precisely the opposite in the ego network. As a node in the network, the storage battery industry is the source node, while the other sector is the target node.

Therefore, when I mention the upstream ego network, I mean that only those industries included in the ego network that provide resources for the storage battery manufacturing industry with direct monetary transactions (connections). On the other hand, when referring to the downstream ego network, I mean that only those industries included in the ego network that need the storage battery industry as a direct resource and connect to storage battery industry manufacturing with direct monetary transactions (connections).

Consequently, for the first ego network approach, I selected the storage battery industry and all its upstream and downstream direct suppliers, including every industry that is directly connected to the storage battery industry. In other words, every industry with which it has a direct transaction, no matter the value of it.

In the second approach, I kept the direct perspective and added the suppliers' transactions. I first put the storage battery industry upstream and downstream direct ego networks together to build the more advanced ego network. After, I excluded all direct transactions with a value of less than 0.005 dollars from the ego network. This cut-off was necessary because otherwise, I would have gotten a fully connected graph impossible to analyse and draw conclusions from in the next step. This threshold was small enough to keep the most transactions in the network and map a more understandable network. After getting the direct, simple ego network with both direct upstream and downstream suppliers, I also included the transactions between the present suppliers. With this method, the storage battery industry was still at the centre, but I could also see the surrounding transactions that don't include the primary sector of the ego network.

Every industry depends on its suppliers; however, those suppliers still depend on their own suppliers. Hence, the industry in question indirectly depends on its supplier's transactions.

A major advantage of the second approach is that I could analyse the whole production environment of the storage battery industry, considering every transaction in close proximity. In this way, I could also identify those industries that stand out and influence the ego network of the storage battery manufacturing industry the most.

4.4 Simple ego network of storage battery industry

The first set of questions aimed to identify the key production relationships of the storage battery industry. For this purpose, I built the storage battery ego network from two dimensions: upstream and downstream. Figure 4.2 shows the results for 2007, while Figure 4.3 for 2012. There are 405 industries in the entire US national production network, from which 378 are present in the upstream and 402 in the downstream storage battery ego network.

4.4.1 Upstream ego network

The **upstream** ego networks are presented in Figure 4.2a and in Figure 4.3a. As stated earlier, there are 378 supplier industries in both upstream networks. For 2007, the values of the monetary transactions range from 0.000001 to 0.22715 dollars. Similarly, for the year 2012, the values are between 0.000001 and 0.25687 dollars. Again, I want to point out that these are normalised values from the national Input-Output Accounts and tell us how many dollars are needed from the supplying industries to produce 1 dollar of storage battery for the final consumer.

In order to highlight even more the essential supplier industries in Figure 4.2b and Figure 4.3b, I show the distribution of the top supply relationships on a treemap chart. While the ego network also displays the key industries, here, they are more visible, and comparing the monetary transaction values can be better.

The range of monetary transactions is extensive, and there are way more smaller transactions than higher, more significant ones. For example, for the year **2007**, in the upstream ego network, there are only two transactions above the value of 0.1 dollars. These are 0.22715 dollars from the *Nonferrous metal (except aluminium) smelting, refining* and 0.12573 dollars from the *Nonferrous metal (except copper, aluminium) rolling, drawing, extruding, alloying* industries, Figures 4.2a and 4.2b. These are exceptionally high values compared to the rest of the ego network, which suggests that these two industries are essential in the production of storage batteries. Even between these two transactions, there is a noticeable difference in the transaction amount. However, the third supply relationship in the row is from the *Iron, steel mills, ferroalloy manufacturing* industry with the value of half of the *Nonferrous metal (except copper, aluminium) rolling, drawing, extruding, alloying* industry monetary transactions, 0.06387 dollars.

Moreover, there are only 20 transactions above the value of 0.01 dollars and 120 above the value of 0.001 dollars. That gives us 258 industries supporting the storage battery manufacturing industry with a monetary transaction value of less than 0.001 dollars. Hence, the high number of industries in the upstream ego network can be explained by the difference in their financial transaction values. Considering only the relatively large monetary transactions, the upstream ego network would have way fewer nodes as industries and connections as transactions.

Likewise, for the year **2012**, the same two transactions remain the most dominant ones, which are also above the value of 0.1 dollars. Figures 4.3a and 4.3b highlight these: 0.256868 dollars from the *Nonferrous metal (except aluminium) smelting, refining* and 0.21735 dollars from the *Nonferrous metal (except copper, aluminium) rolling, drawing, extruding, alloying* industries. An observable contrast is that while both supporting transactions have increased compared to the year 2007, the first one (from the *Nonferrous metal (except aluminium) smelting, refining* industry) with almost 0.03 dollars, the second one (from the *Nonferrous metal (except copper, aluminium) rolling, drawing, extruding, alloying* industry) has almost doubled its value, increasing with 0.09 dollars. The tendency continues with the third supporting transaction that five years later is from the *Other durable goods merchant wholesalers* industry with a value of 0.07095 dollars. Increasing with almost 0.015 dollars and jumping two places to the front, in 2007, this industry supported storage battery manufacturing with a value of 0.0563315 dollars. 2007 third supporting industry: *Iron, steel mills, ferroalloy manufacturing* is now only in the seventh place, decreasing with almost 0.024 to a value of 0.04023 dollars. Another noteworthy industry in the top five supporters of storage battery manufacturing that needs to be mentioned is the *Copper, nickel, lead, zinc mining* industry — from the value of 0.05683 in 2007 reduced to the value of 0.05153, with 0.005.

Furthermore, in the 2012 network, there is one more transaction compared to 2007 above the value of 0.01 dollars, 21, and five more above 0.001 dollars, precisely 125 — the remaining 253 transactions as less than 0.001 dollars.

The numbers on these upstream ego network graphs (Figures 4.2a, 4.3a) and treemap diagrams (Figures 4.2b, 4.3b) show that there has been a slight volume growth between 2007 and 2012 in the resources needed by the storage battery manufacturing industry. It is more perceptible if we add up all transaction values to get the overall resource required by the storage battery industry. For the year 2007, the sum of all supporting monetary transactions is 1.2595 dollars, while for 2012 is 1.4307 dollars. This gives us a 0.1712 dollars rise in supply, from which around 0.12 dollars is the rise of the first two critical industries: *Nonferrous metal (except aluminium) smelting, refining* and *Nonferrous metal (except copper, aluminium) rolling, drawing, extruding, alloying.*

4.4.2 Downstream ego network

The **downstream** ego networks are presented in Figure 4.2c and in Figure 4.3c. In both downstream networks, 402 industries request direct resources from the storage battery manufacturing industry. This slightly higher number than the upstream supplier number indicates that the storage battery industry is needed in a few more industries than vice versa, but as the numbers will tell us later, only with smaller value transactions.

In the same way, as at the upstream network, for the relationships to be more visible quantitatively, I show the distribution of the top inter-industry relationships on a treemap chart in Figure 4.2d and Figure 4.3d.

At the downstream part of the storage battery manufacturing industry, the monetary transaction value range is also significantly wide, however less than at the upstream ego network. For the year 2007, the values of the monetary transactions range from 0.000005 to 0.018948 dollars. Similarly, for the year 2012, the values are between 0.000006 and 0.02268 dollars.

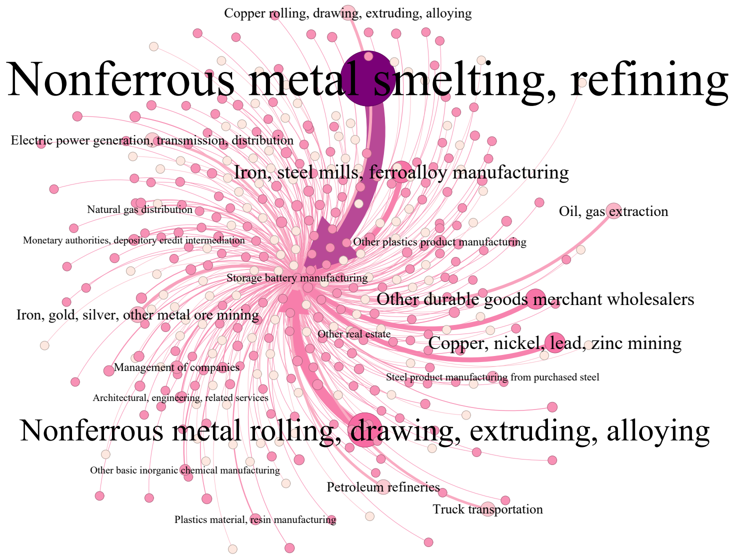
Compared to its upstream pair, the downstream ego network in the year **2007** has zero transactions above the value of 0.1 dollars, only one transaction above 0.01 dollars and 34 above 0.001 dollars. That gives us 368 remaining industries supplied by the storage battery manufacturing industry with a monetary transaction value of less than 0.001 dollars. This figure is 110 more than for the upstream network. The same conclusion is valid here, too, that the high number of industries present in the downstream ego network is explainable by their difference in monetary transaction values. On the downstream side, it is even more striking that most monetary transactions are extremely small, almost negligible. That being the case, we must focus our analysis on the industry transactions that are relatively high compared to the average.

The highest transaction that needed the storage battery manufacturing industry the most in 2007 is from the *Power-driven hand tool manufacturing* with a value of 0.01895 dollars. I need to indicate that if we compare the upstream ego network visualisation in Figure 4.2a to the downstream one in Figure 4.2c, we might notice that the two largest nodes seem to be the same size on both networks. However, they might show a proportional value relative to their own ego network, but they differ extensively compared to each other. On the upstream side, the *Nonferrous metal smelting, refining* industry holds a value of 0.22715 dollars. In contrast, on the downstream side, the *Power-driven hand tool manufacturing* industry is only a tiny fraction of that value, specifically around 8.5%. But in order for the figures to be interpretable, they needed to be scaled to their own size, not relative to the other subfigures. For a finer context, the values can be directly compared next to the ego networks in Figures 4.2b (upstream) and 4.2d. After the *Power-driven hand tool manufacturing* industry transaction, the next three industries need only 30% of the first's value: *Waste management, remediation service* industry with

0.00579 dollars, *Construction machinery manufacturing* industry with 0.00546 dollars and *Fishing, hunting, trapping* industry with 0.00545 dollars.

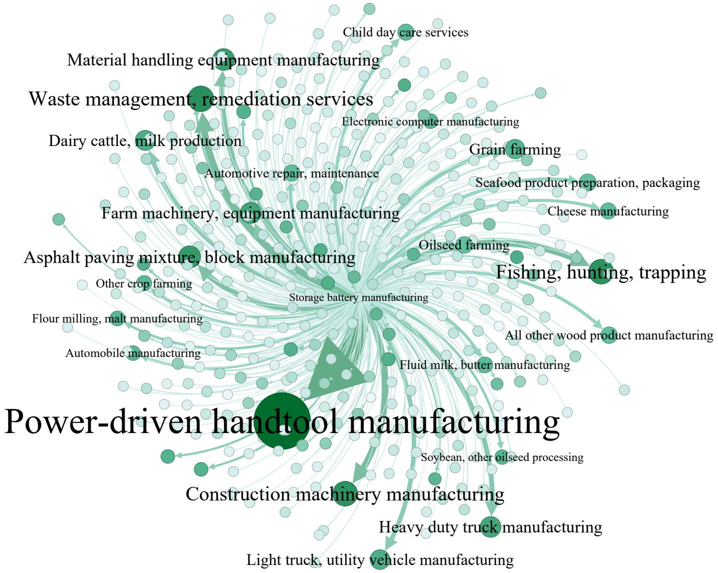
The identical conditions apply for the year **2012**: zero transactions above the value of 0.1 dollars, only one transaction above 0.01 dollars and 41 above 0.001 dollars. Like in the upstream network, there is a slight increase in the downstream ego network too. For the year 2012, the *Power-driven hand tool manufacturing* needs the storage battery manufacturing industry a bit more, the transaction value increasing to 0.02268 dollars. The same happens with the *Construction machinery manufacturing* industry, improving to 0.00723 dollars. Another key industry comes into the picture, the *Farm machinery, equipment manufacturing* industry with a transaction value of 0.00741 dollars.

The growth of the resource provided by storage battery manufacturing for other industries is apparent at the level of the entire ego network. By adding up all supporting transactions from the storage battery industry for 2007, I get 0.1811 dollars, while for the year 2012, 0.1952 dollars. Thus, there is a +0.0141 supply change between 2007 and 2012. While this number is not as high as for the upstream ego network (0.1712 dollars supply boost), it is the same percentage in scalable terms. The entire supply growth in the upstream ego network is 14% compared to the highest value transaction, while in the downstream network, this metric is 8%.



(a) Upstream ego network

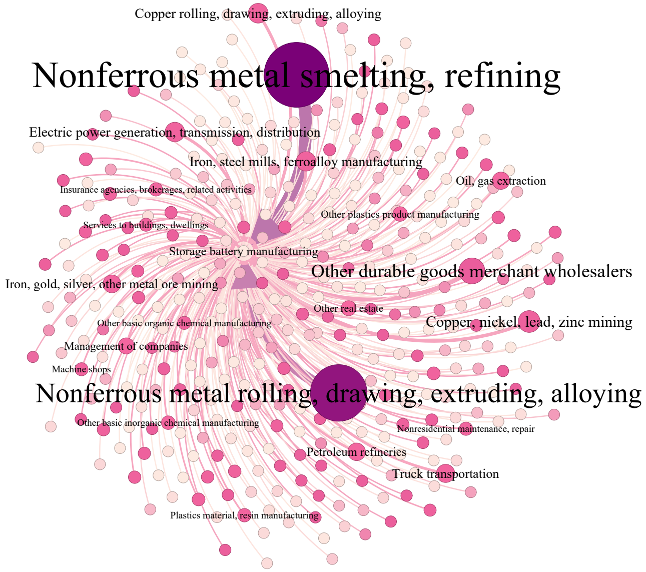
(b) Upstream direct suppliers based on the transaction value required by the storage battery industry



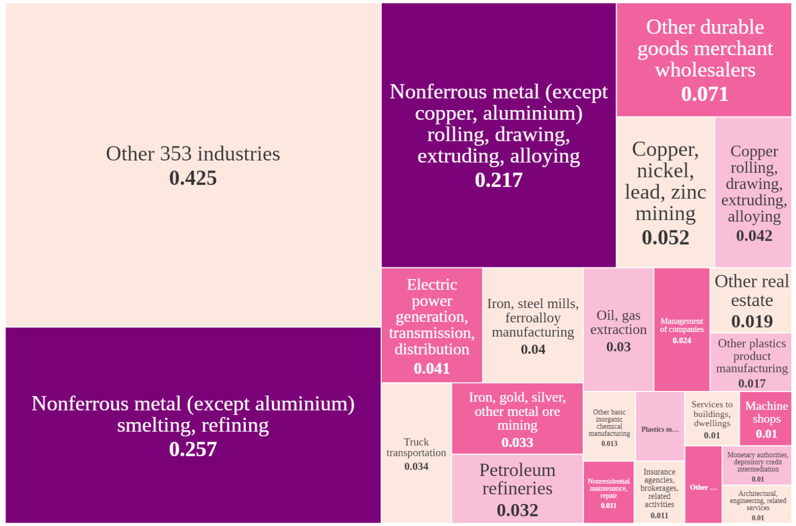
(c) Downstream ego network

(d) Downstream direct industries based on the transaction value provided by the storage battery industry

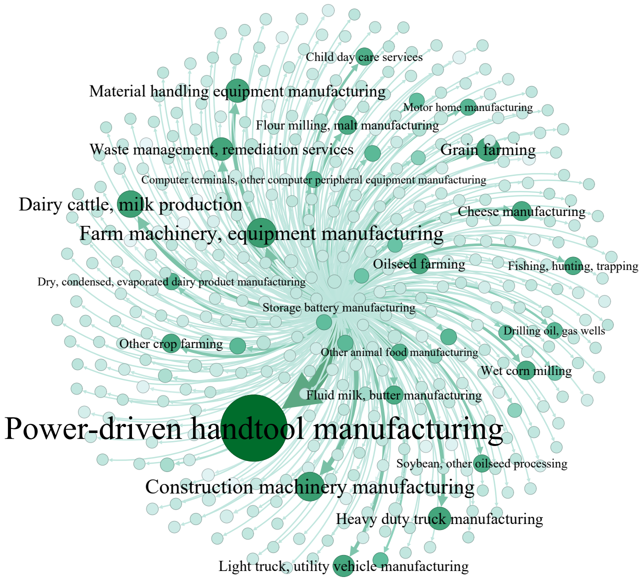
Figure 4.2: Storage battery industry upstream and downstream simple production network (2007)



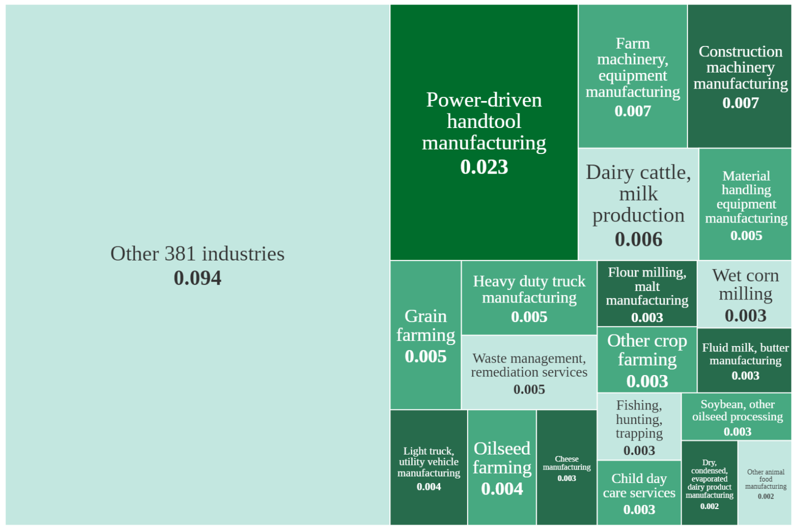
(a) Upstream ego network



(b) Upstream direct suppliers based on the transaction value required by the storage battery industry



(c) Downstream ego network



(d) Downstream direct industries based on the transaction value provided by the storage battery industry

Figure 4.3: Storage battery industry upstream and downstream simple production network (2012)

4.4.3 Discussion and comparison of upstream and downstream network

I summarised all findings from the upstream and downstream network in Table 4.3.

Table 4.3: Upstream and downstream ego network general statistics

	Upstream		Downstream	
	2007	2012	2007	2012
Number of industries	378	378	403	403
Lowest transaction value	0.000001	0.000001	0.000005	0.000006
Highest transaction value	0.22715	0.25687	0.01895	0.02268
Average transaction value	0.00333	0.00378	0.00045	0.00049
Sum of all transaction values	1.2595	1.4307	0.1811	0.1952
Supply change	+0.1712		+0.0141	
Supply change compared	+14%		+8%	

As the table shows, most metrics increased from 2007 to 2012 in both types of networks. So has the highest transaction value, the average transaction value and the sum of all transactions. Supply change is the difference in amount between the sum of all inter-industry transactions between 2007 and 2012. Both for the upstream and downstream network, this is a positive number, referring that overall, the transaction values increase. I also calculated this in percentage, showing in the *supply change compared* row of the table. The transaction values in the upstream storage battery ego network increased by 14% and in the downstream network by 8% in 2012 compared to the 2007 landscape. The **growth** of upstream and downstream supply between 2007 and 2012 can be interpreted as a tendency of increasing interconnectivity between the different industries. If we accept the assumption that the higher the transaction value, the more essential and "closer" the inter-industry relationship is, we can conclude that there is a need and direction for increased proximity between industries.

A clearly visible **asymmetry** needs to be emphasised between the upstream and downstream networks in nominal value. If we examine the key transactions at both dimensions of the storage battery ego network, it is evident that the upstream monetary transactions are way higher in value. This asymmetry can be traced back to the fact that the storage battery manufacturing industry produces a final good; thus, in the upstream ego network, most industries providing the most considerable amount of resources are the raw material industries. Their resource might be more crucial when manufacturing a final product than the resources that the storage battery industry can provide for other durable goods. The key relationships in the downstream ego network are between two durable goods; I might actually say finished goods. However, finished goods is a relative term, and a seller's finished goods may become a buyer's raw materials. Therefore, in the context of the downstream network, the storage battery is a semi-finished good, as it behaves as a component or spare part for the next industry. Consequently, the source of asymmetry is explicable by the difference in the resource type, the difference between the, strictly speaking, raw materials, such as *Nonferrous metal smelting, refining* and final products, such as *Power-driven hand tool manufacturing.*

4.5 Advanced ego network of storage battery industry

The second set of questions aimed to discover the surrounding inter-industry relationships of the suppliers in the storage battery ego network.

The simple ego network earlier showed us that almost 100% of the US national production industries are present in both dimensions of the storage battery ego network (93.3% in the upstream network and 99.5% in the downstream one). Of course, the monetary transaction values are very diverse, and most are negligibly small. Therefore, at this point, all storage battery upstream and downstream suppliers that have reached at least the 0.005 dollars threshold are included in one advanced ego network. Moreover, the transactions between the present suppliers are also included in the advanced ego network.

In this way, I could uncover the indirect dependencies of the storage battery industry and analyse the whole proximity production environment of the storage battery industry; by also identifying those industries that stand out and influence the advanced ego network of the storage battery manufacturing industry the most.

4.5.1 Node-level measures in the advanced ego network

In order to determine the key industries that influence the nearby storage battery manufacturing production cycle, I define three types of actors: **integrator**, **allocator** and **mediator** industries. These definitions are inspired by Kim et al., 2011, but in that article, the authors specify them for firm-level supply networks with both materials flow and contractual relationships. In this thesis, I describe them precisely for industry-level production networks built from the monetary transactions of the national input-output accounts.

I name **integrators** those industries that incorporate the most resources in the ego network. Needless to say, most industries are integrators to some extent, as most industries require different types of resources to function. However, I anoint integrators, those ones that need the most in two different senses of the word.

Firstly, I use the network centrality metric of **in-degree** as a starting point. This measure gives us, for every industry, the count of inward connections, in other words, the *number* of direct suppliers a sector has. This can be interpreted as a quantification of each industry's challenge in effectively managing the inflow of materials from different and numerous upstream industries. Hence, the higher the in-degree value, the higher the challenge of assembling or converting various components into a product with added value and ensuring its proper functioning. Accordingly, the industries with the most in-degree and, thus, upstream suppliers are the integrators (by number) in that ego network production cycle.

Secondly, I consider the **in-degree weighted** network centrality metric. In comparison with the in-degree metric, this is not the count, but the sum of inward connection weights, the sum of the upstream monetary transaction values. In other words, not the number but the amount of resources an industry directly needs to produce one dollar output for the final consumer. This quantifies a slightly similar challenge an industry faces, but also different in some ways. It is the challenge encountered by the sector in effectively handling the *amount* of inflowing material from upstream industries. Integrator (by amount) industries are those that convert high-value transaction components into improved products.

Allocators are those industries that assign the most resources in the ego network. Just as in the case of integrators, most industries are allocators, too, to some extent. Yet, I call allocators those ones that pump the most support in the production cycle again in two different senses.

On the one hand, I use the measure of **out-degree** to define the key allocators. Out-degree is the count of an industry's outward connections, the *number* of sectors an industry directly supplies. I interpret this network centrality metric as estimating the complexity level experienced by the industry when addressing the requests made by different downstream sectors. The higher the number of industries requesting resources from the industry in question, the higher the complexity of meeting the demand. Consequently, the allocator (by number) industries are those ones that distribute scarce resources among numerous industries.

On the other hand, I calculate the **out-degree weighted** metric to get the allocators (by amount). Out-degree weighted, in contrast to out-degree, is the sum of outward connection weights, the sum of monetary transactions directly supporting other industries from the industry in question. Particularly the amount of resources an industry supplies to the ego network. This is the complexity level experienced by the industry when addressing the *volume* of the demand made by downstream industries. At this metric, the attention is focused on the amount of resources requested by the other sectors, not on the count of industries that are soliciting it. For example, the product of one industry may be needed in a larger amount but by fewer industries than the product of another industry, which is required by more industries but in transactions of lower overall value. Thus, the allocators (by amount) are those that distribute resources among industries with high quantity demand overall.

The third key actor group is the **mediator** industries. As their name also says, these industries intermediate the flow of resources between different industries. The first two groups of key industries mainly focused on direct transactions. The difference between them was only the direction of the transactions; allocators were analysed in the opposite direction of integrators. However, when examining the mediator industries, the main idea is to investigate the indirect transactions and see which industries are the key transmitters on the resource chain.

In the first place, I calculated the network centrality measure of **betweenness**. This is the metric of how often an industry is on the shortest path between two other sectors. The shortest path refers to a route connecting two nodes that contain the minimum number of edges. In this case, each edge contributes equally to the overall path length. The logic behind betweenness is that if the shortest path from one node to the other leads through a third, the node from the middle can be determinative. Specifically for our production network, it measures how often an industry plays the role of a mediator in a supply chain and how often the route through this industry is the shortest way to access other sectors. Thus, it is the degree of influence or control an industry possesses in shaping the interactions among other industries within the ego network. I call mediators (by number) those industries that are actors in most supply chains and can regulate the movement of resources throughout the entire ego network.

For the mediator (by amount), I computed the **betweenness weighted** metric. This is also a measure of how often an industry is on the shortest path between two other sectors, but it considers the weights of the connections, the value of the monetary transaction too. I assume that the higher the transaction value, the more essential the connection is; hence the distance is smaller, and the two industries tend to be ”closer”. So, betweenness weighted can also be interpreted as the degree of influence an industry possesses in shaping the interactions among other sectors within the ego network but moving further and even considering that the highest-value transactions and supply chains are the most volatile to this. Hence, mediator (by amount) industries can regulate the movement of the highest-value transaction resources throughout the ego network.

These key industries are important because I can map out a clear picture of the storage battery manufacturing industry's close production cycle by identifying them. In such a manner, I can recognise the critical interdependencies that the storage battery manufacturing cycle is conditional on. Therefore, these crucial industries and inter-industry reliances can be assessed cautiously in policy planning or economic forecasting.

All three types of key actors offer different views and hold other information about the storage battery ego network. Both integrator and allocator industries tell us the fundamental industries in the sense of direct transactions in the network. Integrators are the ones that compress the most resources in the production network cycle of the storage battery industry, while allocators are the ones that pump the most into the ego network. In contrast, mediator industries, focusing on intermediary transactions, show us the important sectors in the middle of the different supply chains integrated into the ego network. Taken together, all three: integrator, allocator and mediator industries are vital to the well functioning of a production cycle. If any of these industries stopped operating, the production cycle would shake overall. By propagating a shock in one of the key industries, the production network would lose important transactions that would, directly and indirectly, affect the storage battery manufacturing industry's performance.

I summarised the definitions of all network centrality metrics and key industry roles in Table 4.4.

Table 4.4: Node-level centrality metrics and their implications for ego production network

Centrality metrics	Definition of the production network metric	Conceptual definition for the ego production network	Implication for central nodes: Role	Implication for central nodes: Description
In-degree	The number of direct suppliers an industry has. (count of inward connections)	The challenge encountered by an industry in effectively handling the inflow of materials from upstream industries.	Integrator (by number)	To assemble or convert various components into a product with added value and ensure its proper functionality.
In-degree weighted	The amount of resources an industry directly needs to produce one dollar output for the final consumer. (sum of inward connection weights)	The challenge encountered by an industry in effectively handling the amount of inflowing material from upstream industries.	Integrator (by amount)	To assemble or convert high-value transaction components into an improved product and ensure its proper functionality.
Out-degree	The number of sectors an industry directly supplies. (count of outward connections)	The level of complexity experienced by an industry when addressing the requests made by downstream industries.	Allocator (by number)	To distribute scarce resources among numerous industries.
Out-degree weighted	The amount of resources an industry supplies to the ego network. (sum of outward connection weights)	The level of complexity experienced by an industry when addressing the volume of the demand made by downstream industries.	Allocator (by amount)	To distribute scarce resources among industries with high quantity demand.
Betweenness	The measure of how often an industry is on the shortest path between two other industries.	The degree of influence or control an industry possesses in shaping the interactions among other industries within the ego network.	Mediator (by number)	To enable or regulate the movement of resources throughout the entire ego network.
Betweenness weighted	The measure of how often an industry is on the shortest path between two other sectors, considering the weights of the links. The bigger the weight, the closer the two industries.	The degree of influence or control an industry possesses in shaping the interactions among other industries within the ego network, considering that the highest value transactions and supply chains are the most volatile to this.	Mediator (by amount)	To enable or regulate the movement of the highest value transaction resources throughout the ego network.

The results for the **integrator** industries are shown in Figures 4.4 and 4.5. Both figures show the values for the year 2007 and the year 2012 too; thus, the change in 5 years is also comparable. Figure 4.4 presents the outcome of the in-degree network centrality metric, hence the integrator (by number) industries, while Figure 4.5 provides the in-degree weighted metric results, thus the integrator (by amount) industries.

Both Figures 4.4a and 4.5a compare the actual value of the metric. Next to them, Figures 4.4b and 4.5b show the ranking of the industries and the rank order number according to the same measure. This differentiation was necessary because the figures showing the actual values are sometimes very dense. Some industries can have very similar values to the same metric and are displayed very close to each other on these figures. These types of figures are exemplary in clearly indicating the magnitude of the difference between the metric values of different industries. However, because of is this dense, the ranking order and the information of which industry comes after which in this order is not clearly visible. Hence, the justification for the other type of figure revealing the hierarchy order and only the ranking number according to that metric. We need both kinds of information provided by different figures to have a transparent image of the top industries in that specific metric.

Table 4.5: Colours of the industry classifications on the advanced ego network and node-level figures

Colour	Industry
Orange	Mining, quarrying, extraction, utilities and construction
Purple	Manufacturing (durable and nondurable goods)
Blue	Wholesale, retail trade, transportation and warehousing
Green	Information, finance, real estate, rental, management of companies, professional, technical and administrative services
Brown	Public administration

The colours on the figures refer to the big sector groups in which every industry is classified, already shown in Table 4.2. In Table 4.5, I present the colours of the most crucial classification groups on the figures.

As shown in Figures 4.4 and 4.5, both integrator (by number) and integrator (by amount) industries are the manufacturing industries, specifically durable goods manufacturing industries. This can be traced back to the fact that durable goods need the most resources as they are already a final product or semifinal for some supply chains.

The results obtained from the **in-degree** centrality metric calculation are presented in Figure 4.4. The first **integrator (by number)** industry is the *Storage battery manufacturing industry*, with a value of 45 for 2007 and 51 for 2012. Once again, the value of the in-degree measure tells us how many direct suppliers an industry has. Logically, the Storage battery industry has the most suppliers as this is its ego production network, and this industry is on the final or semifinal product side of the production. Hence, it is an integrator, and in 2007, 45 industries out of the total of 49 industries and in 2012, 51 out of 57 supported storage battery manufacturing directly. After that, with the second highest value of 44 suppliers for 2007 is the *Construction machinery equipment manufacturing* industry, which falls back slightly to 43 suppliers and the tenth place for 2012. For the year 2012, the *Nonferrous metal rolling, extruding, and alloying* and the *Steel product manufacturing from purchased steel* industry need the most diverse resources from different industries after the storage battery industry with the value of 47 suppliers. They had 39 and respectively 42 suppliers for 2007.

The increasing tendency of interconnectivity between 2007 and 2012 observed on the simple ego networks in section 4.4 can be marked in the advanced ego networks too in Figure 4.4a. Most industries needed resources from more suppliers in 2012 than they did in 2007.

It is also striking in Figure 4.4 that besides almost all manufacturing industries being the top integrator industries, most of the raw materials and the trade and transportation industries are middle on the integrator industries scale, while most service industries are at the bottom.

(a) Value (b) Rank

Figure 4.4: Changes in the integrator (by number) industries in the storage battery ego network between 2007 and 2012, by number of industry contacts (in-degree)

Figure 4.5 presents the results of the **in-degree weighted** centrality metric. There is a clear distinction between the top **integrator (by amount)** and the integrator (by number) industries. The first integrator (by number) industry in Figure 4.4, the Storage battery manufacturing industry here, is only in ninth place in 2012 with a value of 1.19 dollars. The in-degree weighted measure value tells us the sum of all upstream monetary transactions of an industry, thus the amount of the resource integrated by an industry. The amount of resources needed by an industry is not perfectly directly proportional to the number of suppliers it requires it from, hence the difference.

The first integrator (by amount) industries in the ego network are the *Copper rolling, drawing, extruding and alloying* and the *Plastic material and resin manufacturing* industries. They changed places in the five years between 2007 and 2012, but they are ahead of any other industry in the ego production network. The magnitude of the difference is transparently visible in Figure 4.5a. The Copper rolling, drawing, extruding and alloying industry integrated 1.53 dollars in 2007 and 1.56 dollars in 2012. Similarly, the Plastic material and resin manufacturing industry required 1.63 dollars and in 1.54 dollars in 2012.

Two industries, *Plastics packaging materials, unlaminated film manufacturing* and *Communication, energy wire, cable manufacturing*, in third and fourth place with values of 1.39 and 1.37 dollars in 2012, were not even present in the 2007 ego production network. Of course, the industries existed, but they didn't reach the 0.005 dollars monetary transaction threshold with any of the other industries in 2007 in the storage battery ego network to get into the network.

The differences between the in-degree weighted values are a bit higher than for the in-degree values. However, the industry classifications on the integrator (by amount) scale are somehow the same: almost all manufacturing industries are at the top, most of the raw materials in the middle, and most of the trade and transportation industries and service industries are at the bottom of the integrator industries scale.

The first integrator raw material mining group industry, for both the number and amount, is the *Iron, gold, silver and other metal ore mining*, which is ahead of several manufacturing industries too, integrating 39 upstream suppliers and 0.8 dollars resources in 2012. While the first service integrator industry, still for both indicators, is the *Truck transportation* industry with 24 suppliers and 0.74 dollars in resources required in 2012.

(a) Value (b) Rank

Figure 4.5: Changes in the integrator (by amount) industries in the storage battery ego network between 2007 and 2012, by resource volume (in-degree weighted)

As for the **allocator (by number)** indicator, the balance of power changes compared to integrators, and service industries are taking the lead, as shown in Figure 4.6. I measure allocator (by number) industries with **out-degree** centrality metric, that is, the number of sectors an industry directly supplies. This outcome reveals that mainly service industries support most industries.

The general tendency of increasing interdependence is also perceptible in Figure 4.6a, the majority of the individual industry's supporting transactions number rises. This is also due to the fact that more industries reached 0.005 dollars transaction cut-off value in 2012 than in 2007; thus, there are more industries and connections in the 2012 ego production network.

The first allocator (by number) industries are the *Management of companies and enterprises* and the *Other real estate* industries, allocating resources to 48 industries in 2007 and to 56 industries in 2012. That is every sector in both years in the storage battery ego network. There is also in the first place for the year 2012 with supporting all other industries the *Monetary authorities and depository credit intermediation*, *Legal services* and *Employment services*. Monetary authorities and depository credit intermediation and Legal services were also in the front in 2007 by supplying 47 and 43 industries, while Employment services didn't match the threshold limit in the 2007 ego network.

Other fairly significant allocators (by number) in the top ten are from the utility industries: *Electric power generation and distribution*, from the manufacturing industries: *Petroleum refineries* and from the mining and extraction industries: *Oil and gas extraction*. Also, as the leader of the trade and transportation industries, *Truck transportation* must also be mentioned as an essential allocator in the storage battery ego network. Truck transportation also plays a vital role as an integrator industry; hence, it requires special attention when ensuring that the storage battery production cycle is uninterrupted. The importance of this particular industry also for the entire national production network was already explained and highlighted in chapter 2.

(a) Value (b) Rank

Figure 4.6: Changes in the allocator industries (by number) in the storage battery ego network between 2007 and 2012, by number of industry contacts (out-degree)

The outcome of the **out-degree weighted** centrality metric computation that gives us the **allocator (by amount)** industries is presented in Figure 4.7. The out-degree weighted measure shows the amount of resources an industry supplies to the ego network. It is the sum of outward connection weights, hence the aggregate of the direct downstream supplying transactions.

It is apparent from Figure 4.7a, showing the values of this metric, that there are two exceptionally dominant allocators (by amount) in the ego network of the storage battery manufacturing industry. The first one is the *Oil and gas extraction*, which pumps 4.11 dollars in 2007 and 4.01 dollars in 2012 in resources into the ego network. Followed by the *Petroleum refineries* industry, allocating 2.79 dollars in 2007 and 3.03 dollars in 2012 to all sectors in the ego network. This might not seem like high values for an entire production cycle, but let me revive the fact that these transaction values are normalised. They individually show the magnitude of resources needed by one industry from another to produce 1 dollar of its final output. And in these normalised circumstances, these are incredibly increased values, especially compared to other allocator's (by amount) values.

This outcome also explains why I needed two types of figures to present the results of these indicators since these extreme values suppress the other small ones; thus, those might be significant, too, to some extent.

The third allocator (by amount) industry is the *Iron, steel mills and ferroalloy manufacturing* with a value of 1.84 dollars for 2007 and 1.85 dollars for 2012, which isn't even half of the first one.

As is shown later, in the next section, the network-level measures (subsection 4.5.2), the average out-degree weighted for both timestamps is around 0.65 dollars, and in 2007 only 15 of 49 industries, while in 2012, only 21 of 57 sectors are above the mean value. Moreover, in 2007 just 9 and in 2012, 14 sectors reached and exceeded the value of 1 dollar for the out-degree weighted metric. Compared to these statistics, the industries of Oil and gas extraction and Petroleum refineries allocate plenty of resources to the ego network.

This outcome can also be attributed to the fact that these industries are raw materials, and there is a clear difference between raw materials and processed materials and products in the sense of allocating and integrating resources. The asymmetry between allocators and integrators is due to the reality that a few raw materials issue most of all resources. At the same time, the semifinal and final products generally integrate around the same amount of resources. In the case of integrators, there are no such outliers as those for the allocators. This is only true when we consider the magnitude of monetary transactions, too, not just their presence and number. Thus, the distribution of the number of financial transactions between integrators and allocators is mostly the same. However, when comparing the distribution of the amount of resources allocated and integrated, that is very different because these large raw material industries hold a high proportion of resources allocated in the ego network production cycle. In chapter 2, I noticed the same outcome for the entire national production network. This asymmetry and the allocator outliers were the reason for the scale-free nature of the production network from the outward dimension. Hence, concluding that "commanding heights" are the drivers of production and economy.

In this storage battery ego network in 2007, the Oil and gas extraction industry concentrates 13% of all resources allocated. While adding the resources distributed from the Petroleum refineries industry too, the two of them centralise 22% of all resources in the ego network. For the year 2012, this decreases slightly, with the Oil and gas extraction industry allocating 11% and together with the Petroleum refineries 19%, but still being a high ratio of the entire resources. For both years, the first ten allocator industries hold 50% of all resources distributed in the ego production network.

For the allocator (by amount) industries, there is no precise industry classification on the metric scale. The first two can be included primarily in the raw materials industry groups; however, there is no pattern after that. The only thing that stands out is that there are predominantly manufacturing industries at the bottom of the scale.

(a) Value (b) Rank

Figure 4.7: Changes in the allocator (by amount) industries in the storage battery ego network between 2007 and 2012, by resource volume (out-degree weighted)

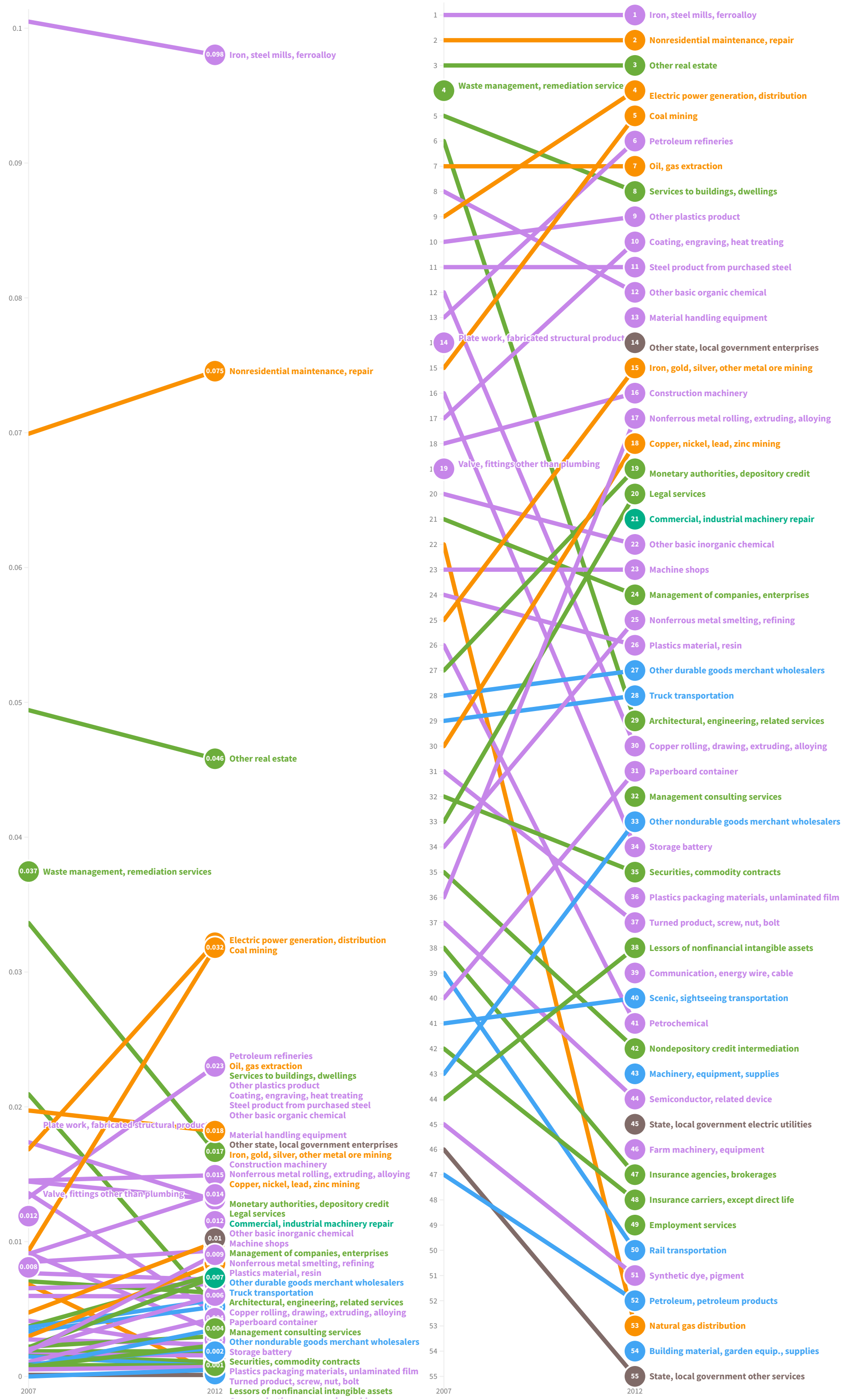
The **betweenness** centrality metric gives us the **mediator (by number)** industries presented in Figure 4.8. Betweenness is a measure of how often an industry is on the shortest path between two other sectors. I use this metric to detect the amount of influence an industry has over the flow of resources in the storage battery ego network.

I calculate both betweenness measures based on Freeman (1977) and on the algorithm later implemented by Brandes (2001, 2008).

The first mediator (by number) industry is *Iron, steel mills and ferroalloy manufacturing* industry with a value of 0.1 for 2007 and 0.098 for 2012. This is a normalised metric with an interval between 0 and 1. Hence, providing us that the Iron, steel mills and ferroalloy manufacturing industry is present on around 10% of the shortest paths in the ego network. The second one was the *Nonresidential maintenance and repair* industry with 7% in 2007 and 7.5% in 2012. The third mediator (by number) industry is the *Other real estate* industry, with 4.9% in 2007 and 4.6% in 2012.

For context, and later expanding more on the subject in subsection subsection 4.5.2, the average value for the betweenness measure is around 1%. Industries are generally present on 1% of the shortest paths in the storage battery ego network. In other words, they hold the power to influence around 1% of all supply chains in the network. However, the variation is relatively high, with the first three industries having the capability to control 10%, 7% and around 5% of the individual flow of resources in supply chains in the storage battery ego network.

Although it is the storage battery manufacturing industry's ego production network, the *Storage battery* industry is only in the middle of the mediator scale, not even reaching the average value. In 2007, it was present on the 0.9%, and in 2012, on the 0.3% of the individual shortest paths between any two other industries in the ego network. This doesn't mean it is not connected to them (it is an integrator with 51 individual suppliers), just it is not on the shortest supply chains in the ego network. I assume that those industries have the most influential power that is present on most short production routes; however, not only these can have a mediator role and influence.



(a) Value (b) Rank

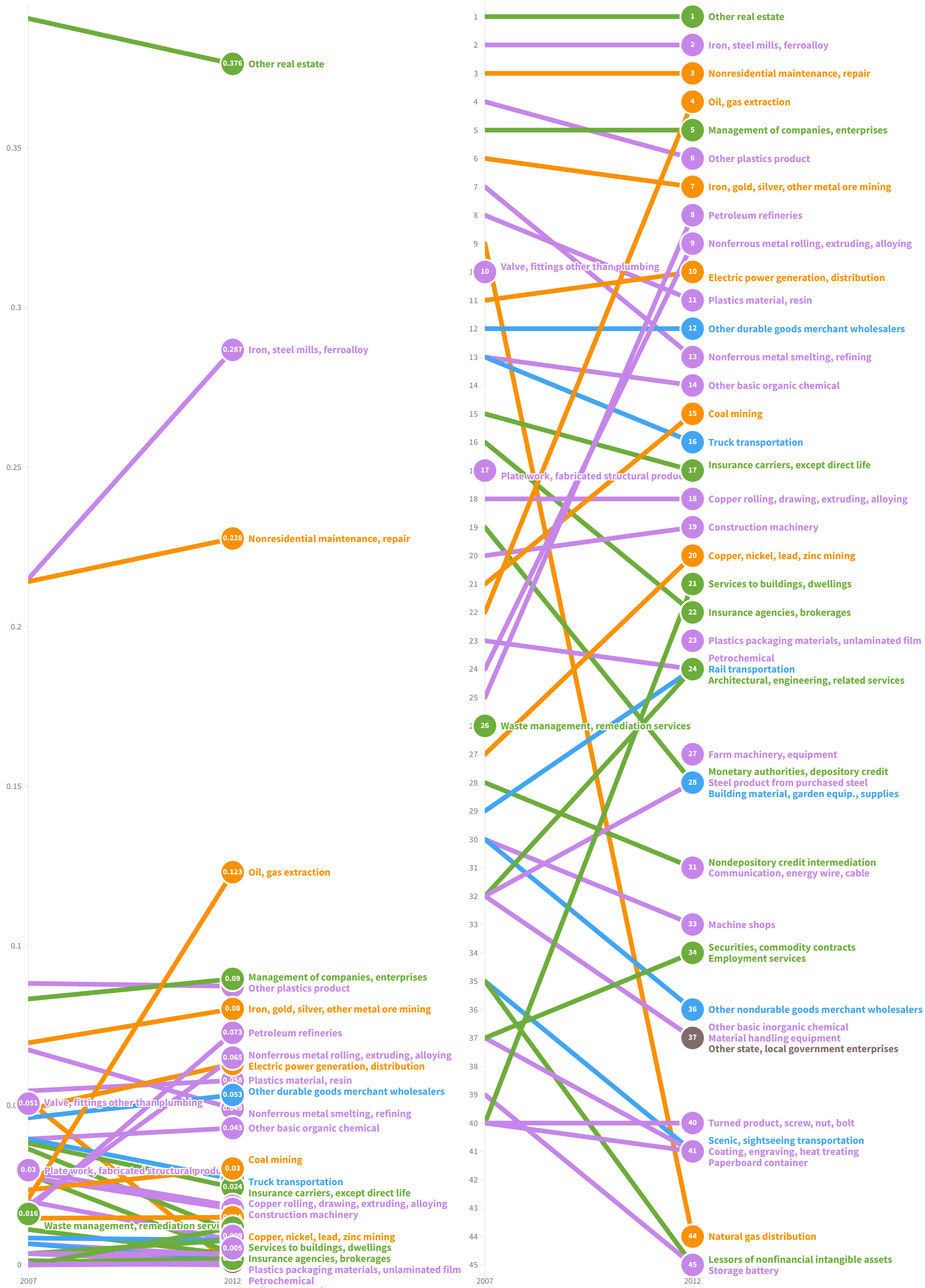
Figure 4.8: Changes in the mediator (by number) industries in the storage battery ego network between 2007 and 2012, by number of industry contacts (betweenness)

Betweenness weighted centrality metric is used to identify the **mediator (by amount)** industries in the storage battery ego network. This also measures how often an industry is on the shortest path between two other sectors. Still, it also takes into account the magnitude of the monetary transactions. I assume the higher the trade value, the closer the two industries are to each other.

My assumption that the transaction value is equally proportional to the connection's essentiality has its limitations. However, at this scale, we don't have any data or indicators that could give us better presumption about inter-industry relationship interdependencies than monetary transactions between them. For example, discovering every industry's product's individual supply chain and mapping essential spare parts or services needed in manufacturing every unique product can be a potential solution. Although to build manually these crucial connections requires an enormous amount of time and effort that this project didn't have. Despite the limitation of this assumption, the research's results are valid and useful. And the mediator role through the betweenness weighted measure can suggest a degree of influence or control an industry possesses in shaping the interactions among other sectors within the ego network, considering that the highest value transactions and supply chains are the most volatile to this.

The first three mediator (by amount) industries are the same as the mediator (by number) industries but in a different order. The first is the *Other real estate* industry, with being present on 39% of the shortest paths in 2007 and on 37.6% of the shortest paths in the 2012 ego network. The second is the *Iron, steel mills and ferroalloy manufacturing* industry with 21.5% in 2007 and 28.7% in 2012, while the third is the *Nonresidential maintenance and repair* being present on 21.4% quickest routes in 2007 and on 22.8% in 2012.

These proportions are higher than for the mediator (by number) industries, also highlighting the fact that if we consider the magnitude of the monetary transactions, the influential power concentrates on fewer industries than if we consider only the transactions by number.



(a) Value (b) Rank

Figure 4.9: Changes in the mediator (by amount) industries in the storage battery ego network between 2007 and 2012, by resource volume (betweenness weighted)

4.5.2 Network-level measures in the advanced ego network

In order to describe and analyse the storage battery ego network at a higher and more general level, I measured the descriptive statistics of the indicators explained above in subsection 4.5.1 (Node-level measures). I calculated the average, minimum, maximum and standard deviation values for the integrator (in-degree, in-degree weighted), allocator (out-degree, out-degree weighted) and mediator (betweenness, betweenness weighted) measures. I also computed the **network density** metric for both the 2007 and 2012 ego networks.

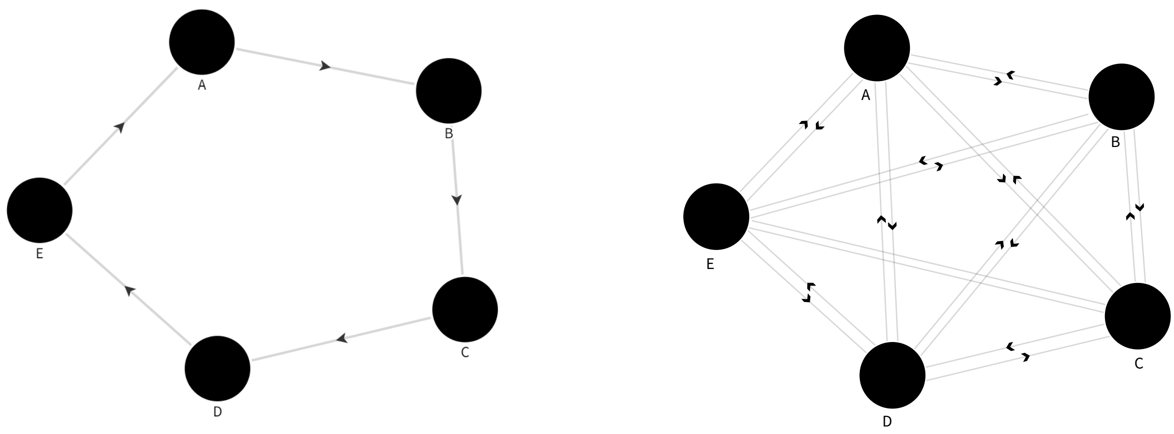
The term network density refers to the ratio of existing relationships in a network to the total number of potential relationships within the network.

A finite number of relationships exists for any given set of nodes within a network; hence, every ego production network is limited to the maximum number of transactions possible. Each industry as a node has the potential to be either the source or the target of a relationship with every other industry/node. Consider an ego network with three industries: A, B, and C. The table below (Table 4.6) showcases all the potential directed relationships between these nodes, thus all possible monetary transactions between industries.

Table 4.6: Possible directed relationships for three nodes without self-loops

Source	Target
A	B
A	C
B	A
B	C
C	A
C	B

Every industry initiates a relationship with the other two sectors. Nevertheless, in reality, not all potential relationships may be established. Certain industries might lack direct connections with others, and some directed relationships may not reciprocate. Even more, for the ego production network, some monetary transactions might be insignificantly small that they are not calculated in the network.



(a) Sparse network (b) Dense network

Figure 4.10: Sparse and dense networks

The network density is the percentage of the relationships present compared to all possible connections. The metric value spans from 0 to 1, where the minimum value indicates networks without any links, while the maximum value signifies networks with all possible relationships. As the value approaches 1, the network becomes denser, indicating greater cohesion among the nodes within the network. For the ego production network, this means that the industries are more interdependent in that subset of the national production network.

In networks with high density, the flow of resources tends to be smoother compared to networks with low density. In high density networks the paths are shorter and multiple paths are possible. Figure 4.10 illustrates two networks comprising five nodes. The sparse network exhibits only five out of the potential 20 relationships between the nodes, resulting in a density of 0.25. Conversely, the dense network encompasses all possible relationships, causing a density of 1.

Table 4.7 summarises the results of the indicator's general statistics for the 2007 ego network, while Table 4.8 for the 2012 network. I also included the **degree** and the **degree weighted** metric. A node's degree is simply the total number of links that it has with other nodes in the network, while the degree weighted is the sum of these links' weights.

For the production ego network, the degree of an industry is the number of other industries with which it has direct transactions, no matter the direction of the transaction. It can be upstream and downstream, too. While the industry degree weighted is the sum of all direct resources received plus all direct resources allocated to the ego network.

In the interpretation of these results, I use the name of the centrality metrics. To see the extensive description of these metrics' implication for production network industry roles, see subsection 4.5.1, and for a short summary of these conceptual definitions, see Table 4.4.

Table 4.7: Storage battery advanced ego network measures descriptive statistics (2007)

Variable	Obs	Mean	Std. Dev.	Min	Max
In-degree	49	25.531	11.358	2	45
In-degree weighted	49	0.641	0.374	0.013	1.630
Out-degree	49	25.531	14.582	0	48
Out-degree weighted	49	0.641	0.732	0	4.112
Degree	49	51.061	10.098	15	74
Degree weighted	49	1.282	0.778	0.340	4.234
Betweenness	49	0.011	0.019	0	0.100
Betweenness weighted	49	0.037	0.069	0	0.391

In the 2007 storage battery ego network, there are 49 different industries, hence, 49 different values for every network measure. The overview of these is shown in Table 4.7.

The mean value of the in-degree and out-degree is the same, with 25.531 average connections (supplier transactions) per industry, as they calculate the same transactions but from different dimensions. Similarly, both the in-degree weighted and the out-degree weighted metric is 0.641 dollars resource on average, computing the same normalised transaction values from the two sides.

The difference is the distribution of these values. For the in-degree and out-degree is more or less similar, the deviation is around half of the average value. 11.358 suppliers for the in-degree, while for the out-degree, slightly higher with 14.582 suppliers. The minimum and maximum values are quite alike, too: the minimum is 2 connections for the in-degree and 0 connections for the out-degree, while the maximum is 45 for the in-degree and 48 for the out-degree.

The most striking contrast is between the in-degree weighted and the out-degree weighted distribution. The standard deviation is also around half of the average for the in-degree weighted with 0.374 dollars. However, for the out-degree weighted is even more than the average value: 0.732 dollars. This is also visible in the minimum and maximum values. The minimums are relatively close: 0.013 dollars for the in-degree weighted and 0 dollars for the out-degree weighted. But the maximums are very distinct from each other: it is 1.63 dollars for the in-degree weighted and, more than a duplicate of this, 4.112 dollars for the out-degree weighted. This rather interesting finding can be explained by the fact that most resources in amount are clustered around a few allocator industries, hence the high out-degree weighted metric value. But when integrating resources, there are no such highly clustered integrator industries. Even the outstanding integrators are not as outliers as the outstanding allocators, as they are way closer to the average than the allocators with maximum value. This outcome already emerged at node-level measures (subsection 4.5.1) and at the national production network (chapter 2) analysis too.

The mean betweenness is 0.011; hence, industries, on average, are present on 1.1% of all the shortest paths between any two sectors when not considering the magnitude of the monetary transactions. When taking into account the weights of the links, this increases to 3.7%, as the number of the shortest routes changes too. The higher the transaction value, the shorter the route is between industries. The standard deviation is around the same for both metrics, close to the duplicate of the average value: 0.019 for betweenness and 0.069 for betweenness weighted. The minimum value is 0 for both. This value can be attributed to the point that some industries don't have mediator roles in this sense, and they are not present in any short supply chains in the ego network. The most important mediator industries are present on 10% of all short paths if we don't calculate the monetary transaction value, only the presence of it (betweenness). And the maximum value for the betweenness weighted is 39.1%.

Table 4.8: Storage battery advanced ego network measures descriptive statistics (2012)

Variable	Obs	Mean	Std. Dev.	Min	Max
In-degree	57	28.544	12.284	6	51
In-degree weighted	57	0.669	0.388	0.113	1.564
Out-degree	57	28.544	17.914	0	56
Out-degree weighted	57	0.669	0.744	0	4.010
Degree	57	57.088	13.419	22	86
Degree weighted	57	1.339	0.825	0.337	4.341
Betweenness	57	0.010	0.017	0	0.098
Betweenness weighted	57	0.033	0.070	0	0.376

The general statistics of the indicators for the 2012 storage battery ego network are shown in Table 4.8. In the 2012 network, there are 57 different industries.

Likewise, the mean value of the in-degree and out-degree is the same, with 28.544 average supplier transactions per sector. Also, both the in-degree weighted and the out-degree weighted metric is 0.669 dollars resource on average. The identical condition applies in the 2012 network, too, both out and in metrics calculating the same transactions and normalised transaction values from the two dimensions.

The difference in the distribution of these values applies in the 2012 network, too. The in-degree and out-degree distributions are more or less similar, the deviation being around half of the average value. 12.284 suppliers for the in-degree, while for the out-degree, slightly higher with 17.914 suppliers. The minimum and maximum values are quite alike, too: the minimum is 6 connections for the in-degree and 0 connections for the out-degree, while the maximum is 51 for the in-degree and 56 for the out-degree.

The pattern of the most notable distinction (between the in-degree weighted and the out-degree weighted distribution) is valid in the 2012 network. The standard deviation is also around half of the average for the in-degree weighted with 0.388 dollars. However, for the out-degree weighted is even more than the average value: 0.744 dollars. The minimum values are relatively close: it is 0.113 dollars for the in-degree weighted and 0 dollars for the out-degree weighted. But the maximums are very distinct from each other: it is 1.564 dollars for the in-degree weighted and, more than a duplicate of this, 4.010 dollars for the out-degree weighted.

The 2012 mean betweenness is 0.010; hence, industries, on average, are present on 1% of all the shortest paths between any two industries when not considering the magnitude of the monetary transactions. When taking into account the weights of the links, this increased to 3.3% in 2012, as the number of the shortest routes changed too. The standard deviation is around the same for both metrics, close to the duplicate of the average value: 0.017 for betweenness and 0.070 for betweenness weighted. The minimum value is 0 for both. The same conclusion applies in 2012, too, that some sectors are not present in any short supply chains in the ego network. The most significant intermediary industries are present on 9.8% of all short paths if I don't calculate the monetary transaction value, only the presence of it (betweenness). And this increases quite a lot, to 37.6% if I consider the magnitude of the present transactions too.

Table 4.9 re-organizes some information from Tables 4.7 and 4.8 and summarizes all complexity metrics for both 2007 and 2012 next to each other to highlight the change in those five years.

In the 2007 network, there were 49 industries present, while in the 2012 network, 57. As I already concluded in the earlier section (section 4.4), there is an increasing tendency in the monetary value of inter-industry transactions. Thus, showing it here too: for the year 2012, there were more industries included in the 0.005 dollars threshold value (16% more than in 2007).

Table 4.9: Storage battery industry advanced ego network complexity change between 2007 and 2012

Network measures		2007	2012	Change
Network size (industries)		49	57	+16%
Network density		0.532	0.510	-4%
In-degree (Integrator)	Average	25.531	28.544	+11%
	Minimum	2	6	+200%
	Maximum	45	51	+13%
Out-degree (Allocator)	Average	25.531	28.544	+11%
	Minimum	0	0	0%
	Maximum	48	56	+17%
Degree	Average	51.061	57.088	+12%
	Minimum	15	22	+46%
	Maximum	74	86	+16%
In-degree weighted (Integrator)	Average	0.641	0.669	+4%
	Minimum	0.013	0.113	+770%
	Maximum	1.630	1.564	-4%
Out-degree weighted (Allocator)	Average	0.641	0.669	+4%
	Minimum	0	0	0%
	Maximum	4.112	4.010	-2%
Degree weighted	Average	1.282	1.339	+4%
	Minimum	0.340	0.337	-1%
	Maximum	4.234	4.341	+3%
Betweenness (Mediator)	Average	0.011	0.010	-9%
	Minimum	0	0	0%
	Maximum	0.100	0.098	-2%
Betweenness weighted (Mediator)	Average	0.037	0.033	-10%
	Minimum	0	0	0%
	Maximum	0.391	0.376	-4%

However, the network density between the two years seems to remain the same. For both timestamps is around 0.5, with the threshold value of 0.005 dollars mentioned earlier. For the year 2007, it is 0.532, and for 2012 it is slightly less: 0.510. Evidently, the no-threshold storage battery ego network density is closer to the value of 1, to a fully dense network. However, that dense network is not useful to draw conclusions from; hence, the satisfactory 0.005 dollars threshold. This network density value of 0.5 tells us that 50% of all transactions are relatively high in value compared to the other half of the ego network, in which monetary transactions are insignificantly tiny.

The change is the most meaningful in the average value of the network centrality indicators. There might be outstanding values in one of the timestamps that distort the change for the maximum and minimum values. If I analyse the difference in the averages, only the role of the mediator industries decreased, as betweenness and betweenness weighted is slightly lower for 2012 than for 2007. It is around the same percentage, -9% for the average betweenness and -10% for the average betweenness weighted, as the average in-degree and out-degree values increased, +11%. While average in-degree weighted and out-degree weighted also improved with +4%.

Taken together, there is no such shocking high change between the two years. That is also due to the fact that 5 years is a small interval for a big transformation. However, what stands out in this table is the growth of almost every indicator between 2007 and 2012. This finding provides some tentative initial evidence of the increasing interdependence of the economic sectors. It may also be related to the fact that more industries are included in the 2012 network. But that also explains the growing interconnection in actually two ways. Firstly, even if there are more sectors in the ego network, the average values of the centrality metrics still rise, indicating that there are more high-value transactions in percentage in the 2012 production cycle than in 2007 one. Secondly, the addition of more industries to the 2012 network is because more sectors and their transactions reached the 0.005 dollar value transaction limit; hence they rely more on each other than in 2007. It can therefore be assumed that the diversity and the magnitude of resources integrated and allocated in the storage battery ego network are on the rise.

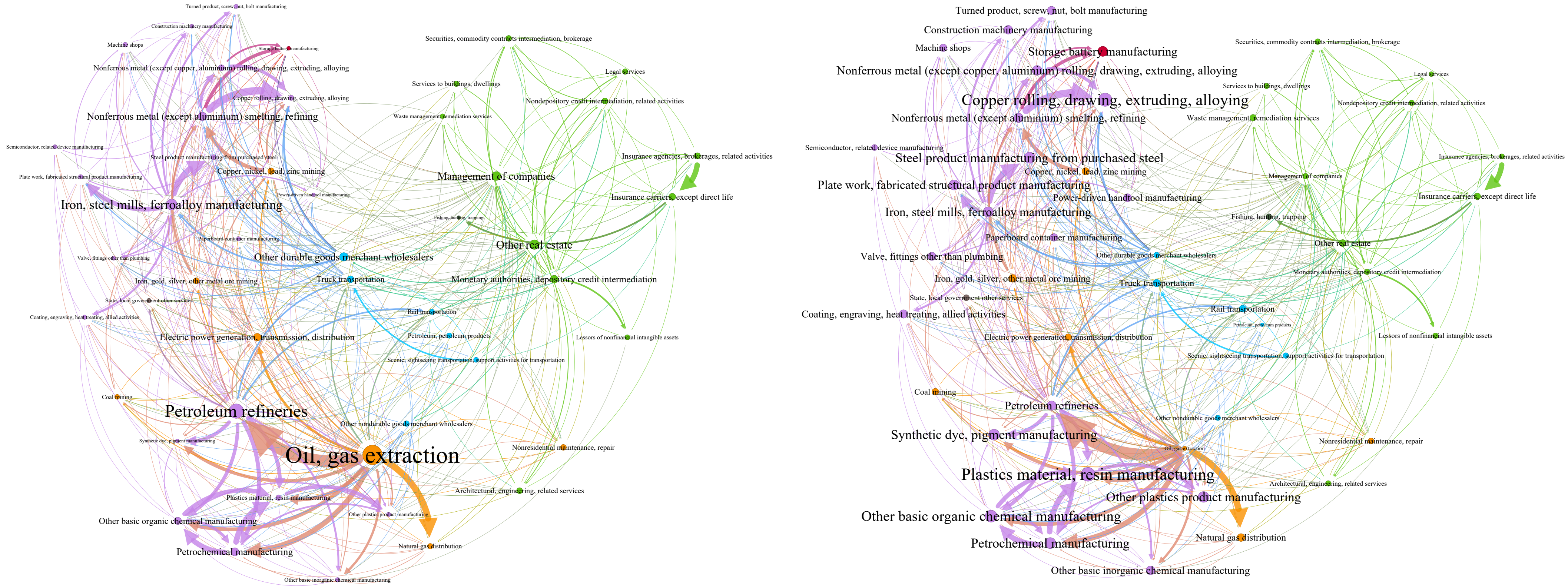
The advanced storage battery ego networks are visualised in Figures 4.11 and 4.12. Figure 4.11 shows the advanced network in 2007, while the 2012 advanced network is presented in Figure 4.12.

This graphical representation serves as an initial overview of the patterns present in the storage battery ego network. By visually mapping the interdependencies and flows of inputs and outputs within the ego production network, it becomes possible to gain valuable insights into resource allocation, supply chains, and economic linkages. It enables the identification of critical nodes, such as sectors with a high number of upstream and downstream suppliers or those facilitating the flow of resources. Also, a visual image is associated with the numbers, with the already explained and analysed complexity and centrality metrics of the ego network.

In both plots, subfigures 4.11a and 4.12a, the size of the nodes, representing individual industries in the ego networks, is proportional to the industry's allocator (by amount) metric value. On the other hand, in subfigures 4.11b and 4.12b, the node is scaled according to the integrator (by amount) metric. The size of the links is equal to their weight, thus, to the inter-industry monetary transaction value. The colour grouping of the industries is the same as presented in Table 4.5.

The patterns observed at the node-level measures in subsection 4.5.1 are perceptible in the network figures too. If the node sizes in the ego network are equivalent to allocator measures (4.11a, 4.12a), then the industries that distribute the largest amount of resources will be the ones that stand out in the network figure, such as the two outliers: *Oil and gas extraction* and *Petroleum refineries*. However, if the size of the nodes is proportional to integrator measures (4.11b, 4.12b), then no industry reaches the same node size as for the two allocators in the other network. However, several sectors are highlighted as essential integrators, mostly all from the manufacturing industry group.

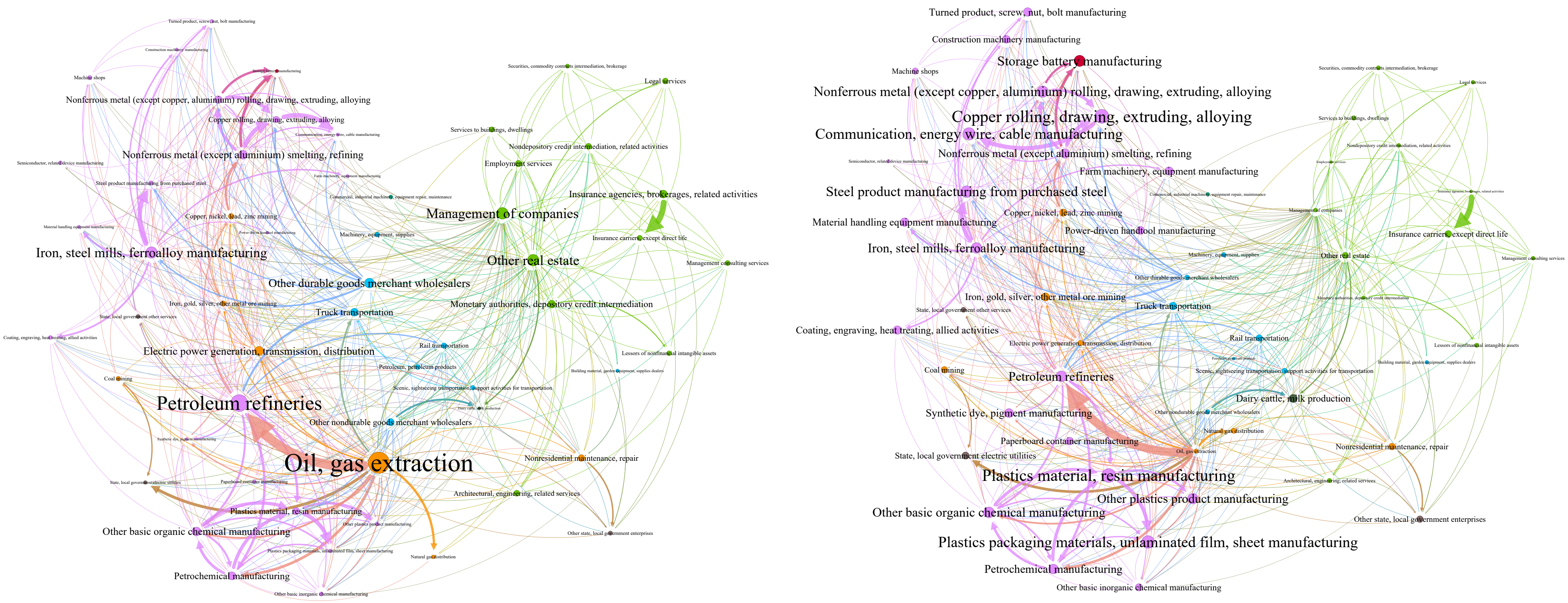
These storage battery ego network figures compared to the simple ego networks shown in Figures 4.2a, 4.2c, 4.3a and 4.3c are denser because include both upstream and downstream direct suppliers and even the transactions between them.



(a) Size of the nodes proportional to allocator industries (out-degree weighted)

(b) Size of the nodes proportional to integrator industries (in-degree weighted)

Figure 4.11: Storage battery industry advanced ego network (2007)



(a) Size of the nodes proportional to allocator industries (out-degree weighted)

(b) Size of the nodes proportional to integrator industries (in-degree weighted)

Figure 4.12: Storage battery industry advanced ego network (2012)

While these figures show the entire advanced ego network of the storage battery manufacturing industry, they are very dense, and the individual connections are hard to see. Some of the key industries and key transactions are observable, but these graphical representations are good only to visualise the magnitude of the difference between different sectors.

In order to plot the main relations in the storage battery ego network, I draw the medium-level and micro-level advanced ego networks in Figure 4.13 and 4.14. In other words, I zoomed in on a subset of the ego network, selecting the highest-value monetary transactions. In all network plots, the node size is proportional to the allocator (by amount), the out-degree weighted centrality metric value.

For the medium-level ego network, I used the 0.05 dollars threshold value. Thus, I only included those transactions in the ego network that reach the value of 0.05 dollars. This is way less than the 0.005 dollars monetary transaction cut-off value at the entire advanced ego network. While for the micro-level ego network, I incorporated only those transactions that reach and are higher than the 0.1 dollars cut-off value. Table 4.10 summarises the node and edge numbers in the medium-level and micro-level ego networks. It also compares these to the original advanced ego network node and edge numbers in the *Ratio* column.

Table 4.10: Number and ratio of nodes and edges in the medium-level and micro-level advanced storage battery ego network (2007 & 2012)

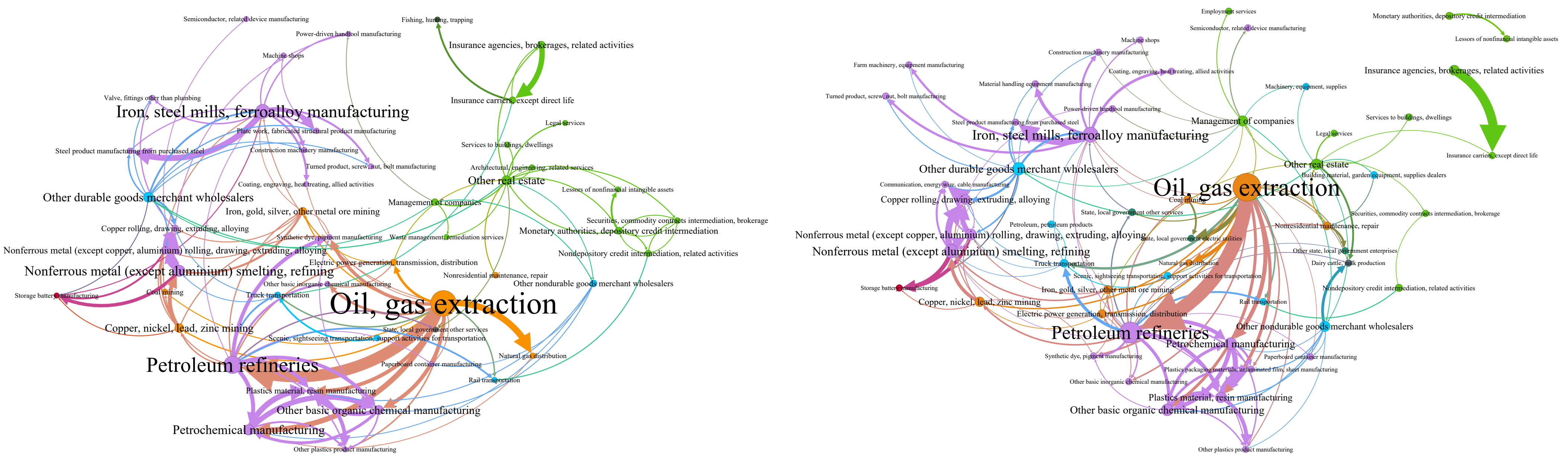
Network fragment		2007	2012	Avg. ratio
Original (0.005)	Nodes	49	57	100%
	Edges	1251	1627	100%
Medium-level (0.05)	Nodes	48	54	97%
	Edges	124	142	10%
Micro-level (0.1)	Nodes	26	30	53%
	Edges	38	50	3%

Figure 4.13 provides the medium-level advanced ego network at the threshold of 0.05 dollars. Sub-figure 4.13a shows the network in 2007, while sub-figure 4.13b in 2012. It is quite apparent that density decreased compared to the original ego network. 97% of the industries from the original ego network are included in the medium-level ego network, whereas only 10% of all monetary transactions.

The micro-level advanced ego network at the threshold of 0.1 dollars is shown in Figure 4.14. The 2007 network is presented in sub-figure 4.14a, while sub-figure 4.14b provides the network in 2012. At this point, only half of all industries from the original ego network are incorporated in the medium-level ego network (53%), while just 3% of all monetary transactions. However, these are the highest in value.

For both figures is apparent that interdependence increases with time. The 2007 network had 48 industries at the medium-level and 26 at micro-level; thus, the 2012 network had 54 sectors at medium-level and 30 at micro-level. These are partially responsible for the growth of the monetary transaction number. In the 2007 environment, there are 124 transactions above the value of 0.05 dollars (medium-level) and 38 above the value of 0.1 dollars (micro-level). While in the 2012 situation, 142 inter-industry connections reached the 0.05 dollars threshold (medium-level) and 50 above the value of 0.1 dollars (micro-level).

Also, what is the most striking difference between the 2007 and 2012 micro-level network figures is that there is a break in 2007 in the middle of the ego network. Consequently, adding to the growing interconnection through time argument.



(a) 2007 (b) 2012

Figure 4.13: Storage battery industry advanced ego network (2007 & 2012) — Medium environment (0.05 dollars threshold)

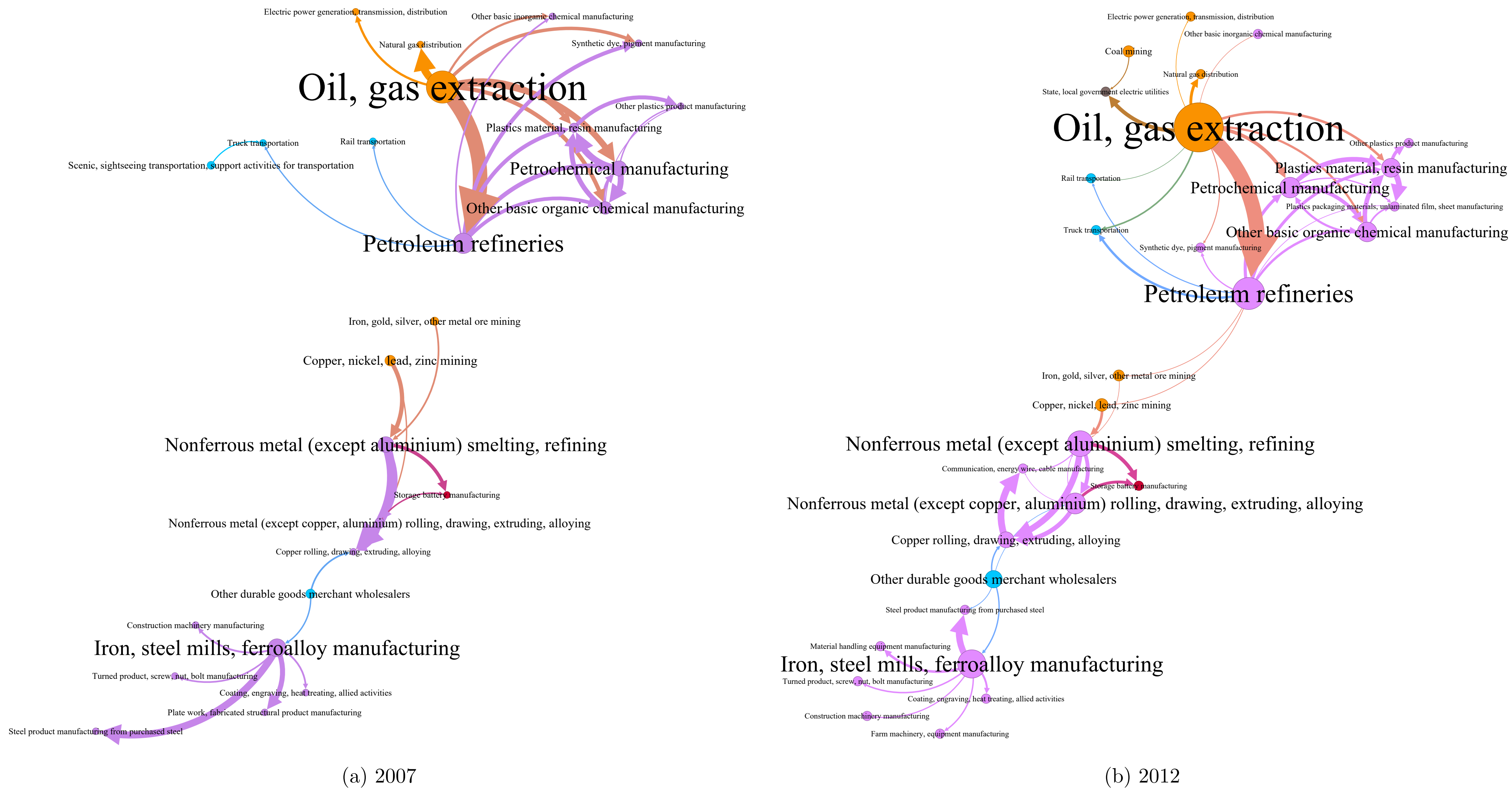


Figure 4.14: Storage battery industry advanced ego network (2007 & 2012) — Micro environment (0.1 dollars threshold)

4.6 Conclusions from the ego network analysis

This chapter examined the storage battery industry using a quantitative approach focusing on the industry's production network interconnections. The aim was to explore the storage battery supply chain through the lens of network science on the basis of systematic data and bring additional perspective to this research area.

I conducted an analysis of the storage battery ego network, which was constructed based on monetary transactions between industries. This investigation utilised the already introduced production network framework in Chapter 2 to identify the significant integrator, allocator, and mediator industries within the storage battery production sector. This section aimed to demonstrate how data-driven approaches can offer novel insights into the interconnectedness between the storage battery industry and other industries within its production network. Moreover, this examination aimed to explore the implications of the ego production network framework, revealing vital information about the storage battery industry's dependencies within its supply chain and identifying the industries that play indispensable roles in this network.

When analysing the *simple ego network*, I found a growing tendency in supply amount. Supply change is the difference in amount between the sum of all inter-industry transactions between 2007 and 2012. Generally, the transaction values in the upstream storage battery ego network increased by 14% and in the downstream network by 8% in 2012 compared to the 2007 landscape. The **growth** of upstream and downstream supply between 2007 and 2012 can be interpreted as a tendency of increasing interconnectivity between the different industries. If we adopt the assumption that higher transaction values indicate a more crucial and indispensable inter-industry supplying relationship, we can deduce that there exists a demand and inclination for increased proximity between industries.

According to monetary transaction values, the most important suppliers for the storage battery industry are the Nonferrous metal (except aluminium) smelting, refining and Nonferrous metal (except copper, aluminium) industries (upstream), and the most notable receiver is the Power driven hand tool manufacturing industry (downstream).

A prominent asymmetry exists between the upstream and downstream networks regarding nominal value. Upon analysing the key transactions within the storage battery ego network across both dimensions, it becomes evident that the upstream monetary transactions hold a significantly higher value. This asymmetry can be attributed to the fact that the storage battery manufacturing industry produces a final product. Consequently, within the upstream ego network, the industries that contribute the most considerable amount of resources are primarily those involved in producing raw materials. The resources provided by these industries are often more vital in the final product's manufacturing process than the resources that the storage battery industry itself can offer for other durable goods. The crucial connections within the downstream ego network occur between two durable, precisely finished goods. However, it is important to note that the term "finished goods" is relative, as what constitutes a seller's finished goods may subsequently serve as a buyer's raw materials. In the context of the downstream network, the storage battery can be considered a semi-finished good, as it functions as a component or spare part for the subsequent industry. Hence, the underlying reason for this asymmetry can be attributed to the disparity in the type of resources involved. Specifically, the discrepancy arises from the contrast between strictly classified raw materials, such as Nonferrous metal smelting and refining, and final products, such as Power-driven hand tool manufacturing.

Moving on to the *advanced ego production network*, each of the three categories of key actors I identified provides distinct perspectives and unique information regarding the ego network of storage battery manufacturing. The integrator and allocator industries play vital roles in revealing the fundamental industries through direct transactions within the network of storage battery manufacturing. Integrators are responsible for consolidating the highest amount of resources in the production network cycle of the storage battery industry, while allocators contribute the most to the ego network. On the other hand, mediator industries, which primarily engage in intermediary transactions, highlight the significant sectors in the middle of various supply chains integrated into the ego network.

Collectively, the integrator, allocator, and mediator industries are essential for the smooth operation of a production cycle. Any disruption in the functioning of these industries would have a profound impact on the overall stability of the cycle. If one of these key industries were to cease operations, it would create a ripple effect throughout the production network, leading to the loss of critical transactions that directly and indirectly impact the performance of the storage battery manufacturing industry.

Similar asymmetry is also evident in this context, and it can be attributed to the distinction between raw materials and processed materials/products when it comes to resource allocation and integration. The asymmetry between allocators and integrators arises because a few raw materials industries account for the majority of resources. Conversely, semi-final and final products generally integrate a similar amount of resources. In the case of integrators, there are no significant outliers like those observed for allocators. This observation holds when considering the magnitude of monetary transactions, not just their presence and frequency. Therefore, the distribution of financial transactions between integrators and allocators is largely similar. However, when comparing the distribution of allocated and integrated resources, a notable difference emerges due to the dominance of large raw material industries that allocate a significant proportion of resources within the ego network production cycle. Hence, a significant portion of resources is concentrated within a small number of allocator industries. However, when it comes to resource integration, there are no highly concentrated clusters of integrator industries. Even the notable integrators do not deviate as outliers to the same extent as the outstanding allocators, as they are relatively closer to the average regarding values.

In chapter 2, a similar pattern was observed for the entire national production network. The asymmetry in the national dimension also could be traced back to the difference between raw materials and finished goods. Raw materials pump resources of higher value to the production cycle. These are the critical sectors that dominate economic activity, the so-called ”commanding heights”. Vladimir Lenin used this phrase in the early 1920s, referring to the control of key segments of a national economy. In the context of the storage battery ego network, these are the allocator industries that stand out from the rest because of their high-value resource supplied. This asymmetry, along with the presence of allocator outliers, contributes to the scale-free nature of the production network from the supplying perspective. Therefore, it can be concluded that the "commanding heights" represented by these industries are the driving forces behind production and the economy.

Consequently, the contribution of this study can be summarised in two key aspects. Firstly, this chapter demonstrates the applicability of the production network framework developed in Chapter 2 when applied to the analysis of a specific industry and its production network. It serves as a case study showcasing how the system and network perspective on the economy and industry expansion, as discussed in the previous chapter, can be effectively utilised to examine a particular segment of the national production cycle, particularly the supportive environment surrounding one industry. Consequently, it highlights the distinctions between analysing a comprehensive national production network and a specific ego network. Secondly, it thoroughly explores the storage battery industry and its supply chain from the network science side on the basis of data from different production network perspectives. Maps the indispensable relations and industries in its production cycle.

Chapter 5

Ego network growth model: the case study of the storage battery industry

5.1 Introduction

This chapter adopts a quantitative approach to analysing the production network of the storage battery industry. I delve into the topic of its economic growth by examining the positioning of the storage battery industry within its fundamental production cycle. The chapter's research question is to what extent the growth of the industries in the storage battery supply chain can be determined solely by their characteristics and role in the storage battery ego network. Hence, concluding what are the most critical relational factors influencing industry growth.

The contribution of this chapter can be summarised in two ways. Firstly, this chapter shows the viability of the network growth framework developed in Chapter 3 when analysing one specific industry and its production network. Secondly, because of the functionality of this interconnected framework on the storage battery ego network, this chapter also contributes to the emphasis on the fact that by understanding how the economic sectors interconnect even in tight surroundings, it is more possible to explain their economic growth. It argues that if we consider the inter-industry relationships in the storage battery ego network and the role of the sectors in this storage battery production cycle, we can have valuable insights into the cause of storage battery industry growth and even into the growth of the sectors in close supplying relationship with the storage battery industry.

This chapter is organised as follows:

I discovered the question of industry growth in relation to their topological place in the storage battery core production cycle, such as the storage battery ego network. I start by explaining the strategy to analyse the industry expansion or reduction in the ego network. However, the model is built based on the regression model in Chapter 3; in this part, I clarify the difference between the growth model for the ego network and the growth model for the national production network. After, I present the overall results from running the storage battery ego network growth model and zoom in on the significant network centralities influencing growth. As the last part of this chapter, I zoom out and retake a look at the national production cycle and present the relevant metrics from the national network growth model for the storage battery industry.

This section aims to answer whether the location of an industry in the storage battery ego network topology can explain the macroeconomic dynamics of its growth and development. This research question has been asked on a national production network scale in Chapter 3, and I built a model to respond to it. It is the first model to approach the question of industry growth based only on the industry's place in the national production cycle. This section is based on that approach, concentrating particularly on the storage battery industry's ego network, a subset of the national one. It is almost equivalent methodologically; however, the goal is different. The purpose of this section is to analyse in what ratio the location and the role of an industry in the storage battery ego network explain the industry's economic increase or decrease.

5.2 Strategy to analyse industry growth in the ego production network

While the storage battery analysis methodology is almost the same as for economic growth examination in the national production network, in Chapter 3, the data used is partly different.

First, I need to build the 2007 and the 2012 storage battery ego network, for which I use the US input-output accounts data source presented earlier in both Chapter 2 and Chapter 3 (*Input-Output Accounts Data — U.S. Bureau of Economic Analysis* 2021). The individual industry growth metrics are calculated based on the data from the National Bureau of Economic Research (NBER) and the US Census Bureau's Center for Economic Studies (CES) Manufacturing Industry Database (Becker, Gray, and Marvakov, 2021), also shown in Chapter 3. While this data set contains extensive information only on manufacturing industries, I use only manufacturing industries for the ego network, too. Precisely industries with 3 as the NAICS code first digit, manufacturing durable and nondurable goods. Until this, the methodological approach is the same as in Chapter 3.

Following the construction of the manufacturing production network from the input-output accounts, I select a subset of the network, exactly the storage battery ego network. The selection process is similar to the one used in the section 4.5, as this will also be an advanced ego network. I selected the storage battery industry and all its direct industry contacts, both upstream and downstream. No matter the direction of the connection, all industries that directly transact with the storage battery manufacturing industry are selected. Afterwards, I add all other transactions between the different sectors in the ego network as links. Hence, the advanced ego network contains all transactions of the suppliers, too.

I get a very dense storage battery ego network with all kinds of transactions with different magnitudes in value. In order to get a more focused ego network, I exclude all very small-value transactions. After trying out several thresholds, I decided on the 0.0005 dollars cut-off value.

In Chapter 3, I got the outcome of the analysis that the threshold value of 0.001 dollars is the best when examining the entire national production network. The storage battery ego network is an ego network containing only a subset of the whole production network. Hence, by default, a big part of all transactions are excluded from the ego network. Therefore, I needed a threshold for the storage battery ego network that is less rough than the 0.001 dollars used in the national network. I also tried smaller thresholds, for example, the 0.0001 dollars cut-off. However, these don't make a difference; they cut only around 3-5%

Table 5.1 provides what ratio the storage battery ego network is of the national production network at the chosen threshold value of 0.0005 dollars.

Table 5.1: Number and ratio of nodes and edges in the storage battery ego network compared to the national network next to the 0.005 dollars threshold value (2007 & 2012)

Network type		2007 Nr	2007 Ratio	2012 Nr	2012 Ratio
National network	Nodes	156	100%	156	100%
	Edges	7133	100%	6803	100%
Ego network	Nodes	55	35%	60	38%
	Edges	1896	27%	2068	30%

When reaching the final status of the storage battery ego network, I compute network metrics that will be the independent variables in my model. These define every industry's location in the ego network in every possible way. By name, the network centrality metrics are: *In-degree, Out-degree, Degree, In-degree weighted, Out-degree weighted, Degree weighted, Betweenness, Betweenness weighted, Closeness, Outward Closeness, Closeness weighted, Outward Closeness weighted, PageRank, Reversed PageRank, PageRank weighted, Reversed PageRank weighted, Eigenvector, Reversed Eigenvector, Eigenvector weighted, Reversed Eigenvector weighted*. For a comprehensive description of every measure individually, see Chapter 3 subsection subsection 3.2.4 or for a summary of every metric, see Table 3.2.

The next step is to compute the growth rates that will be the dependent variables. Here, I also use the change in the VADD (total value added in one million dollars) variable presented in the NBER-CES Manufacturing Industry Database to describe industrial growth and expansion. I use the same approach as in Chapter 3 by calculating four different growth rates for every industry, both in 2007 and 2012 too, in order to choose the best one and check the model's performance at every one of them. For the equations behind the computation, see Equations 3.1, 3.2, 3.3, 3.4.

Growth rate R_1 in Eq. 3.1 is the industry's value-added of the year in question directly compared to the value 5 years before. Growth rate R_2 in Eq. 3.2 compares the year in the inquiry industry's value-added to its average value for the last 5 years. Growth rate R_3 in Eq. 3.3 is the industry's average annual growth rate in the last 5 years, including the particular year. Growth rate R_4 in Eq. 3.4 is the industry's average annual growth rate in 10 years when the year at issue is in the middle of that 10 years.

Finally, the linear regression model was carried out with the storage battery ego network topology metrics as independent variables and industry growth rates as dependent variables, specified in Equation 3.5. Also, Table 3.2 provides all variables and their descriptions: the network metrics as independent variables and the growth rates as dependent variables.

5.3 Results of the storage battery ego network growth model

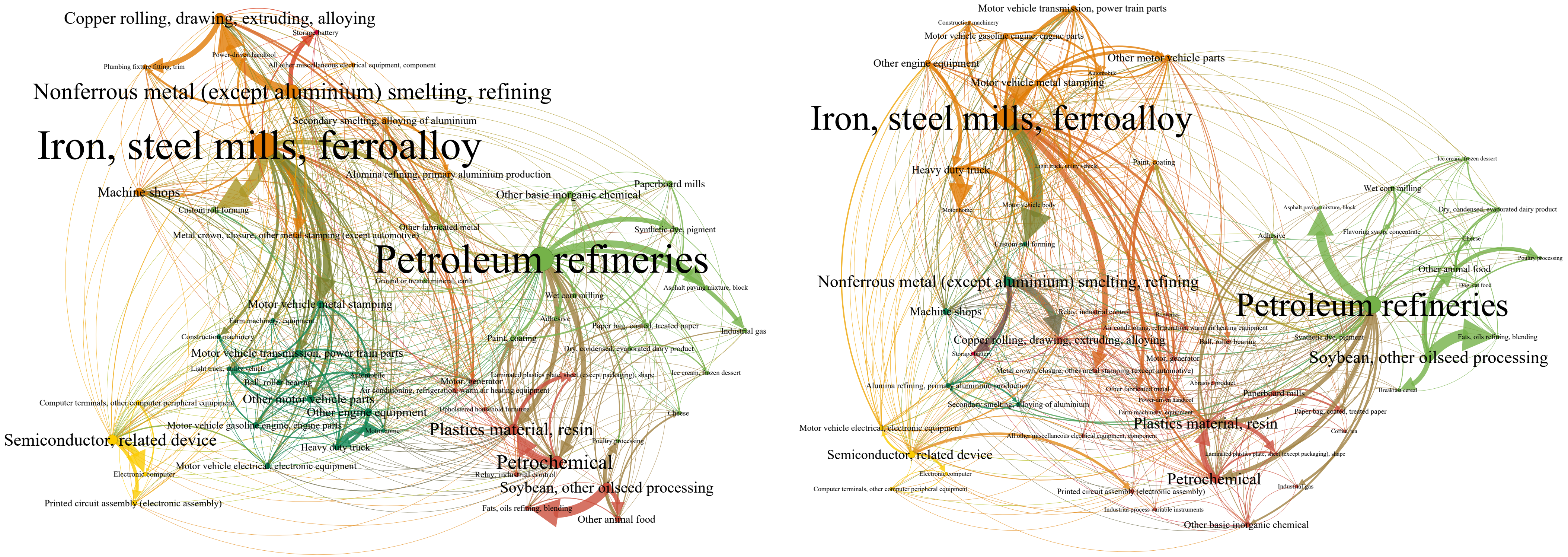
Figure 5.1 shows the storage battery ego network built for the network growth model. Figure 5.1a is the ego network in 2007, while Figure 5.1b is the one in 2012. In both networks, the size of the node represents the out-degree weighted metric. Hence, it is equally proportional to the magnitude of the resource pumped into the storage battery ego network. The colouring of the network illustrates the found modularity and community structure in the network (Blondel et al., 2008, Lambiotte, Delvenne, and Barahona, 2014).

In the 5-year interval between 2007 and 2012, there have been some small changes in the storage battery production cycle at the level of the manufacturing interconnections and manufacturing ego network, too. While some might be visible on the network figures, these changes are minor. Table 5.2 quantifies a few properties of both networks and summarises the change between these values. There is a 9% increase in both the number of industries and transactions. The same pattern was observable when mapping the advanced storage battery ego network in section 4.5. Also, network density decreased slightly but still was above 50%, and moderately higher than for the advanced ego network.

Table 5.2: Storage battery manufacturing ego network for the growth model (2007 & 2012)

Network feature	2007	2012	Change
Industries (nodes)	55	60	+9%
Transactions (connections)	1896	2068	+9%
Network density	0.638	0.584	-8%
Diameter	5	5	0%
Average path length	1.43	1.63	+14%

Having discussed how to construct the storage battery ego network for the growth model and how does this network look like in 2007 and 2012, I will now move on to examine the growth model results. Firstly, as in Chapter 3, I aim to choose the best-performing linear regression model based on Eq. 3.5. Table 5.3 compares the four model accuracy scores that I get when running the model with the four different growth rate types. By the model score, I indicate the R-square of the linear regression.



(a) 2007

(b) 2012

Figure 5.1: Storage battery industry advanced manufacturing-industry ego network (2007 & 2012)

Table 5.3: Ego network growth model accuracy scores based on growth rate type (0.0005 dollars threshold)

Growth rate	Model score
R_1	0.54
R_2	0.38
R_3	0.51
R_4	0.46

The table reveals that all values far surpass zero. This result further supports the idea that regardless of the combination chosen, the industrial expansion or decrease illustrated by different growth rates can be attributed partially to topological metrics. Thus, when focusing only on a portion of the entire national production, specifically on storage battery manufacturing, the location and functions of the sectors in this production cycle play a critical role in their growth.

The best outcome is at the growth rate R_1, 0.54. However, both growth rates R_1 and R_3 give us above 50%. While growth rate R_1 is the industry value-added of the year in question compared directly to the value 5 years before, R_3 is the average annual growth rate for the last 5 years. I choose to stick with the growth rate R_3 because R_1 does not filter out outstanding years in the interval. This result implies that by solely examining the network metrics, we can describe 51% of the industry growth as measured by the change in (real) value-added.

This is a rather interesting and remarkable outcome, stating that generally, half of the industry's success depends on its place in the close surrounding production network, on the near key interfaces of the industry. Perhaps the most striking finding is that these model scores are significantly higher than those of the growth model for the national production network in Chapter 3. When running the model for the entire national production cycle, the explanatory power of network centrality and topological metrics was very close but didn't reach 30%. However, in the storage battery ego network model, even the smallest model score with growth rate R_2 is 38%. For context, the national model with the same growth rate had an outcome of 14%. Table 5.4 shows the best results for the national production network growth model (that is at 0.001 dollars threshold) compared to the storage battery ego network growth model accuracy scores.

Table 5.4: Ego network growth model accuracy scores compared to national production network growth model scores

Growth rate	Ego score (0.0005 threshold)	National score (0.001 threshold)
R_1	0.54	0.29
R_2	0.38	0.14
R_3	0.51	0.29
R_4	0.46	0.27

It seems possible that these results are due to the fact that the independent variables are from an ego network. An ego network is somehow a ”core” production network for a specific industry, in this case, for the storage battery manufacturing industry. It is logical that the relations in the ”core” production network have more explanatory power for that condition than the general ones in the national inter-industry network. Consequently, the closer proximity of an industry influences its growth more than its role in the national production cycle.

5.3.1 Factors influencing growth

As for now, from the network growth model analysis between 2007 and 2012, we know that all together, individual topological features in the storage battery ego network explain more than half of the industries present in the ego network decrease or increase. In order to zoom in on this 51% influential value that topological features have, I must analyse what independent variables stand out in the linear regression. In this way, this 51% will be broken down into actual functions, and this will tell us which network centrality metrics are the most meaningful and, thus, what type of supply role is the most significant when trying to predict industry growth. All network centrality measures as independent variables shed light on different functions an industry has in the ego network, hence, in the storage battery manufacturing cycle.

Table 5.5: Storage battery industry (manufacturing) ego network growth model descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Growth rate R_3	115	0.041	0.083	-0.241	0.456
In-degree	115	34.470	5.463	13	44
Out-degree	115	34.470	21.509	0	59
Degree	115	68.939	20.358	25	99
In-degree weighted	115	0.412	0.194	0.028	0.939
Out-degree weighted	115	0.412	0.691	0	3.651
Degree weighted	115	0.825	0.659	0.148	3.900
Betweenness	115	0.009	0.015	0	0.097
Betweenness weighted	115	0.037	0.047	0	0.199
Closeness	115	0.636	0.070	0.448	0.782
Outward closeness	115	0.729	0.270	0	1
Closeness weighted	115	0.003	0.001	0.002	0.004
Outward closeness weighted	115	0.006	0.006	0	0.042
Eigenvector	115	0.130	0.021	0.050	0.172
Reversed eigenvector	115	0.109	0.074	0	0.184
Eigenvector weighted	115	0.094	0.093	0.005	0.615
Reversed eigenvector weighted	115	0.051	0.122	0	0.787
PageRank	115	0.017	0.015	0.005	0.083
Reversed PageRank	115	0.017	0.010	0.003	0.029
PageRank weighted	115	0.017	0.024	0.004	0.140
Reversed PageRank weighted	115	0.017	0.032	0.003	0.200

Linear regression analysis was used to predict growth based on network centrality metrics (Eq. 3.5). Table 5.5 shows the descriptive statistics for all independent variables in the ego network growth model and for the dependent variable chosen, growth rate R_3.

There are 115 observations, as there are 55 industries with all different topological metric values in the 2007 storage battery ego network, while there are 60 industries in 2012 one.

The independent variable, growth rate R_3, is the industry's average annual growth rate in the last 5 years, including the particular year. By the growth rate, I mean the change in the VADD (total value added in one million dollars). For the industries present in the 2007 ego network, it is the average annual growth rate between 2002 and 2007, while for the 2012 ones is the average annual growth rate in the interval of 2007 and 2012.

The growth rate equation 3.3 is presented in Chapter 3. The mean of this growth rate for the years 2007 and 2012 is 4.1%, meaning that there is, on average, a slow, stable growth between the years 2002 and 2012. However, this is only the average, and there might be outlier values, as this a quite problematic period with having in the middle of it the 2008 economic crisis. The increase in the first period (2002-2007) is higher on average than in the second one (2007-2012). Also, the 5-year interval growth rate in both cases tries to reduce the effect of outstanding years. A consequence of the 2008 crisis in the data is that the minimum outstanding value is from the 2007-2012 period with an extreme value of -24.1% decrease. Also, in this period, there was no possibility to expand that much; therefore, the maximum growth rate is from the 2002-2007 period with a 45.6% increase.

Also, the same asymmetry appears in the descriptive statistics table 5.5 as in section 4.5 between the in-degree weighted and out-degree weighted distributions, also shown in Table 4.7, 4.8 and 4.9.

In order to have a clearer picture of what all these measures mean and which industries could be the leaders in which network centrality metric, I show the sample of the most connected sectors according to all analysed network measures in the storage battery ego network. Table 5.6 provides the top industries for the 2007 ego network, while Table 5.7 for the 2012 production cycle.

Table 5.6: Sample of the most connected industries in the storage battery ego network based on all analysed network measures (2007)

Measure name	Manufacturing industry with the highest value	Highest measure value
In-degree	Motor home manufacturing	44
Out-degree	Petroleum refineries (+11 others equal)	54
Degree	Other motor vehicle parts	94
In-degree weighted	Light truck, utility vehicle manufacturing	0.792
Out-degree weighted	Iron, steel mills, ferroalloy	3.521
Degree weighted	Iron, steel mills, ferroalloy	3.747
Betweenness	Other animal food manufacturing	0.054
Betweenness weighted	Iron, steel mills, ferroalloy	0.135
Closeness	Other animal food	0.782
Outward closeness	Petroleum refineries (+11 others equal)	1.0
Closeness weighted	Ice cream, frozen dessert	0.004
Outward closeness weighted	Petroleum refineries	0.042
Eigenvector	Motor home manufacturing	0.172
Reversed eigenvector	Motor home manufacturing	0.183
Eigenvector weighted	Motor home manufacturing	0.524
Reversed eigenvector weighted	Motor home manufacturing	5.596
PageRank	Cheese manufacturing	0.069
Reversed PageRank	Other motor vehicle parts	0.029
PageRank weighted	Dry, condensed, evaporated dairy product	0.140
Reversed PageRank weighted	Petroleum refineries	0.200

Table 5.7: Sample of the most connected industries in the storage battery ego network based on all analysed network measures (2012)

Measure name	Manufacturing industry with the highest value	Highest measure value
In-degree	Light truck, utility vehicle	44
Out-degree	Plastics material, resin (+9 others)	59
Degree	Other motor vehicle parts	99
In-degree weighted	Motor home	0.939
Out-degree weighted	Iron, steel mills and ferroalloy	3.650
Degree weighted	Iron, steel mills and ferroalloy	3.900
Betweenness	Other animal food	0.097
Betweenness weighted	Machine shops	0.198
Closeness	Cheese + dog, cat food	0.734
Outward closeness	Other motor vehicle parts (+9 others)	1.0
Closeness weighted	Dog, cat food	0.003
Outward closeness weighted	Petroleum refineries	0.040
Eigenvector	Light truck, utility vehicle	0.162
Reversed eigenvector	Other motor vehicle parts (+9 others)	0.184
Eigenvector weighted	Motor home	0.615
Reversed eigenvector weighted	Petroleum refineries	0.739
PageRank	Cheese	0.083
Reversed PageRank	Other motor vehicle parts	0.028
PageRank weighted	Dry, condensed, evaporated dairy product	0.137
Reversed PageRank weighted	Petroleum refineries	0.186

The linear regression results are shown in Table 5.8. I explained comprehensively in Chapter 3 all network metrics and their implications for the national production network. However, to interpret the results of the growth model of the storage battery ego network, I will summarise the meaning of all significant variables again and highlight their implications for predicting industry growth.

The explanatory variable *Reversed PageRank weighted* is the most significant (p-value < 0.01), having a quite high positive effect on industry growth. The PageRank network centrality (Page et al., 1999; Langville and Meyer, 2004) is an eigenvalue-based measure with the reason that an industry can be considered strategically crucial if its immediate neighbours are indispensable or if the neighbouring industries of its neighbours hold significance. This metric takes into consideration the origins of an industry's suppliers, including their popularity, location in the storage battery ego network, and overall importance. It operates under the assumption that the influence of neighbouring industries matters not only in terms of their quantity but also their level of significance. The unique characteristics of trading industries can either amplify or moderate an industry's impact on others and the entire ego network. As a result, industries rely on each other both directly and indirectly. Within this framework, centrally located industries are those that require a diverse range of resources and are supplied by sectors that also rely on a variety of resources. Interestingly, we might intuitively assume that centrally located industries are not the most "dependent." Therefore, I also calculate the perspective of out-edges, known as Reversed PageRank. This perspective allows me to identify the centrally located industries that supply the highest amount of resources to the ego network, particularly to those suppliers that are still crucial within the production cycle. Both PageRank weighted and Reversed PageRank weighted follow a similar logic but also take into account the magnitude of monetary transactions. In PageRank weighted, centrally located industries require the greatest value of resources from demanding suppliers (regardless of diversity). Conversely, in Reversed PageRank weighted, key industries supply the highest amount of monetary transactions to those suppliers that are still essential within the storage battery ego network.

Table 5.8: Regression for industrial growth results using the storage battery ego network framework (Eq. 3.5)

Name	Coefficient
In-degree	-0.002
Out-degree	-0.001
Degree	-0.002
In-degree weighted	0.031
Out-degree weighted	-0.045*
Degree weighted	-0.014
Betweenness	0.233
Betweenness weighted	0.385**
Closeness	-0.485
Outward closeness	0.134*
Closeness weighted	-59.408*
Outward closeness weighted	-12.823**
Eigenvector	2.984
Reversed eigenvector	2.648*
Eigenvector weighted	-0.175*
Reversed eigenvector weighted	-0.025
PageRank	0.754
Reversed PageRank	-20.359*
PageRank weighted	-0.372
Reversed PageRank weighted	4.913***
Year	-0.023***
Constant	47.132***
R-sq	0.511
N	115

Significance level: *p<0.2 **p<0.05 ***p<0.01

Consequently, Reversed PageRank weighted being significantly positive, it is important when considering growth, what other industries an industry supports and by how much. In light of this measure, it is advantageous for the advancement of the industrial sector to provide assistance to other central industries in the ego network that possess significant value. Thus, forecasting the trajectory of an industry becomes more manageable when it occupies a vital position in the storage battery ego network by supporting other high-value industries, which, in turn, support additional industries of high value. In this scenario, the assessment of its transformation or expansion can be more precisely estimated based on its position in the ego network. The key factor is not the number of connections it has but rather the stability derived from supplying a few essential industries, even if those connections are relatively fewer in number. Also, this centrality metric was the most important in the national network growth model too.

The second most significant independent variables are *Betweenness weighted* and *Outward closeness weighted*. Both with a p-value of less than 0.05, however, in opposite directions. Betweenness weighted is significantly positive, while Outward closeness weighted is significantly negative with a weighty coefficient.

Betweenness centrality (Freeman, 1977; Easley and Kleinberg, 2010) operates on the premise that an industry's significance stems from its function as a mediator within supply chains. When the shortest route between two industries passes through a third industry, the intermediary sector becomes influential and influential in determining the outcome. Similarly, Betweenness weighted centrality operates under the same assumption, but it also takes into account the magnitude of monetary transactions. In this case, we assume that the higher the value of the transaction, the more crucial the link between the supplying industries. Consequently, the connection distance becomes shorter, indicating a closer relationship between the two industries in the storage battery ego network. Hence, being a mediator in the ego network is beneficial for an industry's growth.

The other intricate metric, closeness centrality, posits that an industry occupies a central position if others can readily and swiftly access it in the ego network. It quantifies the number of steps required for all other industries in the storage battery production cycle to reach the specified industry relative to a scenario where all sectors are only one step away from the industry in question. The highest value indicates that all sectors are direct suppliers to the industry under examination. When examining the outbound perspective of this metric, outward closeness centrality assumes that an industry holds a central position in the ego network when it can effortlessly and promptly reach other industries. In this case, the maximum value signifies that the analyzed industry directly supplies all other sectors. When considering the weights of the links, closeness weighted centrality and outward closeness weighted centrality also incorporate the magnitude of monetary transactions. Therefore, it negatively affects growth if an industry in the storage battery ego network easily reaches every other sector.

Other somehow significant metrics are the *out-degree weighted*, *outward closeness*, *Reversed eigenvector*, *Eigenvector weighted* and *Reversed PageRank* with p-values less than 0.2.

Eigenvector (Newman, 2018) is a very similar measure to PageRank with the exact reason that industry is systemically vital if its neighbours are crucial and/or the neighbours of the neighbours are noteworthy. I then computed Reversed Eigenvector centrality, Eigenvector weighted centrality, and Reversed Eigenvector weighted centrality with the same logic used at PageRank centralities. In the storage battery ego network growth model Reversed eigenvector is significantly positive; however, Eigenvector weighted is significantly negative. This tells us that supporting other central industries within the ego network contributes to the growth of the supporting industry. Consequently, when an industry occupies a crucial position within the ego network, providing resources to other high-value industries, it becomes easier to predict its expansion. However, if the industry in question is supported by industries with a lot of influence, it can be to the detriment of the industry.

The out-degree weighted measure is the amount of resources one industry supplies directly to other industries. In the ego network linear regression model, it is significantly negative, stating that the more resources an industry provides and allocates, the more volatile that industry is.

5.3.2 Storage battery manufacturing industry in the national production network growth model

For the purpose of connecting even more to the national production network growth model and to see if that model has relevant outcomes for the storage battery industry, I retrieved the growth rates and the network centrality metrics of the storage battery industry. Table 5.9 provides the growth rates for 2007 and 2012, while Table 5.10 the network centrality metric values of the storage battery industry node in the 2007 and 2012 national production network. The national production network uses a threshold value of 0.001 dollars, way higher than the storage battery ego network, because there are much more transactions and industries included in the national production network. This means that all monetary transactions with a value of less than 0.001 dollars are excluded from the network.

Table 5.9: Storage battery industry growth rates in the national production network growth model (2007 & 2012)

Growth rates	2007	2012	Change
R_1	0.702	0.244	-65%
R_2	0.546	0.129	-76%
R_3	0.121	0.053	-56%
R_4	0.086	0.034	-60%

What stands out clearly in the growth rate table above is the striking difference between the 2007 and 2012 growth rates. There is a rapid decrease in the growth rates, and a quite logical explanation for this result is that it is due to the 2008 economic crisis. While the 2012 rates include the 2008 period, the 2007 growth rates end exactly before the crisis. Although even with this enormous difference at both timestamps, the storage battery industry VADD (total value added in one million dollars) has increased at all four types of rates.

Table 5.10: Storage battery industry centrality metrics in the national production network growth model (2007 & 2012)

Network metrics	2007	2012	Change
In-degree	22	24	+9%
Out-degree	15	19	+26%
Degree	37	43	+16%
In-degree weighted	0.431	0.454	+5%
Out-degree weighted	0.059	0.074	+25%
Degree weighted	0.486	0.528	+9%
Betweenness	0.0007	0.0017	+143%
Betweenness weighted	0	0	0%
Closeness	0.347	0.343	-1%
Outward closeness	0.447	0.444	-1%
Closeness weighted	0.0018	0.0019	+6%
Outward closeness weighted	0.0028	0.0027	+4%
Eigenvector	0.053	0.061	+15%
Reversed eigenvector	0.0097	0.0103	+6%
Eigenvector weighted	0.039	0.036	-8%
Reversed eigenvector weighted	0.0008	0.0014	+75%
PageRank	0.0030	0.0031	+3%
Reversed PageRank	0.0026	0.0028	+8%
PageRank weighted	0.0029	0.0029	0%
Reversed PageRank weighted	0.0013	0.0014	+8%

Closer inspection of Table 5.10 showing the storage battery industry centrality measure values in the national production network tells us that there was a development in the role of the storage battery industry in the production cycle between 2007 and 2012. The observed increase in the metrics could be attributed to the assumption that as time passes, storage battery becomes more and more crucial for the well functioning of the economy.

As in-degree increased with 9% and in-degree weighted with 5%, the storage battery industry consolidated slightly more suppliers and higher value upstream transactions in 2012 than in 2007. This contributes to its role as an integrator in the national production cycle, which narrowly increased in this 5-year interval. However, its role as an allocator in the national production network, as well as in the ego network, is way smaller and increased even more from 2007 to 2012. The measures behind the role, out-degree and out-degree weighted, increased by 26% and respectively 25% between 2007 and 2012. While this is still way less in actual real value compared to the in-degree and in-degree weighted. In 2012, the storage battery industry had 24 direct upstream suppliers compared to allocating directly to 19 downstream industries. Also, in 2012, this industry integrated 0.454 dollars in resources; however, it pumped only 0.074 dollars into the national production network.

The highest increase is in its role as a mediator in the production chains. The value of betweenness has risen 143% from 0.0007 to 0.0017 between 2007 and 2012. This means that in 2007, the storage battery industry was present on 0.07% of all shortest paths in the national production network. While in 2012, it acted as an intermediary on 0.17% of all shortest supply chains in the national production cycle. Whereas this might seem a small percentage, it is actually in the middle of the betweenness scale. In the national production network, there are 156 industries with 4651 transactions above the value of 0.001 dollars, and the diameter of the network is 7 transactions, while the average path length is 2.5 transactions.

5.4 Discussion

This chapter examined the storage battery industry using a quantitative approach focusing on the industry's production network interconnections. The general purpose was to explore the storage battery supply chain through the lens of network science on the basis of systematic data and bring additional perspective to this research area.

Moreover, specifically, this chapter of the thesis focused on investigating the economic growth of the storage battery industry through an exploration of its position within the core production cycle. The primary research inquiry was to what degree the growth of industries within the storage battery supply chain can be attributed to their inherent characteristics and their role within the storage battery ego network. As a result, I derived conclusions regarding the most significant relational factors that influence industry growth.

The accuracy score of the linear regression model for various growth rate combinations consistently exceeded zero. This finding strongly supported the notion that irrespective of the selected combination, the observed industrial expansion or decline, as indicated by different growth rates, can be attributed to topological metrics, at least in part. Therefore, when examining a specific segment of the overall national production, such as storage battery manufacturing, the position and functions of sectors within this production cycle significantly influence their growth trajectory.

The notable outcome indicated that by solely analysing network metrics, we are able to explain 51% of the industry growth, as measured by the change in real value-added. This finding is both intriguing and remarkable, as it suggests that a significant portion, approximately half, of an industry's success can be attributed to its position within the immediate production network and its proximity to key interfaces in the industry.

One of the most remarkable discoveries is the substantial advancement between the model scores obtained in this study and those of the growth model for the national production network discussed in Chapter 3. In the analysis of the entire national production cycle, the explanatory ability of network centrality and topological metrics was close to but did not exceed 30%. However, in the case of the storage battery ego network model, the score reached an impressive 51%.

These findings may be attributed to the nature of the independent variables derived from an ego network. In this case, an ego network represents a "core" production network specifically tailored to the storage battery manufacturing industry. It is reasonable to assume that the relationships within this "core" production network hold greater explanatory power for industry growth compared to the broader connections within the national inter-industry network. Consequently, the proximity and connections within the industry's immediate production cycle have a more substantial influence on its growth trajectory than its role within the larger national production cycle.

Consequently, the contribution of this study can be summarised in two key aspects. Firstly, this chapter demonstrates the applicability of the network growth model framework developed in Chapter 3 when applied to the analysis of a specific industry and its production network. Secondly, this chapter contributes to the emphasis on the interconnectedness of economic sectors within the storage battery ego network. By comprehending the interrelationships among these sectors within the close-knit environment, valuable insights can be gained regarding the growth of the storage battery industry itself, as well as the growth of sectors closely associated with it through supply relationships. This highlights the significance of considering inter-industry connections within the storage battery ego network and the roles played by sectors in this production cycle.

Chapter 6

Conclusions

6.1 Overview of principal findings

In this thesis, I have undertaken a comprehensive examination of production networks and their implications on industrial growth, utilising a quantitative approach and data-driven analyses. In Chapters 2 and 3, I explored various facets of production networks, and in Chapters 4 and 5, I focused on the specific case of the storage battery industry. In this discussion section, I will synthesise the key findings and contributions of each chapter, highlighting their significance and broader implications.

In **Chapter 2**, I delved into the influence of threshold values on production network topology. The hypothesis, centred on changing the edge cut-off threshold and its impact on network structure, was confirmed through an analysis of the US input-output accounts data for 2007 and 2012. I found that the production network's topology was highly sensitive to this threshold, with distinct changes in network characteristics and core industries as the threshold value varied.

One critical insight was the exposure of the industry inter-dependence network to the edge cut-off threshold value. Under very low thresholds, the network did not conform to traditional random or scale-free network structures, demonstrating the intricate interplay of monetary transactions within the production ecosystem. However, at slightly higher values, the production network took the shape of a scale-free network from the resource-supplying side (out-degree). Furthermore, the analysis of node-level threshold susceptibility revealed that core industries, which serve as hubs, lost their central positions as the threshold increased. These core industries played pivotal roles in supporting numerous other sectors through small monetary transactions. Neglecting these seemingly minor transactions could distort propagation mechanisms and affect resilience in the face of disruptions.

The chapter's findings emphasise the non-linear nature of these interdependencies. Even industries contributing small percentages to others can have a profound impact on the entire production chain. This chapter's limitations can be addressed in future work by examining resilience aspects and exploring the distribution fitting methods. In conclusion, Chapter 2 underscores the critical importance of carefully selecting the threshold value ζ in production network analysis and calls for a deeper exploration of concrete propagation mechanisms in the context of threshold modification.

Chapter 3 shifts the focus to the relationship between network characteristics and industrial growth. By analysing the production network's role in explaining growth, the thesis made a pioneering contribution to the field. The results revealed that network characteristics can explain up to 30% of industrial growth, highlighting the significance of the production network framework as an explanatory variable. A key finding was that industries benefit more from supplying central and essential industries with high-value monetary transactions rather than simply having numerous connections. A limitation mentioned is that the value of transactions between industries in the production network did not necessarily correlate directly with essentiality. However, at this point, the only inter-industry relational data that we have at this detailed level are the monetary transactions between sectors.

This chapter's contribution lies in shedding light on the distinctive role of production network metrics in explaining industrial growth and enhancing growth research to include the systematic perspective in its research direction. It prompts further research to delve deeper into the specific sectors' growth patterns within the network.

Chapter 4 provided a detailed examination of the storage battery industry through ego network analysis. I explored the supply chain interconnections in the framework of production network and identified integrator, allocator, and mediator industries within this sector. The analysis revealed a growing tendency in supply amount, emphasising the increased interconnectivity among industries.

I also observed an asymmetry between upstream and downstream networks in terms of monetary transactions attributed to the nature of raw materials and final products. This reflects the industry's role in the production cycle, where it serves as a final product in the upstream network but a semi-finished product in the downstream. Additionally, the chapter demonstrated how different categories of key actors (integrators, allocators, mediators) provide unique insights into the ego network, showcasing their roles in the storage battery production cycle. Any disruption in these key industries can have a ripple effect throughout the network, highlighting their importance.

Overall, this chapter highlights the utility of ego network analysis in examining specific industry supply chains and emphasises the critical role played by key industries and key relations within these networks.

In **Chapter 5**, I applied the insights gained from earlier chapters to a case study of the storage battery industry. I explored the degree to which the industry's growth could be attributed to its position within the core production cycle. The results were striking, revealing that network metrics alone could explain 51% of the industry's growth, a substantial improvement over previous models applied to the entire national production network. This finding underscores the importance of examining industry-specific ego networks to understand growth dynamics better. It suggests that proximity and connections within an industry's immediate production cycle have a more significant impact on its growth trajectory than its role within the broader national production network.

In conclusion, this chapter demonstrates the applicability of the network growth model framework to specific industries and underscores the significance of considering inter-industry connections within a production cycle.

In this thesis, I have explored the intricate relationships between production networks and industrial growth. The findings emphasise the critical role of threshold values in network analysis, the impact of network characteristics on growth, and the importance of examining industry-specific ego networks. This thesis advances our understanding of production networks' influence on industrial growth and offers valuable contributions to the field of economic research.

6.2 Implications

Beyond the specific findings within each chapter, this thesis contributes to a broader understanding of the role of production networks in shaping industrial dynamics and economic growth. The implications of the research extend to both academic discourse and policy considerations, providing valuable insights for policymakers and researchers alike.

1. Bridging the gap in production network research

The thesis bridges a crucial gap in the existing literature by shedding light on the impact of threshold values on production network analysis. While prior studies have recognised the significance of production networks, this work emphasises the need for careful consideration of threshold values when studying these networks. This insight has the potential to reshape the way researchers approach and interpret production network data.

2. Rethinking traditional economic metrics and models

The findings challenge traditional economic metrics for assessing the importance of industry connections. The research suggests that industry relations contribute significantly to industrial economic growth. This has implications for policymakers and economists who rely on standard metrics to gauge economic interdependencies.

3. Industry-specific insights

The thesis highlights the importance of examining industry-specific ego networks to gain a deeper understanding of supply chain dynamics. I explore the storage battery industry; however, this approach can be extended to various sectors, offering tailored insights that can inform targeted interventions and policies.

4. Policy Implications

a. Resilience and risk management: The research underscores the importance of maintaining the resilience of core industries within production networks. Policymakers should consider strategies to safeguard key industries and ensure the continuity of essential transactions, especially during times of crisis or disruption.

b. Industrial growth strategies: The thesis reveals that industry growth is significantly influenced by its position within the production cycle, especially in the close-related production cycles. Policymakers can use this insight to develop growth strategies that focus on strengthening industry connections within the immediate production network, potentially leading to increased economic stability and growth.

c. Threshold considerations: Researchers and policymakers should be cautious when selecting threshold values in production network analysis. Different thresholds can yield vastly different results and conclusions. Standardised approaches for threshold selection or sensitivity analyses can enhance the robustness of research findings.

d. Supporting key industries: Identifying and supporting key industries within production networks can be a strategic policy move. These industries play pivotal roles in the production cycle, and targeted policies can ensure their stability and growth.

e. Economic metrics reform: The findings regarding the limitations of traditional economic metrics suggest a need for reevaluation. Policymakers and economists may consider developing new metrics that better capture the importance of industry connections within production networks.

5. Future Research Directions

The thesis opens the door to several avenues for future research. Further analysis can delve into concrete propagation mechanisms, resilience aspects, and distribution fitting methods within production networks. It can explore the stability of production networks and the impact of disruptions on industry dynamics. This analysis can inform policies aimed at enhancing supply chain resilience. Additionally, exploring industry-specific ego networks for other sectors can provide valuable insights for a comprehensive view of supply chain dynamics across the economy and also targeted interventions and policies during economic fluctuations. Evaluating the effectiveness of policies designed to strengthen industry connections and support key sectors within production networks can provide valuable guidance to policymakers.

Notwithstanding the relatively limited and more than ten years old interval of the data used in this thesis, this work offers valuable insights into production network research. However, it would be beneficial to examine longer intervals and newer timestamps as soon as they are available. Also, it would be useful to gather data on transaction costs and market power, specifically when analysing an ego production network. Moreover, to tailor new methods and visualisation techniques to represent and interpret the information in the most accurate way. Also, it would be worthwhile to move towards international standardisation by using the conclusions learned in the detailed level US data.

From specifically the academic side, more extensive studies can investigate the sensitivity of different production networks to threshold values, offering insights into the generalisability of our findings. Also, researchers can develop and test new economic metrics that better capture the nuances of industry connections within production networks.

Exploring further research directions, particularly in the realm of clusters and regional economic development, presents an intriguing avenue for extending the insights gained from the production network framework. The emphasis on proximity and connections within industries' immediate production cycles highlights the potential to reshape regional economic development strategies by leveraging local clusters as drivers of economic growth. This inquiry becomes especially pertinent in the context of cities undergoing critical strategic transformations to emerge as regional hubs of development. By delving into the dynamics of supply chains and expanding the scope to encompass industries beyond traditional manufacturing, such as the cultural industry, a more comprehensive understanding of cluster formation and interaction can be achieved. This approach not only sheds light on the geographical distribution of industries within regions but also explores the intricate relationships that underpin cluster structures and interactions.

Leveraging insights from economic theory on clusters and regional economic growth offers a robust framework for identifying essential industries and activities that contribute most significantly to cluster formation and regional growth. Moreover, by incorporating considerations of urban and national regeneration, as well as resilience, this research direction holds promise for informing policy interventions aimed at fostering sustainable and inclusive economic development.

Another direction could explore network motifs, which serve as the building blocks of complex networks, offering insights into recurring patterns and configurations within the dataset. This approach holds promise for uncovering hidden dynamics and elucidating mechanisms driving industries' economic interactions and growth trajectories.

Further study could explore blockmodeling techniques as key methodological tools for analysing input-output tables. By consulting literature on blockmodeling and structural equivalents, researchers can gain theoretical insights into the network theory surrounding economic growth. This approach may help bridge the gap between the vocabulary used in framing the research related to economic growth and the theoretical concepts underlying network analysis.

Exploring the potential of scenario simulations to analyse various dynamics within production networks could offer practical tools for policymakers and practitioners seeking to optimise for evolving conditions. By simulating different propagation scenarios and understanding nonlinear relationships, researchers can gain insights into how dynamics unfold and circular patterns emerge within the network. Additionally, investigating the impacts of changing technology on production networks could be crucial. Assigning values or labels to links based on technology could facilitate understanding how technological advancements, such as AI upgrades or carbon restrictions, influence linkages within the network. Moreover, exploring how innovation in technologies affects industries and their diffusion could provide valuable insights into industry dynamics and network behaviour, offering alternative measures for understanding technological improvements within the production network framework.

In conclusion, this thesis contributes significantly to our understanding of production networks and their implications for industrial growth. The findings have both academic and policy relevance, prompting a reevaluation of traditional economic metrics and emphasising the importance of network characteristics in economic analysis. As we continue to navigate an increasingly interconnected global economy, a deeper understanding of production networks will be crucial for shaping effective economic policies and strategies.

6.3 Postface and the Material Social Futures Doctoral Training Centre

As I reflect on my journey in the Material Social Futures Doctoral Training Centre (MSF), it's only fitting to look back on the bigger picture that this institution has encouraged us to consider. The MSF was created with a simple yet crucial aim: to place specific techniques within a broader context. In my case, that meant delving into economic network science, a tool I've used to investigate and understand complex economic relationships. However, it's important to remember that economics is just one piece of the puzzle that is our society. Economics, while vital and illuminating in its own right, is but a single facet of our intricate societal structure. It would be short-sighted to believe that economics can solve everything; such thinking limits our understanding of the complex interactions that shape our world. The MSF's mission was to equip us with the tools and techniques necessary to perceive our research through a more expansive lens, transcending the boundaries of our respective disciplines.

Looking back on my research journey, I'm now confident in my ability to conduct network analyses, and with the knowledge and skills I've gained at the MSF, this proficiency allows me to collaborate with experts and others interested in this topic, such as sociologists and environmental scientists. Network science, as I initially employed in economics, can be adapted to explore relationships, not just economic transactions.

In reality, the economic transactions I initially studied are just one part of the complex relationships between industries and companies. Production networks built from monetary transactions, while great at addressing specific questions within a particular scope, are essentially a representation of specific transactional relations. However, the scale of relations is very wide. This encourages us to explore the broader realm of material social futures, where networks don't revolve solely around economic exchanges but represent a framework for understanding our society's intricate connections.

As I place this thesis within the MSF, I'm thankful for the intellectual growth and expanded perspectives I've gained during my time here. It has empowered me to think about future research that embraces the interdisciplinary nature of material social futures, moving beyond economics to explore the vast complexities that shape our ever-evolving society. The MSF has helped me ask deeper questions, explore wider horizons, and contribute to a better understanding of the material and social aspects that shape our world.

Appendix A

Explaining growth: the production network-based prediction of industrial growth

A.1 Variable distribution (production network)

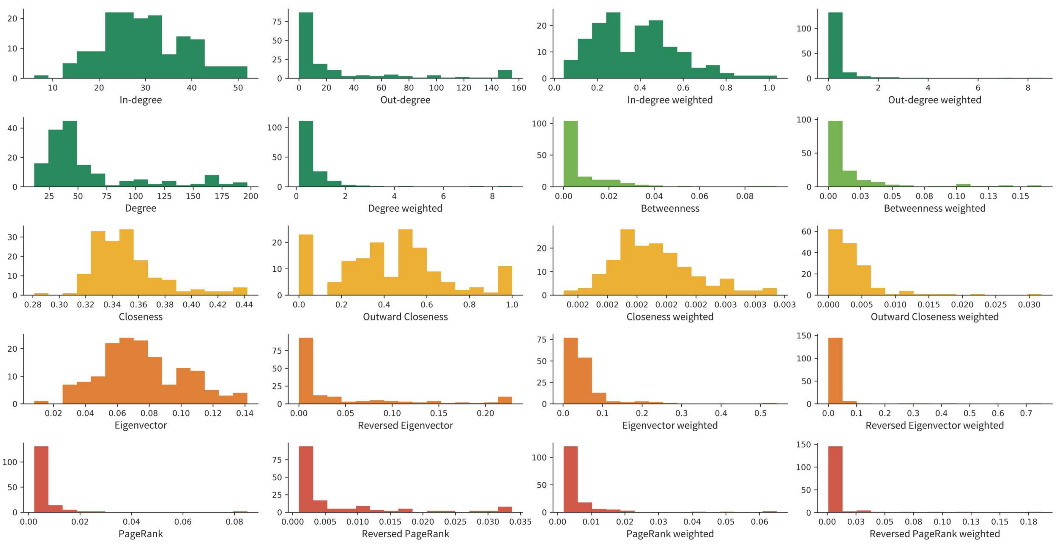


Figure A.1: Distribution of every measure calculated in the production network

A.2 Correlation matrix (production network)

Table A.1: Correlation matrix

	R3	InDeg	OutDeg	Deg	InDegW	OutDegW	DegW	Bet	BetW	Clos	OutClos	ClosW	OutClosW	Eig	OutEig	EigW	OutEigW	Pg	OutPg	PgW	OutPgW
R3	1																				
InDeg	-0.05	1																			
OutDeg	0.15	-0.13	1																		
Deg	0.14	0.07	0.98	1																	
InDegW	-0.04	0.69	-0.05	0.09	1																
OutDegW	0.27	-0.16	0.69	0.67	-0.12	1															
DegW	0.26	-0.03	0.69	0.69	0.06	0.98	1														
Bet	0.12	0.13	0.41	0.44	0.12	0.30	0.33	1													
BetW	0.19	0.06	0.49	0.50	0.03	0.50	0.51	0.64	1												
Clos	0.02	0.80	-0.20	-0.04	0.60	-0.14	-0.03	0.20	0.08	1											
OutClos	0.17	-0.13	0.85	0.82	-0.04	0.56	0.56	0.40	0.46	-0.26	1										
ClosW	-0.19	0.69	-0.29	-0.15	0.55	-0.23	-0.13	0.13	0.03	0.58	-0.35	1									
OutClosW	0.27	-0.22	0.78	0.74	-0.15	0.87	0.85	0.31	0.50	-0.25	0.78	-0.35	1								
Eig	-0.05	0.99	-0.11	0.09	0.68	-0.12	0.00	0.14	0.07	0.77	-0.12	0.73	-0.19	1							
OutEig	0.14	-0.12	0.99	0.97	-0.05	0.68	0.68	0.41	0.50	-0.19	0.84	-0.28	0.78	-0.09	1						
EigW	-0.07	0.73	-0.14	0.00	0.76	-0.12	0.02	0.05	-0.02	0.67	-0.18	0.65	-0.18	0.76	-0.14	1					
OutEigW	0.30	-0.22	0.57	0.52	-0.17	0.90	0.87	0.20	0.35	-0.17	0.45	-0.23	0.87	-0.19	0.57	-0.13	1				
Pg	0.01	0.29	-0.16	-0.10	0.18	-0.08	-0.05	0.25	0.09	0.44	-0.24	0.49	-0.17	0.28	-0.16	0.26	-0.07	1			
OutPg	0.16	-0.13	0.98	0.96	-0.06	0.72	0.71	0.44	0.53	-0.19	0.83	-0.28	0.80	-0.11	0.99	-0.15	0.60	-0.15	1		
PgW	0.00	0.32	-0.17	-0.11	0.20	-0.10	-0.06	0.23	0.08	0.47	-0.24	0.53	-0.18	0.31	-0.17	0.29	-0.08	0.93	-0.16	1	
OutPgW	0.31	-0.23	0.59	0.54	-0.17	0.92	0.89	0.22	0.36	-0.17	0.47	-0.24	0.87	-0.19	0.59	-0.14	0.98	-0.07	0.63	-0.08	1

Appendix B

Ego network growth model: the case study of the storage battery industry

B.1 Correlation matrix (storage battery ego network)

Table B.1: Correlation matrix

	R3	InDeg	OutDeg	Deg	InDegW	OutDegW	DegW	Bet	BetW	Clos	OutClos	ClosW	OutClosW	Eig	OutEig	EigW	OutEigW	Pg	OutPg	PgW	OutPgW
R3	1																				
InDeg	-0.17	1																			
OutDeg	0.09	-0.33	1																		
Deg	0.05	-0.08	0.97	1																	
InDegW	-0.13	0.58	-0.24	-0.10	1																
OutDegW	0.43	-0.36	0.43	0.36	-0.30	1															
DegW	0.42	-0.21	0.38	0.35	-0.02	0.96	1														
Bet	0.06	0.18	0.01	0.06	-0.03	0.02	0.01	1													
BetW	0.21	0.02	0.34	0.36	-0.08	0.45	0.45	0.32	1												
Clos	0.07	0.69	-0.41	-0.25	0.33	-0.26	-0.18	0.27	0.01	1											
OutClos	0.16	-0.37	0.95	0.91	-0.23	0.43	0.38	0.05	0.35	-0.38	1										
ClosW	0.01	0.54	-0.46	-0.34	0.32	-0.30	-0.22	0.06	-0.06	0.87	-0.42	1									
OutClosW	0.40	-0.53	0.58	0.47	-0.35	0.88	0.82	-0.05	0.35	-0.41	0.58	-0.41	1								
Eig	-0.12	0.99	-0.37	-0.12	0.56	-0.35	-0.20	0.20	0.03	0.79	-0.39	0.65	-0.52	1							
OutEig	0.11	-0.35	0.99	0.95	-0.24	0.42	0.37	-0.02	0.31	-0.39	0.94	-0.43	0.58	-0.38	1						
EigW	-0.18	0.58	-0.37	-0.23	0.73	-0.24	-0.03	-0.13	-0.17	0.35	-0.43	0.43	-0.31	0.56	-0.34	1					
OutEigW	0.47	-0.48	0.39	0.29	-0.35	0.91	0.85	-0.05	0.28	-0.34	0.38	-0.32	0.95	-0.46	0.39	-0.24	1				
Pg	-0.04	0.42	-0.56	-0.48	0.09	-0.21	-0.19	0.40	0.01	0.64	-0.55	0.61	-0.34	0.49	-0.57	0.12	-0.21	1			
OutPg	0.12	-0.34	0.99	0.95	-0.24	0.43	0.39	-0.01	0.34	-0.37	0.95	-0.42	0.59	-0.37	0.98	-0.35	0.40	-0.56	1		
PgW	-0.11	0.39	-0.49	-0.41	0.11	-0.19	-0.17	0.14	-0.02	0.50	-0.51	0.56	-0.29	0.44	-0.50	0.22	-0.17	0.81	-0.49	1	
OutPgW	0.49	-0.48	0.41	0.31	-0.34	0.92	0.87	-0.03	0.31	-0.33	0.41	-0.32	0.95	-0.46	0.41	-0.25	0.99	-0.21	0.43	-0.18	1

B.2 Variable distributions (storage battery ego network)

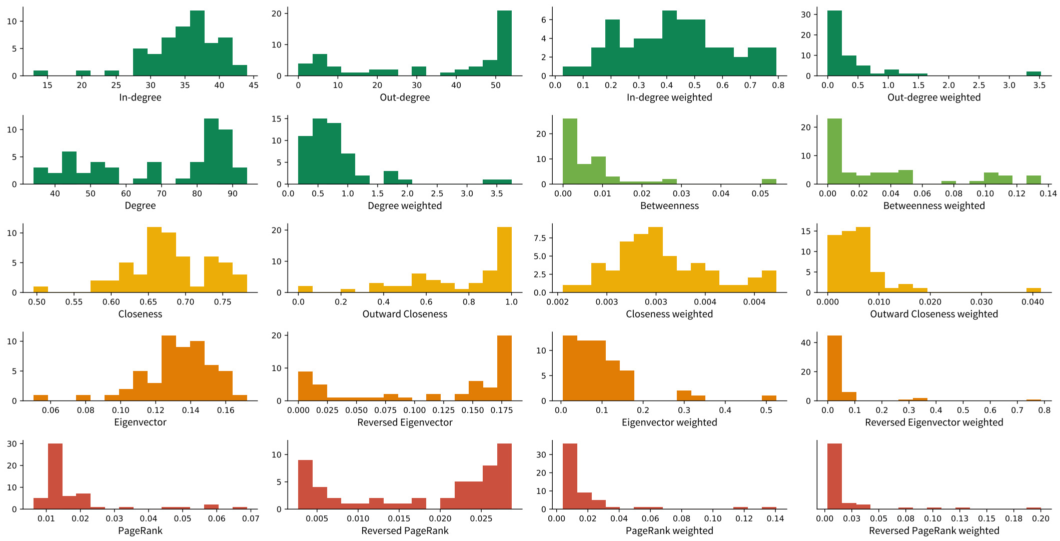


Figure B.1: Distribution of every measure calculated in the storage battery ego network (2007)

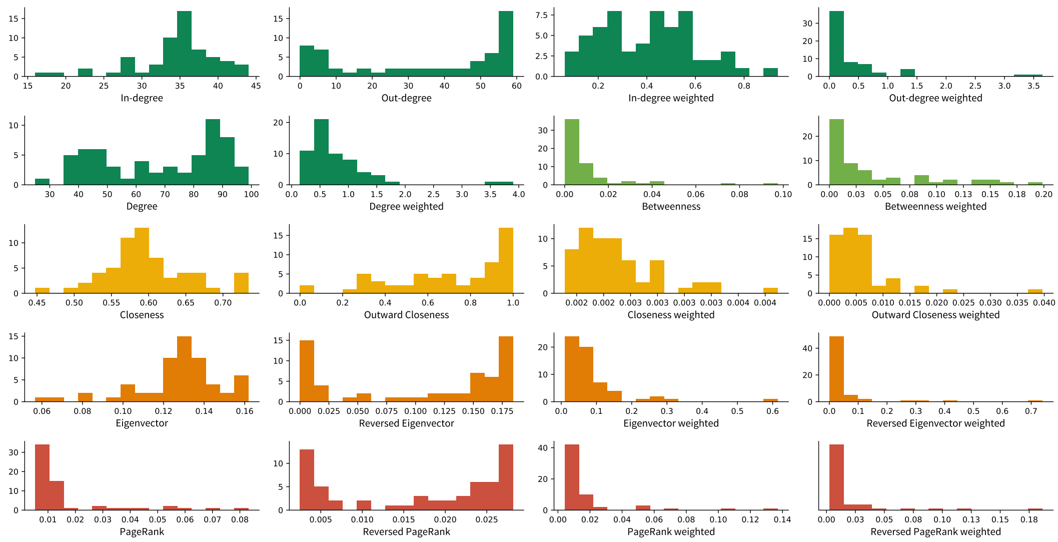


Figure B.2: Distribution of every measure calculated in the storage battery ego network (2012)

References

Acemoglu, Daron et al. (2012). "The Network Origins of Aggregate Fluctuations". en. In: *Econometrica* 80.5, pp. 1977–2016. ISSN: 1468-0262. DOI: 10.3982/ECTA9623.

Agnew, Scott and Paul Dargusch (2017). "Consumer preferences for household-level battery energy storage". In: *Renewable and Sustainable Energy Reviews* 75, pp. 609–617. ISSN: 1364-0321. DOI: https://doi.org/10.1016/j.rser.2016.11.030.

Ahuja, Gautam (2000). "Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study". In: *Administrative Science Quarterly* 45.3, pp. 425–455.

Ajanovic, A. and R. Haas (2019). "Economic and Environmental Prospects for Battery Electric- and Fuel Cell Vehicles: A Review". In: *Fuel Cells* 19.5, pp. 515–529. ISSN: 1615-6854. DOI: 10.1002/fuce.201800171.

Albert, Réka and Albert-László Barabási (2002). "Statistical mechanics of complex networks". In: *Reviews of modern physics* 74.1, p. 47.

Albert, Réka, Hawoong Jeong, and Albert-László Barabási (2000). "Error and attack tolerance of complex networks". In: *Nature* 406.6794, pp. 378–382.

Aqidawati, Era Febriana et al. (2022). "Technology Readiness and Economic Benefits of Swappable Battery Standard: Its Implication for Open Innovation". In: *Journal of Open Innovation: Technology, Market, and Complexity* 8.2, p. 88. ISSN: 2199-8531. DOI: 10.3390/joitmc8020088.

Armaroli, Nicola, Vincenzo Balzani, and Nick Serpone (2013). *Powering Planet Earth. Energy Solutions for the Future.* Wiley-VCH. ISBN: 978-3527334094.

Arnaboldi, Valerio, Marco Conti, Massimiliano La Gala, et al. (2016). "Ego network structure in online social networks and its impact on information diffusion". In:

Computer Communications 76, pp. 26–41. ISSN: 0140-3664. DOI: 10.1016/j.comcom.2015.09.028.

Arnaboldi, Valerio, Marco Conti, Andrea Passarella, and Robin I. M. Dunbar (2017). "Online Social Networks and information diffusion: The role of ego networks". In: *Online Social Networks and Media* 1, pp. 44–55. ISSN: 2468-6964. DOI: 10.1016/j.osnem.2017.04.001.

Arnaboldi, Valerio, Marco Conti, Andrea Passarella, and Fabio Pezzoni (2012). "Analysis of Ego Network Structure in Online Social Networks". In: *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing*, pp. 31–40. DOI: 10.1109/SocialCom-PASSAT.2012.41.

Arthur, W. Brian (2014). *Complexity and the Economy*. en. Oxford University Press. ISBN: 978-0-19-933430-8.

Asif, Amir A. and Rajendra Singh (2017). "Further Cost Reduction of Battery Manufacturing". In: *Batteries* 3.2, p. 17. ISSN: 2313-0105. DOI: 10.3390/batteries3020017.

Backhaus, Richard (2021). "Battery Raw Materials - Where from and Where to?" In: *ATZ worldwide* 123.9, pp. 8–13. ISSN: 2192-9076. DOI: 10.1007/s38311-021-0715-5. (Visited on 06/14/2023).

Baldwin, Richard and Javier Lopez-Gonzalez (2015). "Supply-chain Trade: A Portrait of Global Patterns and Several Testable Hypotheses". In: *The World Economy* 38.11, pp. 1682–1721. ISSN: 1467-9701. DOI: 10.1111/twec.12189.

Barabási, Albert-László (2003). *Linked: The new science of networks.*

— (2013). "Network science". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 371.1987, p. 20120375. DOI: 10.1098/rsta.2012.0375.

Barabási, Albert-László and Márton Pósfai (2016). *Network science*. Cambridge University Press. ISBN: 9781107076266.

Barak, M. (1980). *Electrochemical Power Sources: Primary and Secondary Batteries*. IET. ISBN: 978-0-906048-26-9.

Barrat, A. et al. (2004). “The architecture of complex weighted networks”. In: *Proceedings of the National Academy of Sciences* 101.11, pp. 3747–3752. DOI: 10.1073/pnas.0400087101.

Bartesaghi, Paolo, Gian Paolo Clemente, and Rosanna Grassi (2022). “Community structure in the World Trade Network based on communicability distances”. In: *Journal of Economic Interaction and Coordination* 17.2, pp. 405–441. ISSN: 1860-7128. DOI: 10.1007/s11403-020-00309-y.

— (2023). “Clustering coefficients as measures of the complex interactions in a directed weighted multilayer network”. In: *Physica A: Statistical Mechanics and its Applications* 610, p. 128413. ISSN: 0378-4371. DOI: 10.1016/j.physa.2022.128413.

Basole, Rahul C. (2016). “Topological analysis and visualization of interfirm collaboration networks in the electronics industry”. In: *Decision Support Systems* 83, pp. 22–31. ISSN: 0167-9236. DOI: 10.1016/j.dss.2015.12.005.

Basole, Rahul C. and Marcus A. Bellamy (2014). “Supply Network Structure, Visibility, and Risk Diffusion: A Computational Approach”. In: *Decision Sciences* 45.4, pp. 753–789. ISSN: 1540-5915. DOI: 10.1111/deci.12099.

Battiston, Stefano et al. (2007). “Credit chains and bankruptcy propagation in production networks”. en. In: *Journal of Economic Dynamics and Control*. Tenth Workshop on Economic Heterogeneous Interacting Agents 31.6, pp. 2061–2084. ISSN: 0165-1889. DOI: 10.1016/j.jedc.2007.01.004.

Bavelas, Alex (1950). “Communication Patterns in Task-Oriented Groups”. In: *The Journal of the Acoustical Society of America* 22.6, pp. 725–730. ISSN: 0001-4966. DOI: 10.1121/1.1906679.

Beaudet, Alexandre et al. (2020). “Key Challenges and Opportunities for Recycling Electric Vehicle Battery Materials”. In: *Sustainability* 12.14, p. 5837. ISSN: 2071-1050. DOI: 10.3390/su12145837.

Benham-Hutchins, Marge and Thomas R. Clancy (2010). “Social Networks as Embedded Complex Adaptive Systems”. In: *JONA: The Journal of Nursing Administration* 40, p. 352. ISSN: 0002-0443. DOI: 10.1097/NNA.0b013e3181ee42bc.

Bigio, Saki and Jennifer La'O (2016). "Financial Frictions in Production Networks". en. In: p. 84.

Biswas, Anupam and Bhaskar Biswas (2015). "Investigating community structure in perspective of ego network". In: *Expert Systems with Applications* 42.20, pp. 6913–6934. ISSN: 0957-4174. DOI: 10.1016/j.eswa.2015.05.009.

Blondel, Vincent D. et al. (2008). "Fast unfolding of communities in large networks". In: *Journal of Statistical Mechanics: Theory and Experiment* 2008.10, P10008. ISSN: 1742-5468. DOI: 10.1088/1742-5468/2008/10/P10008.

Bollobás, Béla et al. (2003). "Directed scale-free graphs". In: *Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms*. SODA '03. USA: Society for Industrial and Applied Mathematics, pp. 132–139. ISBN: 978-0-89871-538-5.

Borah, R. et al. (2020). "On battery materials and methods". In: *Materials Today Advances* 6, p. 100046. ISSN: 2590-0498. DOI: 10.1016/j.mtadv.2019.100046.

Borgatti, Stephen P. and Daniel S. Halgin (2011). "On Network Theory". In: *Organization Science*. DOI: 10.1287/orsc.1100.0641.

Borges, Jorge Luis (1998). "The library of Babel". In: *Collected fictions*.

Börner, Katy, Soma Sanyal, and Alessandro Vespignani (2007). "Network science". In: *Annu. rev. inf. sci. technol.* 41.1, pp. 537–607.

Borondo, Javier et al. (2014). "To each according to its degree: The meritocracy and topocracy of embedded markets". In: *Scientific Reports* 4.1, p. 3784.

Boulaire, Fanny et al. (2019). "Benefit assessment of battery plus solar for customers and the grid". In: *Energy Strategy Reviews* 26, p. 100372. ISSN: 2211-467X. DOI: https://doi.org/10.1016/j.esr.2019.100372.

Brandes, Ulrik (2001). "A faster algorithm for betweenness centrality". In: *The Journal of Mathematical Sociology* 25.2, pp. 163–177. ISSN: 0022-250X. DOI: 10.1080/0022250X.2001.9990249.

— (2008). "On variants of shortest-path betweenness centrality and their generic computation". In: *Social Networks* 30.2, pp. 136–145. ISSN: 0378-8733. DOI: 10.1016/j.socnet.2007.11.001.

Brintrup, Alexandra and Anna Ledwoch (2018). "Supply network science: Emergence of a new perspective on a classical field". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 28.3, p. 033120. ISSN: 1054-1500, 1089-7682. DOI: 10.1063/1.5010766.

Bugga, Ratnakumar V. and Erik J. Brandon (2020). "Energy Storage for the Next Generation of Robotic Space Exploration". In: *The Electrochemical Society Interface* 29.1, p. 59. ISSN: 1944-8783. DOI: 10.1149/2.F08201IF.

Burt, Ronald S. (1980). "Models of Network Structure". In: *Annual Review of Sociology* 6.1, pp. 79–141. DOI: 10.1146/annurev.so.06.080180.000455.

— (1995). *Structural Holes: The Social Structure of Competition*. Harvard University Press. ISBN: 978-0-674-26246-1.

Carnovale, Steven and Sengun Yeniyurt (2014). "The Role of Ego Networks in Manufacturing Joint Venture Formations". In: *Journal of Supply Chain Management* 50.2, pp. 1–17. ISSN: 1745-493X. DOI: 10.1111/jscm.12015.

— (2015). "The Role of Ego Network Structure in Facilitating Ego Network Innovations". In: *Journal of Supply Chain Management* 51.2, pp. 22–46. ISSN: 1745-493X. DOI: 10.1111/jscm.12075.

Carvalho, Vasco M (2008). "Aggregate Fluctuations and the Network Structure of Intersectoral Trade". en. PhD thesis. The University of Chicago. URL: https://doi.org/10.17635/lancaster/thesis/1377.

— (2014). "From Micro to Macro via Production Networks". en. In: *Journal of Economic Perspectives* 28.4, pp. 23–48. ISSN: 0895-3309. DOI: 10.1257/jep.28.4.23.

Carvalho, Vasco M. and Alireza Tahbaz-Salehi (2019). "Production Networks: A Primer". In: *Annual Review of Economics* 11.1, pp. 635–663. DOI: 10.1146/annurev-economics-080218-030212.

Chen, Tianmei et al. (2020). "Applications of Lithium-Ion Batteries in Grid-Scale Energy Storage Systems". In: *Transactions of Tianjin University* 26, pp. 208–217. DOI: https://doi.org/10.1007/s12209-020-00236-w.

Chen, Zhuowen et al. (2022). "Safety Concerns for the Management of End-of-Life Lithium-Ion Batteries". In: *Global Challenges* 6.12, p. 2200049. ISSN: 2056-6646. DOI: 10.1002/gch2.202200049.

Choi, Daiwon et al. (2021). "Li-ion battery technology for grid application". In: *Journal of Power Sources* 511, p. 230419. ISSN: 0378-7753. DOI: https://doi.org/10.1016/j.jpowsour.2021.230419.

Choi, Thomas Y, Kevin J Dooley, and Manus Rungtusanatham (2001). "Supply networks and complex adaptive systems: control versus emergence". en. In: *Journal of Operations Management* 19.3, pp. 351–366. ISSN: 1873-1317. DOI: 10.1016/S0272-6963(00)00068-1.

Concepts and Methods of the U.S. Input-Output Accounts — U.S. Bureau of Economic Analysis (BEA) (2023). URL: https://www.bea.gov/resources/methodologies/concepts-methods-io-accounts (visited on 09/27/2023).

Cooke, Philip (2020). "Gigafactory Logistics in Space and Time: Tesla's Fourth Gigafactory and Its Rivals". In: *Sustainability* 12.5, p. 2044. ISSN: 2071-1050. DOI: 10.3390/su12052044.

Crompton, Thomas Roy (2000). *Battery Reference Book*. Newnes. ISBN: 978-0-7506-4625-3.

Cusenza, Maria Anna et al. (2019). "Energy and environmental benefits of circular economy strategies: The case study of reusing used batteries from electric vehicles". In: *Journal of Energy Storage* 25, p. 100845. ISSN: 2352-152X. DOI: 10.1016/j.est.2019.100845.

De Benedictis, Luca et al. (2014). "Network Analysis of World Trade using the BACI-CEPII Dataset". In: *Global Economy Journal* 14.03n04, pp. 287–343. ISSN: 2194-5659. DOI: 10.1515/gej-2014-0032.

Demirel, Guven (2022). "Chapter 20 - Network science for the supply chain: theory, methods, and empirical results". In: *The Digital Supply Chain*. Ed. by Bart L. MacCarthy and Dmitry Ivanov. Elsevier, pp. 343–359. ISBN: 978-0-323-91614-1. DOI: 10.1016/B978-0-323-91614-1.00020-4.

Deng, Weibing et al. (2011). “The exponential degree distribution in complex networks: Non-equilibrium network theory, numerical simulation and empirical data”. en. In: *Physica A: Statistical Mechanics and its Applications* 390.8, pp. 1481–1485. ISSN: 0378-4371. DOI: 10.1016/j.physa.2010.12.029.

Dietzenbacher, Erik and Michael Lahr (2008). *Wassily Leontief and Input-Output Economics*. Cambridge Books. Cambridge University Press. URL: https://econpapers.repec.org/bookchap/cupcbooks/9780521049436.htm.

Duan, Wen-Qi (2012). “Modelling the Evolution of National Economies Based on Input–Output Networks”. en. In: *Computational Economics* 39.2, pp. 145–155. ISSN: 0927-7099, 1572-9974. DOI: 10.1007/s10614-011-9261-3.

Duffner, F. et al. (2020). “Battery cost modeling: A review and directions for future research”. In: *Renewable and Sustainable Energy Reviews* 127, p. 109872. ISSN: 1364-0321. DOI: 10.1016/j.rser.2020.109872.

Dunn, Jennifer B. et al. (2014). *Material and Energy Flows in the Materials Production, Assembly, and End-of-Life Stages of the Automotive Lithium-Ion Battery Life Cycle*. Tech. rep. ANL/ESD/12-3 Rev. Argonne National Lab. (ANL), Argonne, IL (United States). DOI: 10.2172/1177517.

Dykes, Katherine and John Sterman (2017). “Dynamics of Innovation and Diffusion in Large-Scale Complex Technical Systems: the Case of Wind Energy”. en. In: p. 30.

Easley, David and Jon Kleinberg (2010). *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. en. Cambridge University Press. ISBN: 978-1-139-49030-6.

Economic Input-Output Life Cycle Assessment. Carnegie Mellon University Green Design Institute (2008). http://www.eiolca.net. (Visited on 07/06/2021).

Ellis, Nick C. and Diane Larsen-Freeman (2009). *Language as a Complex Adaptive System*. John Wiley & Sons. ISBN: 978-1-4443-3400-5.

Erdős, Pál and Alfréd Rényi (1960). “On the evolution of random graphs”. In: *Publication of the Mathematical Institute of the Hungarian Academy of Sciences* 5, pp. 17–61.

Euler, Leonhard (1741). “Solutio problematis ad geometriam situs pertinentis”. In.

Everett, Martin and Stephen P. Borgatti (2005). “Ego network betweenness”. In: *Social Networks* 27.1, pp. 31–38. ISSN: 0378-8733. DOI: 10.1016/j.socnet.2004.11.007.

Fagiolo, Giorgio, Javier Reyes, and Stefano Schiavo (2008). “On the topological properties of the world trade web: A weighted network analysis”. In: *Physica A: Statistical Mechanics and its Applications*. Applications of Physics in Financial Analysis 387.15, pp. 3868–3873. ISSN: 0378-4371. DOI: 10.1016/j.physa.2008.01.050.

— (2009). “World-trade web: Topological properties, dynamics, and evolution”. In: *Physical Review E* 79.3, p. 036115. DOI: 10.1103/PhysRevE.79.036115.

— (2010). “The evolution of the world trade web: a weighted-network analysis”. In: *Journal of Evolutionary Economics* 20.4, pp. 479–514. DOI: 10.1007/s00191-009-0160-x.

Fagiolo, Giorgio, Tiziano Squartini, and Diego Garlaschelli (2013). “Null models of economic networks: the case of the world trade web”. In: *Journal of Economic Interaction and Coordination* 8.1, pp. 75–107. ISSN: 1860-7128. DOI: 10.1007/s11403-012-0104-7.

Farmer, J. Doyne et al. (2012). “A complex systems approach to constructing better models for managing financial markets and the economy”. en. In: *The European Physical Journal Special Topics* 214.1, pp. 295–324. ISSN: 1951-6355, 1951-6401. DOI: 10.1140/epjst/e2012-01696-9.

Fisher, Eric (2006). “The Linchpins of a Modern Economy”. In: *Annual meeting of the American Economic Association*.

Foerster, Andrew and Jason Choi (2017). “The Changing Input-Output Network Structure of the U.S. Economy”. en. In: *The Federal Reserve Bank of Kansas City Economic Review*. ISSN: 01612387. DOI: 10.18651/ER/2q17FoersterChoi.

Foster, John (Nov. 2005). “From simplistic to complex systems in economics”. In: *Cambridge Journal of Economics* 29.6, pp. 873–892. ISSN: 0309-166X. DOI: 10.1093/cje/bei083.

Franco, Alejandro (2015). *Rechargeable Lithium Batteries: From Fundamentals to Applications*. Elsevier. ISBN: 978-1-78242-098-9.

Freeman, Linton C. (1977). “A Set of Measures of Centrality Based on Betweenness”. In: *Sociometry* 40.1, pp. 35–41. ISSN: 0038-0431. DOI: 10.2307/3033543.

— (1982). “Centered graphs and the structure of ego networks”. In: *Mathematical Social Sciences* 3.3, pp. 291–304. ISSN: 0165-4896. DOI: 10.1016/0165-4896(82)90076-2.

Fri, Robert W. and Maxine L. Savitz (2014). “Rethinking energy innovation and social science”. In: *Energy Research & Social Science* 1, pp. 183–187. ISSN: 2214-6296. DOI: 10.1016/j.erss.2014.03.010.

Gallo, A. B. et al. (2016). “Energy storage in the energy transition context: A technology review”. In: *Renewable and Sustainable Energy Reviews* 65, pp. 800–822. ISSN: 1364-0321. DOI: 10.1016/j.rser.2016.07.028.

García-Algarra, Javier, Mary Luz Mouronte-López, and Javier Galeano (2019). “A stochastic generative model of the World Trade Network”. In: *Scientific Reports* 9.1, p. 18539. ISSN: 2045-2322. DOI: 10.1038/s41598-019-54979-1.

Gell-Mann, Murray (1994). “Complex Adaptive Systems”. In: ed. by G. Cowan, D. Pines, and D. Meltzer. Reading, MA: Addison-Wesley, pp. 17–45. ISBN: 978-0-201-62605-6.

Gintis, Herbert (2006). “The Economy as a Complex Adaptive System”. In: *Journal of Economic Literature* 54, pp. 1018–1031.

Girvan, Michelle and Mark Newman (2002). “Community structure in social and biological networks”. In: *Proceedings of the national academy of sciences* 99.12, pp. 7821–7826.

Giuliani, Elisa (2007). “The selective nature of knowledge networks in clusters: evidence from the wine industry”. In: *Journal of Economic Geography* 7.2, pp. 139–168. ISSN: 1468-2702. DOI: 10.1093/jeg/lbl014.

Gjermëni, Orgeta (2017). “Temporal Statistical Analysis of Degree Distributions in an Undirected Landline Phone Call Network Graph Series”. en. In: *Data* 2.4, p. 33. DOI: 10.3390/data2040033.

Golpîra, Hêriş and Salah Bahramara (2020). “Internet-of-things-based optimal smart city energy management considering shiftable loads and energy storage”. In: *Journal of Cleaner Production* 264, p. 121620. ISSN: 0959-6526. DOI: https://doi.org/10.1016/j.jclepro.2020.121620.

Gómez, Vicenç, Andreas Kaltenbrunner, and Vicente López (2008). "Statistical analysis of the social network and discussion threads in slashdot". In: *Proceedings of the 17th international conference on World Wide Web*. WWW '08. New York, NY, USA: Association for Computing Machinery, pp. 645–654. ISBN: 978-1-60558-085-2. DOI: 10.1145/1367497.1367585.

Gonçalves, R., S. Lanceros-Méndez, and C. M. Costa (2022). "Electrode fabrication process and its influence in lithium-ion battery performance: State of the art and future trends". In: *Electrochemistry Communications* 135, p. 107210. ISSN: 1388-2481. DOI: 10.1016/j.elecom.2022.107210.

Goyal, Sanjeev, Marco J. van der Leij, and José Luis Moraga-González (2006). "Economics: An Emerging Small World". In: *Journal of Political Economy* 114.2, pp. 403–412. ISSN: 0022-3808. DOI: 10.1086/500990.

Granovetter, Mark S. (1973). "The strength of weak ties". In: *American Journal of Sociology* 78.6, pp. 1360–1380.

Grazzini, Jakob and Alessandro Spelta (2015). *An empirical analysis of the global input-output network and its evolution.* DISCE - Working Papers del Dipartimento di Economia e Finanza def031. Università Cattolica del Sacro Cuore, Dipartimenti e Istituti di Scienze Economiche (DISCE). URL: https://econpapers.repec.org/paper/ctcserie1/def031.htm.

Gualdi, Stanislao and Antoine Mandel (Dec. 2016). "On the emergence of scale-free production networks". In: 73, pp. 61–77. ISSN: 0165-1889. DOI: 10.1016/j.jedc.2016.09.012.

Gulati, Ranjay (1995). "Social Structure and Alliance Formation Patterns: A Longitudinal Analysis". In: *Administrative Science Quarterly* 40.4, pp. 619–652. ISSN: 0001-8392. DOI: 10.2307/2393756.

Guo-hong, Wang et al. (2010). "The risk of cascading breakdown in industrial cluster innovation networks: A complex networks perspective". In: *2010 International Conference on Management Science & Engineering 17th Annual Conference Proceedings*, pp. 1445–1455. DOI: 10.1109/ICMSE.2010.5719978.

Hampton, Keith and Wenhong Chen (2021). "On Social Media: Studying Social Media from an Egocentric Perspective". In: pp. 718–733. ISBN: 978-1-108-83997-6. DOI: 10.1017/9781108878296.056.

Hasgall, Alon (2013). "Digital social networks as complex adaptive systems". In: *VINE Journal of Information and Knowledge Management Systems* 43, pp. 78–95. ISSN: 0305-5728. DOI: 10.1108/03055721311302151.

Hausmann, Ricardo et al. (2013). *The Atlas of Economic Complexity: Mapping Paths to Prosperity*. The MIT Press. ISBN: 9780262525428. URL: http://www.jstor.org/stable/j.ctt9qf8jp.

Hidalgo, C. A. et al. (2007). "The Product Space Conditions the Development of Nations". In: *Science* 317, pp. 482–487. DOI: 10.1126/science.1144581.

Hidalgo, César A. and Ricardo Hausmann (2009). "The building blocks of economic complexity". In: *Proceedings of the National Academy of Sciences* 106.26, pp. 10570–10575. DOI: 10.1073/pnas.0900943106.

Hodges, J. L. (1958). "The Significance Probability of the Smirnov Two-Sample Test". In: DOI: 10.1007/BF02589501.

Holland, John H. (1992). "Complex Adaptive Systems". In: *Daedalus* 121.1, pp. 17–30. ISSN: 0011-5266.

— (2006). "Studying Complex Adaptive Systems". In: *Journal of Systems Science and Complexity* 19.1, pp. 1–8. ISSN: 1559-7067. DOI: 10.1007/s11424-006-0001-z.

Horvath, Michael (1998). "Cyclicality and Sectoral Linkages: Aggregate Fluctuations from Independent Sectoral Shocks". en. In: *Review of Economic Dynamics* 1.4, pp. 781–808. ISSN: 1094-2025. DOI: 10.1006/redy.1998.0028.

— (2000). "Sectoral shocks and aggregate fluctuations". en. In: *Journal of Monetary Economics* 45.1, pp. 69–106. ISSN: 0304-3932. DOI: 10.1016/S0304-3932(99)00044-6.

Input-Output Accounts Data — U.S. Bureau of Economic Analysis (2021). https://www.bea.gov/data/industries/input-output-accounts-data. (Visited on 11/09/2021).

IRENA — Energy transition outlook (2023). URL: https://www.irena.org/Energy-Transition/Outlook (visited on 06/14/2023).

IRENA — International Renewable Energy Agency (2023). URL: https://www.irena.org (visited on 06/14/2023).

Jaffe, Sam (2017). "Vulnerable Links in the Lithium-Ion Battery Supply Chain". In: *Joule* 1.2, pp. 225–228. ISSN: 2542-4351. DOI: 10.1016/j.joule.2017.09.021.

Jennings, Helen (1937). "Structure of Leadership-Development and Sphere of Influence". In: *Sociometry* 1.1/2, pp. 99–143. ISSN: 0038-0431. DOI: 10.2307/2785262.

Jeong, Byongug et al. (2020). "Evaluation of the Lifecycle Environmental Benefits of Full Battery Powered Ships: Comparative Analysis of Marine Diesel and Electricity". In: *Journal of Marine Science and Engineering* 8.8, p. 580. ISSN: 2077-1312. DOI: 10.3390/jmse8080580.

Kali, Raja and Javier Reyes (2007). "The architecture of globalization: a network approach to international economic integration". en. In: *Journal of International Business Studies* 38.4, pp. 595–620. ISSN: 1478-6990. URL: https://doi.org/10.1057/palgrave.jibs.8400286.

— (2010). "Financial Contagion on the International Trade Network". en. In: *Economic Inquiry* 48.4, pp. 1072–1101. ISSN: 1465-7295. DOI: 10.1111/j.1465-7295.2009.00249.x.

Kao, Ta-Wei (Daniel) et al. (2017). "Relating supply network structure to productive efficiency: A multi-stage empirical investigation". In: *European Journal of Operational Research* 259.2, pp. 469–485. ISSN: 0377-2217. DOI: 10.1016/j.ejor.2016.11.008.

Keith, David R., John D. Sterman, and Jeroen Struben (2017). "Supply constraints and waitlists in new product diffusion". In: *System Dynamics Review*. DOI: https://doi.org/10.1002/sdr.1588.

Kenett, Dror Y. and Shlomo Havlin (2015). "Network science: a useful tool in economics and finance". In: *Mind & Society* 14.2, pp. 155–167. ISSN: 1860-1839. DOI: 10.1007/s11299-015-0167-y.

Kiehne, H. A. (2003). *Battery Technology Handbook*. CRC Press. ISBN: 978-0-203-91185-3.

Kim, Yusoon et al. (2011). "Structural investigation of supply networks: A social network analysis approach". In: *Journal of Operations Management* 29.3, pp. 194–211. ISSN: 1873-1317. DOI: 10.1016/j.jom.2010.11.001.

Klenow, Peter J. and Andrés Rodríguez-Clare (1997). "Economic growth: A review essay". In: *Journal of Monetary Economics* 40.3, pp. 597–617. ISSN: 0304-3932. DOI: https://doi.org/10.1016/S0304-3932(97)00050-0.

Konig, Denes (1936). *Theorie Der Endlichen Und Unendlichen Graphen.*

Kuijp, Tsering Jan van der, Lei Huang, and Christopher R. Cherry (2013). "Health hazards of China's lead-acid battery industry: a review of its market drivers, production processes, and health impacts". In: *Environmental Health* 12.1, p. 61. ISSN: 1476-069X. DOI: 10.1186/1476-069X-12-61.

Kumar, Pankaj and Akbar Zaheer (2019). "Ego-Network Stability and Innovation in Alliances". In: *Academy of Management Journal.* DOI: 10.5465/amj.2016.0819.

Kwade, Arno et al. (2018). "Current status and challenges for automotive battery production technologies". In: *Nature Energy* 3.4, pp. 290–300. ISSN: 2058-7546. DOI: 10.1038/s41560-018-0130-3.

Lambiotte, Renaud, Jean-Charles Delvenne, and Mauricio Barahona (2014). "Random Walks, Markov Processes and the Multiscale Modular Organization of Complex Networks". In: *IEEE Transactions on Network Science and Engineering* 1.2, pp. 76–90. ISSN: 2327-4697. DOI: 10.1109/TNSE.2015.2391998.

Langville, Amy and Carl Meyer (2004). "A Survey of Eigenvector Methods of Web Information Retrieval". In: *SIAM Review* 47.1, pp. 135–161. DOI: 10.1137/S0036144503424786.

Lansing, J. Stephen (2003). "Complex Adaptive Systems". In: *Annual Review of Anthropology* 32.1, pp. 183–204. DOI: 10.1146/annurev.anthro.32.061002.093440.

Lee, Kyu-Min et al. (2011). "Impact of the Topology of Global Macroeconomic Network on the Spreading of Economic Crises". en. In: *PLOS ONE* 6.3, e18443. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0018443.

Leontief, Wassily (1970). “Environmental Repercussions and the Economic Structure: An Input-Output Approach”. In: *The Review of Economics and Statistics* 52.3, pp. 262–271. ISSN: 0034-6535. DOI: 10.2307/1926294.

— (1986). *Input-Output Economics*. en. Oxford University Press. ISBN: 978-0-19-536522-1.

Leontief, Wassily W. (1936). “Quantitative Input and Output Relations in the Economic Systems of the United States”. In: *The Review of Economics and Statistics* 18.3, pp. 105–125. ISSN: 0034-6535. DOI: 10.2307/1927837.

Leskovec, Jure and Julian Mcauley (2012). “Learning to Discover Social Circles in Ego Networks”. In: *Advances in Neural Information Processing Systems*. Vol. 25. Curran Associates, Inc.

Levin, Simon A. (1998). “Ecosystems and the Biosphere as Complex Adaptive Systems”. In: *Ecosystems* 1.5, pp. 431–436. ISSN: 1432-9840. DOI: 10.1007/s100219900037.

Lewis, Ted G. (2011). *Network Science: Theory and Applications*. ISBN: 978-1-118-21101-4.

Li, Jianlin, Claus Daniel, and David Wood (2011). “Materials processing for lithium-ion batteries”. In: *Journal of Power Sources* 196.5, pp. 2452–2460. ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2010.11.001.

Li, Xin, Konstantinos J. Chalvatzis, and Phedeas Stephanides (2018). “Innovative Energy Islands: Life-Cycle Cost-Benefit Analysis for Battery Energy Storage”. In: *Sustainability* 10.10, p. 3371. ISSN: 2071-1050. DOI: 10.3390/su10103371.

Linden, D. (1984). *Handbook of batteries and fuel cells.*

Litjens, G. B. M. A., E. Worrell, and W. G. J. H. M. van Sark (2018). “Economic benefits of combining self-consumption enhancement with frequency restoration reserves provision by photovoltaic-battery systems”. In: *Applied Energy* 223, pp. 172–187. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2018.04.018.

Liu, Linqing, Mengyun Shen, and Chang Tan (2021). “Scale free is not rare in international trade networks”. In: *Scientific Reports* 11.1, p. 13359. ISSN: 2045-2322. DOI: 10.1038/s41598-021-92764-1.

Lu, Wei et al. (2021). "Scientific collaboration and career stages: An ego-centric perspective". In: *Journal of Informetrics* 15.4, p. 101207. ISSN: 1751-1577. DOI: 10.1016/j.joi.2021.101207.

Luo, Jianxi and Daniel E. Whitney (2015). "Asymmetry in in-degree and out-degree distributions of large-scale industrial networks". In: *arXiv:1507.04507 [physics]*. URL: http://arxiv.org/abs/1507.04507.

MacKay, David JC (2009). *Sustainable Energy - Without the Hot Air*. UIT Cambridge Ltd. ISBN: 978-0-9544529-3-3.

Maiser, Eric (2014). "Battery packaging - Technology review". In: *AIP Conference Proceedings* 1597.1, pp. 204–218. ISSN: 0094-243X. DOI: 10.1063/1.4878489.

McNerney, James et al. (2022). "How production networks amplify economic growth". In: *Proceedings of the National Academy of Sciences* 119.1, e2106031118. DOI: 10.1073/pnas.2106031118.

Mercer, Carolyn et al. (2012). "Energy Storage Technology Development for Space Exploration". In: *AIAA SPACE 2010 Conference & Exposition*. American Institute of Aeronautics and Astronautics. DOI: 10.2514/6.2010-8709.

Milgram, Stanley (1967). "The small world problem". In: *Psychology today* 2.1, pp. 60–67.

Miller, Clark A., Alastair Iles, and Christopher F. Jones (2013). "The Social Dimensions of Energy Transitions". In: *Science as Culture* 22.2, pp. 135–148. ISSN: 0950-5431. DOI: 10.1080/09505431.2013.786989.

Miller, Ronald E. and Peter D. Blair (2009). *Input-Output Analysis: Foundations and Extensions*. en. Cambridge University Press. ISBN: 978-1-139-47759-8.

Moores, Simon (2021). "The Global Battery Arms Race: Lithium-ion Battery Gigafactories and their Supply Chain". In: *Oxford Energy Forum*.

Morowitz, Harold J. (2018). *The Mind, The Brain And Complex Adaptive Systems*. Routledge. ISBN: 978-0-429-96131-1.

Muchnik, Lev et al. (2013). "Origins of power-law degree distribution in the heterogeneity of human activity in social networks". en. In: *Scientific Reports* 3.1, p. 1783. ISSN: 2045-2322. DOI: 10.1038/srep01783.

North American Industry Classification System (2021). https://www.census.gov/naics. (Visited on 12/14/2021).

Becker, Randy A., Wayne B. Gray, and Jordan Marvakov (2021). *NBER-CES Manufacturing Industry Database (1958-2018, version 2021a). National Bureau of Economic Research.* https://www.nber.org/research/data/nber-cesmanufacturing-industry-database. (Visited on 05/04/2022).

Nemeth, Roger J. and David A. Smith (1985). "International Trade and World-System Structure: A Multiple Network Analysis". In: *Review (Fernand Braudel Center)* 8.4, pp. 517–560. ISSN: 0147-9032. URL: http://www.jstor.org/stable/40241006.

NetworkX Python package (2014). https://networkx.org. (Visited on 02/22/2022).

Newman, Mark (2003). "The structure and function of complex networks". In: *SIAM review* 45.2, pp. 167–256.

— (2005). "Power laws, Pareto distributions and Zipf's law". In: *Contemporary Physics* 46.5, pp. 323–351. ISSN: 0010-7514. DOI: 10.1080/00107510500052444.

— (2018). *Networks*. en. Oxford University Press. ISBN: 978-0-19-252749-3.

Nishi, Yoshio (2001). "Lithium ion secondary batteries; past 10 years and the future". In: *Journal of Power Sources* 100.1, pp. 101–106. ISSN: 0378-7753. DOI: https://doi.org/10.1016/S0378-7753(01)00887-4.

Oughton, Edward J. et al. (2018). "Infrastructure as a Complex Adaptive System". In: *Complexity* 2018, e3427826. ISSN: 1076-2787. DOI: 10.1155/2018/3427826.

Page, Lawrence et al. (1999). *The PageRank Citation Ranking: Bringing Order to the Web.* Technical Report 1999-66. Previous number = SIDL-WP-1999-0120. Stanford InfoLab. URL: http://ilpubs.stanford.edu:8090/422/.

Park, Jung-Ki (2012). *Principles and Applications of Lithium Secondary Batteries*. en. John Wiley & Sons. ISBN: 978-3-527-65042-2.

Park, Yousin, Iori Nakaoka, and Yunju Chen (2022). "The R&D Direction and Business Strategy: The case study on the cooperation of EV and battery makers". In: *International Conference on Artificial Life and Robotics*.

Pastor-Satorras, Romualdo and Alessandro Vespignani (2001). "Epidemic spreading in scale-free networks". In: *Physical review letters* 86.14, p. 3200.

Pathak, Surya D. et al. (2007). "Complexity and Adaptivity in Supply Networks: Building Supply Network Theory Using a Complex Adaptive Systems Perspective". en. In: *Decision Sciences* 38.4, pp. 547–580. ISSN: 1540-5915. DOI: `10.1111/j.1540-5915.2007.00170.x`.

Perera, Supun, Michael G. H. Bell, and Michiel C. J. Bliemer (2017). "Network science approach to modelling the topology and robustness of supply chain networks: a review and perspective". en. In: *Applied Network Science* 2.1, pp. 1–25. ISSN: 2364-8228. DOI: `10.1007/s41109-017-0053-0`.

Phister, Paul W. (2010). "Cyberspace: The Ultimate Complex Adaptive System". In: *Int C2 J* 4, p. 33.

Pinch, Steven et al. (2003). "From 'industrial districts' to 'knowledge clusters': a model of knowledge dissemination and competitive advantage in industrial agglomerations". In: *Journal of Economic Geography* 3.4, pp. 373–388. ISSN: 1468-2702. DOI: `10.1093/jeg/lbg019`.

Pistoia, Gianfranco (2010). *Electric and Hybrid Vehicles: Power Sources, Models, Sustainability, Infrastructure and the Market.* Elsevier. ISBN: 978-0-444-53566-5.

Popovich, Natalie D. et al. (2021). "Economic, environmental and grid-resilience benefits of converting diesel trains to battery-electric". In: *Nature Energy* 6.11, pp. 1017–1025. ISSN: 2058-7546. DOI: `10.1038/s41560-021-00915-5`.

Prado, Julio (2017). "Essays on the effects of clusters and networks on industrial growth during the great recession and other periods of industrial shocks". en. PhD thesis. Lancaster University. URL: `https://doi.org/10.17635/lancaster/thesis/1377`.

Radicchi, Filippo, José J. Ramasco, and Santo Fortunato (2011). "Information filtering in complex weighted networks". In: *Physical Review E* 83.4, p. 046101. DOI: `10.1103/PhysRevE.83.046101`.

Responsible Minerals Initiative (2023). URL: `https://www.responsiblemineralsinitiative.org/` (visited on 06/14/2023).

Reyes, Javier, Stefano Schiavo, and Giorgio Fagiolo (2010). “Using complex networks analysis to assess the evolution of international economic integration: The cases of East Asia and Latin America”. In: *The Journal of International Trade & Economic Development* 19. ISSN: 0963-8199. DOI: 10.1080/09638190802521278.

Ruan, Jiageng, Paul Walker, and Nong Zhang (2016). “A comparative study energy consumption and costs of battery electric vehicle transmissions”. In: *Applied Energy* 165, pp. 119–134. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2015.12.081.

Ruetschi, Paul (1993). “Energy storage and the environment: the role of battery technology”. In: *Journal of Power Sources* 42.1, pp. 1–7. ISSN: 0378-7753. DOI: 10.1016/0378-7753(93)80132-9.

S. Ziegler, Micah, Juhyun Song, and Jessika E. Trancik (2021). “Determinants of lithium-ion battery technology cost decline”. In: *Energy & Environmental Science* 14.12, pp. 6074–6098. DOI: 10.1039/D1EE01313K.

Sabidussi, Gert (1966). “The centrality index of a graph”. en. In: *Psychometrika* 31.4, pp. 581–603. ISSN: 1860-0980. DOI: 10.1007/BF02289527.

Sacks, Michael Alan, Marc J. Ventresca, and Brian Uzzi (2001). “Global Institutions and Networks: Contingent Change in the Structure of World Trade Advantage, 1965-1980”. en. In: *American Behavioral Scientist* 44.10, pp. 1579–1601. ISSN: 0002-7642. DOI: 10.1177/00027640121958069.

Sarkar, Epica Mandal, Tanmay Sarkar, and Mridula D. Bharadwaj (2018). “Lithium-ion battery supply chain: enabling national electric vehicle and renewables targets”. In: *Current Science* 114.12, pp. 2453–2458. ISSN: 0011-3891.

Saw, Lip Huat, Yonghuang Ye, and Andrew A. O. Tay (2016). “Integration issues of lithium-ion battery into electric vehicles battery pack”. In: *Journal of Cleaner Production* 113, pp. 1032–1045. ISSN: 0959-6526. DOI: 10.1016/j.jclepro.2015.11.011.

Schilling, Melissa A. and Corey C. Phelps (2007). “Interfirm Collaboration Networks: The Impact of Large-Scale Network Structure on Firm Innovation”. In: *Management Science*. DOI: 10.1287/mnsc.1060.0624.

Schweitzer, Frank, Giorgio Fagiolo, Didier Sornette, Fernando Vega-Redondo, Alessandro Vespignani, et al. (2009). “Economic Networks: The New Challenges”. In: *Science* 325.5939, pp. 422–425. DOI: 10.1126/science.1173644.

Schweitzer, Frank, Giorgio Fagiolo, Didier Sornette, Fernando Vega-Redondo, and Douglas R. White (2009). “Economic Networks: What Do We Know and What Do We Need to Know?” en. In: *Advances in Complex Systems* 12, pp. 407–422. URL: https://papers.ssrn.com/abstract=1596044.

Sgouridis, Sgouris and Denes Csala (2014). “A Framework for Defining Sustainable Energy Transitions: Principles, Dynamics, and Implications”. In: *Sustainability* 6.5, pp. 2601–2622. ISSN: 2071-1050. DOI: 10.3390/su6052601.

Shiau, Ching-Shin Norman et al. (2009). “Impact of battery weight and charging patterns on the economic and environmental benefits of plug-in hybrid vehicles”. In: *Energy Policy* 37.7, pp. 2653–2663. ISSN: 0301-4215. DOI: 10.1016/j.enpol.2009.02.040.

Silva Lima, Lígia da et al. (2022). “The role of raw materials to achieve the Sustainable Development Goals: Tracing the risks and positive contributions of cobalt along the lithium-ion battery supply chain”. In: *Journal of Industrial Ecology*. ISSN: 1530-9290. DOI: 10.1111/jiec.13336.

Smil, Vaclav (2010). *Energy Transitions: History, Requirements, Prospects*. ABC-CLIO. ISBN: 978-0-313-38177-5.

— (2018). *Energy and Civilization: A History*. MIT Press. ISBN: 978-0-262-53616-5.

— (2019). *Energy In World History*. Routledge. ISBN: 978-0-429-72012-3.

Snyder, David and Edward L. Kick (1979). “Structural Position in the World System and Economic Growth, 1955-1970: A Multiple-Network Analysis of Transnational Interactions”. In: *American Journal of Sociology* 84.5, pp. 1096–1126. ISSN: 0002-9602. DOI: 10.1086/226902.

Song, Jiali et al. (2019). “Material flow analysis on critical raw materials of lithium-ion batteries in China”. In: *Journal of Cleaner Production* 215, pp. 570–581. ISSN: 0959-6526. DOI: 10.1016/j.jclepro.2019.01.081.

Sovacool, Benjamin K. (2016). *The history and politics of energy transitions: Comparing contested views and finding common ground*. Working Paper 2016/81. WIDER Working Paper. DOI: 10.35188/UNU-WIDER/2016/124-6.

Soyyiğit, Semanur and Çiğdem Boz (2018). “Global Input - Output Analysis: A Network Approach”. In.

Soyyiğit, Semanur and Yasemin Asu Çırpıcı (2017). “An Input-Output Network Structure Analysis Of Selected Countries”. en. In: p. 24.

Subburaj, Anitha S., Bejoy N. Pushpakaran, and Stephen B. Bayne (2015). “Overview of grid connected renewable energy based battery projects in USA”. In: *Renewable and Sustainable Energy Reviews* 45, pp. 219–234. ISSN: 1364-0321. DOI: https://doi.org/10.1016/j.rser.2015.01.052.

Sun, Xin et al. (2019). “Supply risks of lithium-ion battery materials: An entire supply chain estimation”. In: *Materials Today Energy* 14, p. 100347. ISSN: 2468-6069. DOI: 10.1016/j.mtener.2019.100347.

Tesfatsion, Leigh (2003). “Agent-based computational economics: modeling economies as complex adaptive systems”. In: *Information Sciences* 149.4, pp. 262–268. ISSN: 0020-0255. DOI: 10.1016/S0020-0255(02)00280-3.

Thackeray, Michael M., Christopher Wolverton, and Eric D. Isaacs (2012). “Electrical energy storage for transportation—approaching the limits of, and going beyond, lithium-ion batteries”. In: *Energy Environ. Sci.* 5 (7), pp. 7854–7863. DOI: 10.1039/C2EE21892E.

Timmer, Marcel P. et al. (2015). “An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production”. In: *Review of International Economics* 23.3, pp. 575–605. DOI: https://doi.org/10.1111/roie.12178.

Total Requirements Data — U.S. BEA Input-Output Accounts (2023). URL: https://apps.bea.gov/industry/xls/io-annual/IxI_TR_2007_2012_PRO_DET.xlsx (visited on 07/22/2023).

Total Requirements Definition — U.S. Bureau of Economic Analysis (2021). https://apps.bea.gov/industry/pdf/TotalRequirementsDerivation.pdf. (Visited on 11/09/2021).

Turcheniuk, Kostiantyn et al. (2021). "Battery materials for low-cost electric transportation". In: *Materials Today* 42, pp. 57–72. ISSN: 1369-7021. DOI: 10.1016/j.mattod.2020.09.027.

Vartanian, Charles (2010). "Grid stability battery systems for renewable energy success". In: *2010 IEEE Energy Conversion Congress and Exposition*, pp. 132–135. DOI: 10.1109/ECCE.2010.5618059.

Väyrynen, Antti and Justin Salminen (2012). "Lithium ion battery production". In: *The Journal of Chemical Thermodynamics* 46, pp. 80–85. ISSN: 0021-9614. DOI: 10.1016/j.jct.2011.09.005.

Wang, Lei and Xinchen Wang (2017). "The Influences of Edge Asymmetry on Network Robustness". In: *ArXiv*.

Wang, Yixuan et al. (2019). "From the Perspective of Battery Production: Energy – Environment – Economy (3E) Analysis of Lithium-Ion Batteries in China". In: *Sustainability* 11.24, p. 6941. ISSN: 2071-1050. DOI: 10.3390/su11246941.

Watts, Duncan J. and Steven H. Strogatz (1998). "Collective dynamics of 'small-world'networks". In: *Nature* 393.6684, pp. 440–442.

Weimer, Lucas, Tobias Braun, and Ansgar vom Hemdt (2019). "Design of a systematic value chain for lithium-ion batteries from the raw material perspective". In: *Resources Policy* 64, p. 101473. ISSN: 0301-4207. DOI: 10.1016/j.resourpol.2019.101473.

Wicki, Samuel and Erik G. Hansen (2017). "Clean energy storage technology in the making: An innovation systems perspective on flywheel energy storage". In: *Journal of Cleaner Production* 162, pp. 1118–1134. ISSN: 0959-6526. DOI: 10.1016/j.jclepro.2017.05.132.

Wu, Liuren (2015). "Centrality of the Supply Chain Network". In: DOI: 10.2139/ssrn.2651786.

Xu, Ming, Brad R. Allenby, and John C. Crittenden (2011). "Interconnectedness and Resilience of the U.S. Economy". en. In: *Advances in Complex Systems* 14.05, pp. 649–672. DOI: 10.1142/S0219525911003335.

Xu, Ming and Sai Liang (2019). "Input–output networks offer new insights of economic structure". en. In: *Physica A: Statistical Mechanics and its Applications* 527, p. 121178. ISSN: 0378-4371. DOI: 10.1016/j.physa.2019.121178.

Yan, Jie et al. (2019). "Multi-stage transport and logistic optimization for the mobilized and distributed battery". In: *Energy Conversion and Management* 196, pp. 261–276. ISSN: 0196-8904. DOI: 10.1016/j.enconman.2019.06.001.

Yan, Tingting et al. (2015). "A Theory of the Nexus Supplier: A Critical Supplier From A Network Perspective". In: *Journal of Supply Chain Management* 51.1, pp. 52–66. ISSN: 1745-493X. DOI: 10.1111/jscm.12070.

Yergin, Daniel and Joseph Stanislaw (2002). *The commanding heights: The battle for the world economy*. Simon and Schuster.