PETER VAKHAL

NETWORK ANALYSIS OF GLOBAL VALUE CHAINS

Department of Operations Research and Actuarial Sciences

Supervisor:
Dr. Erzsébet Kovács, CSc

CORVINUS UNIVERSITY OF BUDAPEST
DOCTORAL SCHOOL OF BUSINESS AND MANAGEMENT

NETWORK ANALYSIS OF GLOBAL VALUE CHAINS

PhD thesis

Supervisor:
Dr. Erzsébet Kovács, CSc

Peter Vakhal

Budapest, 2021.

Contents

List of figures

List of tables

List of abbreviations

Abbreviation	Equivalent
IO	Input-Output Table
GVA	Gross Value Added
BoP	Balance of Payments
CPC	Central Product Classification
CSO	Hungarian Statistical Office
DVA	Domestic value added
UN	United Nations
FATS	Foreign Affiliates Trade in Services
FDI	Foreign Direct Investment
FGP	Factoryless goods producer
FTE	Full-time equivalent
GDP	Gross Domestic Product
GMM	Generalized Method of Moments
GNI	Gross National Income
HS (code)	Harmonised System (code)
HVG	Horizontal Visibility Graph
IIP	International Investment Position
IMF	International Monetary Found
IPP	Intellectual Property Products
ISIC	International Standard Industrial Classification
SME	Small and Medium Enterprises
MRIO	Multiregional input-output (database)
NPISH	Non-profit institutions serving households
SITC	Standard International Trade Classification
SNA	System of National Accounts
SPE	Special Purpose Entities
TFP	Total factor productivity
TNC	Transnational companies
VAR	Vector autoregression
VG	Visibility Graph

For the list of country abbreviations see Annex I.

“Measure what can be measured,

and make measurable what cannot be measured.”

Galileo Galilei

Acknowledgements

It was a long road with numerous adversities. There were even times when I considered giving up, although everyone encouraged me not to do so. I was repeatedly mulling over whom shall I be grateful to. First, I like to thank Dr. Erzsébet Kovács for putting trust in me from the very beginning, and I hope that she will be also proud of me for this dissertation. I am indebted to Dr. Erzsébet Czakó for supporting my research and not letting me go astray.

I am immensely grateful to my friend and mentor, Dr. Miklós Losoncz, for seeing the opportunity in me a decade ago and encouraging me throughout the whole process. I am certain that I could not complete this journey without him.

I would like to thank my institute, the Kopint-Tárki Institute for Economic Research, for backing my research. I am obliged to Dr. Éva Palócz, who spurred me from the beginning, Dr. Katalin Nagy, who was highly appreciative, Dr. Ágnes Hárs, whom I could always turn to, Dr. Rozália Jehoda, who was solicitous from the beginning itself, Mr. Zoltán Matheika, who helped me find my way in the world, Mrs. Erika Rózsás, who became my second mother, and Mr. Ferenc Sinkó, who always cheered me up.

I am grateful to Dr. Péter Vékás for a number of common lectures. I must express my gratitude to my former mates in the doctoral school: Dr. Anna Radványi, Dr. Eszter Monda, Dr. László Mohácsi, and Mr. Tamás Tibori, who all made my last couple of years a pleasant memory.

I have much to be thankful to my family – my father, Peter Vakhal, who helped me start this journey and then helped me through it. Thanks, are also due to my mother, Myrtill Plaveczky, who has been holding by me since the very beginning. I am obliged to my aunt, Erika Vakhal, who always supported me. Unfortunately, some did not live that moment; however, their merits are undying: Dr. Gyula Szabó, who always supported my scientific aspirations, my grandmother, Mrs. Peter Vakhal, without whom I would be a different man, my grandfather, Tivadar Plaveczky, who always supported, and to my grandmother Mrs. Tivadar Plaveczky who encouraged me from the beginning.

Finally, I am thankful to Mária Godó and Kálmán Godó, who accepted me in their family. At last, but not least, I am enormously grateful to my fiancée, who stood by me and my aspirations, and to whom I cannot say thanks enough.

Any mistakes and fallacies in this dissertation are liable to me and only me.

1. Introduction

The development of national economies has been at the centre of economic research for centuries. Why some countries, industries, and companies are more successful, while others on the same path are not? Neither economic history nor the modern science can provide a sufficient answer, chiefly because economies are getting increasingly more complex and integrated, and hence studying their development is also challenging. Classical theories (such as Ricardo's foreign trade theory) were markedly reshaped, renewed, and even superseded. Describing the new phenomena requires large amount of data, not to mention that data collection takes place in an ever-changing economic environment.

Global value chains (GVCs) have become the centre of the world economy in a generation. World trade is five times higher since the WWII (in constant prices), and this trend could not be broken by any crises. The countries, industries and companies are directly or indirectly connected, which results in an interdependent, multicollinear world economic system. In that landscape, events such as the coronavirus pandemic in 2020 can easily generate chaos just as Lorentz poetically described the butterfly effect (Lorenz, 1963).

The organisation of companies (and indirectly countries) into production networks has restructured the dependencies, and thus one cannot clearly conclude whether reliance goes from a smaller country to a larger one or the other way around. In such systems, the role of companies or industries in the network has become one of the principal research questions. Most scholars study the links, development paths, and outlooks of modern supply chains.

The answer to the aforementioned questions is akin to a multivariate equation – one cannot solve the same simply using economic tools, because it is essential to incorporate tools and approaches from other disciplines. Statistics, econometrics, and operations research have been integral parts of the economic and social sciences for decades, while network theory is a new member of that group. In the past 10 years, as more and more data became available in good quality (primarily in the field of social networks), the toolbox of graph theory could be applied to real-world data as well. The first tables required for GVC analysis surfaced in the middle of the 2010s; however, analyses applying network theory have been rare, because these data are conspicuously different from the ones used in the classical graph theory research: these networks are heavily dense, complete, and the value of the vertices mainly depend on themselves. Classical algorithms (e.g., cluster analysis) that run on these data are not appropriate for the task in question, and inferences are very limited. Therefore, adjustments and modifications should be applied to these methods, and this was the principal goal of this research.

By utilising these new methodologies, one can identify and reveal the true relative position of a country, industry, or even a group of companies in the GVC. All these can potentially contribute to more accurate analyses and the development of targeted economic policies.

1.1 Objectives

Nowadays, the globalised world trade is totally interdependent. From the network theory perspective, one can state that today, all countries have considerable trade with the other countries across the world. Calculating by the 193 members of the United Nations (UN) and applying the $\frac{n(n-3)}{2}$ formula (where n is the number of nodes in the network), there are 18,335 trade links. Nevertheless, world trade is not only a densely linked network but also a layered one, as production is organised in stages. This successive system determines a semi-strict order, wherein the successive subsequent production stages follow each other in a $a_1 \leq a_2 \ldots \leq a_n$ process. Thus, neighbouring phases are sometimes interchangeable; however, the whole process cannot be changed significantly[1].

The analysis of input–output (IO) tables using graph theory is not a new approach, because the structure of the transaction matrices ($\mathbb{R}_{\geq 0}^{n \times n}$) are similar to an adjacency matrix, and many scholars have already investigated their applicability. Still, the literature is enriched by descriptive statistics only. Analysing the GVCs using network theory can put international competitiveness into a different light, because it can provide a more complex map of relative position of countries, industries, and companies in the international space of trade. The possibilities of value chain research are very limited if one purely relies on official statistics only. The datasets are on gross terms and bias not only the trade statistics but other macroeconomic indicators as well. Owing to the interweaves in labour division,

[1] In the production process, the procurement of some parts and accessories is interchangeable; however, the whole production procedure cannot be changed. Let us take car production, where the wheels can be built on the chassis at the beginning of the assembly process or even at the end.

pricing, and ownerships, the official statistical accounting of transnational companies (TNCs) is accomplished via estimations; it affects not only the level of GDP but all other indicators that are derived from it. Unfortunately, many decisions and strategies rely on these indices.

Heretofore, no study in the pertinent literature has been published that deals specifically with the system of value chains in the field of graph theory. One reason behind this could be the fact that the adjacency matrices of IO tables are very special, and from an economic perspective, they have characteristics that make them difficult to analyse using network theory:

1. The edges of the IO adjacency matrix are weighted and directed. At the same time, these weights have a high correlation with the size of nodes representing the size of the economy[2]. Generally, the weight of a random V_i vertex can be derived from $W(V_i) = \sum_{j=1}^{n} a_{ij}$ (where a_{ij} is the weights of the edges), which results in the following: $cor(W(V_i), \rho_i) > 0$, where ρ_i is degree of node i. In other words, the weight of the node correlates with the number of edges. In the special-case IO matrices, this relation does not hold, because the network is complete (i.e. every node is connected to every other node); thus, the weights of the vertices depend[3] on the size of the country it represents, consequently interpreting the edge weights as distances are inaccurate.

[2] Throughout the dissertation, the term 'size of the economy' represents the total value added created by the country or industry.
[3] This was confirmed by Natarajan Meghanathan (Meghanathan, 2014).

2. The expected values of gross value added produced at different production stages are not equal, that is $E(x \in S) \not\approx E(x \in S')$, where x is the production sequence, S and S' are the different stages of production (Sturgeon et al., 2013). If they were approximately equal, one could apply the classical segmentation algorithms. These methods cannot be utilised directly because the procedures rely on distance metrics, which are not satisfied (see the first point).
3. There are no clearly distinct production processes because there are no producers in the world who would not use any imported value added from a foreign country[4]. Thus, one cannot build a network flow because there is no t_0 point of source. Nevertheless, the end points are known, because once the final good is made in the production process, there will be no more transformation. Knowing only the final points (sinks), the flows cannot be interpreted as a whole but by stages only.
4. The path of value-added flow cannot be simplified using different tools of combinatorics (there is no 'shortest path'), which is a consequence of the previous point. The existence of such a path would also be inaccurate from economic perspective because it would assume that the production can be rationalised if some edges or nodes (countries) are left out.
5. On account of the preceding point, it is worth examining dependency in the GVCs. At the same time, it is beyond the purview of network theory, because besides the dependency on raw materials, there are other (political, cultural, and historical) factors that play a crucial role in the development of trade (for example, in the form of a free trade agreement) (Pratono, 2019).
6. In contrast to the classical models in graph theory, the diagonal of the adjacency matrix (a_{ij}) has a key role. In many cases in the traditional models, the values on the diagonal are zero. For IO matrices, however, it is not true – even more, they are the largest elements in the matrix (alternatively, if there were any larger elements in the matrix, it would mean that the industry exports more than it uses, which is highly unlikely). In dependency analysis, it means that all nodes depend mainly on themselves. It can be concluded that in competitiveness analysis, the effectiveness of domestic production is more important than foreign linkages.

[4] In some international IO tables – like the WIOD – there are elements with 0 value, which indicates no transaction between the units. However, this only means that the value is below a predefined threshold. The Eora database, which is utilised in this dissertation (see Chapter 4), does not apply such threshold values, and thus all elements in the IO matrix are above zero.

By mapping the value chains as a network, one could analyse the time dimension as well. This expands the statistical–econometric framework and could reveal certain aspects of value chains that were hidden. The analysis of production sequences can turn the static approach into a dynamic one, and by that, one can get more accurate position of a country or industry in any supply chain. This dissertation uses both static and dynamic approaches: a network flow model (static) and a time series model (dynamic). Both methods serve the stratification of the network into layers, which provides a punctilious representation of any graphs depicting value chains.

The concept of stratified networks also illustrates another perspective, which deserves particular attention. Sequences are also dependent; however, there are multiple interpretations of it. In this dissertation, dependencies are investigated by the subsequent dimensions:

- In the *short term*, every trade link is also a dependency because all parts and accessories are essential and substituting a supplier can be done only at the price of cost increase. If the competition on the suppliers' market is low, the cost can be higher; however, if the competition is intense and the supply of raw materials is ample, the degree of dependency is low[5].
- There is a long-term dependency towards some suppliers if there is a monopoly in the supplier's market because the cost increase of substitution is infinite (Clelland, 2014).

The primary focus of this dissertation is on long-term dependencies. The peculiarity of this approach is provided by the observation that the level of dependency is independent of the degree distribution (density, size) of the network and the degree of individual nodes. Mathematically, $\xi\left(V_i \middle| E(V_i), P(k)\right) = \xi(V_i)$, where ξ is the level dependency, V_i is the node (country, industry, or good), $E(V_i)$ is the number of edges (partners) of V_i, and $P(k)$ is degree distribution of the network. Consequently, the relative importance of a node (country) does not depend on its size, and a smaller country can also have a critical role in the value chain if it has a quasi-monopolistic market.

[5] See for example the resource dependency in construction (Donato et al., 2015).

Measuring market concentration in GVCs has always been deemed a challenging task because the international IO tables represent industries and not goods or services. These tables are highly aggregated, and one cannot determine which goods were included in the trade between two industries. Connecting trade data with the IO tables only allows particular analyses; however, the scope of goods can be extended. Classical competitiveness indicators can also be examined in case of product base models in value chain analyses (see for example revealed comparative advantages).

This dissertation intends to call the attention to the relationship between market concentration and the relative position in the value chain. This correlation also impacts the application of the classical graph theory algorithms. In particular, the segmentation algorithms require modifications or even a new model, because these procedures partition networks using network degrees. In that case, nodes with low number of edges (or low weights) are pushed to the background. The algorithm then focuses on the vertices that have more edges (or higher weights), and the segmentation is done along with them. These algorithms do not show the real dependencies in the networks. That could be only done if one could measure the market concentration of goods, but this possibility is unfortunately not ensured by the international IO tables.

The principal goal of the research chronicled in this dissertation was to modify the aforementioned methods and put the competitiveness of Hungarian companies into a different light and analyse their relative position in GVCs. A more accurate determination of positions would contribute to studies investigating upgrade possibilities for identifying key points in development. In addition, it is anticipated to support economic policy planning. A better understanding of globalisation is likely to contribute to avoiding the middle-income trap in Hungary and in the region.

1.2 Concepts

This dissertation relies on the concepts of management studies, micro- and macroeconomics, international economics, official statistics, statistical inference, and mathematics, in particular graph theory. The synergy of these sciences forms the unity and complexity of this dissertation; it is therefore of prime importance to first discuss certain fundamental concepts because later they will be used as postulates. This chapter can also be used as a glossary of this dissertation.

1.3 Economic and statistical concepts

A firm is a specialised production unit, which produces industrial goods or services (hereinafter goods) by using labour and other inputs (Demsetz, 1997). The output of a firm is the total products produced measured at market price. Every good can be regarded as an output if it can be used by other firms, consumers, or by the firm itself. Goods can be classified into three groups:

- Intermediate goods: can be used only for production. Users can be other firms or households. It is worth noting that tangible assets are not intermediate goods, even if they are utilised for production (see capital goods).
- Final goods: these products are purchased by customers who are going to consume them. No more physical transformation is done, and they are not used as inputs for any production. However, durable goods may have a secondary market, and the total value of these markets are not significant in the global market.
- Capital goods: those physical goods and tangible assets used for production but not as inputs (e.g. machinery, buildings, tools, computers, vehicles). Capital goods are finished goods – although used for production, they are not intermediate goods and can be purchased not only by the firms but also by households (e.g. real estate).

The product classification is based on the UN's Central Product Classification or CPC nomenclature (UNSTAT, 2015), which allocates the harmonised custom tariff (HS) code to every product. It also provides the industry the International Standard Industrial Classification of All Economic Activities (i.e. ISIC code) to every good and service, which represents the industry that generally produces the product in subject.

One should note that in an IO framework, the products are classified by the nature of consumption. The aforementioned categories are not disjunct sets. There are products that can be both final goods or intermediate goods, depending on the buyer[6].

[6] A good example is purchase of printing papers. If bought by a household, it is a final good, because then it will be fully consumed. However, if purchased by a firm, then it is an intermediate good, because the company will use it as an input for production.

A household consists of (a) person(s), whose purchases and consumption for living are done collectively. An insignificant part of households conducts production. While most consumed products are final goods, households can also purchase capital goods (e.g. real estates). Besides households, there can be such non-profit institutes (NPISH) financed by the government that serve households by providing services free of charge (e.g. the church). In an IO system, their consumption is usually added to the households. Their weight in some non-secular countries can be high.
Besides the companies and households, the government is a separate entity in the economy. Its consumption of final goods is regarded as public consumption, while its investments are public investments. It must be noted that the government itself is not an actor in the market, but it can be present through its public companies (e.g. public transportation firms). These firms must be considered part of the business sector.

Purchasing capital goods, like tangible assets (independently from the entity of the customer) must be considered a capital formation, or in other words, an investment. It also involves the purchase of non-produced assets such as land, mines, or legal rights[7]. The acquisition of these assets usually happens between industries, and thus, they are considered investments in the system of national accounts (EUROSTAT-OECD, 2015).

All types of products (final, intermediary, and capital) can be used by the resident actors of an economy (firms, households, NPISHs, and government) and by the world. The latter is considered as exports and in that case the ownership of the goods is changed between two actors of different residency. From a buyer's standpoint, the transaction is export, and from the perspective of a buyer, it is import. Households usually do not participate in merchandise trade; however, they constitute a large part of the trade in services.

Firms are not obliged to sell the final good to any consumers; this can be done at any time in the future. Until the ownership of a final product is not changed to a final consumer, the good remains in the inventory of the company.

Goods are produced by firms by using inputs such as raw materials and intermediate and capital goods. Inputs are transformed into outputs. This process can be described by the function $q = f(v_1, \dots, v_n)$, where q is the output, v_n is production factor (or input), and f

[7] Generally, the purchase of any kind of capital goods that are part of the assets in the balance sheet shall be considered an investment.

is the production function. Table 1 represents the cooperation of four companies by the nature of the produced goods, the used inputs, and the final output (1. table):

1. Table: Production sequencies in a hypothetic value chain by inputs used, output and nature of the product

<table>
<tr><th colspan="2">Stage 1.</th><th colspan="2">Stage 2.</th><th colspan="2">Stage 3.</th><th colspan="2">Stage 4.</th><th>Stage 5.</th></tr>
<tr><td colspan="2">Industry:
Mining</td><td colspan="2">Industry:
Machining (galvanisation)</td><td colspan="2">Industry:
Machining (cutting)</td><td colspan="2">Industry:
Retail trade</td><td rowspan="6">Consumption</td></tr>
<tr><td>Inputs</td><td>Output</td><td>Inputs</td><td>Output</td><td>Inputs</td><td>Output</td><td>Inputs</td><td>Output</td></tr>
<tr><td>Capital</td><td rowspan="3">Copper, Zinc</td><td>Capital</td><td rowspan="3">Brass</td><td>Capital</td><td rowspan="3">Handle</td><td>Capital</td><td rowspan="3">Retail trade services[8]</td></tr>
<tr><td>Labour</td><td>Labour</td><td>Labour</td><td>Labour</td></tr>
<tr><td>Deposit</td><td>Copper, Zinc</td><td>Brass</td><td>Handle</td></tr>
<tr><td colspan="2">Intermediate good</td><td colspan="2">Intermediate good</td><td colspan="2">Intermediate good</td><td colspan="2">Final good</td></tr>
</table>

Source: own edition

Production functions can assume various forms. The most prevalent version is the so called Cobb–Douglas production function (Cobb & Douglas, 1928)[9]. This function can have an additive or a multiplicative form:

$$q = Av_1^{\alpha} v_2^{\beta} \tag{1}$$

where

q = output;

A = constant;

v_1 = capital used for production;

v_2 = labour used for production;

α = share of capital in production;

β = share of labour in production.

The production function determines the level of output that can be achieved by the combination of inputs. In the simplest case, the firm does not maximise its production but its profit (equivalent to the minimisation of costs at the level of the revenue). In the medium run, it is assumed that the inputs used are substitutable, whereas in the short run, this assumption is constrained, because the elasticity of substitution is fairly low (Jones,

[8] Value-added can be only increased by the cost paid for the retail trade serives.

[9] Besides the Cobb–Douglas production function, there are many other function forms: CES, CET, Translog, etc. (Heathfield & Wibe, 1987).

2003). For the purposes of this dissertation, the fixed proportion production function[10] is crucial. This function assumes that the ratio of inputs required for production is fixed and substitution is not allowed. However, for the sake of simplicity, inputs usually cover the capital and labour only, and this can be easily extended to all required inputs including parts and accessories (Csontos & Ray, 1992; W. E. Diewert, 1971):

$$Q(y) = \{x: f(x) \geq 0\} \tag{2}$$

$$y = \min \{x: (x_1, x_2 \dots x_n) \geq 0\} \tag{3}$$

where

Q(y) = set of production possibilities;

y = possible volume of outputs;

x_n = input factor used for production.

According to the function above, the producer can supply as many outputs as it is possible from the available inputs. Owing to the fact that these inputs are not sustainable in the short run, the maximum output is constrained by the least available input.

If similar amount and quality of production factors are available for several producers, the output created depends on the productivity of the producer. A producer can increase its productivity in different ways (OECD, 2001):

- Technological change: owing to Research and Development and Innovation (R&D&I) developments in an industry, the production can be boosted by technological change in the long run, which can also improve the productivity of the economy. Technological change is the most common factor that can lift productivity; however, it is the most difficult to quantify. In the long term, it can also alter the production function and the set of required inputs (Kortum, 1997).
- Improvement in efficiency: from the set of production possibilities, an element yielding higher output is achieved. Thus, increasing efficiency is not equivalent to technological change, as neither the production function nor the input set is changed (E. Diewert & Lawrence, 1999).

[10] Also known as the Leontief production function.

- Cost saving: usually, the result of the previous two source productivity gains is the increase in output. At the same time, if the producer saves costs, a boost in productivity can be achieved without the expansion of production. One is able to save cost if the firm adopts the advantages of technological changes without any significant investment (e.g. it upgrades its software for free of charge) or the authorities change the rules in a favourable way.

One can measure the change in productivity in several ways; however, there is no universally accepted methodology to do so. Generally, input, labour, and capital productivity changes are measured separately, although composite indicators are also available, and they can measure the total factor productivity in production. These indices are summarised in Table 2.

2. Table: Ways to gauge productivity

<table>
<tr>
<td></td>
<td>Labour
Ways to measure:
Number of employees adjusted for full-time work (FTE)
Hours worked
Wages</td>
<td>Capital
Ways to measure:
total assets
equity
stock of fixed assets</td>
<td>Labour and capital
Ways to measure:
The role of production factors in output and value added is typically estimated by linear regression.</td>
<td>Labour, capital and inputs
energy consumption
services
material inputs
Source: IO tables</td>
</tr>
<tr>
<td>Gross output</td>
<td>Labour productivity
$P = \frac{\Delta output}{\Delta labour}$

Shows how changes in labour productivity affect output. Other interpretation: how many units of labor are needed to produce output per unit.
The disadvantage is that change in labour productivity cannot be separated from the change in capital productivity.</td>
<td>Capital productivity
$P = \frac{\Delta output}{\Delta capital}$

It shows how the productivity of the capital used for production affects output. Its change cannot be decouped from changes in labour productivity.</td>
<td>Multi-factor productivity[11]
$P = \frac{\Delta output}{\Delta(\alpha L + (1-\alpha)K)}$

It shows how changes in the combined use of labour and capital affect output. Rather, it is a micro-level indicator, and at the macro level aggregation can cause significant bias.</td>
<td rowspan="2">KLEMS[12] multi-factor productivity
$P = \frac{\Delta output}{\Delta(\sum_{i=1}^{5} \alpha_i F_i)}$
$F_i: \{K, L, E, M, S\}$

Technological and efficiency indicators are indicators that can show change. Due to its huge data requirements, it is only available on a limited number of grounds.

Not suitable to measure value-added.</td>
</tr>
<tr>
<td>Value added</td>
<td>Labour productivity
$P = \frac{\Delta value\ added}{\Delta Labour}$

It shows how changes in labor productivity affect the production of added value. The indicator is less dependent on changes in other factors (especially input).</td>
<td>Capital productivity
$P = \frac{\Delta value\ added}{\Delta capital}$

It shows how changes in capital productivity affect the production of added value. The indicator is less dependent on changes in other factors (especially input).</td>
<td>Multi-factor productivity
$P = \frac{\Delta value\ added}{\Delta(\alpha L + (1-\alpha)K)}$

It shows how changes in the combined use of labour and capital affect value added production.</td>
</tr>
</table>

Source: (Gullickson & Harper, 1987; OECD, 2001; Schreyer & Pilat, 2001; M. P. Timmer et al., 2007; Vakhal, 2018a)

[11] Also called *Total factor productivity* – TFP.
[12] KLEMS represents the following production function:: $y=f(K, L, E, M, S)$, where K is capital, L is labour, E is energy, M is the parts and accessories used, S is the sign of services (M. P. Timmer et al., 2007)

As evidenced in Table 2, production can be analysed along two dimensions: output and value added. Output is generally quantified in terms of revenue at market prices. Value added is the value that is created during the production stage by the producer itself. The measure can be expressed in both gross and net terms:

- Gross value added (GVA) = output – inputs used for production[13]
- Net value added = GVA – amortisation

GVA plays a central role in the economy. It depicts the total value created by the actors of the economy and constitutes the base of the GDP, the primary income of the residents and consequently a substantial proportion of state budget revenue. Generally, GVA is measured at firm level and is aggregated eventually. At micro-level, GVA is calculated as the following:

3. Table: The calculation of gross value added from company level data in Hungary based on official statistical guidelines

***Output* =** Net sales revenue + Own work capitalized value – - COGS (Cost of goods sold) – - Revenue from intermediary services
***Intermediate use* = Material expenses** – COGS (Cost of goods sold) – Revenue from intermediary services
Gross value added = Output - Intermediate use

Source: (Bella & Kazimir, 2020))

The final goods consumed by consumers (households, NPISHs, government) are built up by value added created during the production process, which can be far from each other in geographical terms. This production chain is also known as value-added chain and is elaborated in Chapter 2 from an economic perspective. Chapter 3 introduces the statistical characteristics. IO tables also have prime importance and are detailed in Chapter 4.

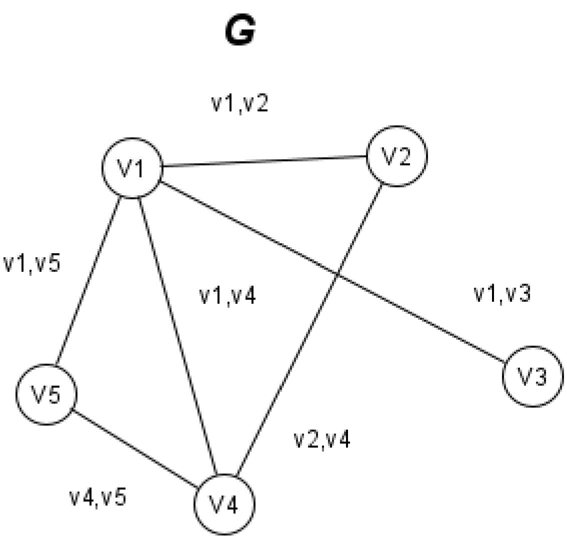
[13] Intermediary use.

1.4 Network theory, graphs[14][15]

Networks are complex systems representing (inter)dependencies. The importance of individual characteristics of the members is limited, while more emphasis is laid on their relations, including visualisation and analysis. That is, two networks can be compared in terms of their dependency structures and not their members. As a consequence, two networks with non-overlapping membership can still be similar. This provides the basis for the exploration of non-trivial structure of networks (Emmert-Streib et al., 2016).

The members of a network are represented by nodes or vertices. Relations between two nodes are illustrated by edges. Graph *G* with nodes *V* and edges *E* can be formalised by *G = (V,E)*, where $V\colon \{v_1, v_2, \dots v_n\}$ is the set of vertices, and $E\colon \{e_1, e_2 \dots e_m\}$ is the set of edges. If there is a relation between v_1 and v_2, it is represented by edge e_1, and notation for that relation is $e_1 = \{v_1, v_2\}$. Figure 1 depicts a simple graph:

1. Figure: A simple graph



Source: own construction

Graph *G* in Figure 1 can be defined as the following: $V(G) = \{V_1, V_2, V_3, V_4, V_5\}\ and\ E(G) = \{v_1v_2, v_1v_3, v_1v_4, v_1v_5, v_2v_4, v_4v_5\}$. E edges are the projection of function $w\colon V \to R^+$, and in a simple case, it can be defined as:

$$w(e_i) = \begin{cases} 1, & if\ \{v_i, v_j\} \in E(G)\ (i \neq j) \\ 0, & otherwise \end{cases} \ \forall e_i \tag{4}$$

---

[14] Throughout the dissertation we use network and graph as synonyms.

[15] This chapter strongly relies on a previous study (Hajnal, 2003).

In case of simple networks, the relations are described by a [0,1] scalar. A more complex weighted graph is drawn if edges receive a $w = [0, \infty]$ $w \in R^+$ random value. The links between the nodes are represented by the so-called adjacency matrix $A^{n \times n}$, where $n = |V|$, that is the count of elements of set V. The adjacency matrix of the network plotted in Figure 1 is the following:

4. Table: The adjacency matrix of the graph on figure 1.

A(w)=

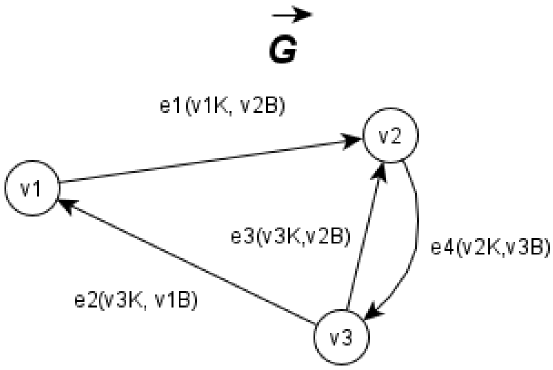
	V1	V2	V3	V4	V5
V1	0	$w_{1,2}$	$w_{1,3}$	$w_{1,4}$	$w_{1,5}$
V2	$w_{1,2}$	0	0	$w_{2,4}$	0
V3	$w_{1,3}$	0	0	0	0
V4	$w_{1,4}$	$w_{2,4}$	0	0	$w_{4,5}$
V5	$w_{1,5}$	0	0	$w_{4,5}$	0

Source: own edition

A salient characteristic of adjacency matrices is that they are always square matrices. However, only non-directed graphs are symmetric, that is $w_{ij} = w_{ji}$ $\forall w \in A$. Edges in such networks represent the existence of the links only, and the relations cannot be ordered.

Directed graphs have vital importance in value chain analysis. A directed graph $\vec{G}$ can defined using four parameters: $\vec{G} = (V, E, K, B)$, where V represents the set of vertices, E the set of edges, while K (origin) and B (destination) represent the relationship between sets of V and E. Two sets of one element are assigned to all edges: $\{v \in V: vKe\}$ and $\{v \in V: vBe\}$ $\forall e \in E$. The interpretation of v_iKe_j: v_i is the origin of e_i. Following that analogy, the interpretation of v_iBe_j: v_i is the destination of e_i. The following graph in Figure 2 is the representation of a simple graph, while Table 5 presents its adjacency matrix:

2. Figure: A directed graph $\vec{G}$



source: own construction

5. Table: The adjacency matrix of $\vec{G}$ directed graph on figure 2.

		V1	**V2**	**V3**
A(w)=	**V1**	0	$w_{1,2}$	0
	V2	0	0	$w_{2,3}$
	V3	$w_{3,1}$	$w_{3,2}$	0

Source: own edition

In case of directed graph, the adjacency matrix is a square but not symmetric, because the relations are depicted by two edges, which can be equal. This peculiarity has crucial importance in GVC analyses, because the inter-industrial use of countries is different in both ways. For example, restaurants immensely rely on agriculture, but agriculture does not use much output from restaurant services.

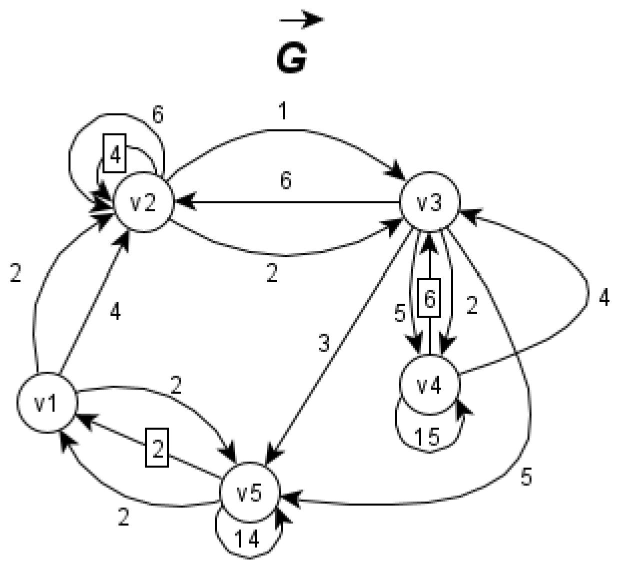
Loops also have high significance in value chain networks. By definition, a loop edge has the same origin and destination node, that is, $e_i = e(v_jK_e, v_jB_e)$. These edges appear on the diagonal in the adjacency matrix. Their economic interpretation is supply of an industry to the same sector (e.g. a firm operating in the machinery industry buys parts of a device from a company also operating in the same industry). Generally, the use from the same sector has the highest volume among all supplies, including households and export. Thus, IO tables are diagonally dominated, that is $a_{ii} > \sum_i^j a_{ij} \ \forall i, i \neq j, a_{ij} \in R^+$. Consequently, IO matrices are always non-singular and positive semi-definite[16]. This ensures

[16] This was proven by Bell (Bell, 1965).

that the matrices will be invertible, which is essential for the Leontief methodology applied in IO analysis (see Chapter 4).

Simple graphs do not hold loops, and therefore, it is necessary to introduce the theory of multigraphs or pseudo-graphs[17], which allows the existence of multiple edges[18] (besides loops). In other words, multigraphs are such networks in which two or more edges are allowed between two nodes, including loops. Parallel edges can be added up if they have no separate identity (weight) but shall be handled individually if they have their own weight. The representation of a multigraph is equivalent to a directed graph. The adjacency matrix is square because the weight must be summed up. In case of value chains, multigraphs serve the purpose of stratified graph visualisation, wherein nodes are countries and edges are industries. In this manner, more dimensions can be represented. Figure 3 and Table 6 depict a hypothetical directed multigraph and its adjacency matrix.

3. Figure: A hypothetical $\vec{G}$ directed multigraph with identities (weights)



source: own construction

[17] According to certain definitions, loops are not allowed in multigraphs and just in pseudo-graphs (Pemmaraju & Skiena, 2003). In this dissertation, the difference between the two is less essential; therefore, hereinafter the multigraph and the pseudo-graph shall be used as synonyms.
[18] In some studies, they are also called parallel edges.

6. Table: The adjacency matrix of $\vec{G}$ multigraph shown on figure 3.

A=	V1	V2	V3	V4	V5
V1	0	6	0	0	2
V2	0	10	3	0	0
V3	0	6	0	7	7
V4	0	0	10	15	0
V5	4	0	0	0	14

Source: own edition

A directed graph will be a network if it has at least one $s \in V(\vec{G})$ source and et least one $t \in V(\vec{G})$ sink node, and if a $c: E(\vec{G}) \to R^+$, the capacity function can be assigned to all edges, which determines the permeability of the links. Let us introduce the $f: E(\vec{G}) \to R^+$ flow function, the $\sum_{E_x^+} f(e)$ sum is the inward flow to node *x*, and the $\sum_{E_x^-} f(e)$ sum is the outward flow from vertex *x*. The expression of $v(f) = \sum_{E_x^+} f(e) - \sum_{E_x^-} f(e)$ is the total value *f* flow. Throughout this dissertation, it is assumed that all nodes have a constrained flow, that is $0 \leq f(e) \leq c(e)$. In other words, the flow running on an edge does not exceed the capacity of that particular edge. In value chain analysis, there are no restrictions concerning the sign of the total value of the flow (can be positive or negative).

According to the economic interpretation (in pure international IO tables), a positive total flow indicates that the country has a trade deficit (more import arrives than the export that leaves the country), while the opposite means positive balance of payments. The following condition holds universally and is also known as the conservation law in graph theory:

$$\sum_{E_x^+} f(e) = \sum_{E_x^-} f(e) \ \ \forall x \tag{5}$$

Not just the intermediate consumption but the final use is also part of the network that gives meaning to the aforementioned equation. This implicitly assumes that globally there is no waste, that is the products do not perish in the flow, and all goods and services are consumed or put into the inventory. Flow networks usually turn up in optimisation assignments in graph theory. However, it has no justification in GVC landscape, because it violates certain assumptions. Among others, graph theory assumes almost full substitutability, which is likely to be very limited in the value chains. Even in the same product groups, there could be major differences in quality; thus, usually the network cannot be simplified by the methods of graph theory, even if the potential substitute has spare capacity, while the other is at 100% or more[19].

1.5 Statistics describing networks[20]

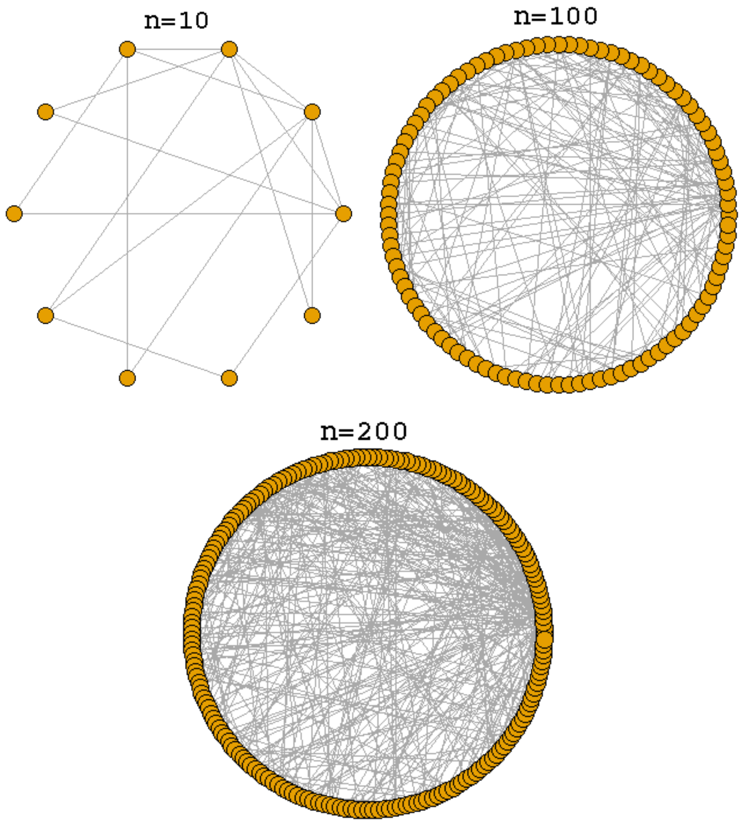
The number of statistical indicators describing networks is manifold. Most of these measures are applied to compare the nodes and the edges within the same network, and there are other indices that analyse the structure of the network. There are only a few indicators that compare two different networks on a scale-free base; however, they are not used in this dissertation, and therefore, this chapter does not discuss them[21].

The density of a network is expressed in degrees, which depict the number links within the graph. In case of non-directed graphs, the degree can be calculated by applying the formula $d(G) = \frac{1}{2}\sum_{i=1}^{N} e_i$. It must be noted that all loops increase the degree by two. In case of directed graphs, inward and outward edges must be handled separately. The degree is given by the sum of these edges: $d(\vec{G}) = d^{+}(\vec{G}) + d^{-}(\vec{G})$. Degree can thus be calculated for all vertices, and a discrete probability distribution, called degree distribution can be defined. This suggests how the nodes are connected, and it provides information concerning the structure of the graph. In some special cases, degree distributions are similar to known probability distributions. For example, in the Erdős-Rényi random graphs, the degree distribution is binomial (Erdős & Rényi, 1960). For high volume of nodes in a random graph, one will get a Poisson distribution[22] (Daudin et al., 2008).

[19] Compared to the planned level.
[20] This subchapter partly relies on the book by László Barabási-Albert (Barabási & Pósfai, 2016).
[21] The reader can find more in Soundarajan et al. (2014).
[22] Poisson distribution is an extension of binomial distribution for large samples and low probabilities (Consul & Jain, 1973).

Owing to its Hungarian relation, it must be noted that for very large random networks (e.g. the Internet), the degree distribution follows the power law. That is $p(x) \propto f(x)^{-\alpha}$, $\alpha > 1$. These networks are often called scale-free networks. The evolutionary model of such complex scale-free graphs was developed by Barabási and Albert (1999). This model simulates the evolution of a non-directed random graph, in a way that it increases the number of vertices by one in every iteration and it links the same to other nodes at the probability proportional to the current degree of the network (this can be calibrated as a hyperparameter). By iteration, a scale-free network will be developed. Figure 4 presents the evolution of such network.

4. Figure: The evolution of a scale-free Barabasi–Albert network



source: generated random graphs by *igraph* package in *R* (code R1)

As one can observe in Figure 4 created using the algorithm of Barabási–Albert, hubs (dense nodes) develop as the network grows. The importance of these vertices in the network is higher as compared to those with fewer edges. Network science uses node degrees to infer to the centrality of a particular vertex. The so-called centrality measures provide three basic measures (Borgatti & Everett, 2006):

- Degree centrality: the simplest centrality index, which depicts the relative importance of a node in a network as compared to the degree of all nodes: $c_{v_i} = \frac{d(v_i)}{d(G)}$. It assumes that the degree of a node is proportional to the information it can access. This indicator also exists for weighted edges; however, it assumes that the nodes represent virtually the same importance (or information) (Bródka et al., 2011). This does not hold in the GVC landscape because the nodes represent countries of different sizes.
- Eigenvector centrality: this indicator measures not only the degree of a node but the degree of those nodes to which is linked. A high rank can be achieved if a node has a lot of edges or it has a limited number of links to nodes with high importance (without saying if a node has many connections to other high significance vertices the centrality measure will be very high). Eigenvector centrality can be derived from the adjacency matrix: $c_{v_i} = x^{-1} \sum_{\{i,j\} \in E} c_{v_j}$, where c_{v_i} is the eigen centrality of node *i* and c_{v_j} are the nodes (*j*) linked to *i*, while *x* is the largest eigenvalue of the adjacency matrix **A**. In other words, c_{v_i} is the solution of the equation $Ac_{v_i} = x^{-1} c_{v_i}$ (Bonacich, 1972). In case of directed graphs, the index is far not perfect, because the inward and outward edges must be handled separately, and if the distribution of these edges is not uniform, it may lead to a conspicuous bias in eigenvector centrality.
- Katz centrality: the aforementioned problem is solved using the Katz centrality index, which assigns unit centrality measures to all nodes: $c_{v_i} = x^{-1} \sum_{\{i,j\} \in E} c_{v_j} + \beta 1$. The derivation of this indicator from here follows the classical centrality measure (Katz, 1953).

Networks can be analysed along smaller subnetworks because clusters can be formed within the graphs. These groups are often called cliques. Every cluster is a clique if it contains at least four nodes that constitute a complete subnetwork. The formation of cliques is typical in case of sparse graphs. In IO framework, all nodes have links to all other vertices, and thus, there could be no clusters. Because of the same reasons, the density indicators also cannot be utilised here. The latter depicts the number of edges of a node to the maximum number of possible nodes.

Being a complete graph is a peculiarity of IO tables, which constrains the applicability of network theory methodology. Thus, the demand for pruning the network usually emerges. In this context, an algorithm finds and eliminates the edges that have very little effect on the characteristics of the network. One of the classical pruning algorithms deletes edges in accordance with the probability, which is inversely proportional to the degree of the nodes linked by the edge in question. Another pruning method is the one that neglects the edges that have a 'shorter' alternative in the Euclidian space (Ahn et al., 2012). This algorithm assumes that the edges are substitutable, which is not plausible from an economic standpoint. Therefore, pruning the network of GVCs is not recommended at all, because it eliminates the existing trade relations that are certainly optimal, and their substitutability is not ensured in the short run.

Segmentation is a popular method in network analysis. Such tools create subnetworks within a graph, and they are fairly similar to the tools applied in multivariate statistics. In a GVC framework, these procedures also have limited useability, as the specificities of these models, the segmentations usually take place along the nodes with high weights. Thus, the world is segmented along the two largest exporters, which suggests that the two largest and most important countries or industries are completely separable, albeit the most intensive trade is most likely to take place between these two countries. This problem is further elaborated in Chapter 5.

1.6 Research questions and hypotheses, the contribution of the dissertation to field

The primary aim of this dissertation is to contribute to the theory of GVCs and enrich the accounting practices through international examples. The literature concerning GVCs is continually developing, and more and more case studies are published, and several international datasets are available for researchers. Consequently, the methodological toolboxes in the hand of the analysts are getting more complex. Despite this, most indicators that are created during the analyses are not robust enough, which can bias inference. This research was focused on the study of these indices and attempted to adjust them in a way that could help better understand GVCs.

The first network analyses on the field of international trade were published in the early 2000s (Garlaschelli & Loffredo, 2005; Serrano & Boguñá, 2003), not long after the research group of László Barabási had published their results concerning the graph theoretical concepts of the World Wide Web and had proven the existence of scale-free networks (Adamic et al., 2000; Albert et al., 1999). In the early 2000s, the collection of trade statistics was burdensome, not to mention that only a few software supported network modelling. Thus, mostly physicists and mathematicians conducted investigations without any meaningful economic interpretation. A decade later, Trade in Value-Added (TiVA) statistics were published and the data could be utilised to conduct network analysis. Since then, several studies of high importance were issued with inference in the field of economics (Amador & Cabral, 2017; Cerina et al., 2015; Criscuolo & Timmis, 2018; Ferrarini, 2013). Notwithstanding these notable works, graph theory has received moderate attention only. One reason behind this could be that the standard indicators used in network science are generally inadequate in GVC research, and visualising the network is not considered a scientific achievement anymore.

Analysing dynamic graph is more than merely drawing the states of the network in different *t* times, but one should also model the flows. For this, the analyst must understand the sequences of the process, which order the events in the network. This is also valid for the flow of value added in the value chain; however because of the over-aggregation, no method is available to reveal the sequences of global trade. Nevertheless, in particular cases (for regions, for a few products), an estimation can be done regarding the order of the sequences, and this will be presented in this dissertation.

Revealing the sequences can be done under the condition of value-added disaggregation from the statistics. Owing to the fact that value chains (in contrast to the classical network-flow models) have no absolute source points. However, they have well-defined destinations (sinks). A method shall be presented in this dissertation, which is suitable to follow the flow of value added between two random points in the value chain.

These innovative tools will help in better understanding the risks within the GVCs. Recent events (natural disasters, epidemics) have directed one's attention towards systematic risk of over-dependencies in the GVCs (Gereffi & Luo, 2014; Lee & Gereffi, 2015). The interpretation of these studies from an economic policy viewpoint was that the length of value chains was shortening. These protectionist trade policy sentiments were enhanced during the COVID-19 pandemic when some scholars called for the return of offshored companies (Baldwin & Tomiura, 2020; Dachs et al., 2019).

However, as the length of value chains has a marginal role in the formation of systematic risks, the market structure may be more important. This could be analysed by quantifying market structure.

The analysis of GVCs necessitates a complex and interdisciplinary approach, which develops the consistency between the estimations and their economic interpretation. This dissertation applies a holistic approach and analyses the GVCs through regional examples, while it puts the Hungarian economy and Hungarian firms into focus. In last decades, Hungary developed the supplier positions that were required to join the international value chain of electronic, machinery, and motor vehicles products. At the same time, there are some indications that Hungary could not fully capitalise on international trade and it is converging towards the middle-income trap (Bod, 2015; Győrffy, 2021). To resolve these problems, certain issues should be discussed: What is the actual relative position of the Hungarian enterprises? How far could the Hungarian value-added get in the value chain? Which domestic and foreign industries provide supplies to the local firms? And finally, how do the comparative advantages and systematic risks look like?

This dissertation poses the following research questions. The methodologies utilised to answer each question are covered in each chapter:

- How do the globalised international trade, production in value chains, and new forms of trade affect official statistics?
- Considering the bilateral trade relations of Hungary, how could the companies and industries be positioned in the GVCs? How could the relations that have particular importance be visualised?
- How far can the Hungarian value-added get in the GVCs? Which routes are the most important? Where are the hubs?
- How do the sequential differences affect the value of the GVC indicators? Is it relevant when countries with similar production profiles enter the production chain?
- How do domestic firms participate in value-added flow in the GVCs?

Regarding the research questions, the following hypotheses can be posed:

Hypothesis I: The current accounting practices of official statistics concerning the transactions between companies in the same value chain can significantly bias macroeconomic statistical indicators.
Hypothesis II: Hungary has strong relationship with other regional countries, while the connection with economies outside the European Union (EU) is rather weak.
Hypothesis III: The Hungarian value-added is circulating mainly in Europe, and typically one cannot measure its presence outside Europe.
Hypothesis IV: Compared to economies of similar profiles, Hungary has joined the value chains later and this has biased the value of GVC indicators downwards.
Hypothesis V: The volume of indirect value-added of the participating Hungarian companies in the value chains is larger than the direct flow, and that amount is mainly produced by the small- and medium-size enterprises.

By utilising the methodology discussed in this dissertation, all research questions were answered and all hypotheses were accepted.

2. Global value chains in a changing global landscape

2.1 The development and importance of global value chains in the global economy

GVCs are one of the most researched fields in economics today. They constitute a research area of macroeconomics, international economics, as well as business economics. GVCs have interwoven national economies without regard to economic development, specialisation, and openness of a given country. The global trade is dominated by GVCs. Nevertheless, our knowledge of the GVCs is quite limited as it is difficult to draw any conclusions from the available data. On the other hand, there is no universal research framework that can clearly define GVCs and make them quantifiable. In absence of the said information and framework, case studies of individual chains provide the most information. See for example Stevens (2001) about agriculture, Campling (2015) regarding the impacts of customs in the trade relationship between the EU and the USA, or Jiménez-Zarco et al. (2019) concerning the role of value chains in the fashion market.

The data scarcity is caused by the rapid and unforeseen changes that eventuated with globalisation, which international organisations and national offices of statistics cannot cope with as the hegemony of the industrial concerns declined. From the beginning of the 20th century, the emergence of mass production enabled the evolution of industrial concerns, which reveal the common pattern of production to be at a central place or at least geographically concentrated. Heavy and light industry districts were emerging (like Detroit), service centres were established especially in the field of finance (like London and Frankfurt), and however, later in other business support service providers (consulting, telecommunication). The large concerns had an overview of the whole production process as they integrated the bigger part of the workflows into the concern. The only exception was the extraction of raw materials. One concern owned the statistical data of almost the entire production chain that the statistical offices required. The international merchandise trade was primarily meant to transport the final goods to the consumers, and thus, the trade was less sophisticated than it is today. The trade of intermediate goods was not common because of the proximity of the suppliers in geographical terms, even if a supplier was not owned by the concern. These satellite companies were highly dependent on the customer, who was their only client in most cases.

In the last decades, the production has been segmented quickly owing to the changing investment and trade policies, especially the emergence of foreign direct investment (FDI) friendly economic policy and the accelerated development of communications and transport technologies. At the dawn of globalisation, the concentrated industrial concerns began to loosen up and subsidiaries emerged at remote places, often abroad. This gave rise to the process of internationalisation. The subsidiary firms remained in the ownership of the parent company, though it was more efficient than the production by the parent company itself. The subsidiaries could use the local competitive advantage, which was mostly lower labour cost and better access to raw materials or markets. It became possible to separate production phases and outsource them to different geographically remote places.

The rapid evolution of the information and communications technology, especially the trend of digitalisation, also known as Industry 4.0, has facilitated the production to be managed remotely not only in areas of management-related matters but also in physical implementation (Czakó et al., 2010; Hayter & Watts, 1983; O. Kovács, 2017).

The ownership structure of the suppliers lost its importance with the transformation of the corporate governance culture. This was fostered by the swift development of the global demand, which claimed new production capacities, and the spread of standard production processes. The latter required rigorous specification and quality control for the suppliers that cannot be changed; it, however, allows that the production can be accomplished at any point of the world. To control the suppliers, contractual agreements became largely prevalent instead of ownership, as this process was also supported by free trade and investment protection agreements in the 1990s. The rapid growth of digitalisation was also added to the process that laid the foundation for new companies, which were already global[23], such as Amazon, Google, and Facebook. The topology of multinational companies thus changed considerably after a specialised company group became a central player with no overview of the production process. However, it has a comparative advantage in the understanding of consumer behaviour, as extensive knowledge has been acquired through data mining operations.

[23] Hence, the born-global naming (Knight & Cavusgil, 2004).

The technological factors, above all digitalisation, play a central role in the development of GVCs as they reduce the production and transaction costs (Lund et al., 2019) and generate a brand new kind of commerce besides those, which emerged parallel to traditional bilateral trade of final goods a few decades ago[24]. The data value chains created not only a new form of commerce but a new industry as well that follows the usage of the goods and with tailor-made services contributing to their value well after the purchase (Kaiser et al., 2019).

Between 1990 and 2018, the volume of the global export of goods at constant price[25] was tripled. Nowadays, all the countries are part of the global trade apart from a few exceptions (countries under embargo). The global spread of multi- or transnational companies, the change in production and inventory management processes, outsourcing, free trade agreements, organisation into value chains, free movement of capital, low interest rates, and liquidity on the money markets contributed to the severe increase in global trade volume. The growth policy of the USA and China built on the increase in consumption provided sufficient demand until the global economic crisis in 2008–2009. The spread of the recession was fast in the world because of the globalised money market and through those real economic channels, which were established by the value chains earlier (Milberg & Winkler, 2010).

Following the financial crisis of 2008–2009, the extremely fast processes called hyperglobalisation[26] were stalling, and a slow-down occurred, which was called 'slowbalisation' by Timmer et al. (2016) or 'deglobalisation' by (Antràs, 2020). Since 2010, it is evident that the global trade became fragmented and regional blocks began to emerge because of protectionist trade policies. They replaced the global production structure (Baldwin & Lopez-Gonzalez, 2015), and consequently, the value chains became shorter (Miroudot & Nordström, 2019), individual companies started to repatriate their production to their own countries or at least closer to the final consumer (Ancarani et al., 2019; Backer et al., 2018). The phenomenon known as 'backshoring' still needs some time to be proven as the latest GVC data include only 2016 at best, and the process of repatriation started after 2018, which was strengthened by the global pandemic in 2020. The latter

[24] See later in Chapter 3.
[25] The prices were adjusted to the level of 2018 by consumer price index of the USA as a deflator (source: OECD).
[26] The expression originates from (Rodrik, 2012) American–Turkish economist, who characterised the period after the 1990s, when the production and business environment factors (regarded as location-specific) were globalised.

started a whole bunch of research projects concerning GVC reshuffles (Gereffi, 2020; Lund et al., 2019; Strange, 2020). Nevertheless, it is important to keep in mind that the pandemic might only cause a realignment but not a return to the global trade trend before the hyperglobalisation. The globalisation has changed the environment of the global economy; China's role has been appreciated, and it has become the world's leading supplier and the second largest consumer (Fernandes, 2020). In such cases, the principle of 'closer to the consumer' assumes another meaning.

2.2 The relevance of value added in the value chains

The examples above are only snippets of the structural changes in international economy and trade in the last two decades. Furthermore, the theory of GVCs is not clear, and recent studies have underlined the disadvantages of integration to the GVCs (McGrath, 2013; Stringer & Michailova, 2018) and shortcomings of current theories. However, one cannot assume that the pertinence of the value chains in the global economy would decline; on the contrary, it is larger than ever.

Hence, it is essential to precisely define the concept of GVC. There are more overlapping definitions in the literature. The first, the most remarkable observer of this phenomenon was *Michael Porter*, who defined the value chain (not global at the time) in his study as decomposition of a company according to its functions. (Porter, 1998) argued that companies specialise in functions (R&D, marketing, production, etc.), in which they have a comparative advantage. Thus, the objectives and strategies of different companies (or their functions) will not contradict each other[27].

The rise of research in value chains started only a few years before the millennium and it is connected to (Gereffi, 1994), although he wrote about global chains of goods. Therefore, the most common definition of value chains today can be linked to Kaplinsky: 'The value chain incorporates all production activities from the concept through the intermediate phase to the shipping of the final goods to the consumer' (Kaplinsky, 2000). Originally, geographical proximity was considered as a decisive factor in the development of

[27] Earlier the collaboration of departments as functions within a company was studded with conflicts.

the chains and in the evolution of the supplier network (Leslie & Reimer, 1999). However, later, (Los et al., 2015) confirmed that the share of foreign value-added is continuously growing along the GVCs.

However, (Sturgeon, 2001) pointed out in his article that the definition by (Kaplinsky, 2000) regarding value chains is not exact. Kaplinky's (2000) interpretation is too broad, and thus, the production network cannot be separated from the value chains. While the former can characterise an intercompany network, which can be an agreement or a cluster, the latter encompasses the division of production functions, which materialises in value-added supplied to one another.

7. Table: Definition of the value chain and network

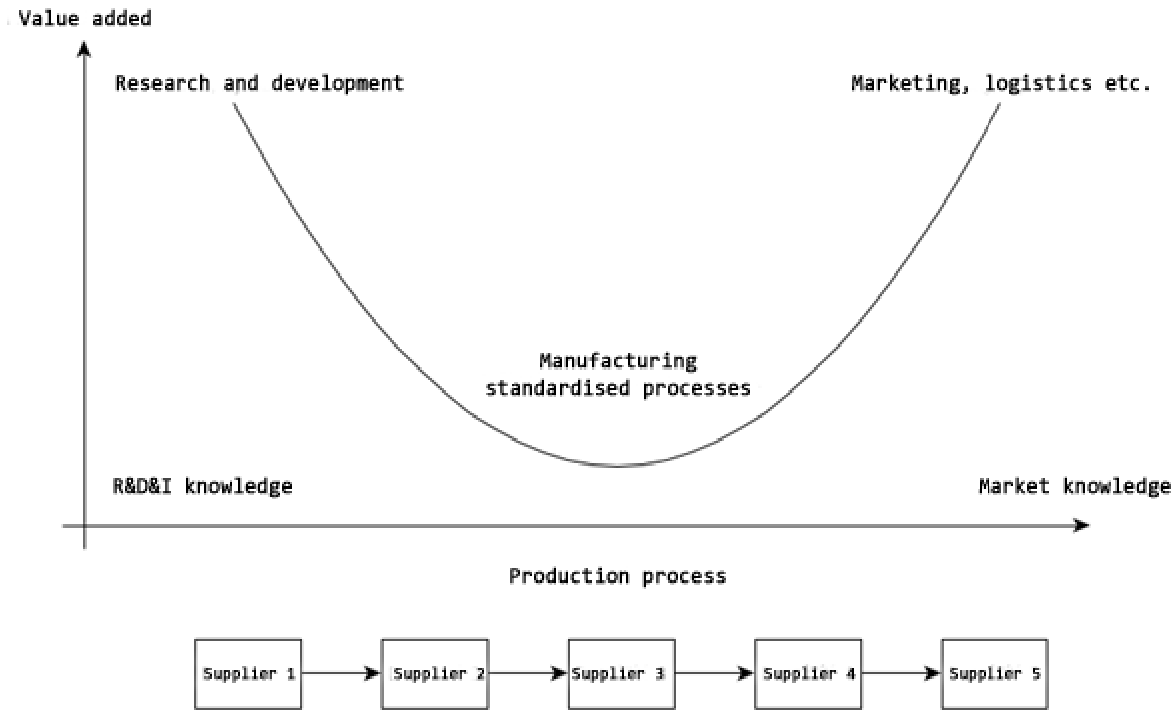
Name	Definition	Way to gauge	Other names
Value chain	A series of production sequencies, at the end of which a final product is produced.	Group of activities in which producers participate.	Supply chain, Chain of goods, production chain, chain of production activities
Production network	A set of intercompany relationships that bind a group of companies into a larger economic entity.	Extent and characteristics of intercompany relations	Value network, supply base

Source: own edition based on Sturgeon (2001)

It has been a considerable milestone in value chain research to include the concept of value-added, as it is a measurable value on micro and macro levels. On the basis of the produced value-added, a relation can be determined between the production sequences.

Therefore, it is possible to evaluate and rank the importance of the companies, company groups, and countries (taking part in the GVC) in the value chain. This enables the connection between business and national competitiveness and GVCs. At that time, the analysis of the value-added content of the different phases was not possible because of lack of data. Eventually, the research of (Mudambi, 2008) based on a case study was published; it analysed vertical integration and showed that manufacturing creates the lowest value-added in the whole production chain, while the design process before physical realisation or the activities following the manufacturing (marketing, distribution, etc.) generates higher value-added. The flowchart drawn by Mudambi (2008) is known as the 'smile curve'.

5. Figure: The smiley curve of value chains



Source: Mudambi (2008)

The production phases shown in the curve and the companies or countries specialised therein are working on manufacturing the same final good and have the same economic interest, though operating under different market conditions. At the two ends of the curve, the intensity of market competition is lower as acquiring the requisite knowledge for these activities is not simple (its owners are in monopolistic situation), whereas the competition is much stronger in the manufacturing activities. In the long run, it could increase the inequalities[28], as the strong competition in the manufacturing phase hinders the growth of value-added, while there are much better possibilities to increase the profits at the end-points (Stöllinger, 2019; Vakhal, 2020).

Before Mudambi (2008), (Heintz, 2006) had already proven that the advantages of the value chain are not equally distributed among the actors. The brand owner determines the consumer price according to the production costs and consumer demand; therefore, *ceteris paribus*, the changes in supplier productivity alter the market price. However, in the end it is not the brand owner's interest to share the profit gained from the lower production cost with the suppliers, especially in case of a strong supplier market competition. As a result, the profit gained from the lower production cost is only shared between the brand

[28] The phenomenon is also known as the Prebish–Singer hypothesis (Harvey et al., 2010).

owner and the consumer. Nevertheless, the supplier is still encouraged to increase its productivity, but it is only beneficial in the middle term, as strong competition pushes all actors in the direction of innovation in the longer term. Heintz (2005) argued that under these circumstances, the export production connected to the value chains leads the economy to the ‘middle-income trap’ in the long run. (Engel & Taglioni, 2017) presented the same reasoning in their analysis of business behaviour. They concluded that the company without productivity growth might be caught in its own trap that could cause its deterioration in a fiercely competitive environment.

Heintz’s (2005) theory is chiefly valid for emerging and middle-income countries; however, its notable shortcoming is that it does not consider the scenario of selling the produced goods in the suppliers’ country (namely reimporting them). In this case, the distribution of production functions has a welfare impact in the suppliers’ country that is proportionate to the size of the market. Nevertheless, these findings direct one’s attention to the vulnerability of the smaller, low-income countries and to the dangers of excessive export concentration.

While the mid-term goal of the companies in the value chain is to improve their productivity, the long-term objective is to cover new production functions. This process is known as ‘upgrade’ in the literature. The upgrade might have considerable welfare impact as it generates higher value-added. Thus, the country is interested in domestic companies covering new functions and moving from the minimum point of the ‘smile curve’. (Knorringa & Pegler, 2006) asserted that the value chain is also a beneficiary of its members’ upgrade as the quality of the produced goods improves. Although they also noted that the large company at the top of the chain has to be more responsible and proceed with ethical procurement. Contrarily, (Gereffi & Luo, 2014) called attention to the existing jobs at the companies at the bottom of the value chain, as they are operating with low wages and are thus insecure and often quite dangerous.

There are only two ways for a company to stay in the competition. Either it establishes a stable position in the strong competition at the current level or it does the upgrade und uses the competitive advantage of innovation as a newcomer. The upgrade in the value chain can be interpreted as an effective combination of production factors (capital and work) that can enable the assumption of new production functions, so that the company can maximise its profit (namely the price is higher than or equals to the marginal cost).

The occurrence of a new production function is a prerequisite; otherwise, there is only increased productivity that would result in a higher market share and no access to other markets. (Szalavetz, 2012, 2013) argued that the upgrade can be functional (increasing the share of white-collar workers increases the value of immaterial goods in the balance sheet) or through manufacturing goods of higher technological standards (which has higher capital intensity).

Since then, several researchers (Baldwin, 2012; Meng et al., 2017; Shin et al., 2012) have empirically confirmed the existence of the smile curve using global data. Companies from developing countries specialised in manufacturing (e.g. China, Vietnam, and Malaysia) can be found at the minimum point of the curve, while company groups resident in developed countries such as Japanese, American, and German firms are at the two ends. It must be noted that measuring value-added in the different phases through case studies is challenging; it is almost impossible to separate inputs from production and management (as the latter ones are not part of the value chain). Therefore, researchers illustrate wage per hour at different production phases. This approach is plausible as material input is not significant in the early stages of production (R&D, design, etc.), while during the manufacturing phase, the increase of workforce input leads to output growth (Vakhal, 2018a).

As the speed of information flow has been increasing, the services have become an integral part of the value chain, and thus, contrary to the early Balassa approach, particular service sectors (business services, commerce, etc.) have become part of the tradable sector. (Miroudot & Cadestin, 2017) affirmed that 40% of value-added in the value chains is provided by services. However, this is not only the privilege of service sectors, as a part of exported services is provided by manufacturing companies linking the service to the product. Additionally, a considerable part of services is linked to foreign investments that enable financing (financial services) and manage (IT, telecommunications services) the project and helps the flow of products through transport and storage services (Heuser & Mattoo, 2017).

Furthermore, the proportion of those companies increased that are only indirectly connected to the value chains via a direct exporter (see Chapter 8). These implicit connections spread out the risks from the value chain and hide them, as the supplier network of a domestic export company cannot be detected in foreign trade or production statistics. The different shocks widely affect the economy through multiplier effects hitting companies in the third or deeper supplier circle. According to the calculations of (Baldwin & Weder di Mauro, 2020) and based on the model of (Bems et al., 2010) 20%–30% economic decline in Europe or in the USA is linked to the economic output of other countries, and the connections are built by the value chains.

All this has considerably softened the definition of the GVC (Koopman et al., 2010): GVCs '… are a system of value-added sources at different locations in a globally integrated production network'. Therefore, the latest definition does not cover product types, distribution channels, companies, or companies organised in a chain but organisation into networks is the central focus. This transforms the image of GVCs (in some studies a new phrase global value network – GVN shows up) in several points:

- The role of the parent company has changed – earlier the firm that owned the highest value-added tracked all the phases of the production. The central actor in the chains is still the owner of the highest value-added; however, they do not have the same control over the suppliers as earlier.
- Entering the network is easier – it is possible to appear with low value-added, and companies are thus allowed to enter the international market with a small investment.
- In case of value chains, the most crucial development question is: 'how is it possible to upgrade and generate higher value-added'. In value networks, this aspiration is more nuanced, as the upgrade in value chains is followed by the change of the function (the manufacturing function is complemented by R&D), while in the network, it is possible to increase value-added within the same function (e.g. through growth in production with a new supplier agreement).

The second point explains why the developing countries could reach a considerable export growth from the end of the 20th century. The organisation into production networks and the international division of labour enabled low capital companies (and countries) specialising in smaller processes to get involved in the production. Earlier, this was not possible because of the geographical concentration, and establishing a concern needed larger time, knowledge, and capital investment. Specialising in a production sequence represents a much smaller cost of entry, and the parent company contributes to the connection through knowledge transfers[29]. This contributes to the growth of economy and welfare in the home country of the supplier company.

In accordance with the third point, the emphasis is shifting from the value-added to the jobs. (Baldwin, 2014), a long-time researcher of value chains, stresses that jobs, more precisely good jobs, are important in the value networks. He further argues that the production process is no longer linear and is rather a network chain of suppliers organised around several centres like a satellite. In such a chain, specialisations do not change but the earlier satellite supplier can grow. The story of Foxconn in Taiwan is a perfect example. The company was founded in 1974 with the goal of manufacturing electronic parts (supplying intermediate products), and today, it is a supplier of all large electronic companies. Although Foxconn has been developed to a monopoly under special circumstances, its example still illustrates that a smaller supplier can become a network centre without altering its original function in the GVC (low value-added intermediate products) (Ngai & Chan, 2012). The structure of parent company suppliers has been softened, and nowadays there are different types of governance systems (Gereffi et al., 2005). The upgrade is not only possible through changing the functions but also through altering the organisational structure. The opportunities of the suppliers are constrained but there is still a possibility to change the position with a fast reaction to industry transitions.

Since the millennium, the structure of production has been changed according to the purpose of consumption. (Wang et al., 2017a) resolved the structure of produced goods and services on the basis of a new approach and suggested a new classification system:

1. Goods produced for pure domestic consumption.
2. Goods produced for classical foreign trade: final products consumed in other countries.
3. Supply to a simple GVC: In this case, the produced intermediate good leaves the country and does not return. It is turned into final good and consumed at the place of destination. These activities are usually the results of cooperation in border regions.
4. Supply to a complex GVC: In this case, the intermediate good crosses several borders before turning into a final good, and it can also return as intermediate or final good to the manufacturing country.

[29] This means procedures, machines, equipment that are leased by the suppliers and involve huge costs.

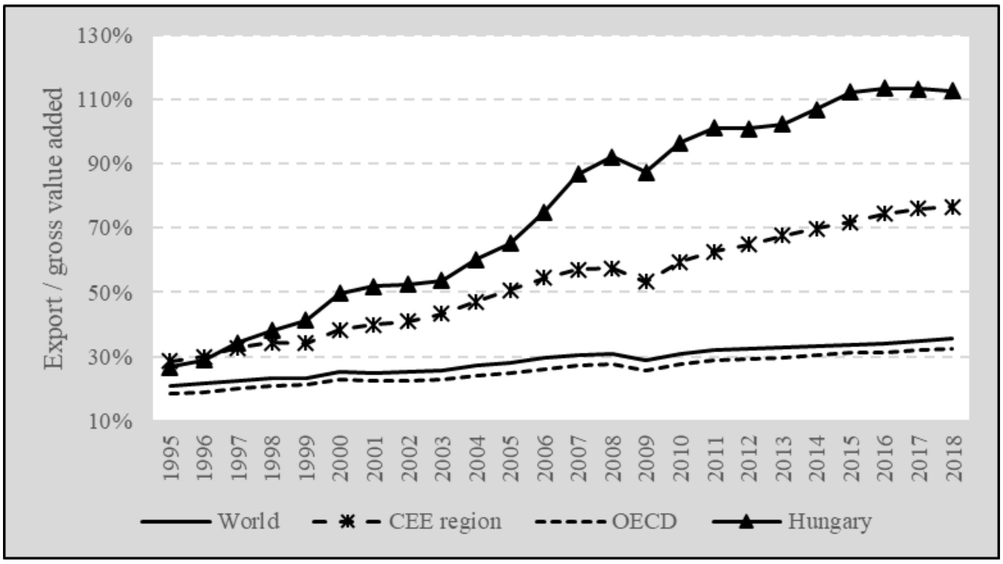
Formally:

GDP = Pure domestic products + Classical foreign trade + Simple GVC + Complex GVC

According to the calculations of Wang et al. (2017), the proportion of goods produced for pure domestic consumption decreases (from 85% to 80%), while the supply into GVCs increased by some percentage points. The highest value-added per product unit or work-hour is provided by supply into complex networks, while the lowest value-added is generated in domestic consumption. However, after the global financial crisis in 2008–2009, the growth of value-added markedly decelerated in all categories because of the downturn of domestic demand and global production (Degain et al., 2017), apart from the phenomena mention earlier.

The slowdown of the value-added growth and the uneven profit distribution of global growth called attention again to the paradox operation of the value chains. The export value is continuously increasing in the less developed countries; however, the growth rate of domestic value-added lags behind considerably. In Central- and Eastern-Europe, the proportion of export according to gross value-added increased by 45 percentage points in 11 years, whereas the change was only 13 percentage points in the OECD countries. The trend of the latter one shows relative stability; however, the growth continued in their region. With a shift in focus, the export of the Central- and Eastern-European region made 4% of the world export, while the generated value-added was only 2% of the world production.

6. Figure: The ratio of gross export in gross value added in the world, in the CEE region, in the OECD member states and in Hungary, between 1995 and 2018 (%)



Source: World Bank

The explanation of the processes shown in Figure 6 is very complex; on the one hand, the export growth is much faster than the increase in value-added that refers to the diminishing weight in the production network; on the other hand, the exposure to export is continuously increasing.

The main emphasis is shifting to the countries and companies that control the production. According to the calculations of (Aguiar de Medeiros & Trebat, 2017), foreign value-added in the export of developed countries is approximately half compared to the developing countries. (M. P. Timmer et al., 2014) demonstrated that the value-added generated in the GVCs is coming from 21 high-income countries[30], and thus, the distribution of profit originating from the value chain is likely not even.

The growth of export dynamics exceeding the economic growth bears dangers as foreign demand for domestic goods gives the companies and policymakers the illusion that the country is highly competitive; moreover, its competitiveness is continuously increasing. Current economic theories and indicators often support this false interpretation. However, classical competitiveness theories are transforming with the emergence of GVCs, and the statistics relying on classical theories present a much different picture.

[30] This category also includes Hungary according to World Bank's classification.

2.3 Summary and conclusions

This chapter gave a brief insight into the history of GVC development and provided a definition and allocated the phenomenon in the literature and of international and business economics. It can be concluded that the rapid growth during the financial crises of 2008–2009 (‘hyperglobalisation’) was replaced by a moderation. Case studies and research reports concerning developed economies[31] draw the shortening of value chains and regression of the global economy (‘deglobalisation’).

GVCs exist in diversified forms depending on which production sequence the investigated country or industry is located in. This determines the volume of value-added it can contribute to income generation, to government budget revenues, and to the growth of the industry and the national economy. The length of the value chain determines the number of produces among which the value-added must be distributed. It also identifies dependencies and strength in the value chain. The emergence of dependencies entails severe risks because a production hiatus at one or two producers can halt the production in the short-term or make the operation much pricier. This is the result of production structures dominated by the Leontief-type production function. One can observe this in the graph topologies (see Chapter 7), which cause bottlenecks in the supply and the production.

Intense competitiveness for the current positions in the value chains pushes the firms towards upgrade, making them interested in innovations. However, the practice does not confirm the theories. One cannot find a certain example for upgrade in international environment but for diversification only. In other words, the event of firms with high productivity penetrating into a market of low-productivity firms. Thus, countries with low productivity are still at the risk of middle-income trap (Eichengreen et al., 2013).

[31] The international IO tables for 2020 are likely to be available in 2026–2027 only.

This directs one's attention towards the risks hidden in GVCs, which are not uniformly distributed among the actors. Moreover, via the multiplier effect, they have an impact on other economic actors that are indirectly linked to GVCs. In order to evaluate the pros and cons of value chain participation, one must know the actual position of the countries, industries, and companies in the supply network, just as their relationship with each other. This dissertation aims to study this while focusing on the Hungarian and regional economy.

3. Accounting global value chains in official statistics

There are two ways to analyse GVCs: through the international IO tables (see Chapter 4) or via case studies. Both sources have their own flaws: for the latter, a bias can arise from the weakness of the sampling or the low number of samples, while for multiregional IO tables (MRIO) analysis, the over-aggregation can cause problems besides the strict theoretical assumptions, which typically fail in practice (see Chapter 5). The gap between the two methods should be bridged by the official statistics and by that all studies would become controllable and verifiable. Nevertheless, the statistical agencies lagged behind in measuring globalisation, and thus, there is significant data scarcity that could credibly demonstrate the role of companies, industries, and countries in GVCs.

The lack of data not just impedes the investigations but also induces severe biases in those statistical indicators that are highly affected by the actual position in the GVCs. This chapter shall present the changes in globalisation of the past decades that substantially influence the development of the macro statistical indicators and bias the inference from them.

3.1 Accounting the ownership structure of the resident companies

The foundation of a subsidiary in a foreign country is considered to be a milestone in the classical economic literature of corporate internationalisation (Johanson & Vahlne, 1990). These firms were generally 100% owned by the parent company, because in that way, they could have control over the production, not to mention investment protection. The statistics of Foreign Affiliates Trade in Services (FATS) register foreign investments into resident companies (inward FATS) and investments of resident forms into foreign companies (outward FATS). With these statistics, one can acquire an insight into the share of subsidiaries owned by foreign affiliates in the domestic economy. Though the data provider communicates the industry of the investor (if it is a company, not an individual) to the data collector (in Hungary it is the National Bank of Hungary), still only the country of origin of the investor is published in the statistics. However, it would be burdensome to provide more details in the official statistics.

8. Table: The share of foreign affiliates owned by parent companies of the top 5 investor countries in total value added at factor prices in 2018

	Czech Rep.	%	Hungary	%	Poland	%	Slovakia	%
1	Germany	15%	Germany	14%	Germany	8%	Germany	14%
2	USA	6%	USA	8%	USA	5%	USA	6%
3	United Kingdom	3%	Austria	4%	France	4%	South Korea	3%
4	France	3%	United Kingdom	3%	Netherlands	3%	Netherlands	3%
5	Austria	2%	France	3%	United Kingdom	2%	France	3%
Domestic	Czech Rep.	57%	Hungary	53%	Poland	63%	Slovakia	53%

Source: own calculation based on Eurostat data (table *fats_glb_08*)

Table 8 illustrates that almost half of the domestic value-added at factor cost is generated by foreign affiliates in the Visegrad countries. The role of Germany and the USA in the region is very clearly outlined, virtually fourth-fifths of the value-added produced in the economy can be linked to companies owned by residents of one of these countries.

If one intends to investigate the role of subsidiaries in the GVCs, the FATS statistics do not provide sufficient information. The roots of this are in the methodology; one of the key factors of the change global economic landscape is the altered preferences of corporate governance, in which the role of subsidiaries is considerably less significant. Instead, the weight of contract base supplier transactions is surging (Nicita et al., 2013), and thus the partnership between two countries is only registered in the bilateral trade statistics. On the other hand, it must be noted that FDIs, albeit being clearly linked to GVCs, give an insight into the ownership structure of the TNCs, rather than providing a picture of the vertical supply chains. This is because the activity of the subsidiary is usually the same as that of the parent company. In such cases, the role of the subsidiary is not just a function of production but to supply the local or regional market as well. In the case of the latter, the subsidiary operates in almost all functions of production; however, this is typical mainly in case of low value-added production (like production of food and beverages). As compared to national accounts, the FATS statistics provide a more flexible definition of subsidiary corporations. Still, it is not common that a parent company would acquire a dominant share of ownership in the supplier's company. Therefore, the FATS statistics show only the satellite system of supplier (subsidiary) network, and it does not provide information regarding the whole vertical production chain. Thus, one cannot reveal the ownership of the firms that supply raw materials and other inputs into the final good.

The statistics of production and FDI are considerably affected by the changes of ownership structures, the emergence of TNCs and special-purpose entities (SPEs) in the GVCs and by transfer pricing. One of the main challenges in GVC accounting is the inaccurate definition of resident companies that consider subsidiaries of TNCs as separate entities (Rassier, 2017). It is a common phenomenon in GVCs that parent companies at the top of the supply chains have subsidiaries all around the world through which they capitalise on local comparative advantages. In such cases, the parent company legally has a presence in the country of subsidiary (it is its owner); however, it does not conduct any production (as it done by the subsidiary). Any transaction (capital or goods) within the TNCs are considered to be investments of trade. Even the input supplied by the parent company is considered as imports in the subsidiary's country (Lipsey, 2006). As a consequence, the role of the country hosting the subsidiary is upward-biased in the statistics because the capital as well as the net profit of a subsidiary (if it is 100% owned by a foreign affiliate) is owned by the parent company. If the latter is left in the subsidiary's country (re-invested), then it occurs in the value-added produced by the subsidiary. At the same time, the resident company contributed to the value-added only by the labour cost and the taxes. This bias is not adjusted by the FDI statistics, as the capital flows and stocks do not include the contribution of labour and local fundraising (e.g. a resident bank finances an investment) to production (OECD, 2008). In other words, during the evaluation of FDI investments of parent companies, the role of subsidiaries remain hidden, while during the evaluation of subsidiaries' performance, the importance of the parent company is downward-biased.

The aforementioned problems are just partly solved by the utilisation of the gross national income (GNI). According to the definition the GNI is the sum total of GDP and net receipts from abroad of compensation of employees, property income, and net taxes, subtracted by subsidies on production, of which only the net property incomes can be linked to the parent companies. In other words, from GNI perspective, the net property income is owned by the parent company; however, the contribution of the subsidiaries (through labour costs) is unknown. It is not marginal, that one of the main aggregates of the SNA considers subsidiaries as foreign residents; however, this is not applied in the foreign trade statistics.

9. Table: Differences in GDP and GNI in the Visegrad countries in 2018

	GDP (current prices)	GNI (current prices)	GDP/GNI
Czech Republic (CZK million)	5279.1	4962.7	1.06
Hungary (HUF million)	41480.6	39676.1	1.05
Poland (PLN million)	2108.3	2022.1	1.04
Slovakia (EUR million)	90.5	88.7	1.02

Source: own calculation based on AMECO data

Evidently, Tables 8 and 9 contradict each other, because while half of the income is produced by foreign affiliates in the region, there is only a few percentage point difference between the GNI and the GDP. The methodological reason behind this is that the GDP is adjusted only by income transactions, while the contribution of subsidiaries is neglected. Besides, the additional tax and subsidy content is also different:

$$\text{GDP} = \Sigma \text{ value-added (base price)} + \text{taxes on products} - \text{subsidies on products} \quad (7)$$

$$\text{GDP} = \Sigma \text{ value-added (at factor prices)} + \text{taxes on products} - \text{subsidies on products} + \text{taxes on productions} - \text{subsidies on production}$$

Thus, the value-added at factor cost does not include those contributions that are not related to the actual product, such as taxes on labour and property taxes. Therefore, the GNI blurs the importance of subsidiaries in the domestic economy because the taxes increase the GDP of the resident country. Nevertheless, it can be stated that the foreign-owned companies have a crucial role in the Visegrad countries, which also means that the relative development compared to the EU15 member states is lower if measured in terms of the GNI.

3.2 The impact of globalisation on statistical data collection

Changes in ownership structures have also impacted the foreign trade statistics. The foreign trade statistics[32] as used today accounts for the trade of 'more closed' economies 'better'. The assumption that the production relies on domestic inputs is valid only in the countries wherein the share of import is very low in production; thus, the export of the goods at CIP or DAP parity can be fully accounted to the exporter's country. In contrast

[32] This is one of the oldest statistical data collections in the world. Its methodology (except the updates nomenclatures) is mostly the same since the 1920s.

to that, the local base of production in more open economies is usually much smaller, such that the volume of domestic value-added in the free-at-frontier value is generally low (Vakhal, 2016b).

The GVC management structures that manage transactions via a holding in a third country (partly because of tax optimisation intentions) require particular attention. These SPEs do not execute any production activities, but they are a crucial part of TNCs, as they operate as accounting and administration units. The activities of such companies can significantly bias some items of balance of payments. Since 2006, the statistics of external balance have improved a lot because the authorities have been filtering out the impact of the SPEs form the balance of payments. However, this is only one aspect of the problems caused by globalisation in the world of statistics. The definition of an SPE is too strict, and hence, the filtering process is imperfect (Koroknai & Lénárt-Odorán, 2011). The reason behind this is that according to the definition by the IMF, SPEs do not execute any real economic activity at all and are thus practically considered as offshore companies. However, it is not necessary that only the SPEs can effectuate the aforementioned activities, as because of tax exemptions, any unit can generate additional income.

However, these companies carrying out pure SPE functions may give rise to a bias in GVC accounting, because they induce such capital movements that have utterly no impact on the real economy. Moreover, it can indicate high-volume flows of investments between countries that have very weak trade relations. It is particularly true in case of groups of companies that are not in the same currency area. For example, a parent company can hedge its foreign exchange risks if a subsidiary is not directly financed (which would be an FDI inflow in the subsidiary's country) but through another subsidiary in the same currency area and in a third country. This transaction is also considered to be FDI in- and outflows but without mentioning the owner of the capital.

Another recent development is that the subsidiaries also found their own subsidiaries – if it is a favourable measure because of taxation, legal, or corporate management reasons. This is a typical structure of those value chains that are managed from a European hub but owned by a parent company from the USA. In such cases, the owner of the resident subsidiary is the European holding, which is owned by an American parent company, but carries out all of its transactions with the European holding. The statistics register the bilateral capital movements only, and even if it may have an impact on the real economy, the US affiliation is not accounted for, because the beginning of the chain is usually unknown.

The foundation of logistical or supplier centres through which the regional value chain can operate, is also a GVC-specific issue (Manders et al., 2016). These centres conduct the international (and domestic) procurements, which decreases the cost of negotiations in international environment. These centres are often located in a third country and merchanting is performed. That is, intermediate or capital inputs are purchased by the supply centre. The goods never psychically enter the country but are directly shipped from the importer to the producer[33]. This intermediary trade is registered as re-export in the balance of payments of the intermediary's country. The export of services is only increased by the difference between entry and selling prices.

The existence of supplier centres calls the attention to another aspect of international labour share and value chain operations. The so-called factoryless goods producers (FGPs) are resident companies that offshore those production activities in which the goods are psychically transformed (usually the process carried out by the manufacturing industry) abroad in the form of tolling agreements. A peculiarity of the FGPs is that subsidiary ownership between the partner companies, which is gaining prevalence within the GVCs. The contractor in the tolling agreement owns all immaterial goods, while the owner of the materials and parts varies, but usually the subcontractors ensure the required inputs for the production. All partners in the process are clearly visible for the statistical offices, as such transactions are well known for a long time, and thus tolling agreements usually do not bias foreign trade and production statistics.

The FGPs are peculiar because they make 100% of their manufacturing production in form of a tolling agreement. However, this often causes a problem of activity code mismatch, because the FGPs usually do not carry out a particular industrial activity under which they are registered in the statistics. In other words, if an automotive company offshores 100% of its manufacturing production into a third country, but it keeps all services in the country (e.g. R&D, sales, marketing, distributions), and then in reality it does not produce any cars and only provides services (mainly retail trade). In such circumstances, the accounting of the transactions in foreign trade statistics is ambiguous because the

[33] This process is also called triangular trade and it is a very typical activity of e-commerce companies (such as Amazon).

products imported by the contractor in the tolling agreement are final goods, while the outward transaction is export of services (it provides the know-how to the subcontractor).

Consequently, the statistics can realise the FGPs as intermediary companies that resale the products. If it is done by another company in a third country, then it is a triangular trade. If the re-export activity is psychically not realised, then it can happen that in the balance of payments it is registered as standard foreign trade, which increases the volume of merchandise trade (IMF, 2017). Thus, resident FGPs upwardly bias the import statistics of the domestic economy, because they account for transactions as classical foreign trade in which no change of ownership took place (Doherty, 2015).

In bilateral trade, the transaction could be registered as international tolling agreement (in this case, the price of the agreement is the import). However, in complex production networks, this is much more difficult, because the subcontractors can also receive inputs that are owned by the FGP. The statistics oversee the last chain only, and the price of the inputs provided by the FGP to the subcontractor are usually much lower as compared to the price of goods that are re-imported by the FGP.

Final goods remaining in the subcontractor's country require special attention. It is hard to estimate their volume because the largest subcontractor in tolling agreements is also one of the largest consumers in the world. Thus, the volume remaining in China is assumed to be considerable. If a share of final goods produced in tolling agreement remains in the subcontractor's country and sold to resident households, then the FGP did not conduct re-export but rather an export (which must be added to the export in the balance of payments of the FGP's country). At the same time, the total value of the production cannot be added to the subcontractor's national accounts because the GDP is increased by the value of the agreement only.

The accounting of international financial transactions relies on the implicit assumption that the prices reported to the statistical offices reflect the normal market conditions. That is, the negotiated prices are results of bargains under competitive conditions. This does not always hold in the supply chains, in particular if the two partners are a subsidiary and a parent company, or if the production took place under a tolling agreement. It is common that owing to tax optimisation reasons, companies apply transfer prices within the GVCs.

This phenomenon is not a concept and there are adjustment methods integrated into system of statistical data collection by which the authorities can tackle this problem[34].

Despite all efforts, transfer pricing is still an issue in statistical accounting. Imputing the true market price is not always possible in international transactions, and it does not adjust the biases coming from the income distribution strategy of the TNC groups (however, it is not aimed by the adjustment). One faces the same dilemmas in transfer pricing, as in the accounting of capital movements, because it is not certain that the income is accounted for in the country where it was formed. In extreme cases, the bias can be severe, like it happened in Ireland in 2015, when the annual GDP volume grew by 26.3% because of the transfer prices and income distributions (OECD, 2016).

The specific internal pricing of transnational firms causes the revaluation of the income generated in the resident economy and it biases the foreign trade price indices. In an economy free of transfer pricing, if a firm switches its supplier or offshores a part of its production abroad, the producer price index and the import price index are also impacted. This should not cause any problems during the estimation of national accounts because the producer price index has no role in the GDP. However, any changes in the import price index directly impact the level of GDP; however, it does not bias if there is a real economic activity behind the alteration.

At the same time, the bias could be severe if the trade within the company group has no impact on the real economy, or the prices do not reflect the real values. The main channel of income transfers in the supplier networks of GVCs is the trade of intellectual property products (IPPs). Adjusting the price of IPP transfers is challenging because usually there are no other alternative products priced under normal market environment and could be utilised as samples for imputing. If the parent company overprices the IPPs for its suppliers or subsidiaries, the import price index in recipient countries will be upward-biased. Consequently, the estimation of the performance of real economy will also be upward-biased. Similarly, if the subsidiary exports overpriced goods within the company group, the GDP volume of the resident economy will be downward-biased (Dridi & Zieschang, 2004; Mead, 2014; Nakamura et al., 2015).

[34] This process is known as the arm's length method (OECD, 2017).

However, its volume is altering economy by economy, and it is worth to summarising the case of work abroad and remittances. In 2009, almost 88 million jobs could be linked to GVCs (Jiang, 2013). The economy relies on the value-added produced by domestic and foreign workforce depending on its actual position in GVCs. (Jiang, 2013) approached this question from employer's side and concluded that the export of the countries on the top of the value chain requires more jobs abroad than at home. Only a few countries could achieve that international trade required more domestic workplaces: China, India, Indonesia, and Brazil. For Hungary, the international trade (both export and import) impacted 2.5 million jobs, while the foreign demand of the country required 1.1 million workplaces.

GVCs have a role in remittances only if they employ a non-resident labour force. Resident workers who regularly transfer remittances are seemingly out of the purview. The international movement of workers and their remittances are covered by the new balance of payment statistics of the IMF (BP6 – Balance of Payments and International Investment Position Manual, sixth edition) and the problem of resident companies also turn up here. If the non-resident employee is the resident of parent company's country (secondments), the statistics register the flow of income between the two countries, although the employee works for the parent company. At the same time, the volume of these amounts is assumed to be small as compared to other labour incomes. It is still an issue (also in case of capital transfers) if the income transfers of subsidiaries (100% per cent owned by foreign affiliates) to the parent company can be considered as true capital outflows.

It can become even more complicated if the non-resident employee is the citizen of a third country. If the company transferring the wages is a 100% domestic firm, then the (secondary) income flow reflects a true transaction. However, if the transfer is made by a foreign subsidiary, the bilateral balances are necessarily revealing the true values, in particular if the subsidiary is an offshore company founded on tax optimisation reasons[35]. The role of temporary work agencies is peculiar in remittances. Outsourcing non-resident employees via agencies is considered a trade in services from the point of view of the recipient firm. In other words, the payments for the agency are not income transfers, and thus, resident transfers are downward-biased. Meanwhile, the agency compensates the outsourced workers. If the employee is resident, then it is a 'simple' payment of wages, but if the employee is non-resident, the compensation is considered to be (secondary)

[35] In Hungary, according to the National Tax and Customs Authority, the share of total wage cost finances by 100% foreign owned companies was 31% in 2017.

income transfers. If the agency is in a third country, it generates income outflow of its country, while there is no real economic activity behind the transfer (just the service fee). This is very similar to the triangular transactions already mentioned.

Challenges posed by globalisation are immensely burdensome for national and international statistical offices. All of these are aggravated because the changes happened extremely fast in the world. Not to mention that a considerable fraction of international transactions is expressly made to avoid the control of the authorities, but the actions are still lawful (like tax optimisation). The real problem is insufficient knowledge concerning these transactions between partner countries and companies; it is also difficult to be estimated on the basis of the available data.

International statistical authorities are struggling not just with the data collection and classification of companies and transactions but also with the new forms of trade. Besides the earlier mentioned triangular trade, two other forms of trade must be highlighted (UNECE, 2011):

Quasi-trade or quasi-transit trade: it is a (relatively) new trade form. In this case, the import or export of the good is registered in an intermediary country usually where the port is located. EU import from a third country should be registered in that member state where it entered the community regardless of its final destination. The goods entered are psychically there, but the legal owner is not a resident of the intermediary country. A typical example of this is the Rotterdam effect that increases the foreign trade of the Netherlands to a notably high level, although much of the imported goods leave the country. This impacts the level of international trade in goods but not the GDP (as it is balanced). At the same time, export/GDP or gross exports are seriously biased. These two indicators are frequently used to describe the competitiveness of any nation. According to estimations, trade through the port of Rotterdam between the EU and the UK increases the volume of the Dutch re-export by 10 billion euros (Lemmers & Wong, 2019).

Internet trade: this form of trade resembles the triangular trade because it often happens that the intermediary is located in a third country. The difference occurs when the financial transactions are made with a fourth country (because of tax optimisation). In this case, the true seller and the intermediary are hidden, while the psychical transfer of goods is conducted between two countries without direct financial compensation. To make the situation more complicated, there is a third country in the transaction – the country of the financial intermediary (which receives the amount of payment in exchange for the goods), which is not equivalent to the country of the intermediary seller. The following example demonstrates the process: A consumer purchases a good on the Internet via an intermediary. The good is not owned by the intermediary. The financial transaction is made between the consumer and a financial intermediary (owned by the intermediary) located in a fourth country because of tax optimisation.

3.3 A fictive example of accounting a global group of companies

The following example illustrates an extremely complicated situation for international statistical data collection, which may often occur in the real economy (Wall & van der Knaap, 2011). A US parent company found a European centre in Germany to control the production on the continent, financial management, and distribute the good in the European market. The European hub creates two factories in Central and Eastern Europe in Hungary and Slovakia. See the process in the following Table 10:

10. Table: The statistical aspects of a hypothetic chain evolution

Event	Participating countries	Statistics affected	Note
The parent company invests into the European hub	USA, Germany	US, German FDI and IIP[36]; German investments, GDP	Green or brown field FDI, the connection is clear
The hub found a factory in the CEE region	Germany, Hungary, Slovakia	German, Hungarian, Slovakian FDI and IIP; Hungarian, Slovakian investments. GDP	The ultimate investor in the U.S., if the relationship can be decrypted, then there is no problem
The centre leases the machines to the factories below market price	Germany, Hungary, Slovakia	German-Hungarian, German-Slovak trade in services; foreign trade price index, IIP	The machines are owned by the US parent company.
The two factories also trade among themselves through transfer prices	Hungary, Slovakia	Hungarian-Slovak trade in goods; foreign trade price index, GDP	It is not clear to what extent a product coming out of U.S./German-owned factories counts as a Hungarian or Slovak product
The factories deliver the products at transfer price to a Logistics Centre in Slovakia, from where they are delivered to consumers	Hungary, Slovakia, Austria	Trade services of Goods in Hungary - Slovakia, and Slovakia - Austria; foreign trade price index, GDP	The Hungarian side is not related to Austrian consumers, there is no real economic performance behind a part of Slovak foreign trade
The centre also commissions a Swiss online intermediation company to sell and pays it a commission	Switzerland, Slovakia, Germany, Austria	trade in Slovak-Austrian goods; Switzerland-Slovak trade in services; Switzerland-German trade in services; Switzerland-Austrian trade in services	The Austrian consumer concludes a contract with the Swiss company, the international transfer takes place here, and then the goods are shipped from Slovakia, which also has a Hungarian part. The Swiss intermediary accounts for the Slovak factory and the German headquarters, transactions are entered in the statistics several times.
The Hungarian factory performs wage work in China, some of the manufactured products go to Chinese and Japanese consumers, some of them are returned to Hungary	Hungary, China, Japan	Hungarian-Chinese trade in goods, Chinese-Japanese trade in goods, Chinese industry, GDP, GNI	If the Chinese party also has a local supplier, the conditions of the wage work are not fulfilled, the Chinese factory carries out industrial production on Hungarian orders. The product is of Chinese origin in Japan, local in China, Chinese imports in Hungary.
The Hungarian factory is contracted with a Croatian labour agency, who also rents out Serbian and Croatian workers for it	Hungary, Croatia, Serbia	Hungarian-Croatian trade in services, Serbian-Croatian secondary income transfer, remittances, GDP, GDI	The income transfers of Serbian workers are considered Croatian for work in Hungary.

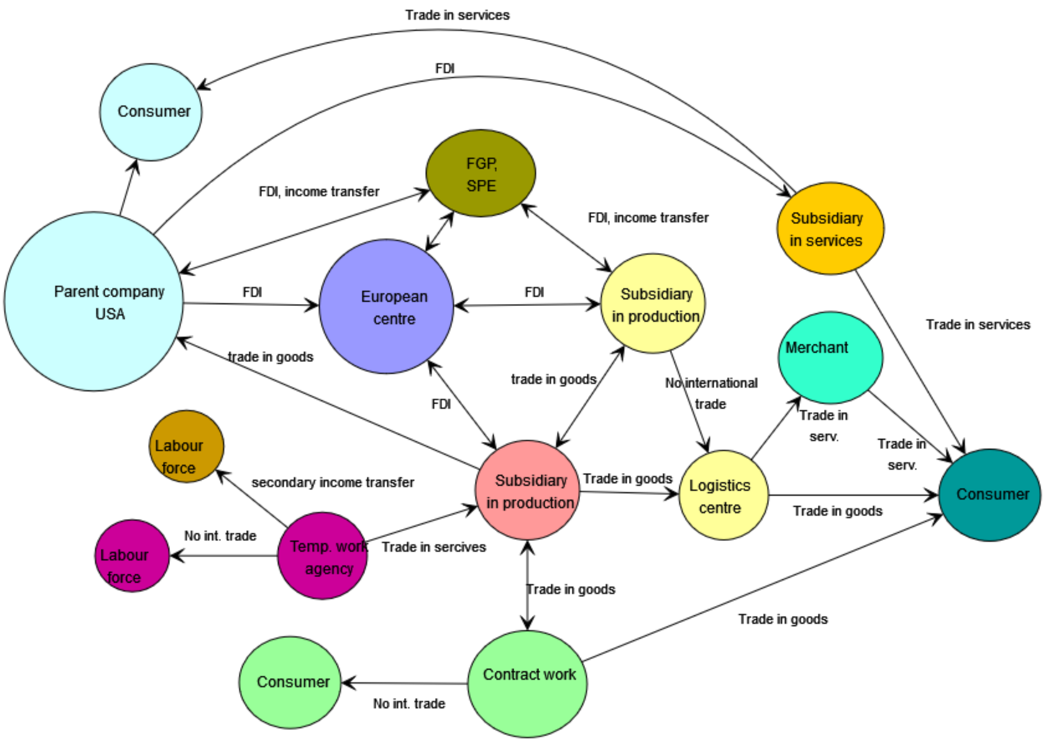
[36] International investment position

Event	Participating countries	Statistics affected	Note
U.S. parent company sets up service support company in India to handle consumer-related financial transactions	Austria, India, USA	Austro-Indian trade in services, U.S.-India primary income transfer, GNI	In reality, the Austrian consumer pays to the Indian service centre of an American company.
The German centre establishes a clearing centre in Luxembourg and through it accounts for regional turnover and profits.	Germany, Hungary, Slovakia, Luxembourg	Primary income transfers and trade in services in the countries concerned, FDI, IIP, GDP, GNI	The Luxembourg company does not produce real economy, although it can be filtered as SCV in FDI, but real traffic will not be traceable
Part of the profits will be transferred back to the USA by the Luxembourg company	Luxembourg, U.S.	Luxembourgish-American FDI and IIP, service trade, GNI	There is no real economic performance behind the Luxembourg capital transfer and is American-owned.
Hungarian factory supplies to U.S. parent company at transfer price to consumers	Hungary, U.S.	Hungarian-American trade in goods; foreign trade price index, GDP, GNI	The product is considered to be Hungarian-made in the USA, which is the product of a Hungarian company owned by the United States and founded by a German headquarters.

Source: own edition

Figure 7 depicts the process described in Table 10:

7. Figure: Flowchart of the statistical aspects of a hypothetical chain organisation for the main accounting items



Source: own edition
Note: Colours represent different countries

The process above illustrates well that not only the international statistical data collection is affected, but it can also cause bias in the system of national accounts. In particular, the balance of payments, current accounts, and primary incomes are impacted. The GDP can also be biased to a certain extent through the inaccurate foreign trade price indices. The ambiguous definition of the resident companies the volume financial transaction accounted to the capital and brand owners could be much lower, while they could be higher in the subsidiaries' countries.

The most crucial challenge for statistics is the accounting of foreign trade in gross terms, which is aggravated by the new forms of trade. Gross exports conceal the true contribution of the participant countries to the value of the goods. This upwardly biases the export and import of trade partners. However, balances with the rest of the world do not get altered, but the bilateral trade balances can be severely biased. This has an extensive literature owing to the publication of TiVA statistics (Banga, 2013; Gereffi & Fernandez-Stark, 2011; Koopman et al., 2014; Nádudvari, 2013). Earlier, the publication of international IO data was hindered by data scarcity, because the demand was to cover the widest possible range of countries. The TiVA database filtered out the bias caused by gross exports and revealed a new, more accurate picture of international trade. However, today the estimation of TiVA data is less challenging in terms of methodology, and the statistical offices still cling to account in gross terms. Thus, the following crucial issues with regard to GVCs are still unknown:

- What is the origin of intermediary inputs used for production? Though the statistics of international trade apply several nomenclatures to register the goods and services, the statistics of industrial production are still not linked to foreign trade.
- What happens to the inputs that are transformed by the firm? It is still unknown what transformation is done on raw materials and parts. Only the code of the main activity suggests what sort of operations are taking place in the background.

The TiVA statistics are a fundamental data source of GVC analyses; the picture of value chains is not complete without the aforementioned issues. Statistical data collection needs to be adjusted, which cannot happen in a way that puts more burden on the shoulders of firms. It must be noted that modern technology provides a solution to follow the lifecycle of any good. The blockchain technology ensures that the psychical transactions can be monitored without the risk of fraud and all information can be shared with no burden on statistical agencies (Tröster, 2020). The operation of the system is cost-effective; however, it is uncertain if the data could be used for research purposes because of the encryption (Goldstein & Newell, 2020). A possible extension of data collection should gather the following records to track GVCs:

- Partner's NACE code in trade.
- The use of product code in invoicing (today, only the name must be included in the invoice).
- Production statistics, in particular how the intermediate input is used, at least at the level of investment, intermediate or final good (the blockchain technology can help here).
- Linking the tax refund with production and trade data.

3.4 Summary and conclusions

Though the globalisation has been ongoing for many decades, the official statistical authorities could not cope with the developments. Owing to the internationalisation of the companies, the owner structures, the network of trade, and capital transactions became more sophisticated, in which the corporate partnerships turned more complex. Thus, one cannot rely on official statistics only – the field can be analysed through case studies and international IO tables. This chapter summarised the bias that can occur in official statistics owing to the altered global economic landscape, and these changes reform GVCs.

Undoubtedly, the biggest challenge for statistical data collection is the lack of transparency on the ownership structure. The resident firm approach, which assigns all transactions to the countries in which the company is registered creates a false illusion of economic power and corporate competitiveness. It is unclear how large is the actual contribution of value-added produced by 100% foreign-owned firms to the GDP? In such cases, the local units contributed to the national income only with the labour costs, and everything else was ensured by the foreign parent company through its FDIs; thus the actual owner of the value-added is unclear. Approaching the question via the GNI improves the estimations to a certain extent; however, the volume of bias can still be large.

It was proven in this chapter that owing to the ambiguous ownership structure of the resident companies, severe bias can occur in the statistical accounting of trade and financial transactions, which can also have an impact on the level GDP. Through a simplistic fictive, however very realistic example, it was illustrated which statistical indicators are affected by GVCs and what do the statistical agencies actually see. If the ownership relation entails just a single step (i.e. the parent company and the subsidiary can be clearly identified), then there is no bias in the system. However, if the ownership structure has multistage relations, the assignment of transactions to the actual owner is not ensured.

New forms of trade within the value chains pose further challenges to the statistical offices. The trace of these transactions by the currently applied data collection methods is almost impossible. The international IO tables, by the fact that they consider the global economy as a closed system, seemingly provide an alternative solution; however, these databases are also based on official statistics and thus conserve these biases that were a priori included in the statistics. Despite these concerns, there is no better tool for the analysis of GVCs from a macroeconomic perspective. In parallel, case studies provide the base for the investigations from microeconomic perspective. The mezo-level is missing, though, and the base of this could be established by the statistical offices.

This chapter presented the current theoretical framework, accentuating its flaws. The posed hypotheses could only be validated through a fictive example, and thus the amount of bias caused by the TNCs is still unknown. However, studies analysing the capital transactions of multinational companies registered in the USA (including the subsidiaries owned by foreign parent companies) showed that level of the US GDP is likely to be increased by 1.5%, while the volume of inward capital is most probably lowered by 33.5% (as these are the transfers of US companies from abroad) (Bruner et al., 2018). This leads to the question that if globalisation caused such a high bias in a net FDI investor country, then what could be the effect on net FDI recipient countries such as Hungary? To examine this, a global database of companies and their financial transactions would be required, which is unfortunately not available at the time of writing this dissertation.

Consequently, the official statistics lagged in tracing globalisation, which severely limited the scope and possibilities of GVC research. In the forthcoming chapters, several methodologies will be introduced that can be applied to map the actual position of companies and countries in GVCs more accurately.

4. Measuring the value-added produced in the global value chains[37]

It was the international IO tables published in the early 2010s that opened the way to the measure of GVCs. This was also induced by the fact that international trade statistics, which was accounted merchandise trade in gross terms, showed a severely biased picture concerning the relative position of the countries in the GVC. The emergence of multinational IO tables did not rise without trace, as the classical version of IO tables, which measured the domestic transactions, flow of value-added, and national cooperation, was available well before the 2000s. On a pilot basis, interregional domestic tables were also published. However, these tables were far away from global coverage.

4.1 Theoretical background of IO tables

The availability of international IO tables is a prerequisite for producing trade in value-added measures. The framework of IO tables was developed by Wassily Leontief, an American economist of Russian origin, in the 1930s, for which won the Nobel prize. In its simplest form, the IO model is a system of linear equations describing the economy. These models cover at least one region, and it accounts for local production. The tables record the bilateral flow of goods and services between the domestic industries and the consumers. In the following, the theoretical background of IO tables will be summarised in accordance with the book by Miller and Blair (Miller & Blair, 2009). The standard IO table is depicted in Table 11:

[37] This chapter relies on the work of Peter Vakhal (Vakhal, 2016a, 2016b).

11. Table: Input-Output transaction table

<table>
<tr><td></td><td></td><td colspan="8">Producers as consumers</td><td colspan="4">Final demand</td></tr>
<tr><td></td><td></td><td>Agriculture</td><td>Mining</td><td>Construction</td><td>Manufacturing</td><td>Retail trade</td><td>Transport, storage</td><td>Services</td><td>Egyéb</td><td>Private consump-tion</td><td>Private investment</td><td>Community con-sumption</td><td>Net exports</td></tr>
<tr><td rowspan="8">Producers</td><td>Agriculture</td><td colspan="8" rowspan="8"></td><td colspan="4" rowspan="8"></td></tr>
<tr><td>Mining</td></tr>
<tr><td>Construction</td></tr>
<tr><td>Manufacturing</td></tr>
<tr><td>Retail trade</td></tr>
<tr><td>Transport, storage</td></tr>
<tr><td>Services</td></tr>
<tr><td>Other</td></tr>
<tr><td rowspan="3">Value added</td><td>Employees</td><td colspan="8">Compensation for employees</td><td colspan="4" rowspan="3">GDP</td></tr>
<tr><td>Equity holders</td><td colspan="8">Profit-type income and capital consumption al-lowances</td></tr>
<tr><td>Government</td><td colspan="8">Indirect business taxes</td></tr>
</table>

Source: (Miller & Blair, 2009)

The interpretation of the table is the following:

Rows (producers): the distribution of output of the given industry in the economy.

Columns (producers): the inputs required for productions from the same or other industries.

Grey cells: inter-industrial trade

Columns (final demand): the part of output produced for final use. The consumers make no further transformation on the goods and use them in their current firm (i.e. the goods were not purchased as intermediate consumption).

Rows (value-added): non-industrial inputs required for the production.

Let z_{ij} the transaction between industries *i* and *j* and suppose that the economy consists of *n* industries. x_i indicates the total output of industry *i*, while f_i is the final demand for the goods of industry *i*. Then, the total output industry *i* is the following:

$$x_i = \sum_{j=1}^{n} z_{ij} + f_i \tag{8}$$

The total output (including $i = j$) can be depicted with matrix notations:

$$\mathbf{x} = \mathbf{Zi} + \mathbf{f} \tag{9}$$

The final demand of industry i is given by $f_i = c_i + i_i + g_i + e_i$, where c_i is the household consumption, i_i is the private gross capital formation, g_i is the public gross capital formation, and e_i is the net export.

In the value-added row, let l_i be the wages, n_i represent the non-wage expenses including the profit. m_i denotes the import in industry i. On the basis of these notations, the two-sector IO matrix can be represented as the following:

12. Table: Transaction table for two sectors

		Processing sectors		**Final demand**				**Total output**
		I	**II**					
Processing sectors	**I**	z_{11}	z_{12}	c_1	i_1	g_1	e_1	x_1
	II	z_{21}	z_{22}	c_2	i_2	g_2	e_2	x_2
Payments	**Value added**	l_1	l_2	l_C	l_I	l_G	l_E	L
		n_1	n_2	n_C	n_I	n_G	n_E	N
Import		m_1	m_2	m_C	m_I	m_G	m_E	M
Total outlays		x_1	x_2	C	I	G	E	X

Source: (Miller & Blair, 2009)

In the next section, the technological coefficient shall be introduced, which is defined by $a_{ij} = z_{ij}/x_j$. The indicator shows the share of inputs produced by industry i in the output of industry j. These coefficients are fixed in the base model, that is, it fails to include the aspects of economic of scale, or in other words, the return to scale is constant. With the help of the technological coefficient, the objective function of the production can be determined:

$$\boldsymbol{x_i} = \min_{\boldsymbol{i=1,..n}} \left(\frac{z_{nj}}{a_{nj}}\right) \tag{10}$$

The objective function can assume multifarious forms. In the IO methodological framework, in particular, if the analysis covers a short period, the utilisation of the Leontief production function is appropriate, because using that, one can evaluate the short-term risks in the system.

By taking advantage of the technological coefficient's definition, one can derive the function of total output and the final demand:

$$x_n = a_{n1}x_1 + \cdots + a_{ni}x_i + \cdots + a_{nn}x_n + f_n \tag{11}$$

$$x_n - a_{n1}x_1 - \cdots - a_{ni}x_i - \cdots - a_{nn}x_n = f_n \tag{12}$$

$$-a_{n1}x_1 - \cdots - a_{ni}x_i - \cdots + (1 - a_{nn})x_n = f_n \tag{13}$$

The equations can be arranged into matrix form, wherein $\hat{\mathbf{x}}$ denotes the diagonal matrix:

$$\hat{\mathbf{x}} = \begin{bmatrix} x_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x_n \end{bmatrix} \tag{14}$$

Using $(\hat{\mathbf{x}})(\hat{\mathbf{x}})^{-\mathbf{1}} = \mathbf{I}$, we get $\hat{\mathbf{x}}^{-\mathbf{1}} = \begin{bmatrix} 1/x_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1/x_n \end{bmatrix}$. If the **Z** transaction matrix is multiplied by $(\hat{\mathbf{x}})^{-\mathbf{1}}$ from the right, we get the technological matrix denoted by **A**:

$$\mathbf{A} = \mathbf{Z}(\hat{\mathbf{x}})^{-\mathbf{1}} \tag{15}$$

The output function in matrix form is $\mathbf{x} = \mathbf{Ax} + \mathbf{f}$. From this, one is able to derive $(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{f}$, which is the final demand. It must be noted that the aforementioned equations can be solved if and only if $(\mathbf{I} - \mathbf{A})^{-1}$ exits, that is, $|\mathbf{I} - \mathbf{A}| \neq 0$. If the inverse matrix exits, the system of equations have a solution, which can be expressed as the following:

$$\mathbf{x} = (\mathbf{I} - \mathbf{A})^{-\mathbf{1}}\mathbf{f} = \mathbf{Lf} \tag{16}$$

where $(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{L} = [l_{ij}]$, or in other words, the Leontief inverse.

The model introduced above can only manage one region, and thus the whole national economy can be described by that. This also means that none of the analysed regions in the subject can be attached to that model, that is, the interregional transactions cannot be studied. The existence of domestic (within the region) and interregional (between regions) transaction tables is a prerequisite for interregional analyses.

Let us examine the case of two regions. Let us denote regions by *r* and *s*, while the industries are indicated by *i* and *j*. The transaction matrix is denoted by **Z** as usual:

13. Table: Intersectoral, interregional transaction table

		Buyer				
		Region *r*			**Region *s***	
Seller		**1**	**2**	**3**	**1**	**2**
Region *r*	**1**	z_{11}^{rr}	z_{12}^{rr}	z_{13}^{rr}	z_{11}^{rs}	z_{12}^{rs}
	2	z_{21}^{rr}	z_{22}^{rr}	z_{23}^{rr}	z_{21}^{rs}	z_{22}^{rs}
	3	z_{31}^{rr}	z_{32}^{rr}	z_{33}^{rr}	z_{31}^{rs}	z_{32}^{rs}
Region *s*	**1**	z_{11}^{sr}	z_{12}^{sr}	z_{13}^{sr}	z_{11}^{ss}	z_{12}^{ss}
	2	z_{21}^{sr}	z_{22}^{sr}	z_{23}^{sr}	z_{21}^{ss}	z_{22}^{ss}

Source: (Miller & Blair, 2009)

The transactions of $\mathbf{Z}^{rs}$ and $\mathbf{Z}^{sr}$ are in the focus of the research, which is the intermediary export for region *r* and at same time is the intermediary import for region *s* from region *r*. The output of the first region can be expressed as:

$$x_1^r = \underbrace{z_{11}^{rr} + z_{12}^{rr} + z_{13}^{rr}}_{\text{Transactions within the region}} + \underbrace{z_{11}^{rs} + z_{12}^{rs}}_{\text{Interregional transactions}} + \underbrace{f_1^r}_{\text{Production for final demand within the region}} + \underbrace{e_1^r}_{\text{Export of final goods}} \tag{17}$$

The technological coefficients can be calculated by the analogy of the one-region case. At the same time, the model can be extended to n regions if all the data and the storage capacity are available for the researcher.

4.2 International input-output tables

Several instates publish international IO tables; the compilation methodology relies on the harmonisation and unification of national IO tables. Although there are recommendations concerning the compilation of national IO tables (United Nations, 1999), just a few countries publish them on a regular basis. Owing to methodological reasons, these tables cannot be integrated (Inklaar et al., 2007). Thus, joining the international tables cannot be done directly, just by estimations. In case of such estimations, it is assumed that all data are available for all units, by which the sides (sum of the rows and columns) can be calculated. These values make it possible that the model can be closed. Values within the tables (if not available) can be estimated using optimisation algorithms (see Chapter 8). Owing to the lack of a universal database that could provide official data, there are numerous MRIO in the market. Gáspár and Koppány (2020) furnished an overview about these tables, which is summarised in the following table (Gáspár & Koppány, 2020):

14. Table: List of available international IO databases

Database	Institution	Number of countries	Regional coverage	Number of sectors	Availability
WIOD, WIOT	EU	43	OECD + developed countries	56	Annual
ICIO	OECD	64	OECD + large-scale global economies	36	Annual
EXIOBASE	EU (financing)	44	OECD	163	Temporal
Eora	University of Sydney	189	global	26	Annual
GTAP	Purdue University	121	global	65	Temporal
FIGARO	EU	29	EU + USA	64	Annual
ADB MRIO	Asian Development Bank	62	Asia	35	Temporal
AIIOT	IDE-JETRO (Japan)	17	Asia	76	Temporal
South American IO tables	ECLAC-IPEA	10	South America	40	Temporal
Project Réunion	Global MRIO Lab	220	global	6357	Annual
UIBE-GVC	UIBE	44	OECD + large-scale global economies	56	Annual

Source: (Gáspár & Koppány, 2020), own collection

The aforementioned databases describe the same phenomenon, albeit estimations based on these datasets can still yield considerably different results. The differences mainly arise from the discrepancies in the Leontief inverse, which can give rise to a severe bias in the final demand (Owen et al., 2014). This was also confirmed by other studies. It was also reinforced that there exist significant biases in the classical GVC indicators derived from the Eora and ICIO database (Arto et al., 2014; Czakó & Vakhal, 2020). Generally, estimations based on the former are smaller as compared to those based on the latter. The values showed a 15% relative variation when the EXIOBASE, the Eora, and the ICIO was compared (Giljum et al., 2019; Hambÿe et al., 2018). Altogether, 60% of that variation can be linked to forward suppliers, which suggests strong dependencies in some blocks of the transaction matrices.

Here it is worth to mention in a few words the database developed by UIBE, which, unlike the previous databases, has developed GVC indicators with a new methodology. The essence of this is that in practice it is extremely difficult, but also impossible, to estimate, on the basis of trade and industrial statistics alone, how much of the production will be exported and how much of it will be used domestically. Therefore, (Wang et al., 2017b) has developed a method that focuses on the source of value added, the factors used for production (land, labour and capital), as they are much easier to measure and trace from the national accounts of the economy it themselves. National IO tables were also used for indicators and WIOT databases for international flows in particular. The data thus generated, as they are based on national accounts, can be considered more credible, but still contain estimates.

As the datasets are estimated by complex optimising algorithms, it is not surprising that they contain some bias. Only the Eora database provides values for variance that describe the quality of the estimations. As other MRIO tables do not contain that, their reliability is also uncertain. As a consequence, this dissertation utilises the Eora MRIO table, which was created by University of Sydney (Lenzen et al., 2013). Of course, there are also disadvantages of using Eora; however, these could hold true for any other dataset as well:

- Values are given in 1,000 dollars at current prices. No constant price version is available. It must be noted that the WIOD database contains values in constant prices; however, the geographical coverage of that is narrower as compared to Eora tables. The bias caused by changing prices level could be avoided if the values are deflated by the GDP.
- Optimisation algorithms are operating with errors, which embody the volume of gross output that cannot be distributed among the countries. Thus, despite the fact that Eora contains almost all countries of the world, it still has a 'Rest of the

World' (RoW) category, which is used as 'supplementary' country (without industries). That RoW row sums up all the error terms. This 'country' has transactions with all other countries and thus can be used as a proxy variable in the models. However, it is not utilised in this dissertation.

4.3 The methodology of trade in value-added statistics

This subchapter provides insight into the GVC-related indicators. It must be noted that the right interpretation of the following measures is essential to evaluate the role of a country in GVCs. One must emphasise that the compliance of MRIO tables relies on such assumptions that are almost impossible to confirm in any form (Sturgeon, 2015):

- The technological coefficients of the firms are constant along size, field of operation, and location. This contradicts with the trade theory of Krugman (Helpman & Krugman, 1985).
- The models do not differentiate the products by their characteristics of use (final or intermediary good) and their main market (export or domestic use).

As a consequence, the data in the transaction matrices are generally inconsistent with the bilateral trade data.

One well-known anomaly of trade in value-added statistics was revealed by Robert Johnson and Guillermo Noguera (Johnson & Noguera, 2012), who also pioneered the framework of TiVA statistics. In their study, they proved that the value-added-based trade deficit between the USA and China is 30% less as compared to the standard trade balance of gross terms.

Introduction to the indicators

Basic indicators

Direct intermediary import: intermediary inputs purchased by the producer from abroad.

Indirect intermediary import: intermediary inputs purchased by the producers from a local supplier, but the origin of the product is abroad. One must note here that this category is not equivalent to re-import.

Import content of export: direct + indirect import. It takes into account the total import demand of the resident firms, by which the intermediary suppliers are filtered out. Note that if the output is exported, then the indirect importer shall become an indirect exporter, because its product, which was sold to a domestic company, will leave the country.

Utilising the methodology of IO models, the import content of the export can be expressed as the following:

$$\text{import content of export} = m * (I - A)^{-1} * e \tag{18}$$

where
m: the import content of total output in the industry;
e: the total export in the industry;
A: the matrix of technological coefficients.

These values are required for estimating the bilateral trade in value-added between two countries. The value-added can thus be calculated as the following:

$$VA_k = \sum_i v_i^k * l_{(kn+i)(kn+j)} \tag{19}$$

where
v_i^k: value-added produced in industry *i* and country *k*;
l: the elements of the Leontief inverse;
n: the number of industries.

After that, the export can be disaggregated as the following:

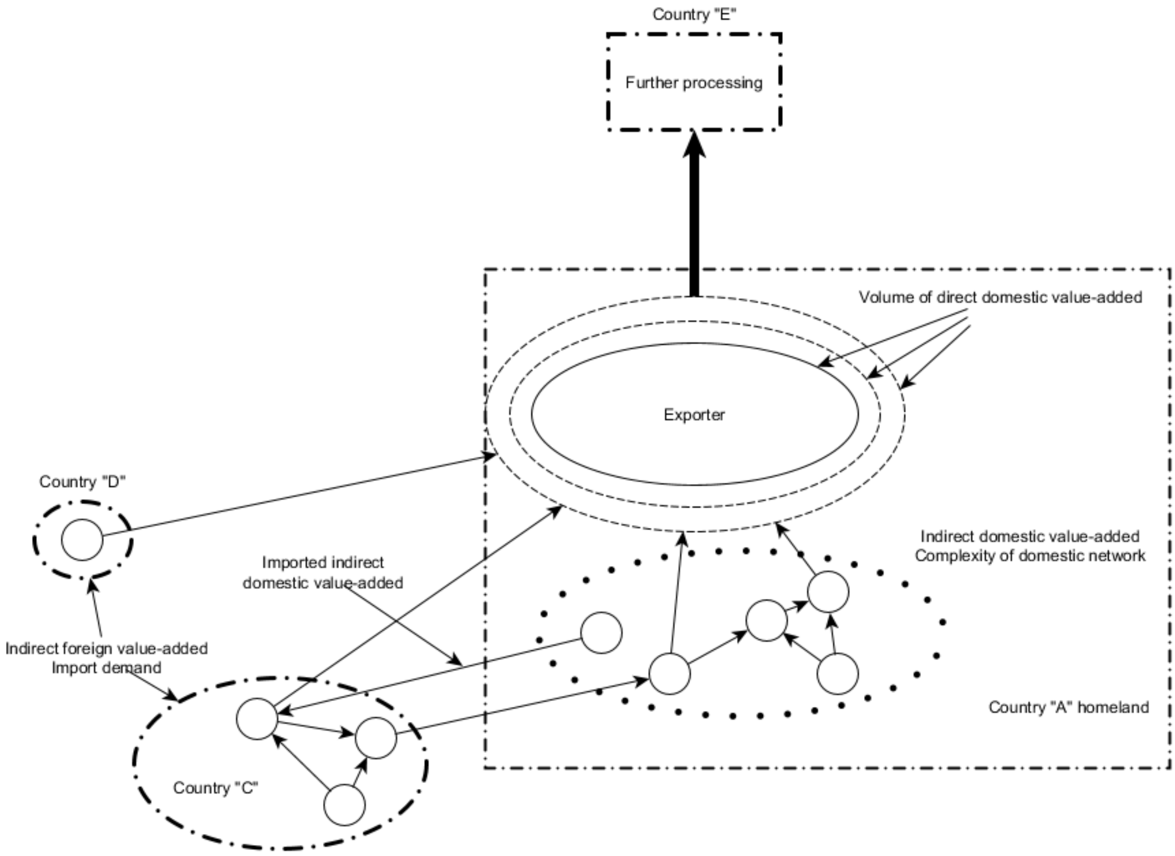
- **Direct domestic value-added**: value-added created by the industry itself. The source of that can be business operation that increases the value of the given product or self-produced investments (like building a storage). This indicator mainly refers to the importance of local production. The higher the value, the more important the role of the domestic producers in the value chain.
- **Indirect domestic value-added**: value-added used by the industry but created by another domestic sector. It represents the depth of the inter-industrial relations. The more the number of industries participating in the production, even directly

or indirectly, the more complex is the creation of goods. Note that the suppliers who are indirectly involved in the production are not direct but indirect exporters. This a key issue for the concerned companies, as they are also going through internationalisation. Therefore, the volume of indirect domestic value-added refers to the diversity of the domestic supplier network.

- **Indirect imported value-added**: value-added by a foreign industry. This indicator represents the dependency of the industry on foreign industries. By this, one can infer the position of the country in the value chain. High indirect imported value-added in the export refers to the fact that industry is very close to the final stage of production chain that produces the final goods.
- **Imported indirect domestic value-added**: value-added that is imported from abroad but also contains domestic value-added. It also shows the rate of feedback, because all domestic value-added that is re-imported will have multiplier effect.

The categories mentioned above are summarised in Figure 8:

8. Figure: Value-added channels in the value chain



In the following section, the most common TiVA indicators will be discussed. Most of the indices measure the integration into the GVCs in different aspects. It often happens that both scholars and politicians depict the indicators differently, particularly in terms of economic competitiveness; however, this way of apprehension is often ambiguous. It is without doubt that every success in the export markets is also a form of success in international competitiveness if the value-added content is high. Nevertheless, the evaluation of this is not straightforward, as the share of value-added in gross export depends on several other factors that can indicate an industry with high value-added ratio as uncompetitive or vice versa. The position in the GVCs and the achieved success are only a part of national competitiveness. Other factors, such as institutional background, human development, and healthcare are also essential (Palócz & Vakhal, 2018a). These determinants are not discussed in this dissertation, and only their raw interpretation of competitiveness shall be covered herein.

Share of value-added in gross output

Dimension: countries and industries

Price: free on board (fob)

Definition: defines the share of value-added produced by industry *i* in the total output of industry *i* in country *c*.

$$PROD\_VASH_{c,i} = \frac{VALU_{c,i}}{PROD_{c,i}} \tag{20}$$

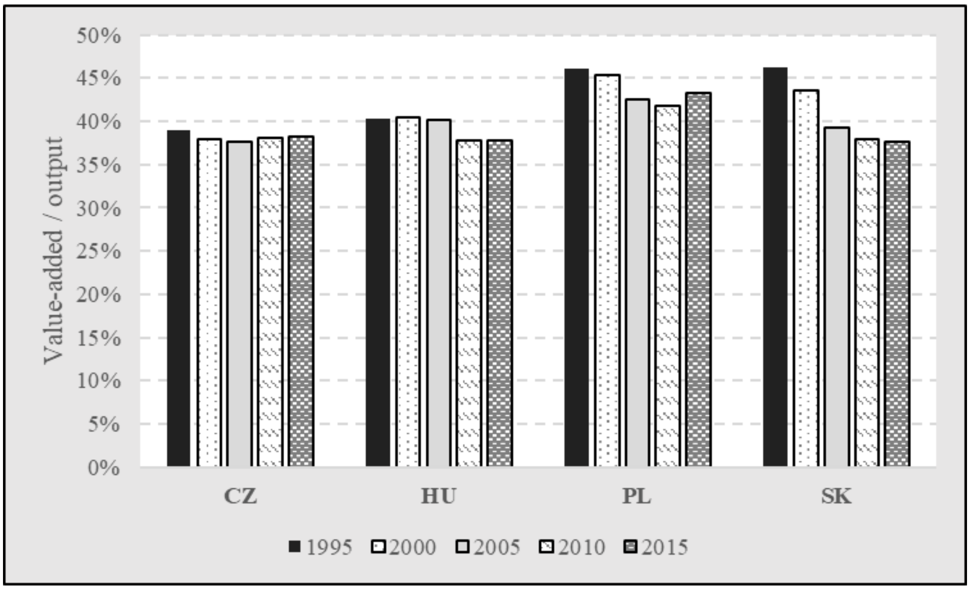
where

$VALU_{c,i}$: value-added produced by industry *i* in country *c*;

$PROD_{c,i}$: total output produced by industry *i* in country *c*.

This indicator signifies the importance of the industry in the national economy, or in other words, how much the industry relies on domestic resources, thereby increasing the GDP. Generally speaking, the larger value is more favourable; however, one shall not regard the higher value-added content as an indicator of competitiveness. The complexity of the goods produced by the industry in subject strongly affects the value of the indicator. If the products are relatively simple and require less inputs (like services), the share of value-added in the output can be higher. The more complex is the product, the more inputs are needed, which are usually purchased from other companies (possibly from abroad). Therefore, paradoxically, the lower share of value-added in the output indicates higher integration in GVCs. From lower PROD_VASH value, one should not infer lower competitiveness.

9. Figure: Change in value-added ratio in the output in the Visegrad countries between 1995 and 2015



Source: own calculation based on Eora database

Figure 9 depicts that Slovakia is the most dependent on foreign inputs, while Poland is the least dependent. In case of the latter, one should take the size of the domestic market into consideration. Supplying the Polish market with (Polish) products requires wider and deeper domestic production base, while the goods produced in Slovakia are not principally consumed by Slovakian firms and households. Thus, Slovakia is most likely more integrated into GVCs than Hungary and it might be more competitive; however, the higher dependency on the domestic market further reduces the risks, which are also factors of international competitiveness.

Share of domestic value-added in gross export

Dimension: countries and industries

Price: fob

Definition: this indicator depicts the value-added created by industry *i* in country *c* during the production of goods exported to country *p*. It also includes the value-added created by industry *j* ($i \neq j$) in country *c* in the case of the same product.

$$EXGR\_DVA_{c,p,i} = V_C B_{c.c} EXGR_{c,p,i} \qquad (21)$$

where

$V_c = \frac{VALU_c}{PROD_c}$: the share of value-added in output in country *c*;

$B_{c,c}$: the Leontief inverse in country *c*;

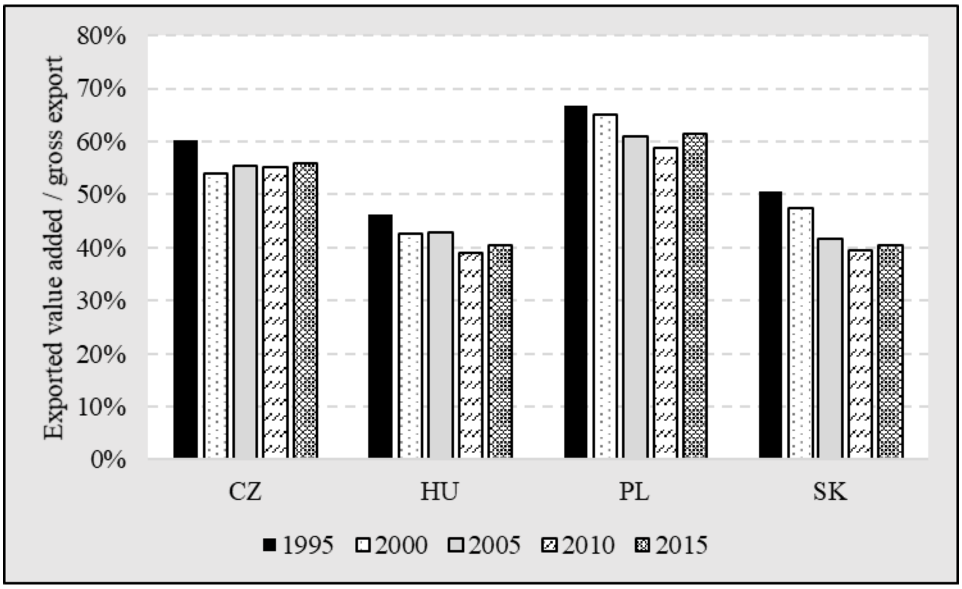
$EXGR_{c,p,i}$: the gross export of industry *i* in country *c* to country *p*.

This value is usually abbreviated as DVA and is one of the most preeminent indicators of trade in value-added statistics. It is usually utilised for illustrating the position of a country or industry in the GVCs: one can differentiate forward and backward positions depending on how much the industry or the country is dependent on the inputs of foreign firms. The participation of such countries that are rich in raw materials is usually forward because their outputs exported are used as inputs in other countries. In contrast, countries that are devoid of raw material must sting on imported inputs, thus their participation is rather backward. Consequently, the value of DVA represents the embeddedness in the GVCs. However, the dependency on the industry and product structure is still strong. Although this measure can indicate the level of competitiveness on the export markets, one should be cautious when interpreting its value.

While the volume of value-added truly depends on productivity, the volume of export is independent of that. The latter mirrors the sequence the production process in which the industry is participating. Recalling the convex curve presented in Figure 5, which illustrated the association between sequence of production and the value-added created, it is already known that the closer the company is to the final consumer, the higher the volume of value-added it exports (except the first sequence, which includes planning). Meanwhile, the value of gross export is permanently increasing, which inflates the share of value-added as we get closer to the final consumer. Therefore, by filtering out the sequences, the DVA can be simplified to a sort of productivity indicator – which is also not an indicator of competitiveness (see Chapter 7).

The DVA indicates the reliance on the domestic production base and is thus rather a measure of revealed comparative advantage than an indicator of competitiveness. From the viewpoint of social welfare, the higher volume of value-added is favourable; however, it does not assume high value of DVA in gross export.

10. Figure: Exported Domestic Value Added (DVA) as a proportion of gross exports in the Visegrad countries between 1995 and 2004



Source: own calculation based on Eora database

To interpret Figure 10, it must be assumed that the product structure of Visegrad countries is almost the same. One explanation of the figure is that the importance of the regional countries in the GVC is higher than in Hungary. The decreasing DVA values refer to the widening value chains, which indicates that more and more producers are participating in the global division of labour. The production has been fragmentising since the 2000s; however, it moderated during the global financial crisis, and since then, it is continuously slowing, and thus the decrease of DVA/export value also halted. The decline in DVA share in gross export is a global process that has a strong impact on the manufacturing industry of emerging countries (Johnson & Noguera, 2017).

Indirect domestic value-added in gross exports

Dimension: countries

Price: base price

Definition: value-added indirectly produced in country *c* in the gross export of country *c*.

$$EXGR\_IDC_c = \widehat{V}_c \text{offdiag} B_c EXGR_c \tag{22}$$

where

$\widehat{V}_c = \frac{VALU_c}{PROD_c}$: the share of value-added in output produced in country *c* in matrix form, where the values are on the diagonal;

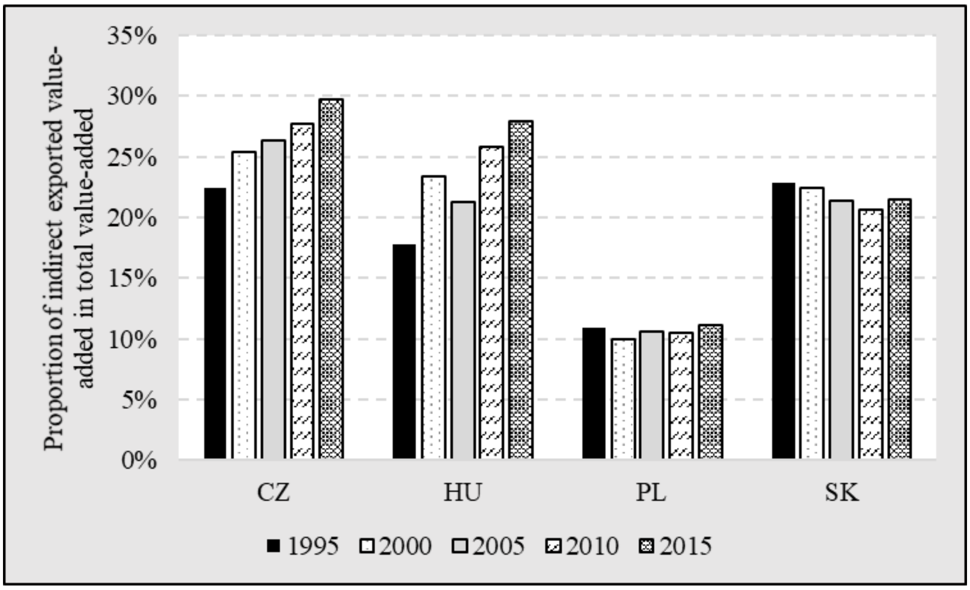
$B_{c,c}$: the Leontief inverse of country *c*;

$EXGR_c$: the gross export of goods and services in country *c*.

Interpretation: this indicator measures the value-added produced indirectly in the gross export. Instead of bilateral relations, this indicator takes the global effects into consideration. As it was mentioned earlier, this indicator depicts how far the domestic production base extends. Companies in the indirect value chain do not export directly and thus do not register any outbound trade in their income statement or balance sheet. They are hidden participants of the value chains; however, their importance is just as high as that of the direct exporters. This measure is also an approximate indicator of competitiveness, because its value can highly correlate with the number of enterprises in the economy.

Consequently, it is worth deflating with total output or total value-added. It must be noted that indexing with the exported value-added is less expedient, because in such cases, the correlation between the numerator and the denominator is high and the dependency is clearly linear.

11. Figure: Indirect exported value added in proportion to the total value added produced in the Visegrad countries between 1995 and 2015



Source: own calculation based on Eora database

Figure 11 depicts how slow integration into GVCs affected the domestic supplier network. There is a group of resident companies in Czechia and Hungary that could become indirect exporters. The case of Slovakia is interesting, because the diminishing share of indirect exports could suggest that these companies became direct exporters. However, the DVA/gross export value for Slovakia (shown in Figure 10) implies that this cannot be the case. A more likely solution can be that some indirectly exporting companies in Slovakia fell out of the value chain and they were substituted by foreign firms. The fluctuation of the value in Hungary refers to the lack of robust supplier position. The decrease in 2005 can be linked to the fast wage increase in the previous years, which could transitorily ‘push out’ Hungarian firms from the international value chains. Later, the comparative advantage in wages was regained and the Hungarian companies could integrate again into the GVCs. This, however, took one complete decade.

Re-imported domestic value-added in gross exports

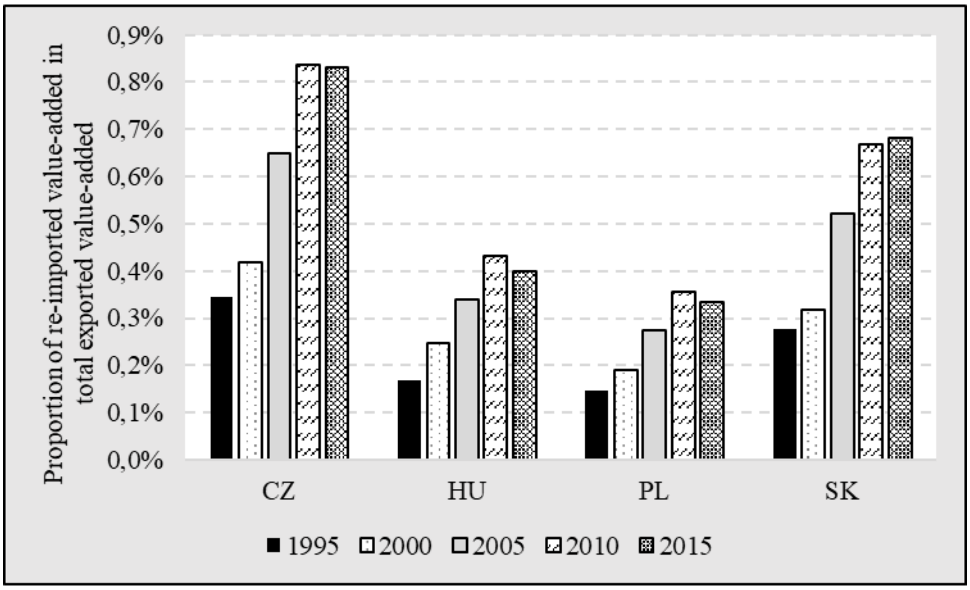
Dimension: countries
Price: base price
Definition: value-added produced in country *c* and exported to country *p* and then re-imported into country *c*.

$$EXGR\_RIM_c = \widehat{V}_c B_{c,c} EXGR_c - EXGR\_DDC_c - EXGR\_IDC_c \qquad (22)$$

where
$\widehat{V}_c = \frac{VALU_c}{PROD_c}$: the share of value-added in output produced in country *c* in matrix form, where the values are on the diagonal;
$B_{c,c}$: the Leontief inverse of country *c*;
$EXGR_c$: the gross export of goods and services in country *c*;
$EXGR\_DDC_c$: direct domestic value-added produced in country *c*;
$EXGR\_IDC_c$: indirect domestic value-added produced country *c*.

Interpretation: this indicator represents the feedback of a country into international trade taking global effects into consideration. It indicates a sort of development in the value chain; the closer is the company to both ends of the value chains, the more value-added can be re-imported. It also refers to high level of function in the production that the company performs. This is chiefly the privilege of the brand owner that produces the lower value-added assignments in low-wage countries, but the higher value-added jobs are assigned to them. The level of this indicator strongly correlates with the size of the domestic market, and with the global demand for the product. It is a flawed indicator of competitiveness because owing to the existence of triangle merchandise trade forms, the products usually do not return to country of the brand-owners but are rather transferred directly to the consumers. Naturally, the parent company receives its share from the profit, but this is not foreign trade but financial transaction. Thus, it is not recorded in the gross export. Consequently, this indicator does not take those companies into consideration that are positioned at the two ends of the value chain. It depicts only the flow of intermediary goods, which usually contain the lowest value-added. It is worth the index it with the exported value-added.

12. Figure: Domestic value added re-imported as a percentage of exported value added in the Visegrad countries between 1995 and 2004



Source: own calculation based on Eora database

The figure above shows that the value of re-imported value-added is very low in the region, thus the products produced in the Central-Eastern European countries generally do not return for further manufacturing. Hungary and Poland have extremely low value, which suggests that the domestic companies do not manage any higher functions in the value chain.

Domestic value-added in foreign final demand

Dimension: countries

Price: base price

Definition: the share of value-added produced by country c in the final demand of country p.

$$FFD\_DVA_{c,p} = \left(\hat{\mathbf{V}}\,\mathbf{B}\,\mathbf{FD}\right)_{c,p} \tag{23}$$

where

$V_c = \frac{VALU_c}{PROD_c}$: the share of value-added produced in country c in the total output;

$B_{c,c}$: the Leontief inverse of country c;

FD: the global final demand matrix.

This indicator moderately connects to competitiveness because it shows how the domestic value-added is globally distributed in the final demand of partner countries. The higher the share, the more embedded is the country into the final demand of the partner country. This also corresponds to the market share of a country. Owing to the diversity in size of the economies, comparison is interpretable only for the competitor countries. Table 15 depicts the share of Visegrad countries in the final demand of their top five partners:

15. Table: V4 countries' share of final demand from partner countries in 2015 (top5)

Ranking	Czech Republic	%	Hungary	%	Poland	%	Slovakia	%
1	Slovakia	11,5	Slovakia	2,8	Slovakia	6,3	Czech Republic	6,7
2	Poland	2,9	Austria	1,7	Czech Republic	5,6	Russia	2,0
3	Austria	2,0	Romania	1,6	Lithuania	5,0	Poland	1,6
4	Hungary	1,9	Czech Republic	1,4	Ukraine	4,8	Hungary	1,6
5	Germany	1,7	Slovenia	1,2	Hungary	3,8	Austria	1,1
Σ		**20**		**8,7**		**25,5**		**13**

Source: own calculation based on Eora database

Table 15 shows that the regional countries have the highest share in the final demand of the neighbouring economies. For Slovakia and the Czech Republic, the common history is a strong factor; however, it is surprising that the supplier network between the Czech Republic, Poland, and Slovakia is denser than in Hungary. A reason behind this can be that the major trading partner for Hungary is Germany, but owing to the large size of the German final demand (which is in the denominator), the Hungarian value-added is ‘lost’.

4.4 Summary, conclusion

In this chapter, the structure of international IO tables was presented. The tables can be derived from the national IO datasets, and all indicators describing GVCs can be obtained from the same. These measures are often interpreted as quantification of international competitiveness, in which the higher exported value-added is the confirmation of right economic policy. However, all the indices were scrutinised and demonstrated to be highly sensitive to the size of the economy as well as to position of the country in the value chain.

All indicators were analysed in the dimension of the Visegrad countries (i.e. Poland, the Czech Republic, Slovakia, and Hungary). All these countries began to integrate into GVCs in the beginning of the 1990s, but they achieved totally different positions. While Poland and the Czech Republic could get better positions and supply higher value-added, Hungary and Slovakia had lagged behind.

When discussing integration into GVCs, it is a key issue that how do firms that are directly exporting integrate other domestic companies into their supply chain network. The Czech Republic has been improving since the beginning, and the local exporters can rely more and more on the domestic suppliers. Meanwhile in Slovakia, this trend is evidently diminishing, which suggests that the domestic firms are being substituted by foreign companies. This is also a risk for Hungary. In the early 2000s, the fast increase in wages in Hungary caused some loss in the economic competitiveness, thus breaking the trend, and some firms fell out of the indirect value-added chain. After a decade, most enterprises could re-integrate into the supplier network, although the trend is still lower, and Hungary lagged behind the Czech Republic.

The case of Poland is peculiar because the size of the domestic market is much larger as compared to the other investigated countries, and thus the firms also have a domestic market. As a consequence, less companies might decide to integrate into the international value chain as an indirect supplier. Although this hedges some risks (during the financial crisis of 2008–2009 Poland could rely much more on the domestic demand and became the only country in the region that did not suffer from recession), it may cause over-specialisation or lost opportunities, assuming that companies can gain higher return from the export market directly or even indirectly.

In summary, one can conclude on the basis of the classical GVC indicators that regional Visegrad countries had integrated deeply into the local value chains, in particular if the indicator of domestic value-added in foreign demand is analysed. The case of Hungary is somewhat different, because its integration into the German economy is larger as compared to other countries. Although the Hungarian value-added in the German final demand is high (in Hungarian-scale), it is relatively low if deflated by the German final demand or by the gross export. This, of course, does not undermine the importance of Germany as it is certainly the most significant partner to Hungary.

5. The position of Hungary in the global value chains

5.1 Graph theory approach

Since the publication of the first IO statistics, GVC scholars have been researching the true position of a country or an industry in GVCs. The main focus of these analyses is how much value-added does a country export and how large it is compared to the competitors. Up- and downstream indicators that can be calculated from the IO databases can only partly answer this question[38]. By further disaggregating the IO data, one can reveal the bilateral trade relations, which by nature can be arranged into a symmetric adjacency matrix and the analyst can map the network of trade in the value-added (Diakantoni et al., 2017). The output of such analyses is often very dense networks, and therefore, the researchers prune them to have a more manageable map (Caldarelli et al., 2012; Xiao et al., 2020).

Transforming value chains as complete graphs to relevant subnetworks

One of the peculiarities of the IO tables is that the $A^{n \times n}$ matrix of technological coefficients is complete, which is also true in case of the international IO databases[39]. This makes it possible to run and evaluate the multiplicator models. However, it also makes the network analysis difficult. There are data for 26 industries in 189 countries in the international IO table published by Eora. In a network, these are individual vertices (total 4,914). In a complete adjacency matrix, there are more than 12 million bidirectional edges. The visualisation and analysis of such a large and a dense network require an extremely large computational capacity. However, for a regional analysis, the study of the whole global graph is not necessary, because no additional information can be gained from low-weight edges. Therefore, the decrease in the size of the network is a rational demand; however, one should optimise it to the country in focus.

[38] One must take into consideration that these are composite indicators, that is, one variable contains all information. Consequently, these variables are aggregated in which the components can be heterogeneous, and therefore the correlation with other indicators of the economy can be weak (Criscuolo & Timmis, 2017).
[39] Values under a predefined threshold are usually neglected in some database (like WIOD), while in other datasets (like Eora), they are not filtered out.

The position of a country in GVCs can be analysed in three ways:

- Forward integration, which reflects how much value-added is exported directly and indirectly into the export of the partner country.
- Backward integration, which shows how much value-added is imported directly and indirectly for the export.
- Integration based on the flow, which is the sum of the direct and indirect value-added export (positive sign) and the direct and indirect value-added import (negative sign) from the same partner.

These three approaches have their own economic interpretation. However, if one studies the position in GVCs, the first option is the appropriate one, while the second is rather for the analyses of raw-material dependency, and the third represents the bilateral relations. This chapter investigates the first approach.

The extant literature offers several ways to segment a *G* graph, or in other words to divide it into coherent subsets[40]. In network science, these algorithms are known as community detection. Within a graph, a community is characterised by a central node with high centrality, and a relatively long distance from other communities. These algorithms are based on iterations and are 'greedy', that is, the result of an iteration is not revisited later (Newman & Girvan, 2004). The steps of this process resemble the agglomerative segmentation analyses applied in multivariate statistics (Ágoston et al., 2019) because the algorithm always begins to extend the community from the two closest nodes. In each iteration, the degree distribution of the new partition is matched to a random graph. In a community, the degree distribution is heavily skewed to the left (i.e. most of the vertices are linked), and thus the difference to a random graph is significant.
By increasing the steps, the algorithm involves more distant nodes. These will have less and less edges, and thus the left skew of the degree distribution will diminish and begin to match with a random graph. When the difference between the partition and a random graph will be smaller than a predefined threshold value, the algorithm will stop.

Most partition algorithms are based on the aforementioned process. Although the underlying methodology might be slightly different, the objective functions (creating coherent,

[40] See Lancichinetti & Fortunato (2009) for more.

homogenous partitions) are the same. To segment GVCs, these algorithms are less appropriate because of the interpretation of distances, which are utterly different in case of GVCs. Consequently, the objective function is not well determined for that use.

The network of GVCs is peculiar because it is complete, and the distribution of edge weights is wide and adjacency matrix is not symmetric. In case of a complete graph, the degree distribution is the same in all vertices[41], because all industries or countries have the same number of edges. If some edges have such a low weight that they can be neglected, the cluster analysis can be performed. Algorithms utilised in network theory try to maximise the modularity[42] of the partitions found:

$$Q = \sum_{i=1}^{k} (e_{ii} - a_i^2) \rightarrow max \tag{24}$$

where
Q: the modularity index;
e_{ii}: the ratio of edges in the partition compared to the whole network;
a_i: the ratio of those edges that have at least one end in the partition.

According to the objective function, the modularity will be maximal where a_i is minimal, or in other words, it will create such communities wherein the members are linked with the largest possible number of edges, while the number of edges between the partitions is the least possible. However, in international trade, the capacity constraints of economies delimit the volume of the export, and thus, smaller economies have less chance to be members of a partition involving a large economy. At the same time, the trade between small economies can constitute a cluster with a high probability. Owing to the fact that the most intensive trade relations are generally conducted with neighbouring countries, partition algorithms based on modularity search will create segments that are based on geographical positions instead of value chain relations.

Consequently, clustering GVCs cannot rely on community modularity. In the next section, a method called *vertical and horizontal detection* will be proposed for which one must define some crucial characteristics of value chains[43]:

[41] In non-weighted case.
[42] The modularity depicts the strength of network division into segments (Newman & Girvan, 2004).
[43] Some letter symbols are both used in graph theory and in also in IO methodology. Changing these could confuse those readers who have knowledge regarding only one of the fields. Thus, we sticked to the coding

1. Definition: Let $G = (V,E(w))$ be a complete graph, where V denotes set of the vertices, E the set of edges, while w denotes the weights of the latter.

We are looking for $\gamma_i(G|_{Si})$, which is the value of S_i partition of the complete graph G, which depends on the subgraphs of $V(G|_{Si})$ and $E(G|_{Si},w)$.

2. Definition: The value of a random V node can be determined by $\gamma(V) = \sum_{i=1}^{k} E(w_i)$, that is the sum of all edge weights[44]. Consequently, the total value of graph G is $\gamma(G) = \sum_{i=1}^{v} \gamma_i(V|G)$, which is the sum of all edge weights in the network.

During segmentation, the S_k partition is compared to S_i $(\#E(G|_{Sk}) > \#E(G|_{Si}))$ in a way that we examine how much the value of Sk decreases if a random edge $e_i(G|_{Sk},w_i)$ is cut from the graph.
The outcome of the cut is the node set of $V_e = \{v \in V: vIe_i\}$, that is the set of those vertices that are still members of the subgraph after the cut of edge e_i.

3. Definition: Let C be the cost of cut and $C_{i,k} = \frac{\gamma(S_k) - \gamma(S_i)}{\gamma(S_k)}$ be the normalised cost of the transition from S_k to S_i ($0 \leq C_{i,k} \leq 1$). It shows how much value did the subgraph lose after pruning edge e_i. Owing to operationalisation requirements, the minimum value of C shall be fixed in the form of $C \geq \varepsilon$ to decrease the computational demand of the algorithm.

In the first step, only the most important edges remain for the node in focus. In the second step, the algorithm maps the further links, which is the most important part of the process. In this round, the addition of those nodes happens that is not a member of the S_k subgraph, that is $V \not\in S_k$. The value of the new vertex can be evaluated in two ways:

1. How much value does the new node add to the previous graph?
2. What is the relationship of the new vertex with the other nodes in the network?

The importance of the second point is that the network should be extended by those nodes that share strong links with the vertices that are already members, and the ‘old’ members

of the fields, always stressing which one is on the subject. There is no intersection of the two methodologies in this dissertation.

[44] Only the direct and indirect export of the value-added is analysed, thus only the outward edges are considered in the model.

are also important for the new one. In that way, it is ensured that the complexity (modularity) of the subgraph increases, while the degree distribution is not skewed a lot towards the dominant nodes in the original complete graph. The result is not a star-structure network, but it more resembles a scale-free network[45]. This can be characterised by the degree distribution of the network.

4. Definition: The degree of $V(G)$ is given by $d_G(V) = |\{e \in E(G): vIe\}|$, which is the number of nodes in G. Consequently, the complete degree of graph G is $\sum_{i=1}^{V} d(v_i) = 2|E(G)|$. The mass probability distribution is given by $f(x) = P(\{d_G \in D: X(d_G) = x\})$.

The extension of the network relies on the condition that the new node is valuable for the vertices that are already members of the subgraph, and the bias of the degree distribution is minimised.

5. Definition: Let $S_1(G) - S_2(G) = \{V_p(G)\}$ true for $S_1(G)$ $S_2(G)$ subgraphs, that is, they differ in one $(V_p(G))$ node only. Then, the bias caused in the degree distribution by the inclusion of node $(V_p(G))$ is $b = 1 + \sqrt{(Gini(S_2(G)) - Gini(S_1(G)))^2}$, $(1 \leq b \leq 2)$.

To choose the optimal $V_p(G)$ node, the algorithm should search for the optimal trade-off between the cut cost and the bias, which implies the following solution:

$$argmax\left\{2\frac{c_{i,k}}{b_{i,k}}: \frac{C}{b} \in \mathbb{R}^+, 0 \leq \frac{C}{b} \leq 2\right\} \tag{25}$$

The value of the C/b fraction is 2 (theoretical maximum) if the new node adds the largest value to the network, while the degree distribution is not changed at all. This ensures that the largest and most crucial nodes in the GVC are involved in the subgraph only if it is important for all other members of the group.

As a consequence, this method can manage the upper or lower triangle adjacency matrices of directed or non-directed graphs only. Keeping node $V_f(G)$ in the focus, the algorithm

[45] It is easy to see that in the GVC, the role of some countries (nodes) is crucial for the small economies. Still, it would be wrong to indicate these smaller economies as full members of value chain of the aforementioned dominant countries. For example, it is certain that Eastern Europe exports value-added to South America; however, it would be disproportionate to infer that the region is just as important for Argentina or Brazil as the other South American smaller economies.

can collectively map the flow of direct and indirect value-added. The reason behind this is the following: let $h_{f,i} = \{v_f, v_i\}$, the value-added flowing from *$V_f(G)$* to *$V_i(G)$*. When the graph is extended by a new $V_j(G)$ node and an $h_{i,j} = \{v_i, v_j\}$ edge, then $h_{i,j} = h_i + \alpha h_f,\ 0 \leq \alpha \leq 1$, where α is share of value-added that was produced by *$V_f(G)$* and remains in node *$V_i(G)$*, that is, $h_{i,i} = F(h_{f,i}, \alpha)$. In other words, the value of $h_{i,j}$ depends on the value of $h_{f,i}$ and α, which cannot be estimated directly. The path of value-added can only be revealed by IO methods which will be elaborated in Chapter 4.

There are two optional stop criteria for the algorithm:

1. Exceeding the threshold value ε. That is, the ratio of $d_S(V_f)$ degree corresponding the *$V_f(G)$* node in focus in partition *$S(G)$*, and the $d_G(V_f)$ degree in the complete graph *G*: $d_S(V_f)/d_G(V_f) \leq \varepsilon,\ (0 \leq \varepsilon \leq 1)$. In other words, of the number of edges of the node in focus, if the partition is equal to ε ratio the algorithm stops.
2. Built on the same analogy, if the ratio of the sum of edge weights of *$V_f(G)$* node in focus in partition *$S(G)$* ($\sum_{w=1}^{n} E_V^w(S) = \{e \in E(S)\colon vIe\}, n = |\{e \in E(S)\colon vIe\}|$), and the total sum of edge weights of the same node on the complete graph G ($\sum_{w=1}^{n} E_V^w(G) = \{e \in E(G)\colon vIe\}, n = |\{e \in E(G)\colon vIe\}|$) exceeds the threshold value $\frac{\sum E_V^w(S)}{\sum E_V^w(G)} \leq \epsilon, (0 \leq \epsilon \leq 1)$. In other words, if some of edge weights of the vertex in focus in the partition are larger than the ε ratio of the sum of the edges in the complete graph, the algorithm stops.

The difference between the two criteria is that the first controls the length of the chains, while the second controls the depth.

The peculiarity of the adjacency matrix $\boldsymbol{A}$ of the value chains is that it is not symmetric, and because of the domestic intermediate use, the diagonal is $diag(\boldsymbol{A}) \geq 0$. This also means that the graph contains loops; therefore, the adjacency matrix is a special one in which the trace is usually larger than the sum of elements out of the diagonal[46], that is, $tr(\boldsymbol{A}) > \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}, i \neq j$. This structure makes the mapping of the international value chain extremely difficult and diminishes the applicability of standard methods for cluster analyses.

[46] It would otherwise mean that the country exports more than its domestic industries use. This is not impossible but rather unlikely.

6. Definition: Let $\boldsymbol{A}(n*n)$, that is, the adjacency matrix of graph *G* be square and for the sake of simplicity symmetric. Let $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{A}$, the so-called *Laplace matrix*, where $\boldsymbol{D}$ = *diag(Σ{e∈E(G): vIe})*, and $\boldsymbol{D_{ii}} = \sum_{j=1}^{n} \boldsymbol{A_{ij}}$, $n = |\{V(G)\}|$, that is, the sum edge weights of all nodes. In that case, **L** is a square and symmetric matrix (because **A** is also symmetric).

It can be demonstrated that graph *G* can be partitioned to $S_1(G)$ $S_2(G)$ subgraphs in accordance with the sign of the $v_{\lambda 2}$ eigenvector corresponding to the second smallest eigenvalue of **L**, that is, λ_2. The sign criterion can be altered to the median. This method is also known as the spectral segmentation (Fiedler, 1973).

In the special case of $a_{ii} >> a_{ij}$, that is, the elements on the diagonal are much larger than the elements out of it, the eigenvalues will also be very large (except the smallest one, because $0 = \lambda_1 \leq \lambda_2 \ldots \leq \lambda_n$)[47], because $\sum_{i=1}^{n} \lambda_i = tr(\boldsymbol{L})$. In other words, the elements on the diagonal dominate the matrix. In that case, the algorithm based on the spectral segmentation method cuts the graph along the two nodes with the largest edge weights. In case of GVCs, these algorithms[48] build the subgraphs around the largest vertices, and thus the value chains of smaller nodes (countries) are much difficult to map.

5.2 Partitioning global value chains

This chapter introduces several algorithms suitable for clustering complete graphs. In the end, segmentation based on the above-mentioned methodology will also be conducted. The data are from the Eora database (see Chapter 4) and cover the year 2015. The space of segmentation is given by the direct and indirect exported value-added into the export of the partner country. That is:

$$dva_{ij} = VAsh_i \times B_i \times \langle X_{ij} \rangle \qquad (26)$$

where

[47] $\lambda_1 = 0$, if and only if there is no such node in the graph that has no edge (in that case, all corresponding elements in the adjacency matrix are zero).
[48] It can be showed that the spectral segmentation is equivalent to the method of k-means kernel functions (Dhillon et al., 2007).

dva_{ij}: the value-added in the export of country *j* produced in country *i*;

$VAsh_i$: value-added produced by country *i*;

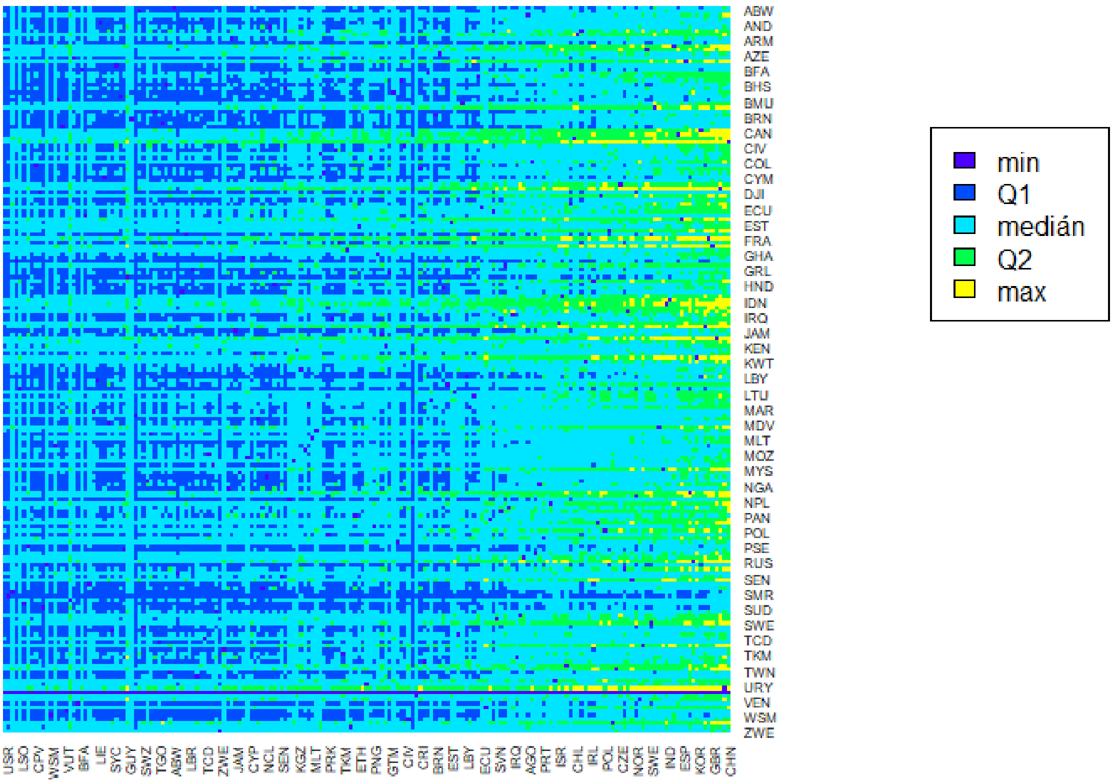
B_i: the block matrix of the Leontief inverse corresponding to country *i*;

X_{ij}: the intermediate export of country *i* to country *j*;

<·>: diagonal matrix in which all off-diagonal elements are zero.

The heat map of such a matrix provides additional information about the structure of the network. On the horizontal axis of Figure 13, one can see the logarithm of value-added in ascending order, while the vertical axis shows the countries in alphabetical order. As evidenced in the map, the largest value-added exporters share a lot of edges with other countries. The colours on the left suggest that the largest value-added importers are importing from more than one large value-added export, which implies the existence of multiple value-added hubs.

13. Figure: Heatmap of the value added transaction matrix



Source: own calculations based on Eora database (R17)

Hierarchical segmentation

At first, a classical community segmentation algorithm is run on the data, which is a simple greedy hierarchical clustering method. Let $C\colon\{C_1, \dots C_n\}$ be the partitions of graph *G* (subsets), that is, $C \subseteq G$. Let $f_{ij}(C) = 1$ if $V_i, V_j \in C_n$, that is, nodes *i* and *j* are in the same cluster in configuration *C* and $f_{ij}(C) = 0$ if not. The *Q(C)* modularity of *C* is given by the following equation (Newman & Girvan, 2004):

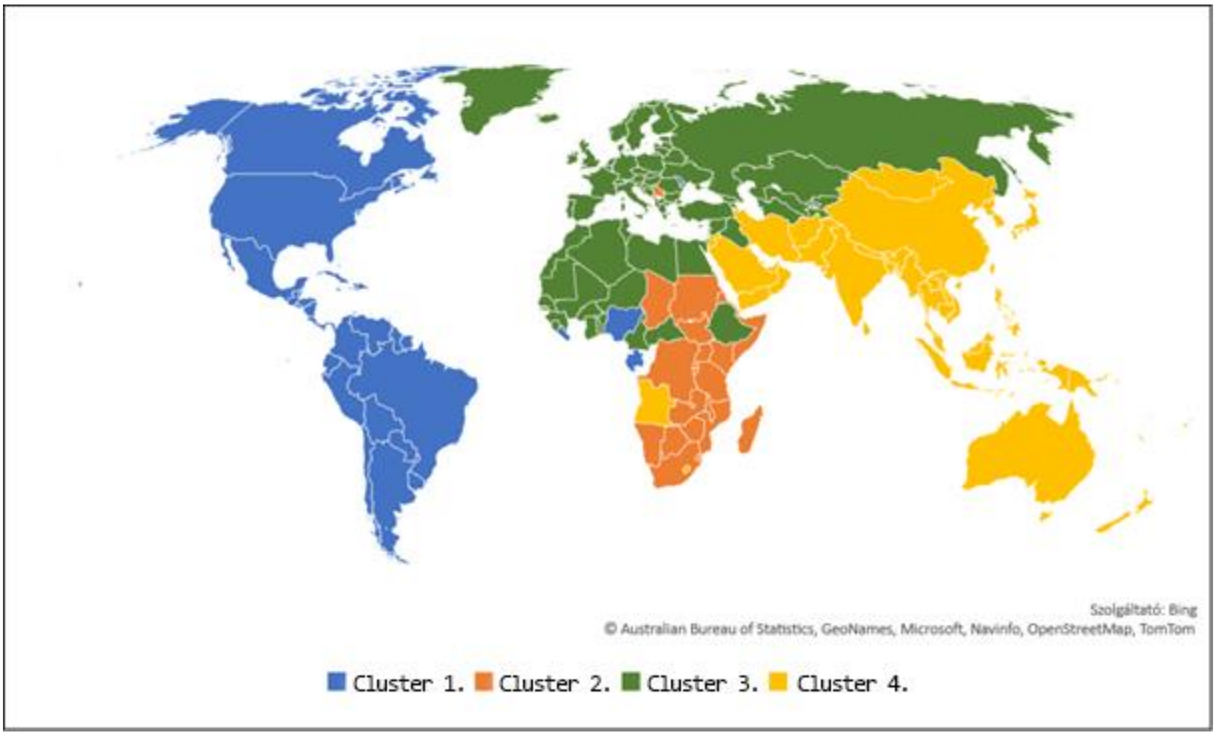
$$Q(C) = \frac{1}{w}\sum_{i,j\in V}\left(A_{ij} - \frac{w_i w_j}{w}\right) f_{ij}(C) \qquad (28)$$

where
A_{ij}: is the weighted adjacency matrix;
w: the weights.

The algorithm maps all possible configurations[49] and finds the most optimal clustering in which the value of *Q(C)* is the maximum. In accordance with the results, one could create a dendrogram and a graph; however, the interpretation in case of such a large data would be extremely difficult. The algorithm introduced above identified four clusters that are shown in Figure 14:

[49] There are also fast heuristic solutions (Clauset et al., 2004).

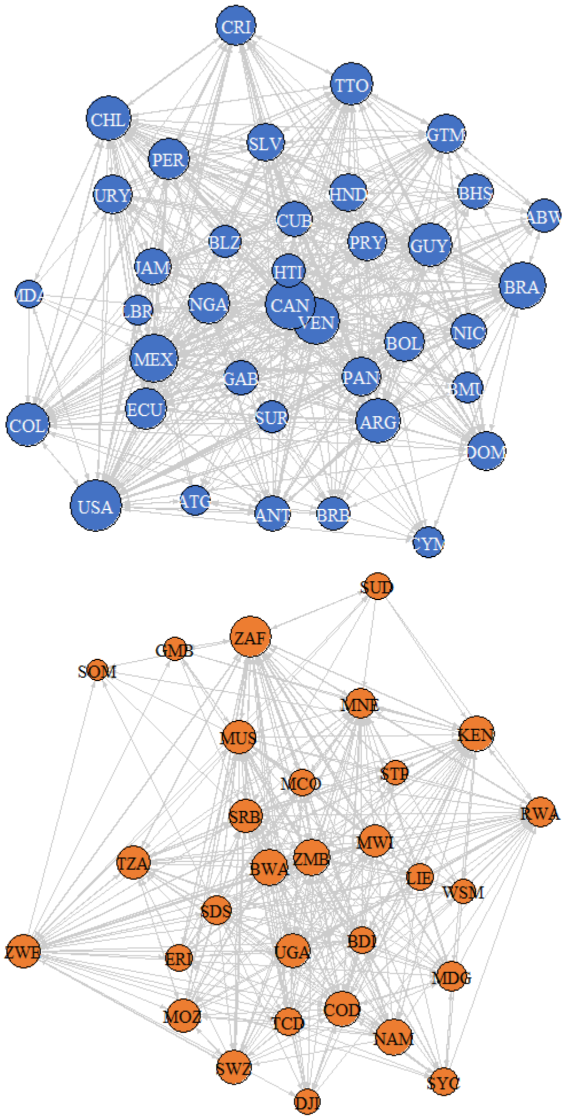
14. Figure: Segmentation based on modularity-based hierarchical cluster algorithm of value-added export (2015, countries with value only)

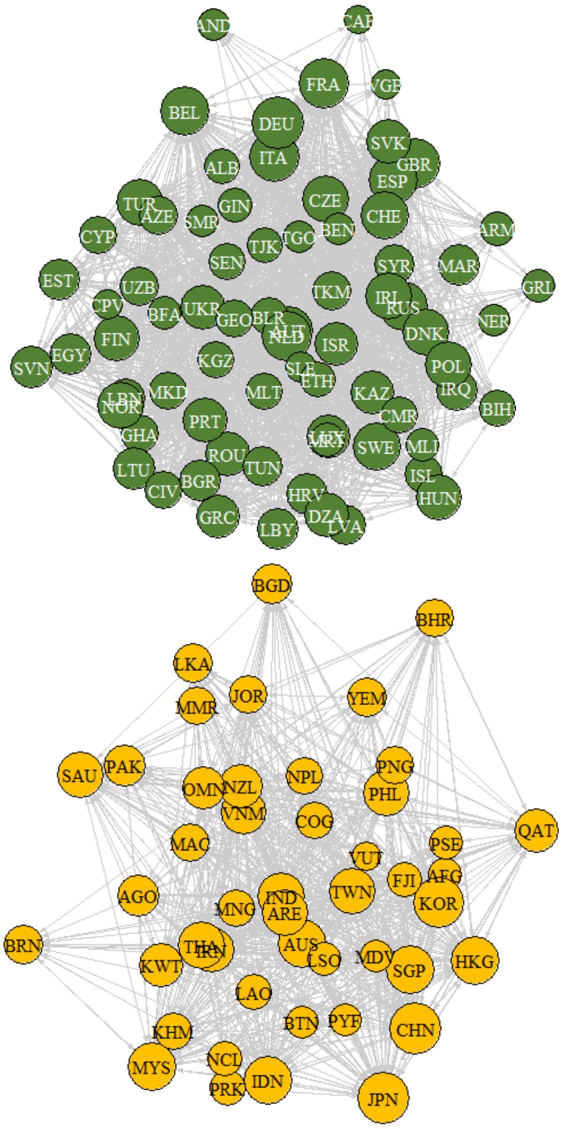


Source: own calculations based on Eora database (R18)

From Figure 14 it is clear that the algorithm clustered the value chains on a regional basis, and thus the EU, North Africa, and Western Asia are in the same partition. The importance of proximity and neighbourhood in international trade is well known and proven by gravity models (Anderson & van Wincoop, 2003; Bergstrand, 1985). Therefore, the results can be well interpreted by distance and other variables of economic policy (Campbell, 2010). In the Figure 15, the networks of the four clusters are visualised. For the sake of simplicity, the diagonal of the adjacency matrix is set to zero, the edges below the median weights are pruned, and the sizes of the nodes are equal to the logarithm of weighted node degrees.

15. Figure: Graphs of clusters generated by hierarchical segmentation

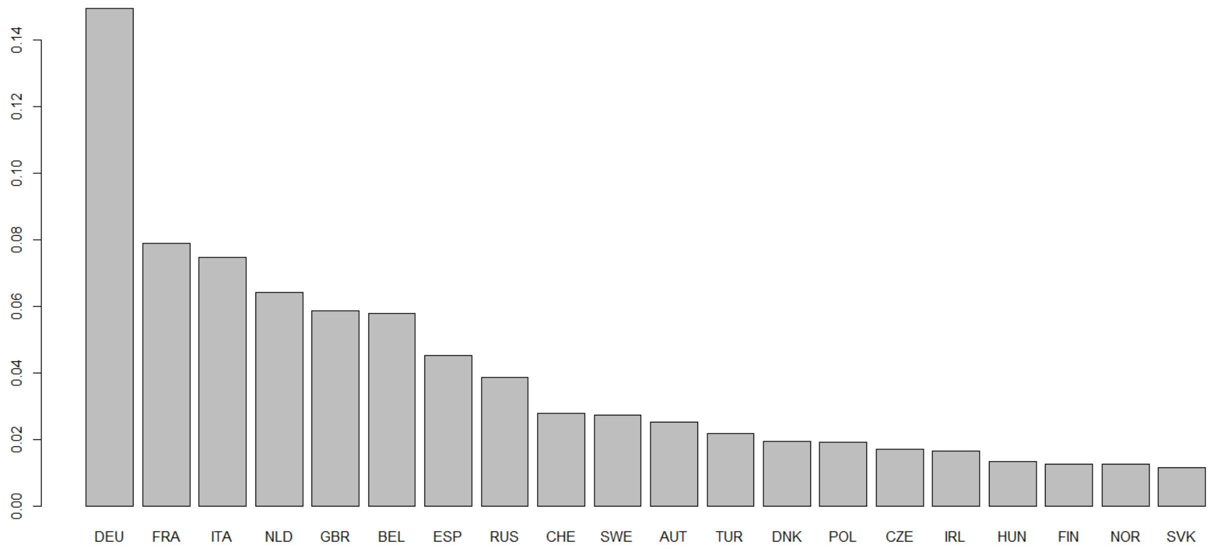




Source: own calculations based on Eora database (R19)

By analysing the PageRank centrality value (see Chapter 7) of cluster 3, which contains Hungary, one can observe that Germany plays a hub role in the region. Although Hungary has the third-largest centrality among the Visegrad countries, its role is still moderate in the cluster (e.g. as compared to Baltic states).

16. Figure: PageRank centrality of third cluster (top 20 countries)

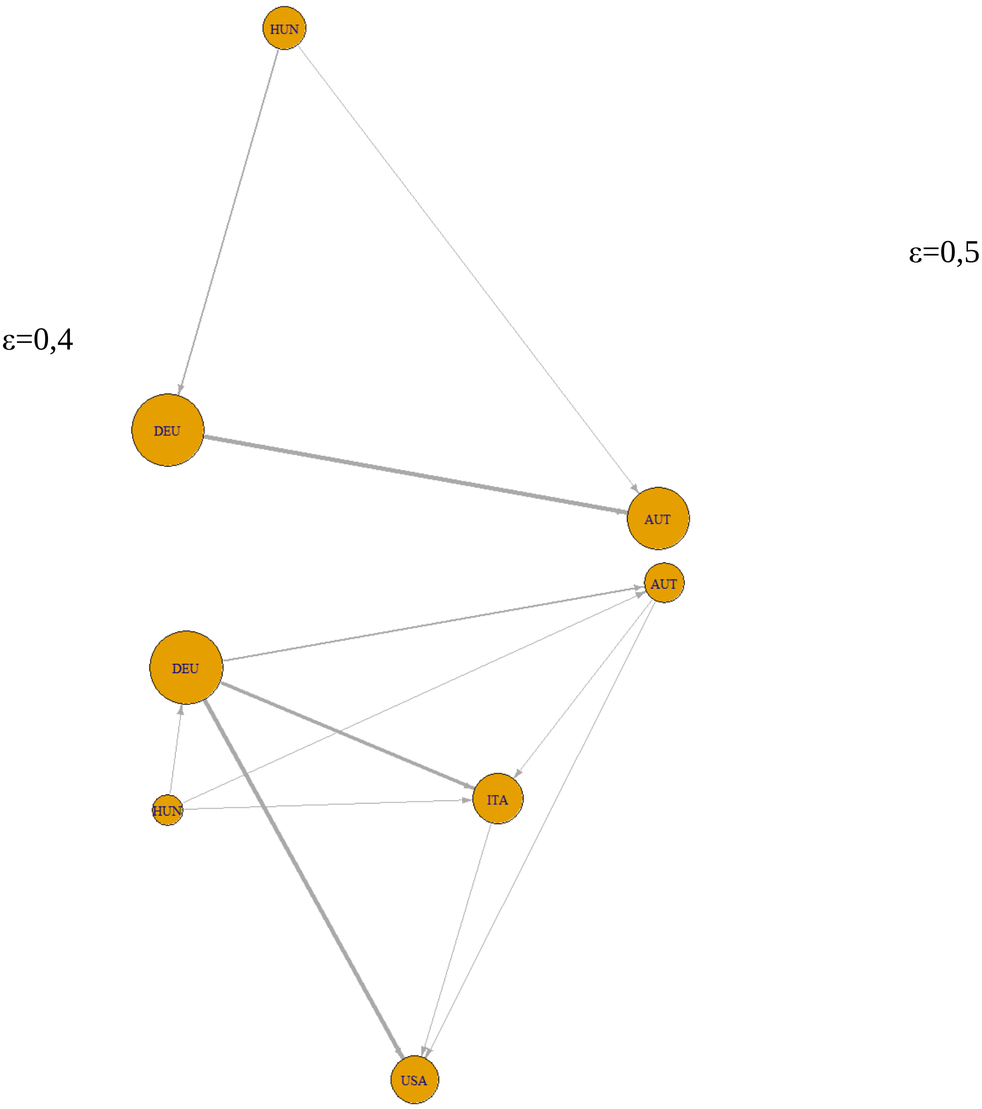


Source: own calculations based on Eora database (R20)

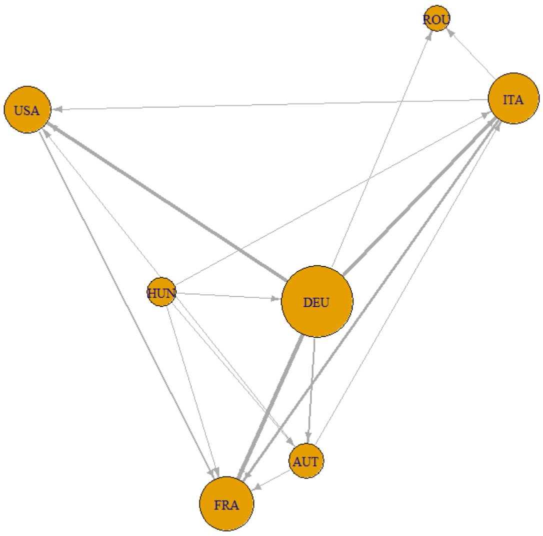
5.3 Vertical and horizontal detection

This subchapter presents the results created by the self-developed vertical and horizontal detection method. Hungary is at the focal point of the graph, and the algorithm grows the network until the stop criterion. The data cover the intermediate goods only, because the buyers of final goods are the final consumers, who do not export the product. To be able to run the algorithm, one must fix the threshold value for cut cost ε. This basically gives the number of nodes in the first iteration. The algorithm was run along several ε values. Figure 17 presents the steps of mapping the value chain of Hungary for the series of $\epsilon_n - \epsilon_{n-1} = 0{,}1$ where $\varepsilon_1 = 0{,}4$.

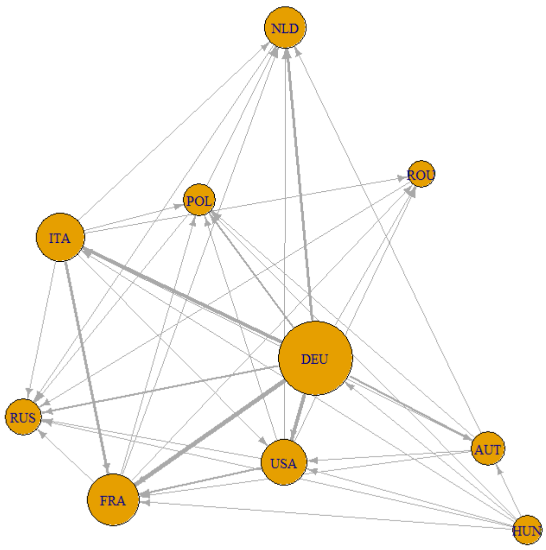
17. Figure: Value added value chain vertical and horizontal exploration with Hungary in focus (2005, with different ε values)

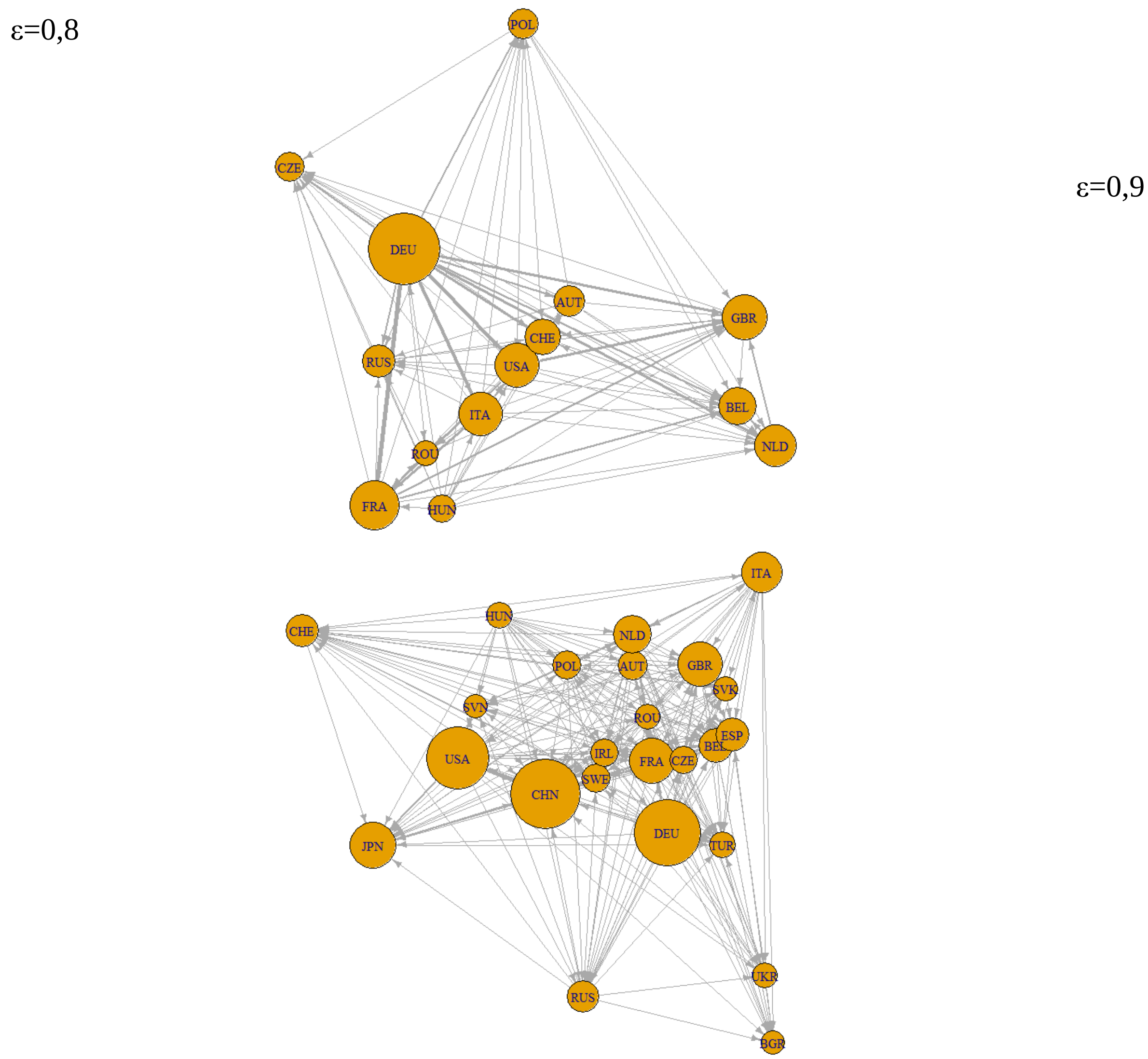


ε=0,6


ROU
ITA
USA
HUN
DEU
AUT
FRA

ε=0,7


NLD
ROU
POL
ITA
DEU
RUS
AUT
USA
FRA
HUN



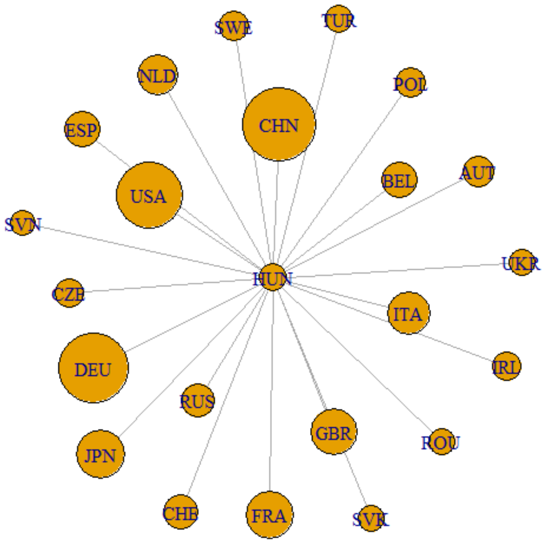
Source: own calculations based on Eora database (R21)

Graphs in Figure 17 clearly depict the network in which Hungary trades is not limited to Europe, as the algorithm connects the USA to the graph at a fairly low ε value. However, it is mainly because of Germany, as the trade of value-added between two countries was already intense in 2015. Both countries play a crucial role in the value chain, but Germany has the highest PageRank centrality measure[50]. The value chain of Hungary is clearly in Europe, but China is also part of it; however, its role is much lower as compared to other non-European countries, such as Japan.

The last (ε = 0.9) network is presented in an unwrapped form in Figure 18:

[50] Hungary has the second highest value, which is not surprising as the algorithm was calibrated to keep it in the focus of the network.

18. Figure: Added values flowing directly and indirectly to Hungary based on vertical and horizontal exploration



Source: own calculations based on Eora database (R22)

Countries in Figure 18 cover 90% of all value-added exported by Hungary into the export of partner countries. The role of China is ambiguous because Hungary exports value-added in a large volume. However, through an analysis of its relations with other countries, one can state that China is more important to Hungary than to other countries in the network.

The evolution of the six networks presented in Figure 17 can be shown on a dendrogram. This plot depicts the aggregation sequence and is frequently utilised during cluster analysis. The theory behind it is based on the fact that one can give the $d(x_i, x_j) = |x_i - x_j|$ Euclidian distance between the observations of $X: \{x_1, x_2 \dots x_i\}, x_i \in \mathbb{R}^n$. At the same time, as it was mentioned earlier, the calculations of weighted distances would give rise to a severe bias, and thus, an alternate method is required to study the sequence of aggregation.

Consider random graphs $G: \{G_1(V_1 E_1), G_2(V_2 E_2) \dots G_z(V_z E_z)\}$, for which $G_1 \subseteq G_2 \subseteq \dots \subseteq G_m$ is true, that is, $|V(G_i)| < |V(G_j|, |E(G_i)| < |E(G_j)|$ if $i < j \leq z$. Another condition is that $V(G_i) \in V(G_j)$ and $E(G_i) \in E(G_j)$ if $i < j \leq z$. In other words, during the evolution of the network, all previous graphs can be found in the consecutive graph in the same form.

Let $A^{n\times n}$ be a symmetrical distance matrix, where $n = |V(G_z)|$ is the number nodes in the last matrix. Find the $E_{j\setminus i}\colon \{E(G_j)\setminus E(G_i)\}, i < j \leq z$ complement, that is, those edges that were not part of graph G_i but are member of the consecutive G_j graph. Define the $a_{kl} \in A^{n\times n}$ elements as following:

$$a_{kl} = \begin{cases} j & ha\ E_i(v_k, v_l)\setminus E_j(v_k, v_l) \neq \{\emptyset\} \\ z+1 & ha\ E_i(v_k, v_l)\setminus E_j(v_k, v_l) = \{\emptyset\} \end{cases} \forall (i < j \leq z) \qquad (29)$$

where z is the index of the last graph G_z. In other words, the value of a_{kl} is equal to the index of the graph in which the edge was created. If there is no edge between two nodes, the distance is the index of the last graph + 1. Thus, it will be further apart of all pairs in the network that share a common edge[51]. One should transform the distance matrix into a symmetric form[52], that is, $a_{kl} = a_{lk}$. One can then visualise the dendrogram, which can be done by any hierarchical clustering algorithm (E. Kovács, 2014).

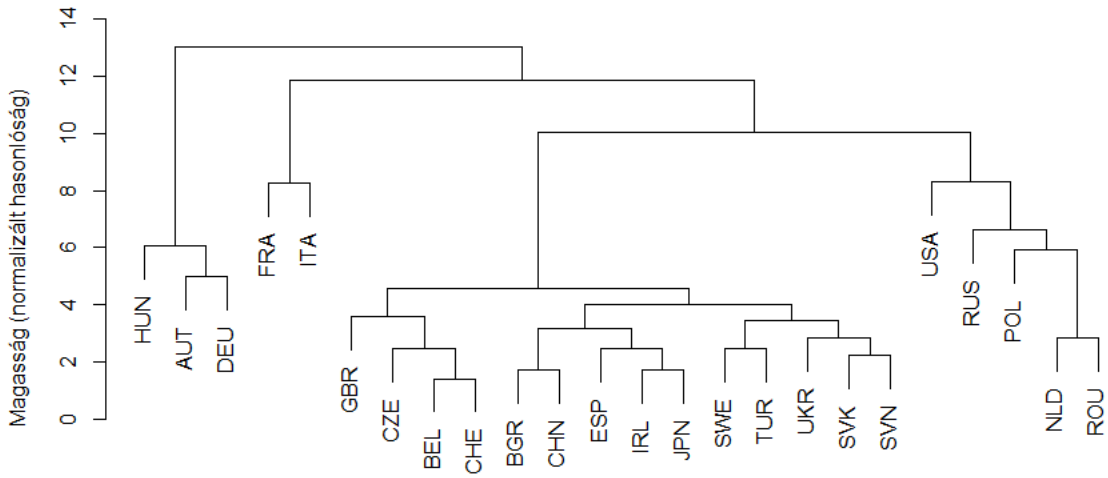
The dendrogram in Figure 19 visualises the sequence of aggregation during the evolution of the network. It shows which edges formed in which step and how the structure was built. The figure should be viewed top-down. It must be noted that all links must be interpreted from the point of view of the country in focus. Based on that, Hungary exports value-added into the export of Austria and Germany in a considerable volume. That amount will be significant during the evolution of the network. Hungary is ranked ‘higher’ on the plot, because while both Austria and Germany are significant partners for Hungary, these two countries are also crucial partners for each other (in AUT->DEU and DEU->AUT direction). Thus, the trade relation between the two is more important than the trade relation with Hungary.

In the next step, the network is extended by the USA and Italy, then France. Links were created with all other countries in the network, and thus, these countries are ranked higher in the dendrogram.

[51] If there is no edge between two nodes, the distance is infinite; however, one cannot indicate such value in a distance matrix.

[52] This operation is not necessary for the distance matrix is an upper or lower triangular matrix.

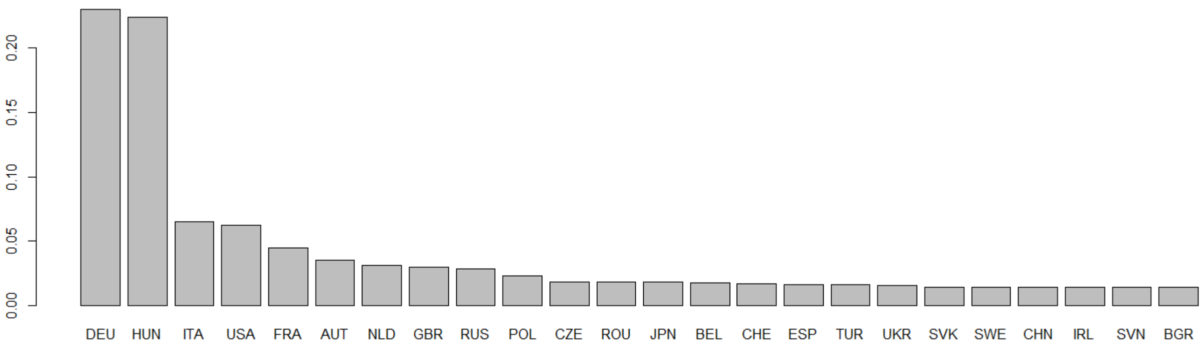
19. Figure: Evolutionary dendrogram of graphs presented in Figure 17



Source: own calculations based on Eora database (R23)

The sequence of aggregation is also reflected on the PageRank centrality of the last graph (Figure 20):

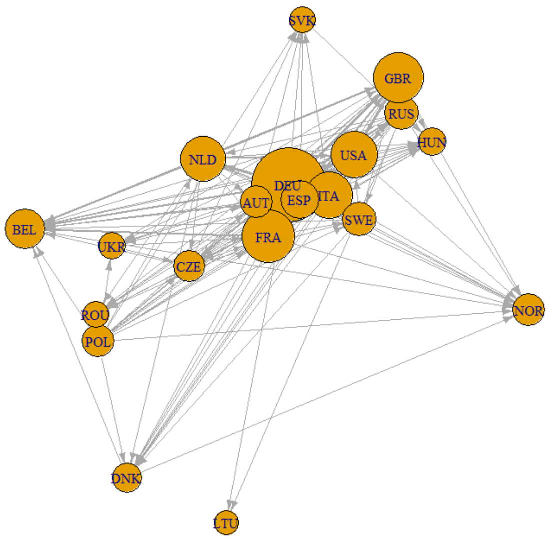
20. Figure: PageRank values of the sixth graph shown in Figure Group 17

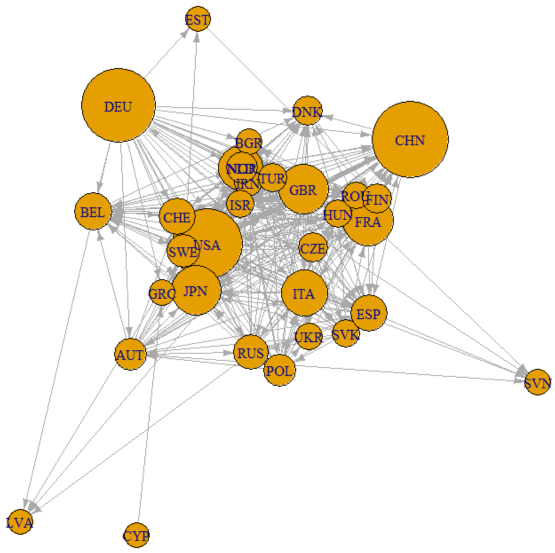


Source: own calculations based on Eora database (based on code R21)

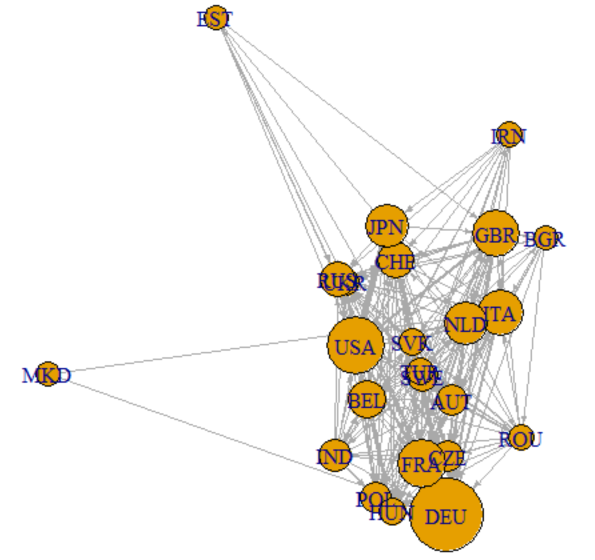
The analysis can be easily extended to other countries, while the differences can also be measured by $G_d = G_i \backslash G_j = \{V(G_i), E(G_i)\} \backslash \{V(G_j), E(G_j)\}$. The network of other countries in the said region was also examined using horizontal and vertical detection method. All figures are created at threshold value ε = 0.9. The importance of the complement nodes and edges is evaluated by PageRank centrality. The lower the value in comparison to the maximum, the less important the node is in the network. The results are summarised in Figure 21.

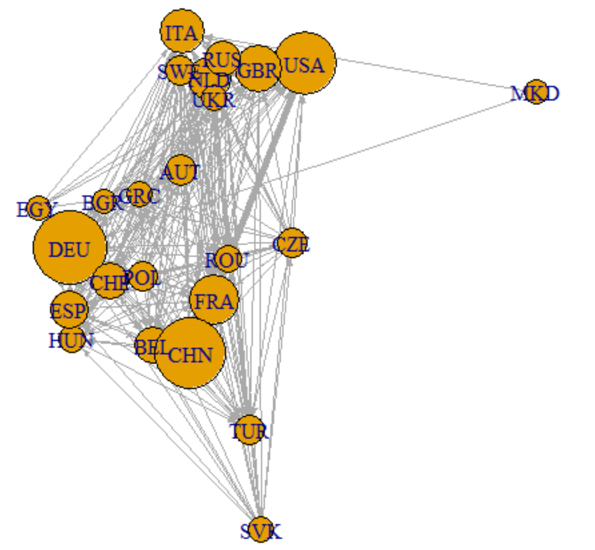
21. Figure: Results of vertical and horizontal exploration in the region (ε=0.9)

<table>
<tr><th>Country</th><th>Difference from Hungary (PageRank scores)</th></tr>
<tr><td colspan="2">Poland</td></tr>
<tr><td>
</td><td>Lithuania (0.018)
Norway (0.018)
Denmark (0.019)

Highest centrality:
Poland (0.238)</td></tr>
<tr><td colspan="2">Czech Republic</td></tr>
<tr><td>
</td><td>Latvia (0.015)
Norway (0.0105)
Finland (0.0107)
Iran (0.0107)
Estonia (0.018)
Israel (0.108)
Cyprus (0.0111)
Denmark (0.0111)
Greece (0.0114)

Highest centrality
Czech Republic (0.209)</td></tr>
</table>

<table>
<tr><th colspan="2">Slovakia</th></tr>
<tr><td>
</td><td>North Macedonia (0.0142)
Estonia (0.0143)
Iran (0.0146)
India (0,0154)

Highest centrality
Germany (0.229)</td></tr>
<tr><th colspan="2">Romania</th></tr>
<tr><td>
</td><td>Egypt (0.0142)
North Macedonia (0.0142)
Greece (0.0170)

Highest centrality
Germany (0.230)</td></tr>
</table>

Source: own calculations based on Eora database (R24)

On the basis of the networks presented in Figure 21, one can conclude that the value chain of Hungary is similar to the value chain of other countries. The differences in the nodes and edges are not significant. Poland and the Czech Republic have more intense relations with the Baltic and Nordic countries owing to their geographic position. It is prominent that the regional countries have intense trade relations with countries on the Balkan, such as North Macedonia, while none of these economies are part of the network of Hungary. This suggests that the value-added export of the regional countries are more diverse than Hungary's.

The diversification can be evaluated by many methods such as the Gini index, area under the Lorenz-curve, Herfindahl index, or entropy index. If the data contain a number of elements with low value, it is worth using the Herfindahl index, because their weight will be also lower in contrast to the Gini index, which assigns the same weight to every element. The Herfindahl index H can be calculated by $H = \sum_{i=1}^{n} \left(\frac{x_i}{\sum x_i}\right)^2, \forall x \leq 0$, where x_i is the value-added exported by the country in focus into the partner country i. The larger the H, the higher is the concentration. Table 16 depicts the Herfindahl values for the investigated countries:

16. Table: Herfindahl index of value-added imports from countries in the region

Country	Herfindahl-index
Poland	0,0127
Czech Republic	0,0191
Romania	0,0192
Hungary	0,0200
Slovakia	0,0349

Source: own calculations based on Eora database

The Herfindahl index shows high concentration for Hungary and Slovakia as compared to the other countries in the region. The absolute differences may seem low; however, it must be noted that the degree of freedom of the Herfindahl index is 1. One can simply calculate how much decrease in the concentration is required to achieve the value of a preceding country. For Hungary, the Herfindahl index would be the same as in Romania if the value-added export to Germany decreased by 8%.

5.4 Summary, conclusion

This chapter highlighted that the application of classical community segmentation algorithms in the field of GVCs is very limited, especially if the economy in focus is much smaller than the larger value-added exporters. It was found that the structure of GVCs resembles the network of the Internet; some nodes have a crucial importance with large weights and these vertices are well connected. In other words, large countries in the world mainly trade with each other. In such a constellation, Hungary lies at a node of very low importance with edges of low weights, and thus the community detection algorithms often connect it to larger nodes (chiefly to Germany), which undermines its role in the value chain.

Hungarian companies are integrated into the GVC, and their importance is assuredly larger than it is suggested by the community detection algorithms. A methodology was proposed that can map the network of any countries in focus, ensuring that the network is extended only by those countries where the relations are as deep as possible while it is extended by the least necessary amount. In a network with Hungary in focus, almost all EU countries are members together with the USA and China. That network is dominated by Germany, just as in case of Romania and Slovakia. In case of Poland and the Czech Republic, they are the dominant nodes (measured by PageRank centrality). This suggests that the latter two countries have their 'own value chain', while Hungary, Romania, and Slovakia were a bit far from that in 2015.

The networks of the Visegrad countries and Romania are very similar to each other. The only differences are in the local trade relations (for example Poland's relation with the Baltic and Scandinavian countries is much stronger than Hungary's). The concentration of exported value-added is also slightly different. It was confirmed that the Polish value chain is more diverse, while the Slovakian is more concentrated. This could bear considerable risks, as there could be high dependency in the supply chain (Koppány, 2017).

The current analysis focused only the upstream value chain, that is, the export of value-added. The reason behind this is that in terms of GVC positions, it is more crucial how far the exported value-added gets. Analysis based only on the bilateral trade statistics does not yield a satisfactory answer to this question. With regard to the import side, it is less important that which countries were 'visited' by the value-added before it arrived at Hungary.

6. Disaggregating the value-added flow in the value chains

Mapping the flow of value-added in IO tables is not possible without transformations, as the value chains do not contain any producer who did not use any inputs from another producer[53]. Thus, the source of the value-added flows is unknown. Nevertheless, the last chain in the process is usually known, as it is the consumer (household or government) who purchases the final good. Still, one cannot analyse value-added flow without a source point.

However, a partial analysis could be conducted if one studies a section of the value chain. In this case, the flow from a random point until the final consumer can be mapped and constrained by the fact that the previous flow of the good remains unknown. A further limitation is that the path could be infinitely long[54]. At the time being, there are estimations only for the number of production sequences (Wang et al., 2017a) and border crossing (Muradov, 2016). As per these studies two to five, producers participate on average in the production process and two to four border crossings take place in the supply chain.

In the following subchapter, a methodology will be introduced that is appropriate to map the direct and indirect flow of value-added export along multiple countries. This method reveals the path of intermediate goods only. These are the goods that are definitely part of the value chains, while in the end, producer will create a final good, which will be purchased by a final consumer. The latter type of goods are not covered here; however, the model can be easily extended. With the help of this method, the flow of value-added into the export of Hungary will be mapped. This method is also known as the structural-path analysis (Miller & Blair, 2009).

6.1 Methodological summary

In the system of GVCs, the countries are linked by trade channels. These can be considered as the graph edges that connect nodes. Values in the international IO tables represent the gross export flow between two countries. Owing to the gross accounting principles,

---

[53] To such a producer, one might go back to the first man in history who made a tool.

[54] It is certain that every intermediate good will be a final good at last; however, there are no limits for the time frame of the process. In an extreme case, the intermediate good can be stuck in the value chain 'forever'.

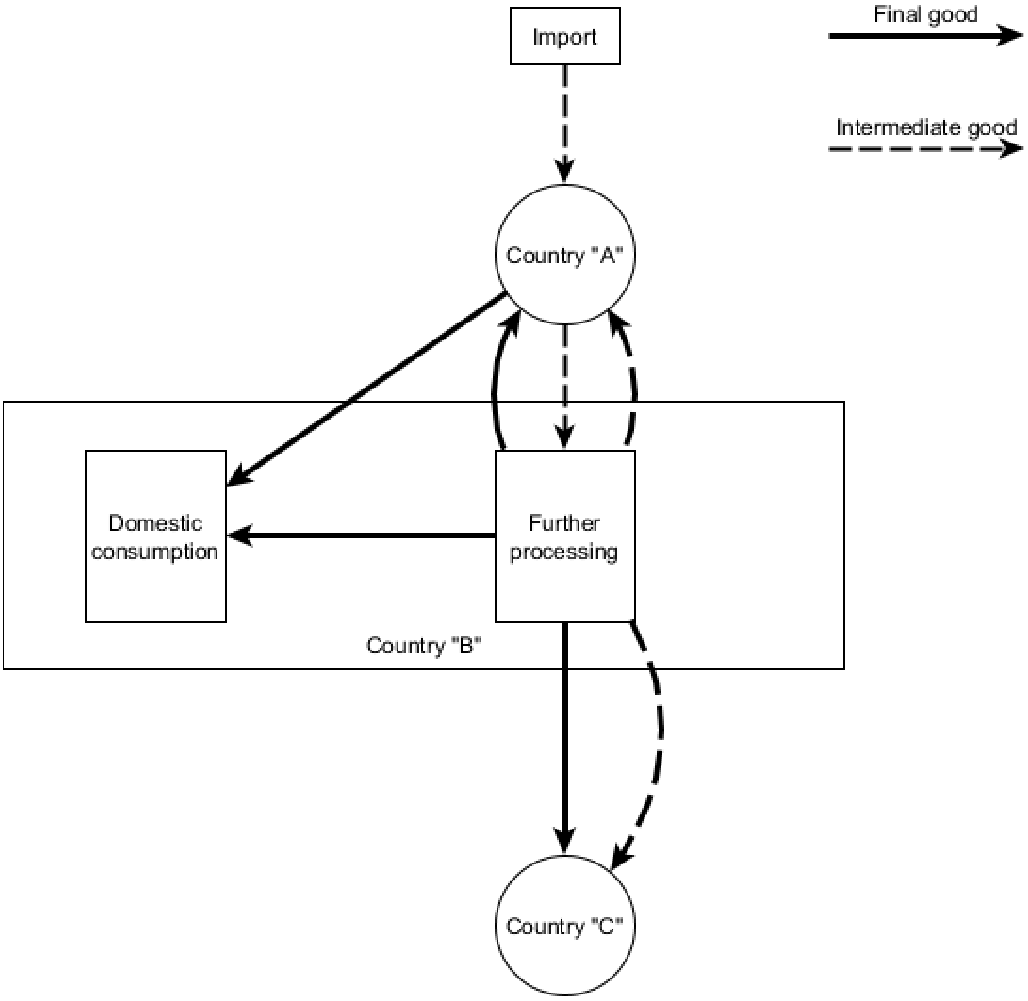
one cannot map the real position of a country[55]. The flow of value-added passing through the edges has eight components, which is depicted in Figure 22, which represents the trade flow between three economies.

1. The final goods of the exporter, which are consumed in the importer country without any further transformation.
2. The intermediate good of the exporter country, which the importer country can use in three ways:
    a. After processing, the intermediate good will become a final good and will be consumed in the importer's country.
    b. After processing, the intermediate good will become final a good and will be exported to a third country.
    c. After processing, the type of the good will be still intermediary and will be exported to a third country.
3. The exporter country processes the imported intermediate good and produces a final good that will be exported to its final destination where it will be consumed.
4. The imported intermediary good will be processed, but the type will not change (still intermediary). Here, the producer faces three choices:
    a. After processing, it will become a final good, which will be consumed in the importer's country.
    b. After processing, it will become a final good and will be exported to a third country where it will be consumed.
    c. After processing, the type of the good will not change and will be exported to a third country.

As one can observe, there are many types of flows depending on the type of the good. The sum of all these flows is equal to the sum of gross export. However, only the intermediate goods are part of value chain because the exported final goods will be consumed by the resident households or government in the importer country without any further transformation[56]. Almost all industries provide inputs to a final good, and thus the exported product contains the direct value-added produced by the exporter's industry and the indirect value-added created by other industries.

[55] This bias is caused by the gross accounting principles covered in Chapter 3.
[56] It must be noted the final goods purchased by the households (or the government) are the products of the industry that produced them and are not related to the retailer. The latter increases the value of the product by the sale services only and by production. This induces the following equation: (Household purchased consumption) = (output of the industry that produced the good) + (output of retail services), where the (retail services) = (sales) – (cost of goods sold) – (change in inventories).

22. Figure: Flow of value added in a three-player value chain according to the nature of use



Source: own edition

7. Definition: The intermediate export of country *i* into country *x* is the following:

$$int.e_{i,x} = int.DVA_{i,x} + REII_{i,m_1,x} + REII_{i,m_2,x} + \cdots + REII_{i,m_n,x} \quad (30)$$

$$int.e_{i,x} = int.DVA_{i,x} + \sum_{m=1}^{n} REII_{i,m,x}\,, \quad m = \{1,2 \ldots n\} \quad (31)$$

where

int.e: the intermediate export;

int.DVA: the domestic value-added in the intermediate product;

REII: re-imported domestic value-added;

i: exporter’s country;

m: importer’s country;

x: export partner.

The first part of this equation represents the value-added contribution of country i. The second part depicts the further part of the value chain to which country i is connected.

The original definition of the re-imported value-added covers the aggregated multilateral $REII_i$ value instead of the bilateral one (Koopman et al., 2014). Thus, it only depicts how large is the re-imported domestic value added in the gross export of country i. To map the value-added flows, one needs bilateral values, which would require the disaggregation of the REII values.

8. Definition: Let $\boldsymbol{B} = (b_{i,j}) \in \mathbb{R}^{n \times n}$ global Leontief inverse, where i represents the supplier, and j the user. For i = j, the elements are on the diagonal of **B** and represent the domestic intermediate use. If **B** is known, the flow of direct and indirect value-added between two random points can be disaggregated in the following way:

$$int.e_{i,x,m} = \underbrace{<\boldsymbol{VA_i}> \boldsymbol{B}_{i,i} int.e_{i,x}}_{\text{domestic } VA \text{ in i}} + \sum_{m_{m \neq i}}^{n} \underbrace{<\boldsymbol{VA_m}> \boldsymbol{B}_{m,i} int.e_{i,x}}_{import\ VA \text{ in i}} + \underbrace{<\boldsymbol{VA_x}> \boldsymbol{B}_{x,i} int.e_{i,x}}_{reimport\ VA \text{ in x}} \tag{32}$$

where

$int.e_{i,x}$: the intermediate export from country i to country x;

VA_i, VA_m, VA_x: the VA/output ratios in the exporter country i, the importer country m and in the export partner country x;

<•>: diagonal matrix;

$\mathbf{B}_{i,i}$: the final direct demand country i;

$\mathbf{B}_{m,i}$: the Leontief inverse of the indirect import partner m in country i;

$\mathbf{B}_{x,i}$: the Leontief inverse of indirect import export partner x in country i.

Consequently, equation (8) can be interpreted as:

gross export = domestic value-added + re-imported domestic value-added from import partners + re-exported value-added from the export partners

The value-added from the import partners can be further disaggregated, as it contains all direct and indirect value-added from all other trade channels:

$$REII_{i,m_1,m_i,x} \underbrace{=< VA_{m_1} > B_{m_1,m_1} REII_{i,m_1,x}}_{domestic\ VA\ in\ m_1\ directly\ to\ i} + \underbrace{\sum_{i=2}^{n} < VA_{m_i} > B_{m_i,m_i} REII_{i,m_1,x}}_{domstic\ VA\ in\ m_1\ via\ m_i\ to\ i} + \sum_{i=2}^{n} \underbrace{\varepsilon_{i,m_1,m_i,x}}_{reimported\ domestic\ VA\ in\ m_1\ via\ m_i\ to\ i} \tag{33}$$

$$i = \{2,3 \dots n\}$$

where

$REII_{i,m1,x}$: the value-added produced by country m_1 and traded directly to country i, and then exported to country x;

$REII_{i,m1,x}$: the value-added produced by country m_1 and indirectly traded to country i via country m_i, and then exported to country x;

$\varepsilon_{i,m1,mi,x}$: value-added re-imported by country i from countries m_1, m_i, and then exported to country x.

Owing to the complete graph, the aforementioned equation can be drawn infinite times. However, after a few times, any further disaggregation does not provide significant additional information. The ε of re-imports is the close of the model, and its volume is insignificant as compared to the other parts of the equation.

The disaggregation of the REII values provides information regarding how far the investigated country (or its industry) can ‘get’ in the value chain. The largest volumes of value-added are certainly on the direct edges, while it is smaller on those edges that are indirectly connecting the country to another one, because the intermediate country will produce its own value-added, and other sources will also be added from third countries. Thus, the relative importance of the domestic value-added diminishes as the network grows.

A practical example is introduced in the II. Annex. Four fictive regions (R, S, T, and U) with industries are trading. For the sake of simplicity, only the intermediate transaction matrix is indicated. One can observe that because of an embargo, there is no direct trade between regions R and T. After calculation of classical measures (transaction matrix and the Leontief inverse), the direct and indirect paths of R's import (equivalent to the sum of the exports of all other countries) can be mapped. In accordance with this, example it can be shown that, albeit no direct trade between regions R and T, the import of R contains indirect value-added from T via the import from S and U. The values with a grey background indicate the direct and indirect import of R (taking advantage that sum of value-added is equal to the gross export). Summing up these values, one can check that the equality is valid and the path of value-added can be mapped. The mapping can cover multiple rounds, with the only difference that one should use the adjusted gross export.

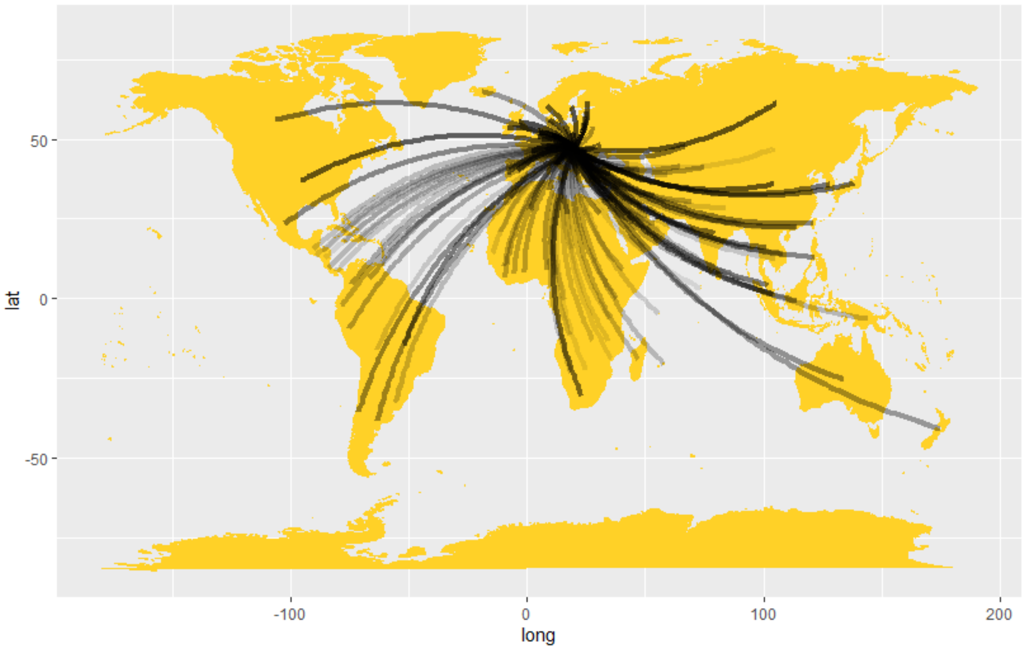
6.2 Mapping the flow of the value-added by Hungary

The value-added produced by resident firms can be traced at national and industrial levels. In case of the latter, it must be noted that an industry uses value-added from all other industries, such that the volume of data to be analysed exponentially increases. In the Eora database, altogether 26 industries can be investigated, which means 10,000 data points even in the first round of the path analysis. Therefore, the following analysis covers national data only and separately covers the automotive industry.

The flow of Hungarian value-added will be mapped with the largest partner, Germany. For a better visualisation, the edges pointing to Germany are neglected. The largest volume of value-added flow is of course transferred via the direct edge between Hungary and Germany.

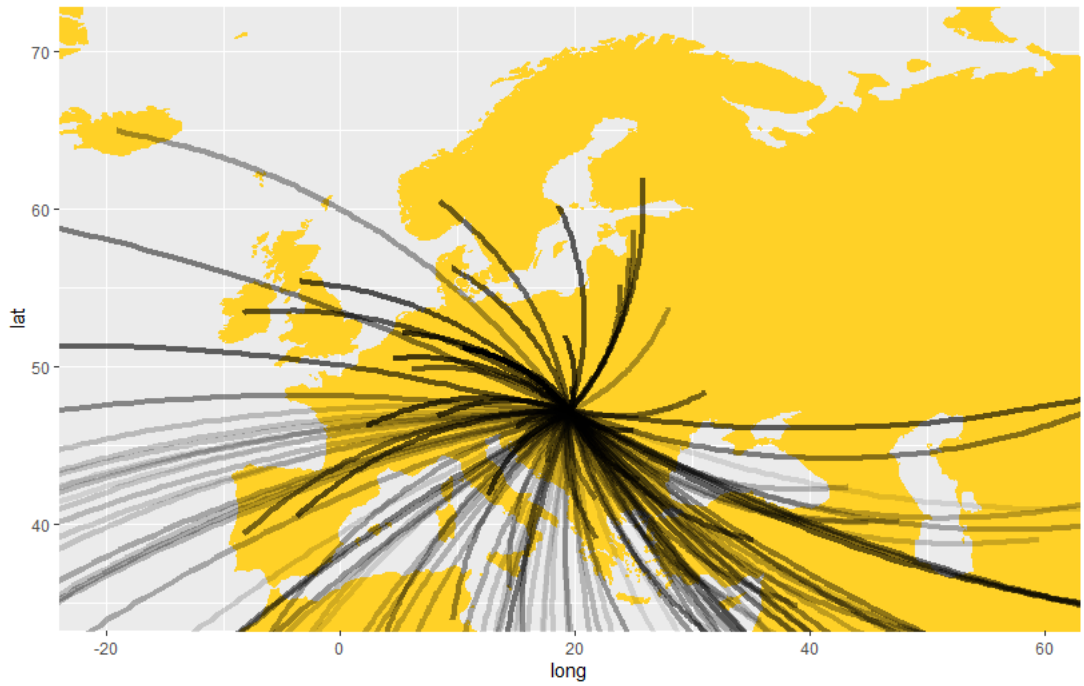
After that, the largest value-added from Hungary to Germany is transferred through Austria (it must be noted that the Hungarian value-added flowing from Austria to Germany contains not only the directly imported value-added from Hungary to Austria but all other sources from third countries). After Austria, the order is the Netherlands, the Czech Republic, Poland, and Belgium. Slovakia and Romania are the ninth and tenth largest partners transferring direct and indirect Hungarian value-added to Germany. Interestingly, the least important indirect partners are Burkina Faso, Myanmar, and Afghanistan. Figures 23–25 depict the flow of direct and indirect value-added from Hungary to Germany:

23. Figure: The flow of added value produced by Hungarian companies and exoirted to Germany* (world map)



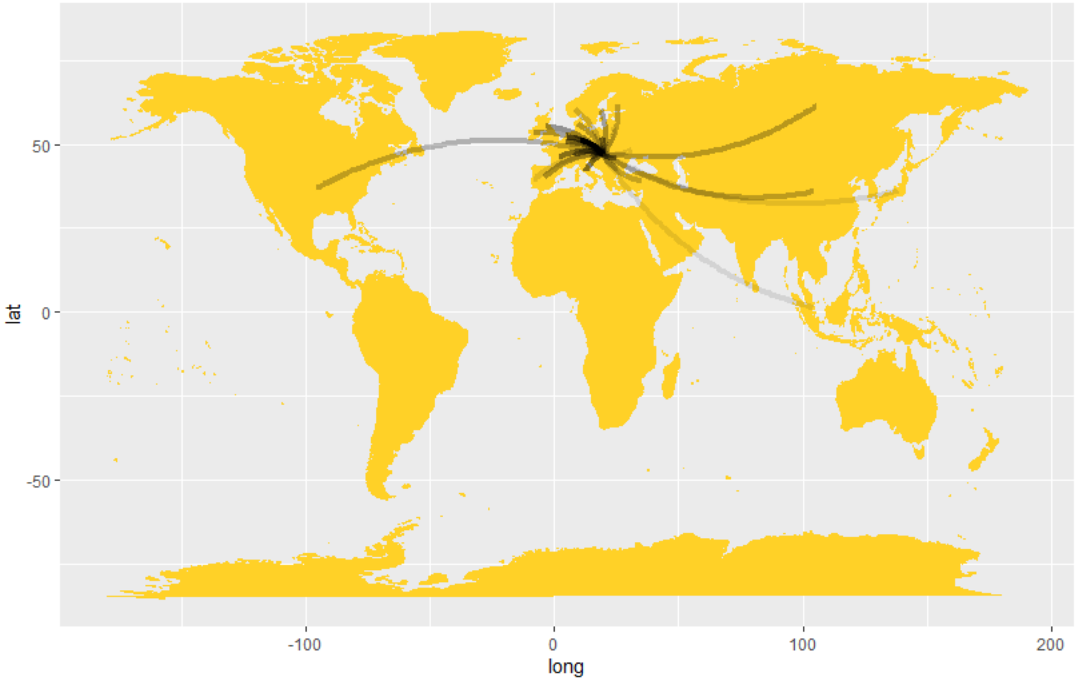
* The edge connecting the partner country with Germany is not displayed for transparency. The thickness of the curves is proportional to the amount of added value exported. Source: own calculations based on Eora data (R25)

24. Figure: The flow of added value produced by Hungarian companies and exoirted to Germany* (Europe)



* The edge connecting the partner country with Germany is not displayed for transparency. The thickness of the curves is proportional to the amount of added value exported. Source: own calculations based on Eora data (R25)

25. Figure: The flow of added value produced by Hungarian companies and exoirted to Germany* (top30, world map)



* The edge connecting the partner country with Germany is not displayed for transparency. The thickness of the curves is proportional to the amount of added value exported.
Source: own calculations based on Eora data (R25)

Mainly European countries participate in the network of value-added flow between Hungary and Germany (in the first round). Among the top 30 countries only Russia, China, the USA, Turkey, Japan, and Singapore are the non-European economies. Investigating the partners in the second round, the largest partners are not located on the continent: Hungary → USA → Canada → Germany is the first, then the path of Hungary → USA → China → Germany ranked the second. The main reason behind this could be that the majority of the value-added flowing from the direct partners are going to Germany, and only a small amount remains in the system for the second round (e.g. in the path of Hungary → Austria → USA → Germany).

26. Figure: The flow of added value produced by Hungarian companies and exoirted to Germany* (top30, second round, world map)

Source: own calculations based on Eora data (R26)

17. Table: The flow of added value produced by Hungarian companies and exoirted to Germany* (top30, second round)

Forráspont	Partner1	Partner2	Desztináció	Hozzáadott érték (ezer USD)
HUN	USA	CAN	DEU	9343366,0
HUN	USA	CHN	DEU	602997,1
HUN	FRA	CHN	DEU	334644,9
HUN	JPN	CHN	DEU	326674,1
HUN	USA	MEX	DEU	298616,6
HUN	USA	DZA	DEU	295293,1
HUN	USA	HKG	DEU	252959,8
HUN	USA	GBR	DEU	240086,6
HUN	USA	SAU	DEU	239771,0
HUN	NLD	CHN	DEU	234214,2
HUN	TUR	RUS	DEU	209328,3
HUN	SWE	NOR	DEU	169149,4
HUN	USA	AUS	DEU	165779,9
HUN	CHN	AUS	DEU	165300,7
HUN	USA	NOR	DEU	141905,3
HUN	FRA	CHE	DEU	137363,4
HUN	ITA	CHN	DEU	137003,0
HUN	NLD	GBR	DEU	135062,9
HUN	RUS	KAZ	DEU	97862,4
HUN	ITA	CHE	DEU	96469,8

Source: own calculations based on Eora data (R6)

6.3 The flow of value-added produced by the Hungarian automotive and other industries

The aforementioned results are highly aggregated, as they depict the flow of value-added at the level of the national economy and thus contain the value-added produced by all actors of the economy. For a more accurate analysis, it is worth studying the automotive industry separately. No surprise that the indirect relation between Hungary and Germany is the most crucial (Hungary → Germany). While the first round is dominated by Austria, the Czech Republic, and Slovakia, almost all European countries can be found among the top 30 partners. Non-European economies in the first round are the USA, and in small amount Japan, China, South Korea and Thailand.

27. Figure: Map of Hungarian value added produced by the automotive industry exported to Germany*

* The edge connecting the partner country with Germany is not displayed for transparency. The thickness of the curves is proportional to the amount of added value exported. Source: own calculations based on Eora data (R25)

In industrial dimension, the partners of the Hungarian automotive industry are generally also carmakers; however the electronic industry, machinery, and the chemical industry can also be found in the list. The occurrence of the industry of financial services (e.g. in case of Switzerland and Luxembourg) is not unusual and suggests that some form of finances can have a larger amount than the value of the exported goods themselves[57]. Besides Germany, it is worth checking the automotive supply network of other economies. The regional supply system is generally dominated by Germany, as it is the largest intermediate export partner of the value-added produced in Hungary.

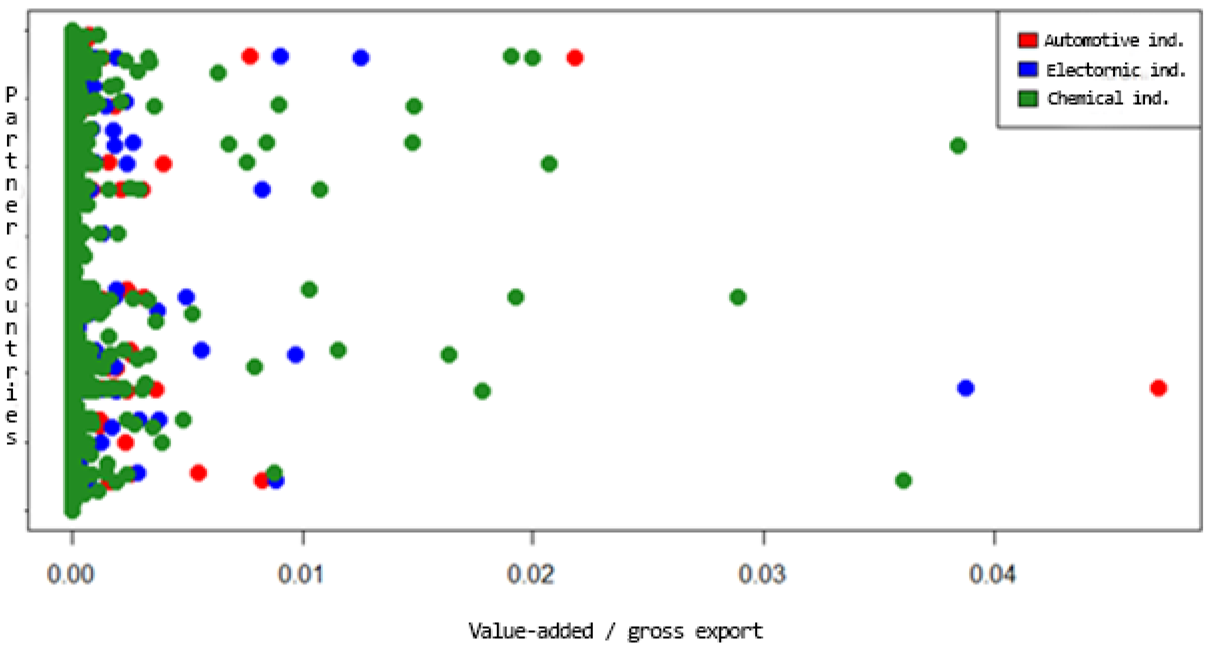
28. Figure: Flow of value added produced by Hungarian automotive companies through indirect partners to the V4 countries and Romania

Source: own calculations based on Eora data (R25)

[57] The statistical accounting issue of this phenomena is discussed in Chapter 3.

Notwithstanding the fact that the network of the Hungarian vehicle industry can be regarded as global, it is still extremely concentrated because the trade is dominated by European countries. In contrast to that, there are several paths in the network of chemical goods where the value-added produced in Hungary 'live longer' (in the share of gross export), particularly in the paths towards Germany, Italy, the USA, Romania, or Austria. Figure 29 illustrates the distribution of value-added in the share of gross export produced by three Hungarian industries (automotive industry, electronics, and chemicals) flowing in all possible directions (countries and industry – more than 35,000 options). Indexing with gross exports adjusts the differences in the export volumes and adapts the volumes to the supplier's industry[58].

29. Figure: The first round of value-added flows of some Hungarian industries by all partner countries



Source: own calculation based on Eora database

Figure 29 shows that there are strong trade relations in case of all three industries. These are usually the same partner industries in some regional countries (e.g. the Austrian and Slovakian automotive). Besides that, the importance of the chemical industry is high because its exported value-added 'remains' longer in the value chain as compared to the other industries.

[58] The index was the Hungarian export and not the export of the partner country because in case of the latter, severe bias could occur because of the differences in the size of the economies.

6.4 Summary and conclusion

This chapter proposed a method (based on structural path analysis) to map the flow exported value-added of Hungarian enterprises. The volume of exported value-added was already known; however, the proposed process one could disaggregate it into direct and indirect values. In accordance with that, the flow of exported or imported value-added between two random partners can be mapped. However, it must be taken into consideration that owing to the high computational requirements, the number of detectable trade relations is limited, and thus, only one or two rounds of flows can be mapped.

On the basis of the findings, one can conclude that the exported value-added of Hungarian produces is circulating chiefly in Europe and in small volume the USA, Japan, and China before reaching its destination. The network of countries in this region is very similar. Austria is clearly the main intermediate partner in the Hungarian–German value-added flow and has a crucial role in case of the other Visegrad countries. According to the analysis, the role of China is not highly significant, but the USA is a major intermediate partner not just for Hungary but also for other regional partners.

The automotive industry is more globalised than the average, but the effect of European counterparts is large. Austria dominates the regional networks for Hungary and other Visegrad countries, in addition to Romania, which suggests interdependence in Central and Eastern Europe.

Contrary to the vehicle industry, the value-added exported by the chemical sector ‘lives longer’ in the export of the partner countries and is thus more globalised than the carmaker firms or even the electronics industry.

This analysis can be further developed to include final goods, the true end of the value chain. However, this chapter aimed to investigate the pure part of the supply chains and covers the intermediate goods only.

7. Mapping the sequences in the regional trade of automotive industry

The number of stages in a complex production process is manifold. Some of these stages follow a strict order, while other sequences or substitutable. Unfortunately, the production and trade statistics conceal these sequences, as they cannot be determined from the aggregated data. To acquire a comprehensive picture of the production network, one also needs information concerning the dynamics. Mapping production and merchandise process is crucial, and its importance shall be introduced in three points.

7.1 Estimating the dynamics of static network

The available data of GVCs are appropriate for static analysis only. If time series were accessible, one could dynamise the networks. Unfortunately, such IO data are not at one's disposal in suitable a frequency domain. The supply chains can be described by the methods introduced in Chapter 5: one can tell how the nodes are integrated (centrality), how strong are the links (degree), and what is the direction of the edges of all nodes in any network. In the possession of the adjacency matrices, the networks can be visualised in any form; however, the process can only be estimated if the nodes have the characteristics that determines the sequence of the flows.

It must be noted that the estimation of the sequences happens retrospectively of an already existing network, in contrast to the classical analysis of graph flows. In case of the latter, the goal is usually the simplification of the path, which assumes that some nodes and edges can be neglected if a favourable path exists (favourable in terms of cost, effectiveness, speed, etc.).

9. Definition: Let $T\colon \{t_1, t_2, \dots t_i\}$ be the members of the production set in which t_i and t represent the production activity, while i is the order of the activities in the production process. Assume that the T production set can be partitioned into $T_i,\ i \in \mathbb{N}^+$ subsets, that is, $T_i \subseteq T$ and $\bigcup_{i=1}^{m} T_i = T$. In such a case, $T_i\colon \{t_i, \dots t_k\}, k \leq i$ contains all production stages of which the order of production is interchangeable, while the set $T_n, \dots T_m \subseteq T$, $n < m$ contains those activities that follow a strict order. Then let $K\colon \left(\{t_i..t_j\} \in T_i, \{t_k, \dots t_k\} \in T_j, \{t_l, \dots t_m\} \in T_k\right), (T_i, T_j, T_k) \subseteq T$ be the constraint (Maher et al., 2008).

The interpretation of K is the following: the production can be split into sequences in which the order of activities is interchangeable but the order of sequences is fixed.

Mapping the sequences of GVC positions

Most indicators introduced in Chapter 4 are fixed base indices in which the index is an aggregate that represents the value of the goods and services determined by the actual market price. Consequently, the base index in each sequence contains the partial cumulative sum of value-added of other producers. The value of this base index is inversely proportional to the order of sequence, that is, $\frac{c}{B} \propto B^{-1}$, where c is the value-added produced by the actual producer in the sequence, and B is the partial cumulative sum of the value-added created by other producers in the previous sequences: $B\colon\{T_1, T_2 \ldots T_N\}$ and $T_i = \sum_{i=1}^{n} t_i$ and $B = \sum_{j=1}^{N} T_j$ where T_i indicates the subset of production, and $t_i \in T_i$ refers the value-added created during the sequence.

It should be noted that if a producer that operates in the sequence of T_i in which the production activities are interchangeable, the value of the base index will be different according to the order of the sequence (it will be small at the beginning but large at the end, while the activity is the same). Owing to the interchangeability of the production process, one should adjust the base index by the order of sequence. The simplest method would be to choose the price of the final product as the deflator, but this value is unknown because of aggregation. This chapter introduces the methods that can reveal the sequence in multivariate time series. At the same time, the detected sequences in themselves are not likely to be suitable for adjustments, because the nature of the data shall be discrete, while the indices are continuous[59]. Thus, it is rather a supplementary data, which provide very valuable information.

A more realistic visualisation of GVC positions

All industries in GVCs are traditionally treaded in an unweighted form (Cappariello et al., 2020). There are studies in which the nodes are weighted in proportion to the value-added they produce (Li et al., 2019). Neither of the above reflects the hierarchy of the value chain, while it favourably visualises the industries with larger output. What these

[59] The data class of sequences is ordinal (or nominal) that are perfect for mapping the relations; however, no operations can be done in the vector space of continuous data.

anticipations have in common is that all of them plot the transactions as a one-layer network, in which all industries are aggregated along the countries. While some basic network indicators such as communalities (Barigozzi et al., 2011) and clusters (Sturgeon et al., 2008) can be determined from these structures, the loss of information could be high because of aggregation. Disaggregating the industries results in multilayer networks, which suggests that the role played by the networks does not correlate with output volume (Alves et al., 2019). In a study published by Alves et al. (2018), the international trade among industries was represented using a multilayer network in such a way that the domestic trade was depicted by a single graph, while international trade was visualised in a second dimension. They observed that the higher the number of layers, the higher the entropy of the system, which had its maximum value during the financial crisis in 2008–2009. They also noted that the growth of randomness in the system diminishes the stability and mangles the structure that leads to the collapse of the world trade.

The disadvantage of the aforementioned analyses is that the frequency of the data was low (annual), and therefore, the models could not depict the dynamics of the system. Unfortunately, no data with higher frequency are available for the IO values[60]; however, the periodicity of merchandise trade data is much higher (monthly), but these data are not suitable to be utilised as perfect substitutes of IO tables. Nevertheless, significantly restricted high-frequency data can be suitable for mapping dynamism. The literature offers numerous methodologies to accomplish this, but unfortunately none of them is universally appropriate.

7.2 Sequences in multiple time series[61]

There are plenty of methods that can be applied to map the phase shifts between multiple time series. One of the simplest tools is the method of cross-autocorrelation, which measures the correlation between two variables shifted in time. For time series (when the autocorrelation function for both time series is not zero), the cross-correlation is a normalised cross-covariance function with expected value in a [-1;+1] interval. Its interpretation is equivalent to the classical Pearson correlation measure. In signal analysis, it

[60] It must be noted that the compliance of higher frequency (like quarterly) data would not be impossible; however, it definitely requires high computational power. The run of the algorithms in a standard computer would take longer time.[61] This subchapter relies on the work of Peter Vakhal (Vakhal, 2017).
[61] This subchapter relies on the work of Peter Vakhal (Vakhal, 2017).

measures the phase shift between two signals of equivalent or very similar frequencies. Its formal equation is the following:

$$\rho_{XY}(\tau) = \frac{1}{\sigma_X \sigma_Y} E[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)] \tag{34}$$

where

$\rho_{XY}(\tau)$: the cross-correlation between X and Y at τ time shift;

σ_X: standard deviation of X;

σ_Y: standard deviation of Y;

$E[\cdot]$: expected value;

μ_X: expected value of X;

μ_Y: expected value of Y;

t: time;

τ: shift in time.

It is assumed the X and Y variables are multidimensional stationary, that is:

$$F_{XY}(x_{t1}, x_{t2}, \dots x_{tn}, y_{t1}, y_{t2} \dots, yt_n) = F_{XY}(x_{t1+\tau}, \dots x_{tn+\tau}, y_{t1+\tau}, \dots y_{tn+\tau})\ \forall \tau, t_n \tag{35}$$

The most likely phase shift is at the maximum value of $\rho_{XY}(\tau)$ in the investigated frequency domain. Cross-correlation is a bivariate indicator that is only applicable for measuring the association between two variables and is highly sensitive to the non-linear relationships. The main problem with the index is that it can map the permanent, deterministic relations only. Despite these issues, it can be still suitable for revealing sequences if one can detect those (preferably short) intervals in which the deterministic functional relation exists.

In other words, phase shifts in the value chains are not permanent, as because of the aggregation, more than one value chains are combined in the data. Thus, one is unable to separate the effects of the different chains. On the other hand, the technological development and the permanent and intense competition also induce changes in the sequences. The production can be shifted to another location, or it can be simply altered (e.g. the factory begins to produce a new type of the same good, while the production of the old version gradually ceases); thus, the whole value chain can significantly change in the medium run. However, if one can detect narrow segments, in which the suppliers' structure is relatively robust, the cross-correlation is a good tool to map the sequences between two countries.

Spectral analysis offers another approach in mapping the phase shifts in a time series. A concise summary of the methodology is the following (Koopmans, 1995): Let $X(t)\colon \{x_1, x_2, \dots x_t\} \in \mathbb{R}$ be a stationary time series. Based on Fourier transformation, $x(t) = \sin \lambda t$, where $\lambda \in \mathbb{R}^+$ and $t \in \mathbb{N}$. Let $\lambda T = 2\pi$, where T is the time required for a complete cycle. It is known that the frequency is $f = T^{-1}$, and thus, $\lambda = 2\pi f$, which is the angular frequency. Introduce the amplitude of $A \in \mathbb{R}^+$ and the phase of $\varphi \in \mathbb{R}, -\pi < \varphi \leq \pi$ (dimensionless scalar, which controls the sinusoidal shift). With the help of these variables, one can generate any number of *y(t)*[62] monochromatic time series in a desired length. That is, $y(t) = A \sin(\lambda t + \varphi), -\infty < t < +\infty$ for all cases. One can mix x(t) time series from many monochromatic functions and by that can generate a harmonic function:

$$x(t) = \textstyle\sum_\lambda A_\lambda \sin(\lambda t + \varphi), -\infty < t < +\infty \qquad (36)$$

which is spectral representation of *x(t)*. Taking advantage of the trigonometric identity of $\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta$:

$$x(t) = \textstyle\sum_\lambda (a_\lambda \sin \lambda t + b_\lambda \cos \lambda t), -\infty < t < +\infty \qquad (37)$$

where $a_\lambda = A_\lambda \cos \varphi_\lambda$ and $b_\lambda = A_\lambda \sin \varphi_\lambda$. To estimate the value of φ_λ phase, one can utilise the $C(\tau)$ autocovariance function:

$$C(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t + \tau) x(t) dt \qquad (38)$$

where τ is a random time-delay parameter. The spectrums can be generated by the Fourier transformation of $C(\tau)$[63]. In the possession of length waves, one can estimate amplitude

[62] In order to avoid any misunderstanding, y(t) is used here.
[63] For more details, please see Welch (1967).

A using linear regression. One of the most valuable advantages of spectral analysis is that it can map the phase shift between two time series.

7.3 Network representation of time series

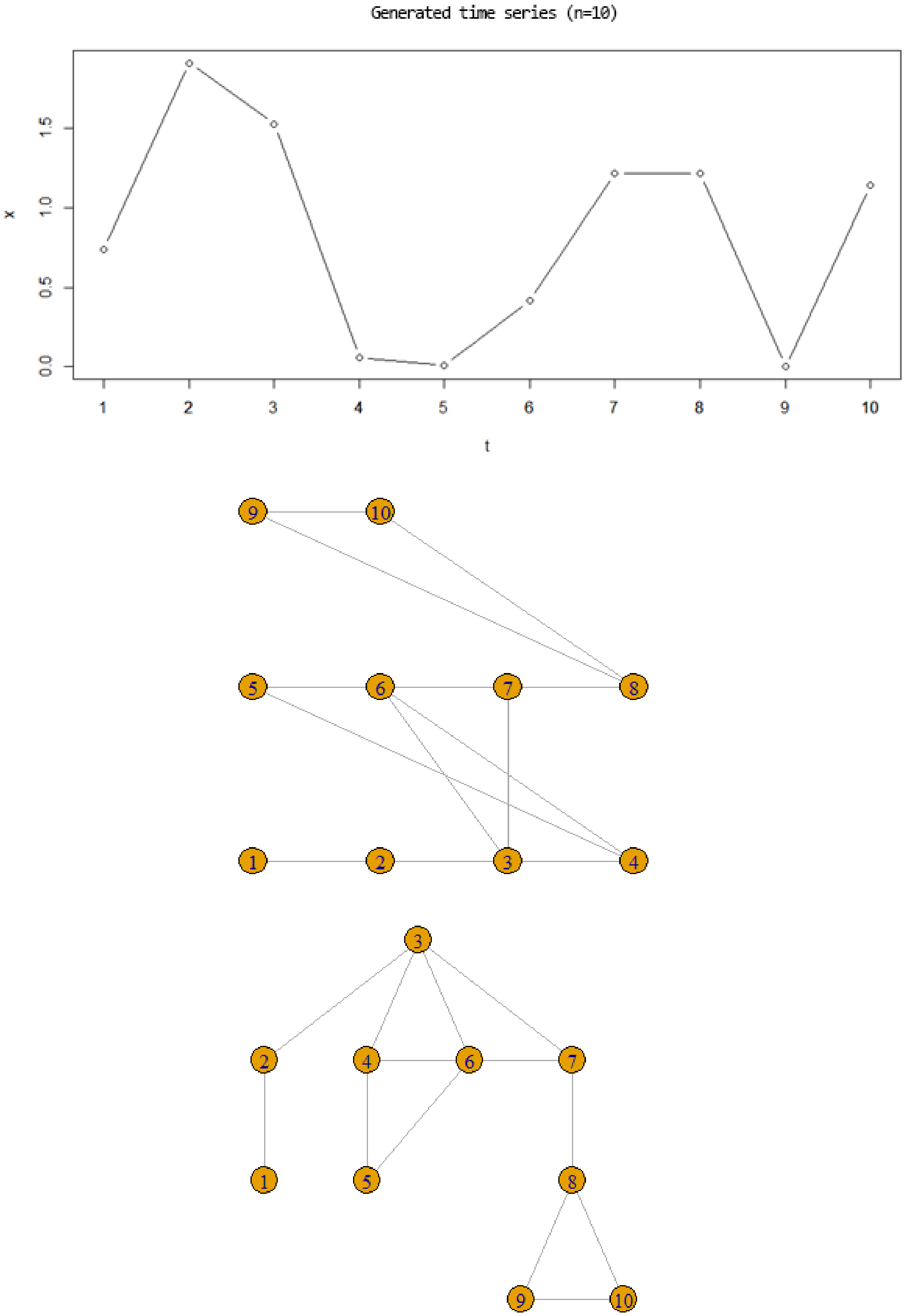
Univariate time series may contain a large amount of information. However, extracting information from multivariate time series is beyond the scope of mapping correlation, as more sophisticated methods are available like cointegration (E. Kovács, 1989), vector-autoregressive models (Lütkepohl, 2006), or Gaussian mixture models (Arellano & Bond, 1991).

The network representation of univariate time series was studied by Lacasa et al. (2008), who proposed a new approach called the visibility graphs (VG). In the $X_t\colon\{x_1, x_2 \dots x_n\}$ time series, let two $\{x_i, x_j\} \in X_t$ elements in any two $\{t_i, t_j\} \in T$ time. Let two $x_k \in X_t$ elements at the same time of $t_k \in T$, which satisfy the $t_i < t_k < t_j$ inequality. Then, x_i and x_j are visible for each other if the following inequality is true:

$$x_k < x_j + (x_i - x_j)\frac{t_j - t_k}{t_j - t_i} \tag{39}$$

If the inequality expressed in equation (39) is true, then the two elements at different times, represent two nodes that can be linked by an edge. The interpretation of the connection is that two time markers ‘see each other’, because all elements between the two markers are smaller. Figure 30 is a randomly generated time series and its graph (the latter visualised in two ways).

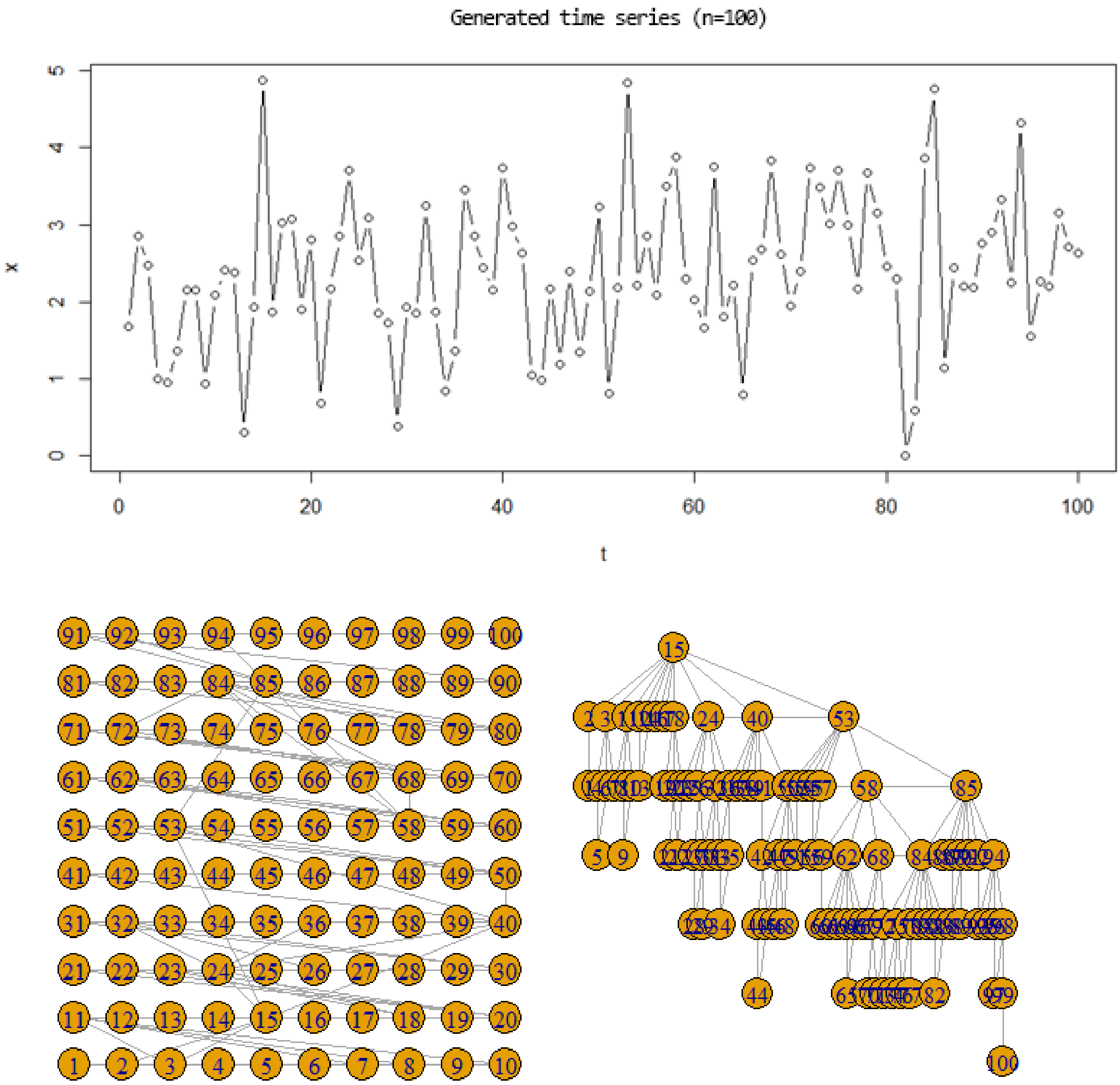
30. Figure: Visibility graph of a generated time series



Source: generated time series (R2)

The interpretation of Figure 30 is the following: Although the second element is the largest one, it is still the third that constitutes the first node in the network, because the second element 'sees' only the first and the third datapoints. All the other values are 'hidden' for the second element, as they are covered by the third one. The third node 'sees' the second, fourth, sixth, and seventh datapoint. The fifth cannot be 'seen' from the third because it is covered by the fourth node. Figure 30 only has 10 elements, which is quite short; however, one can get the substance of the methodology. The power of this algorithm can be seen in case of longer time series; however, the interpretation is going to be more difficult. Figure 31 depicts the network of a random time series of 100 elements.

31. Figure: Visibility graph od a generted time series



Source: generated time series (R3)

The visibility graph of the time series clearly shows the structure of the network. However, it is far too complex to interpret the dynamism[64]. One must put the network representation into a higher dimension to map the sequential shift between two or more time series.

Network representation of multivariate time series

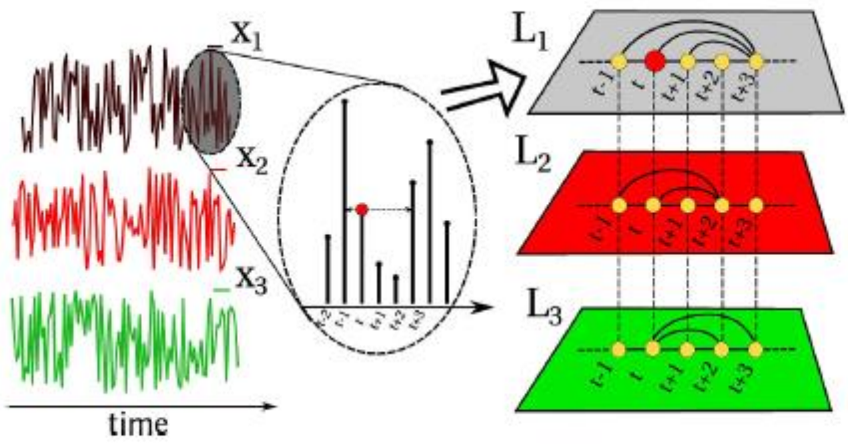
The visibility graph defined over the univariate time series can be readily extended to the multivariate case (Luque et al., 2009). Let $\{x_i\}_{i=1\ldots N}\ x_i \in \mathbb{R}$ be N time series. However, as there are no constraints for the moments of the time series, it is worth standardising the values into a [0;∞] scale. The standardisation does not change the distribution of the data. In the network representation of two time markers in two different $x_i \in X_k$ and $x_j \in X_l, k \neq l$ univariate time-series, two nodes can be linked only if the following geometrical inequality is satisfied:

$$x_i, x_j > x_n \ \forall\, n: i < n < j \tag{40}$$

One can generate the visibility graphs for every M univariate time series, which can be horizontally connected if the inequality described by equation (40) is satisfied. Thus, a multidimensional, so-called horizontal visibility graph (HVG) can be created. Let $A = \{A^1, A^2 \ldots A^M\}$ the set of adjacency matrices for the *1, 2, … M* univariate time series, where $a_{ij} = 1$ if and only if the nodes of *i* and *j* are linked with a unidirectional edge. Figure 32 represents the $M = 3$ case.

[64] All univariate time series contain its own sequence, which is determined by $t_i \in T$, where *i* is the index of the variable.

32. Figure: A representative case of the horizontal connection of visibility graphs



Source: Lacasa et al. (2015)

On the basis of the HVG diagrams, one can draw meaningful conclusions concerning the dynamics of the multivariate time series. Let x_i^1 be a random value from a three-dimensional $X: \{x_t^1, x_t^2, x_t^3\}$ time series. The algorithm examines the environment of $[x_{i-\varepsilon}^2, x_{i+\varepsilon}^2]$ and $[x_{i-\varepsilon}^3, x_{i+\varepsilon}^3]$, where $\varepsilon \in \mathbb{N}$ represents the time shift and creates the $A^{|X|t\times|X|t}$ adjacency matrix, where $|X|$ denotes the number of elements of set *X*. In this example, the $A^{3t\times 3t}$ adjacency matrix is a symmetrical block matrix (for simplicity, let us denote it by $\boldsymbol{A}$):

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = A^T \tag{41}$$

In the lower triangular matrix[65], let $s_i = \sum_{j=1}^{t} a_{ij}\,, i \neq j$ be the sum of the *i*th row in the block of $A_{ij}^{t\times t}$, that is, the number of those connections[66] through which the particular node is linked to nodes in other time series. Let us define variable d_{ij}, which depicts the time relations of linked nodes[67]:

$$d_{ij} = \begin{cases} i-j & ha\ a_{ij} = 1 \\ 0 & ha\ a_{ij} = 0 \end{cases} \tag{42}$$

If the value of d_{ij} is negative, the node in row of the adjacency matrix is earlier in time compared to the node in the column. The value is positive if the row element happened

[65] Owing to the symmetrical adjacency matrix.

[66] Because the network is undirected and $a_{ij} = [0,1]$ the sum of rows and columns is equal.

[67] As the network is undirected, it does not matter if one utilises the sum of the columns or the rows. However, in compliance with the algebraic conventions, the rows are used as the base of the analysis.

later than the column element, and it is zero if the row node has the same time marker as the column or there is no edge between the two. The following equation defines the average time shift between the time series:

$$d^{ij} = (\sum_{i=1}^{t} s_i)^{-1} \sum_{j=1}^{t} d_{ij}\,, \forall d: i \neq j, s_i > 0 \tag{43}$$

As the length of the average shifts bears less importance, one can simplify the results by utilising indicator functions:

$$I\left(d^{ij} > 0\right) = 1 \tag{44}$$

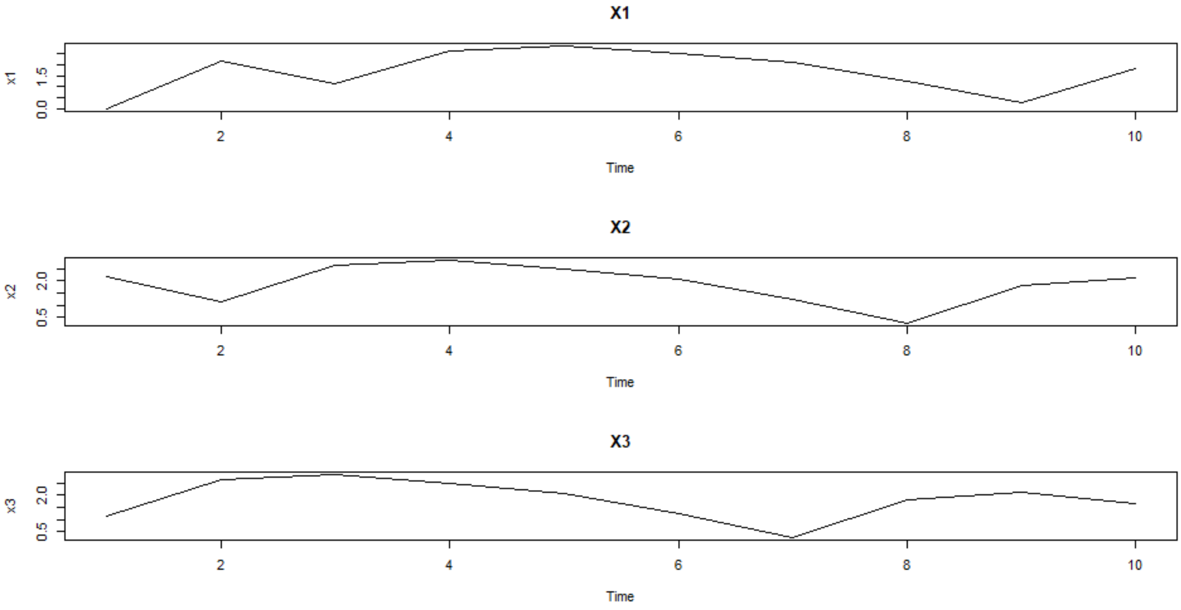
$$I\left(d^{ij} < 0\right) = -1 \tag{45}$$

$$I\left(d^{ij} = 0\right) = 0 \tag{46}$$

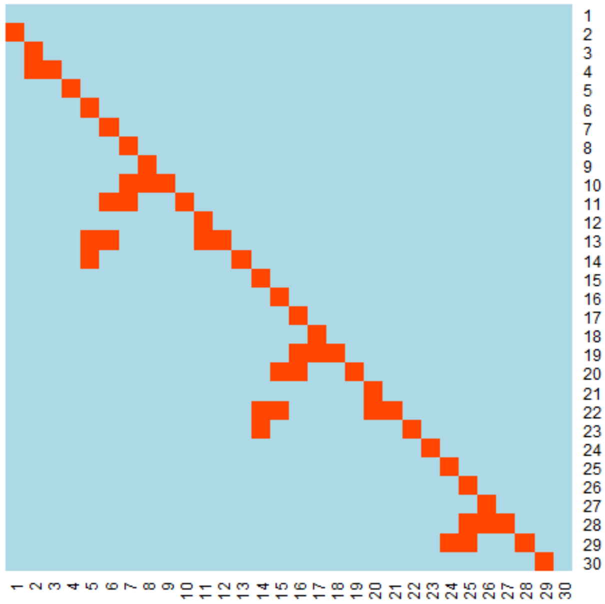
On the basis of the indicator function, one can map the sequential shifts between the time series.

The following practical example presents the methodology for three time series. Each one contains 10 elements only, and they are shifted by one time frame. Their sequence is the following: *x3*, *x2*, and *x1*.

33. Figure: Three generated time series shifted in time, their adjacency matrix and graph[68]

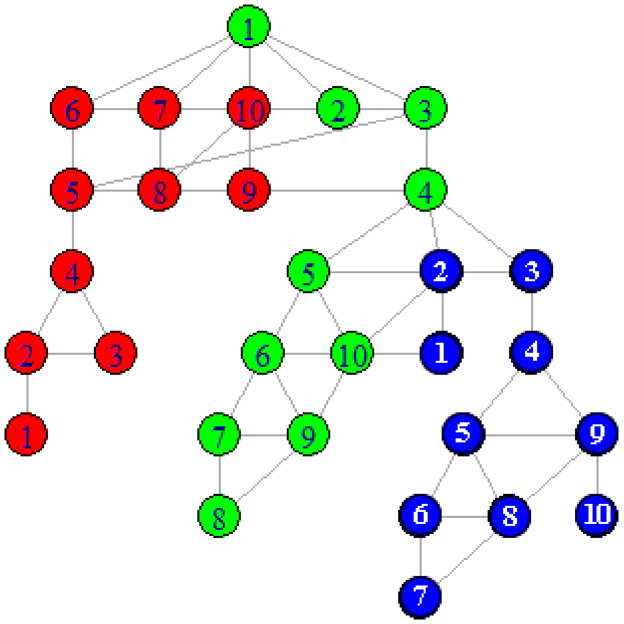


The heatmap of the *30 x 30* adjacency matrix is illustrated in Figure 33b:



The network based on the adjacency matrix is depicted in Figure 33c (x1 is represented by red, x2 by green, and x3 by blue):

[68] The figures are created by code R4.



Source: own edition

One can conclude the potential order from the heatmap and the horizontal graph: the blue time series connects only to the green, while the green is linked only to the red one; in addition, there is no connection between the blue and the red time series. This indicates the following sequence: blue (x3), green (x2), and red (x1). One can calculate $d^{i,j}$ as the following:

$$d^{x3,x2} = (2-4) + (2-10) + (1-10) + (3-4) = -20 \tag{47}$$

$$d^{x2,x1} = (1-6) + (1-7) + (1-10) + (2-10) + (3-5) + (4-9) = -35 \tag{48}$$

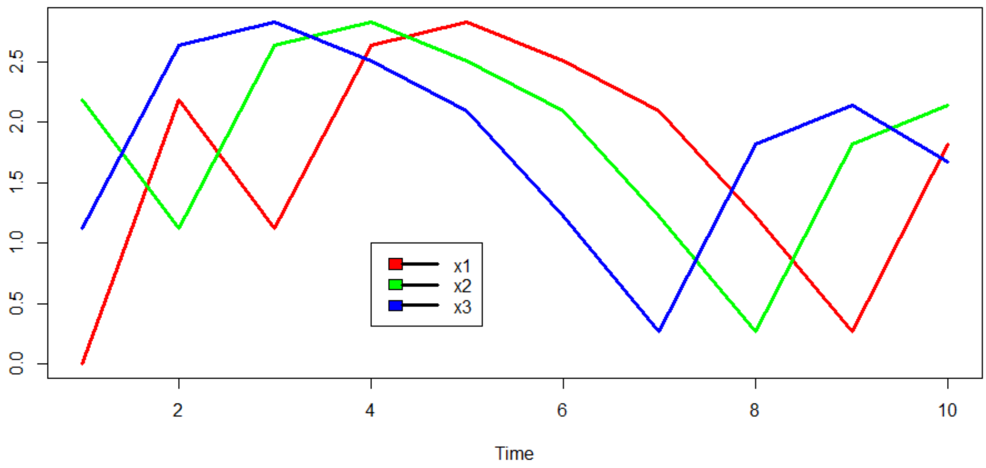
$$I(d^{x3,x2}) = -1 \tag{49}$$

$$I(d^{x2,x1}) = -1 \tag{50}$$

$$I(d^{x3,x1}) = 0 \tag{51}$$

According to the triangular inequality theorem, the sequence is truly $x_3 \rightarrow x_2 \rightarrow x_1$. This can be reinforced by the plot of the three time series presented in Figure 34.

34. Figure: Plot of time series x1, x2, x3



Source: own edition

7.4 Determining the order of sequences for high-dimensional time series

The previous example of the three time series well reflected the theory behind the methodology. One could easily determine the order of sequences in case of three time series; however, it is much harder to do the same in a higher dimensional space. PageRank method proposes a possible solution to the problem.

This method is an eigenvector-based algorithm, which takes the relative importance of the nodes into consideration. Generally, if a node is linked to another vertex that has a central role, then this link is 'worth more' than a link to a node that is less important in the network. This can be measured through centrality (Newman, 2018):

$$c_i = \alpha \sum_j A_{ij} \frac{c_j}{k_j^{ki}} + \beta \, , i \neq j \tag{52}$$

where

c_i: is the centrality of node I;

c_j: is the centrality of node j;

α, β: are constants (see later);

A: is the adjacency matrix[69];

k_j^{ki}: is the number of outgoing edges from node j (where $E(j) = 0$, $k_j^{ki} = 1$).

[69] The indication of A_{ij} ensures that the sum operation will cover the neighbours of node i only.

In matrix form, $c = \alpha A D^{-1} x + \beta 1$, where $\boldsymbol{D}$ is a diagonal matrix, where $D_{ii} = \max(k_i^{ki}, 1)$ and $\boldsymbol{1}$ is the unit vector. After rearranging, $x = \beta(I - \alpha A D^{-1})^{-1} 1$, where $\boldsymbol{I}$ is the unit matrix. As the role of β is marginal in terms of centrality, it is conventionally set to $\beta=1$. This defines the so-called PageRank centrality:

$$c = (I - \alpha A D^{-1})^{-1} 1 \tag{53}$$

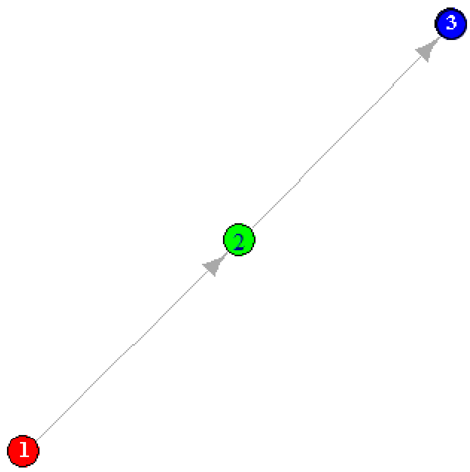
Equation 53 includes the parameter α, which can be set freely between the interval of $0 < \alpha < 1$. Generally, the value of 0.85 is advised (Newman, 2018); however, this has no theoretical basis[70].

Returning to the previous example, the PageRank method cannot be applied directly, as it is only applicable to networks wherein the nodes represent the time series. Thus, let $a_{ij} \in A$ be an adjacency matrix in which $a_{ij} = 1$ if there is a link between x_i and x_j; otherwise, it is $a_{ij} = 0$. In the example of the three time series the 3 x 3 adjacency matrix and network based on that is the following:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{54}$$

[70] The value of 0.85 is a reiteration of the original study (Brin & Page, 1998) by Sergey Brin and Larry Page, the founders of Google. Their article is a crucial fundament of the search engine.

35. Figure: The modified graph of the generated time series



Source: generated time series (R5)

The PageRank centrality vector of the network above is[71]:

$$c = \begin{bmatrix} 0.1844168 \\ 0.3411710 \\ 0.4744122 \end{bmatrix} \tag{55}$$

As the first node in the network has no inward edge, its centrality value will be the minimum. However, it is connected to the second sequence, which is connected to the third one. The second node is valuable for the first, and the third is valuable for the second. The largest centrality belongs to the node that has the most valuable inward edges (Fortunato et al., 2008). That is, in a serially linked network, the last element will have the highest PageRank value.

7.5 Detecting sequences in the automotive industry in the CEE region

Companies operating in the machinery and the vehicle industry are clearly the largest employers in the Central and Eastern European region. The share of employment in the business sector is estimated to be 4%–6%[72], while the share in the output is between 5%–15%. Despite these figures, the multiplier effect of these industries is very small, especially in Hungary (Koppány, 2017). Nevertheless, the automotive industry is the only one that operates as a pure value chain (Vakhal, 2018a):

---

[71] In *R* environment, one can calculate the PageRank centrality with the help of the page.rank() function of the *igraph* package (Kolaczyk & Csárdi, 2014).
[72] Based on 2019 Eurostat data (*nama_10_a64_e*).

- All goods produced by the vehicle industry have demand in all countries of the continent. The market of other producers is much narrower, and they are rather regional (like food industry, services). The pandemic in 2020 necessitated the attention that the production is markedly affected by the demand on the other side of the world (Stubnya, 2020).
- The supplier network is very deep and wide – there are numerous actors at the top of the value chains that are competing each other, and the weight in the global economy is considerable. This cannot be stated in case of any other industries because of the type of goods they produce (perishable, large differences in operation, etc.).

Because of the aforementioned reasons, the study of the production sequences in case of the automotive industry is crucial. The analysis encompasses the market of intermediate goods only because the final product (a functioning vehicle) is part of the value chain for a very short period[73]. A vehicle consists of nearly 3,000 parts, and all these parts can be merchandised. As the descriptions of these parts are very detailed, the analysis takes advantage of the SITC nomenclature (Amighini, 2012):

[73] It must be noted that the vehicles can be final goods, investment goods, or even intermediate goods at the same time if the user is the corporate of government sector. Only the vehicles purchased by households can be considered as final goods; however, the legal status of the owner is not registered in the statistics, and thus, the functioning (final) motor vehicles are not part of the analysis.

18. Table: Intermediate and final goods used in automotive industry

Product code (based on SITC)	Short name
625	Rubber tyres
699.15	Other mountings, fittings and similar articles suitable for motor vehicles
713	Internal combustion piston engines and Parts thereof
762.12	Reception apparatus for radio-broadcasting (for vehicles)
778.12	Electric batteries
778.23	Sealed-beam lamp units
778.31	Electrical ignition or starting equipment
778.33	Parts of the equipment of heading 778.31
778.34	Electrical lighting or signalling equipment, windscreen wipers, defrosters and demisters, of a kind used for cycles or motor vehicles.
784.21	Bodies for the vehicles of group 781.
784.25	Bodies for commercial vehicles
784.31	Bumpers
784.32	Other parts and accessories of bodies (including cabs)
784.33	Brakes and servo-brakes
784.34	Gearboxes
784.35	Drive-axles with differential
784.39	Other automotive components
821.12	Seats of a kind used for motor vehicles

Source: (Amighini, 2012)

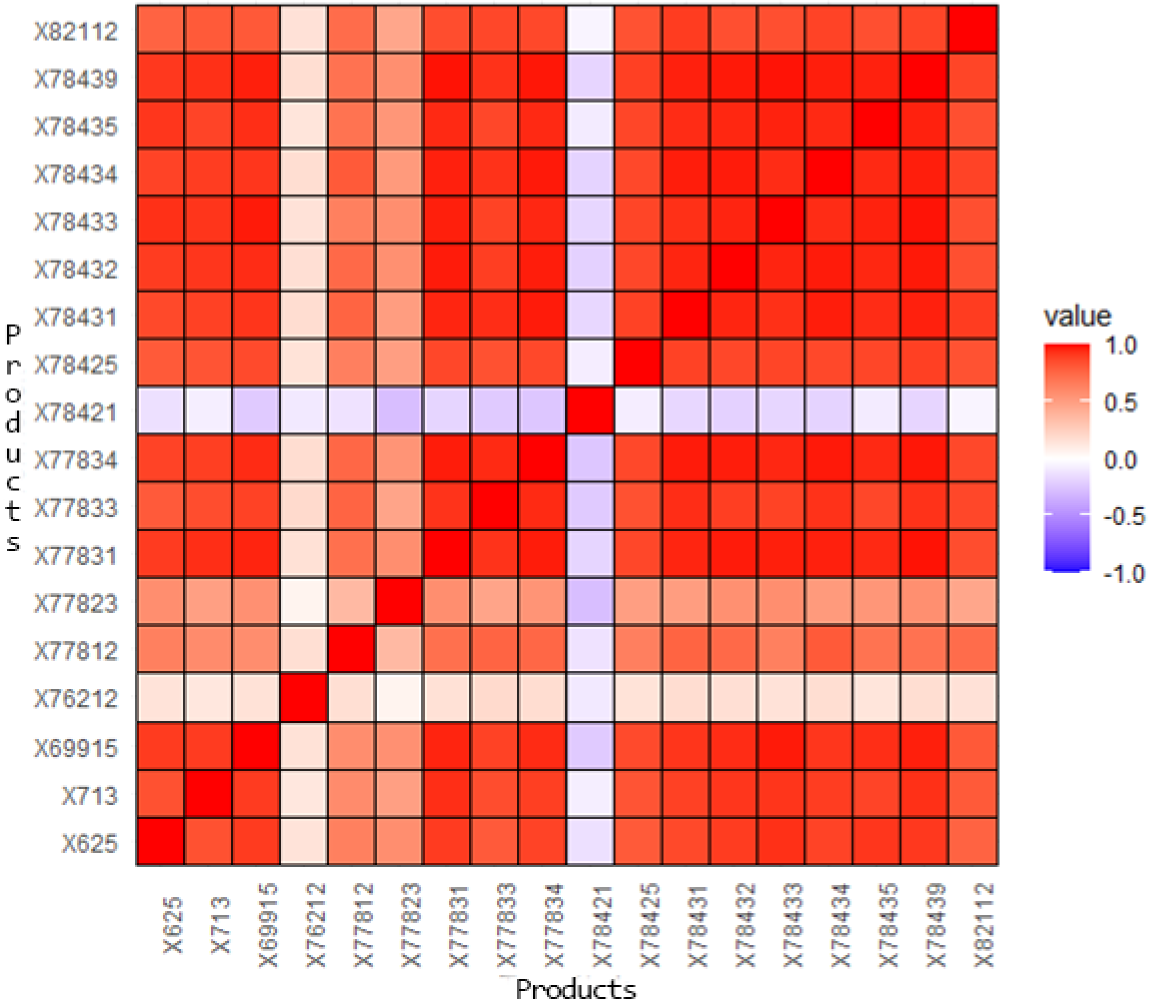
The source of the trade statistics of the goods above is the Comext database of Eurostat. All the time series are from 2004 until October of 2020 (202 observations). The database contains only a few missing data, which considers random rather than structural, thus they were imputed by the process of $x_t = E(x_{t-1}, x_{t+1})$. The data cover five countries: the Czech Republic, Poland, Hungary, Romania, and Slovakia, because these countries have considerable automotive industries. Only the import of merchandise trade was considered because the output of these sectors in the investigated countries is mainly final products, such that the buyers are generally not the companies[74]. In case of imports, the suppliers can only be firms. The number of observations in the time series of 202 elements of five countries and 18 products is 18,180. The unit of the values is current price denominated in euros.

[74] Usually, the consumers cannot purchase directly from the carmaker, only via a retailer. At the same time, the retailers do not make any transformation in the vehicles.

Adjusting the current prices is not required here, because in case of import, the seller's price index is the deflator. Assuming that all importers face the same price fluctuation in the world market, the deflator would be the same for all countries, which would not alter the data.

The correlation between the trade volume of the aggregated product groups is very high. This is reinforced by Figure 36. Only the product number 762.12 (radios) cannot be associated with other parts, while the car bodies (product 784.21) correlate negatively.

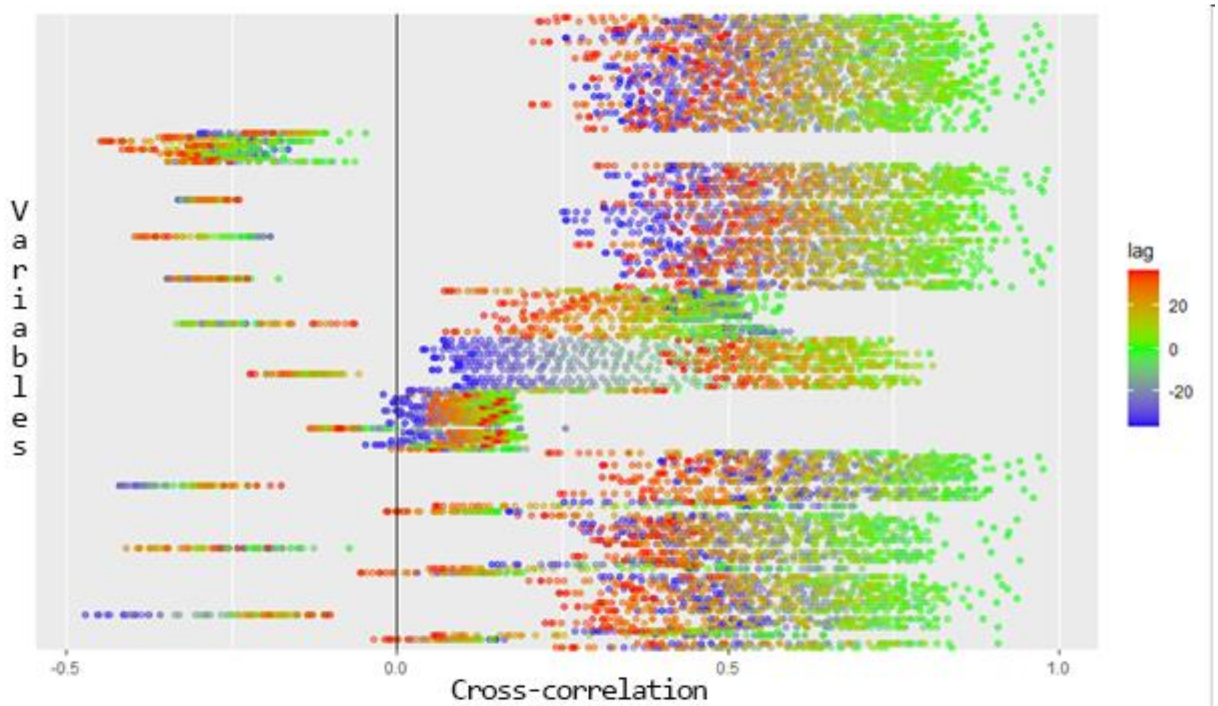
36. Figure: Correlation heatmap of aggregated product groups



Source: Eurostat (R6)

The correlation in the same frequency domain suggests that there are no sequences in the time series; however, on the basis of the cross-correlations, one can observe a forward shift in the data, especially if the dimension of countries appears in the analysis. Figure 37 shows the two dimensions of cross-correlations (time shift and correlation) for every country–product combination (total 153 combinations). The maximum shift was 36 months (three years).

37. Figure: Values of cross-correlations by time shifts according to all country and good dimension



Source: Eurostat (R7)

As per Figure 37, the largest correlations can be measured at 0 or 1–2 time shifts. However, for certain product combinations, high correlations can be observed in case of longer shifts even in positive (red) or negative (blue) direction. Some products must be stressed: The import of sealed-beam lamp units (product 778.23) advances the import of other parts in the group of 784. Car body parts should also be stressed, because the largest correlation to other goods is around the 34 months. For further details, see Table 19:

19. Table: Highest cross-correlations in absolute terms with imports of motor vehicle bodies (784.21)

Lag (months)[75]	Cross-correlation	Lagged variable (product code)	Base variable (product code)	Interpretation in point of view of product group 784.21[76]
34	0,47	784.21	625	pro-cyclical, follower
-27	0,41	784.21	713	pro-cyclical, leader
29	0,42	784.21	699.15	pro-cyclical, follower
20	-0,25	784.21	762.12	anti-cyclical, follower
-36	0,22	784.21	778.12	pro-cyclical, leader
-5	0,33	784.21	778.23	pro-cyclical, leader
34	0,35	784.21	778.31	pro-cyclical, follower
-27	0,40	784.21	778.33	pro-cyclical, leader
27	0,33	784.21	778.34	pro-cyclical, follower
14	-0,35	784.21	784.25	anti-cyclical, follower
29	-0,32	784.21	784.31	anti-cyclical, follower
27	-0,36	784.21	784.32	anti-cyclical, follower
34	-0,42	784.21	784.33	anti-cyclical, follower
34	-0,32	784.21	784.34	anti-cyclical, follower
27	-0,45	784.21	784.35	anti-cyclical, follower
27	-0,36	784.21	784.39	anti-cyclical, follower
-35	-0,30	784.21	821.12	anti-cyclical, follower

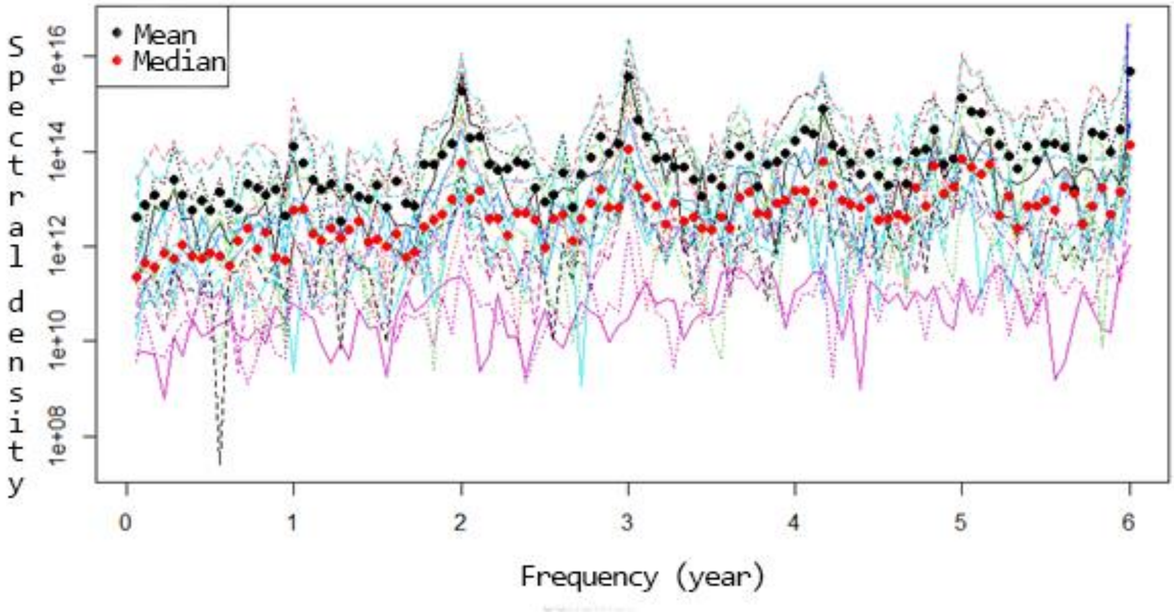
Source: Eurostat (R8)

Although the lengths of the shifts are long, this suggests the existence of long-term trends, which is presumable in relation with the business cycle, inventory management, and technological changes (Chikán et al., 2018).

The spectral analysis of the differentiated[77] time-series detected a 12-month long cycle in the data, in which the spectrums vary by three years; however, the standard deviation is fairly large (see Figure 38). The one and three years long cycles correspond to the planned and permanent technological cycles of the carmakers and reiterates the existence of cross-correlations around 36 months. On the basis of the data, one cannot conclude that the producers have and inventory enough for one to three years, but it depicts long-term trends.

[75] The maximum lag was 36 months.
[76] p is the interpretation of cross-correlation, $p = cor(x_{t+i}, y)$where y is a base variable and x is a variable delayed by time *i*. For $p = cor(y^{784.21}, x^{625}_{t+34}) = 0{,}47$ example, the value in the first row should be read so that a 34-month offset of product 784.21 results in the highest absolute cross-correlation with product 625, which is 0.47. This suggests that imports of product 781.21 are moderately well correlated with product 625 imported 34 months later, the value is positive, which indicates pro-cyclicality.
[77] To ensure stationarity.

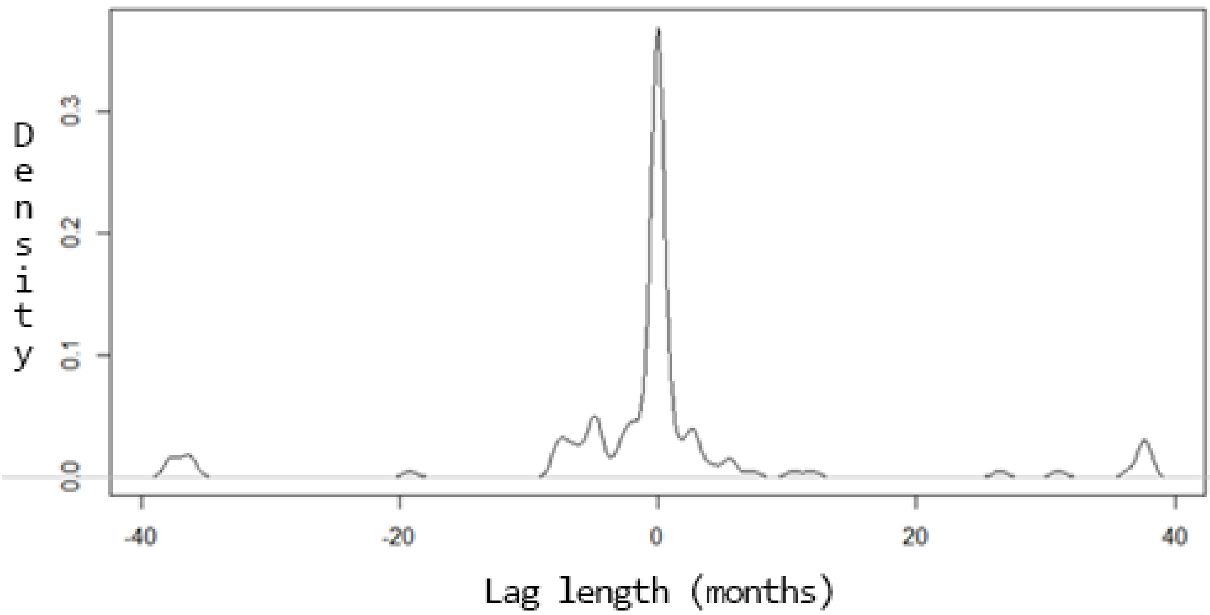
38. Figure: Spectral analysis of the designated automotive import time series



Source: Eurostat (R9)

One can utilise spectral analysis to reveal the phase shifts when two cycles of the same length are analysed. As the number of combinations is very large, it is worth narrowing the frequency band to a domain in which the spectral density is the largest, that is, $\max\left(s_{f_t}^{(i)} + s_{f_t}^{(j)}\right), (i \neq j)$, where s_i and s_j are the spectral densities of two time series of the same f_t frequency domain. At this point of the analysis, the length of the cycles in the domain is marginal. Phase shifts reveal the possible sequences in the import. Figure 39 presents the phase shifts in the same frequency domain:

39. Figure: Density of phase shifts in the same frequency range



Source: Eurostat (R10)

As evidenced in Figure 39, the average value of phase shifts is 0, such that there are no phase shifts between most of the cycles. However, for a few cases, there is a considerable amount of forward and backward shifts. Cross-correlations called the attention to car bodies (product 781.24), and thus, it is worth giving an insight into this car part.

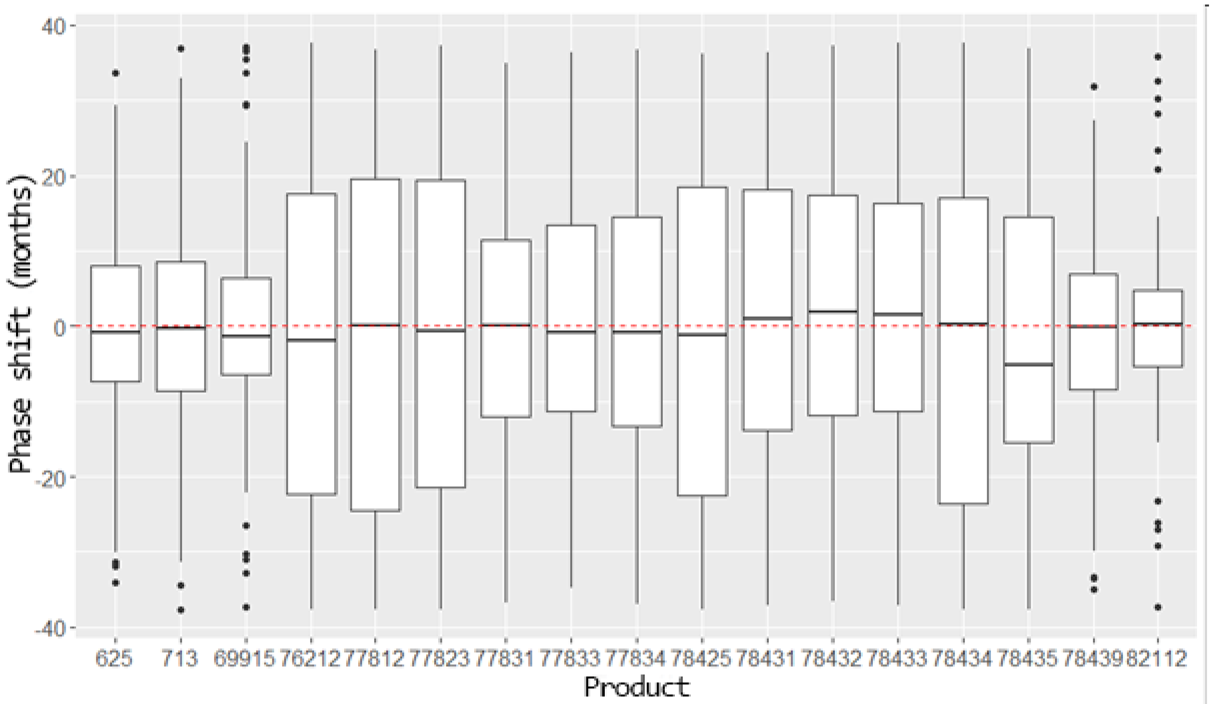
Spectral analysis on the full spectrum (i.e. not just on those domains in which the spectral density is large) showed that there are few months shifts between the import of car bodies and the other parts. The total length of the phase shifts is presented in Figure 40. Though the average shift is still zero, the large variance suggests both forward and backward fluctuations. However, if the phase shifts are analysed only in the frequency domain with the largest spectral density (i.e. one-fourth month or 0.25), it is found that import of five products advances the import of car bodies. This is presented in Table 20:

20. Table: Phase shift against product group 784.21 (vehicles bodies) in the frequency domain of 0.25 in the regional import of the automotive industry

Product group	Name	Mean phase shift (months)
762.12	Reception apparatus for radio-broadcasting (for vehicles)	-2,84
625	Rubber tyres	-0,25
778.23	Sealed-beam lamp units	-0,24
784.25	Bodies for commercial vehicles	-0,21
821.12	Seats of a kind used for motor vehicles	-0,06
784.21	Bodies for the vehicles of group 781.	0,00
778.12	Electric batteries	0,13
784.32	Other parts and accessories of bodies (including cabs)	0,16
713	Internal combustion piston engines and parts thereof	0,17
784.31	Bumpers	0,17
784.33	Brakes and servo-brakes	0,19
784.39	Other automotive components	0,21
699.15	Other mountings, fittings and similar articles suitable for motor vehicles	0,27
784.35	Drive-axles with differential	0,30
778.33	Parts of the equipment of heading 778.31	0,32
778.31	Electrical ignition or starting equipment	0,38
784.34	Gearboxes	0,40
778.34	Electrical lighting or signalling equipment, windscreen wipers, defrosters and demisters, of a kind used for cycles	0,40

Source: Eurostat (R11)

40. Figure: Phase shift against product group 784.21 (vehicles bodies) in the whole frequency domain in the regional import of the automotive industry

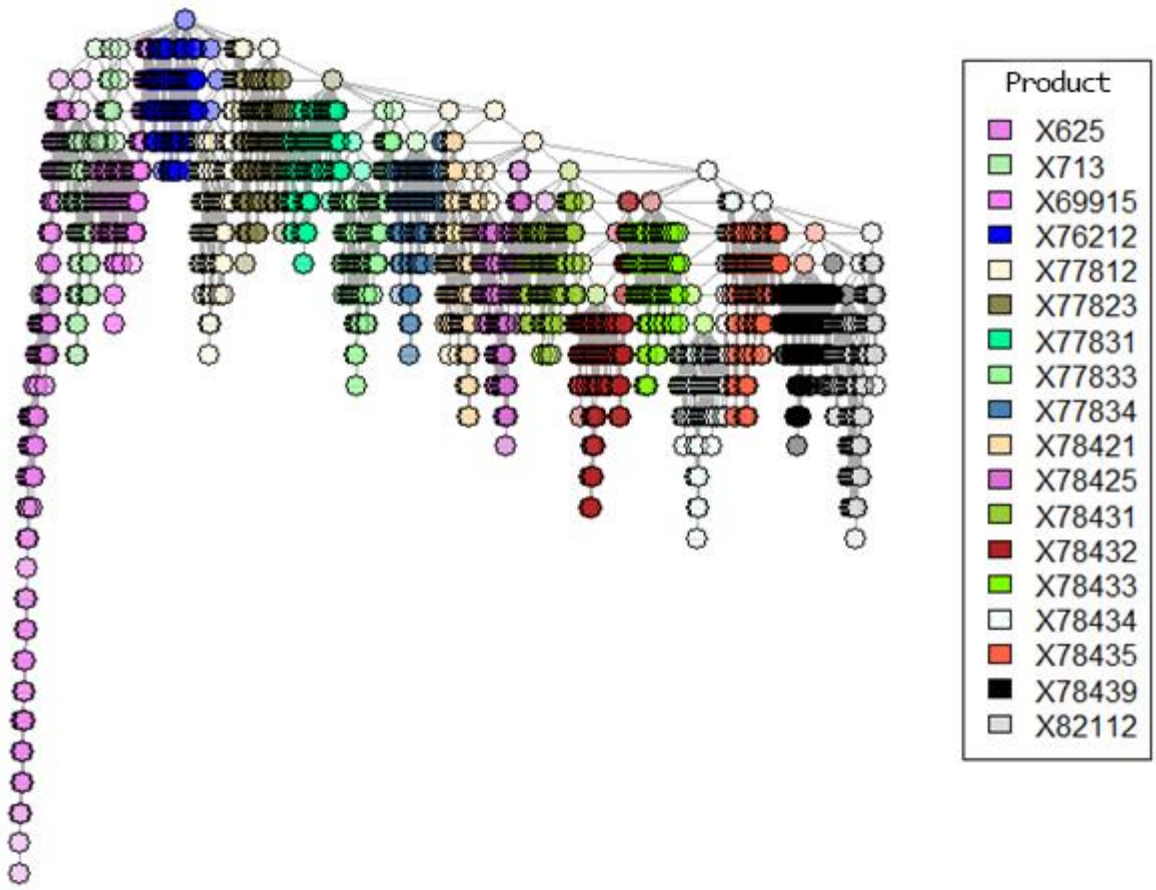


Source: Eurostat (R12)

Recent studies in network theory have proposed a new approach which was introduced earlier in this chapter. The following subchapter utilises this method and includes one additional dimension – the countries.

The analysis relying on network science utilises all observations, while spectral analysis uses only two variables at the same time. Thus, the results of graph theory and spectral analysis cannot be directly compared with each other. The analysis can be made in three ways: graph containing the monthly data, graph with the aggregated product groups, the combination of the two. All these options have a large number of nodes in the network (more than 3,600) in a complex system, which is fairly hard to visualise and analyse. Sequences of the same rank are generally plotted next to each other, and thus, in a tree-structure, the nodes will be so small, such that they are almost impossible to print on a paper.

41. Figure: The HVG graph of the regional automotive import between 2004 and 2020

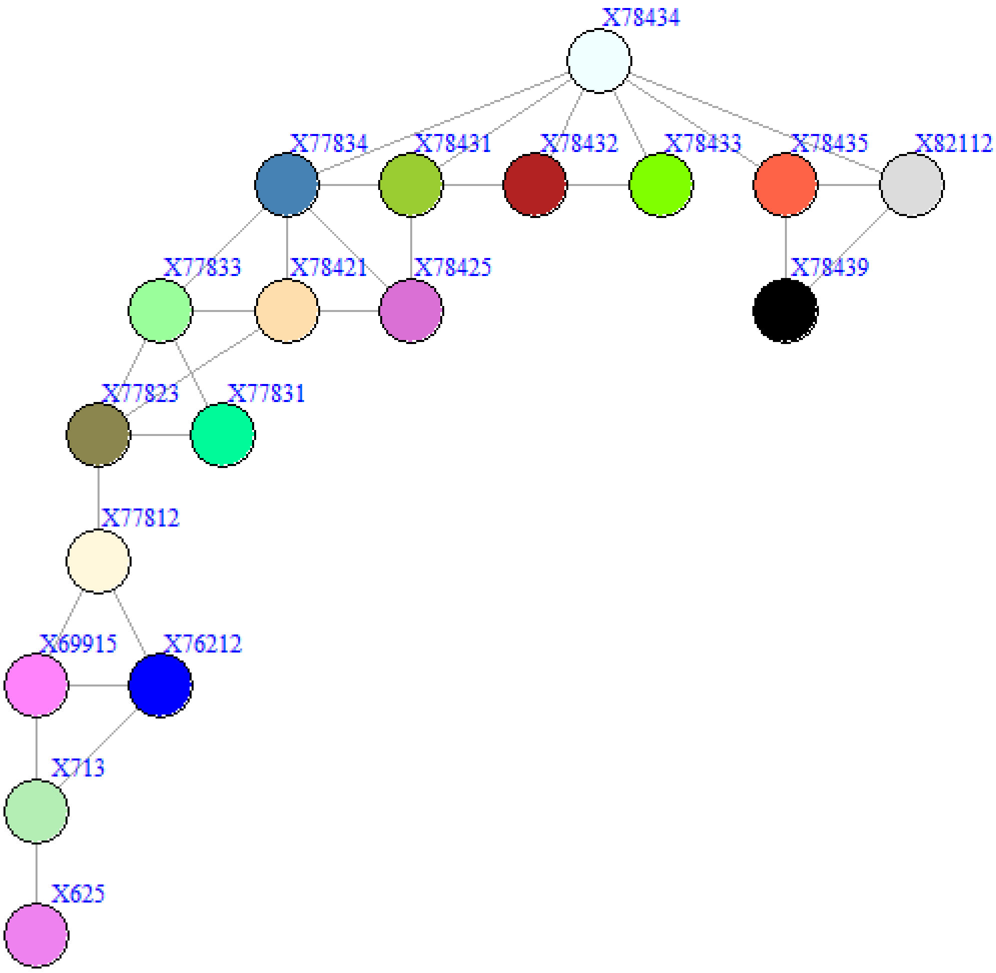


Source: Eurostat (R13)

Figure 41 presents the network of car part imports aggregated by products. All vertices indicate a month between 2004 and 2020. The topology of the network is based on the Reingold–Tilford algorithm, which is adequate to visualise deep graphs (Wetherell & Shannon, 1979). Stages on the graph reflect the role of the nodes on that level, that is, how far they are from the 'peak' (the lowest point) of the graph. In the first stage, those nodes can be found that constitute the base of the network; in the second, the vertices connect directly to the previous group. The highest point in the graph is the root of the tree (or the top of the network) and it generally depicts the node with the highest degree. If there are nodes with the same highest number of degree, all of them constitute the root, as it can be seen in Figure 41.

The structure depicted in Figure 42 is easier to interpret. The network is aggregated by products, as one can also apply the PageRank algorithm.

42. Figure: The aggregated HVG graph of the regional automotive import between 2004 and 2020



Source: own calculation (R14a)

21. Table: The PageRank scores of the adjusted HVG graph of regional automotive import

ID	Product	Name	PageRank score
1	713	Internal combustion piston engines and parts thereof	0,099
2	625	Rubber tyres	0,097
3	778.23	Sealed-beam lamp units	0,095
4	778.12	Electric batteries	0,094
5	778.33	Parts of the equipment of heading 778.31	0,080
6	699.15	Other mountings, fittings and similar articles suitable for motor vehicles	0,075
7	778.34	Electrical lighting or signalling equipment, windscreen wipers, defrosters and demisters, of a kind used for cycles or motor vehicles.	0,070
8	784.31	Bumpers	0,057
9	762.12	Reception apparatus for radio-broadcasting (for vehicles)	0,053
10	778.31	Electrical ignition or starting equipment	0,047
11	784.34	Gearboxes	0,043
12	784.32	Other parts and accessories of bodies (including cabs)	0,041
13	784.25	Bodies for commercial vehicles	0,037
14	784.35	Drive-axles with differential	0,031
15	784.21	Bodies for the vehicles of group 781.	0,029
16	784.33	Brakes and servo-brakes	0,022
17	784.39	Other automotive components	0,017
18	821.12	Seats of a kind used for motor vehicles	0,013

Source: own calculation (R13)

As one can observe in Figure 42, the time-series of the products can be well segmented. Goods in the same row reflect the same hierarchical order, while the edges indicate the links in the import of goods. Vertices on different stages refer to sequential connections. The lower is a node positioned, the longer is the path to the node with the highest centrality value.

Car parts are well separated in the figure. Electrical parts (778) and car bodies (784) clearly constitute two distinct groups. In terms of dynamics, other car parts (784.39) and car seats (821.12) are also in two separated groups. The graph depicts well that the import of tyres (625) is far from the most central node, just as the engines (713) and the mountings (699.15) relating to it. These parts are functionally different from the electrical parts, but they connect together to car body (784.21 and 784.25). At this point, the graph connects other car parts and breaks (784.3). Other miscellaneous car parts (784.39) and seats

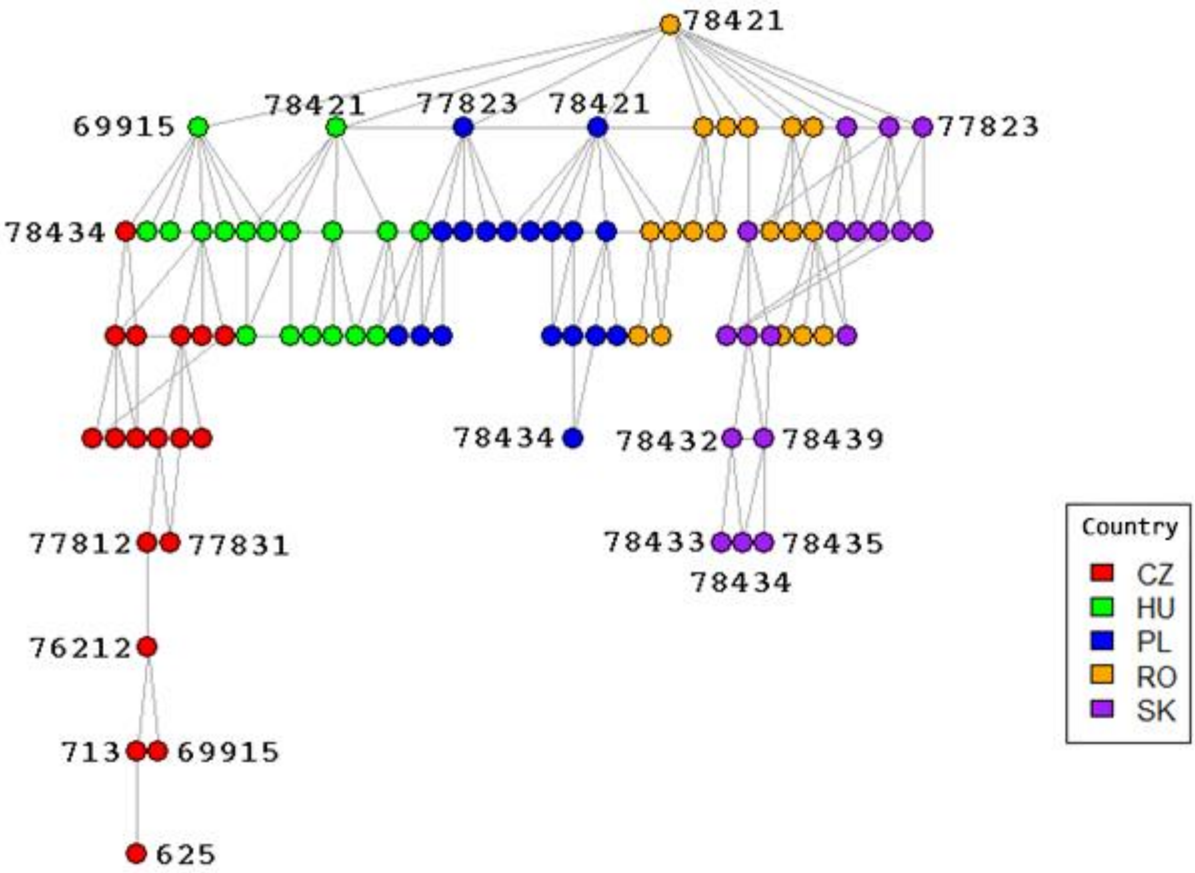
(821.12) constitute a separate group together with drive axles (784.35). At the end, gearboxes (784.34) joins the network.

The most sequential differences are between electronic parts (778) and car bodies (784.3). Companies producing these goods can specialise without assembling the final vehicle, that is, the import of the whole car is not needed to produce these items. Thus, the countries specialising in engines or tyres may have a higher domestic value added in the ratio of exports as compared to those who join the network a bit later with the electrical parts or other car body parts, as presumably they need to import both the engines and the tyres which later increase value of the export but not the domestic value-added.

If the data are aggregated by countries (see Figure 43) the HVG shows that the aggregated regional import mainly resembles the Czech data[78]. The Hungarian and Czech time series connect each other at several points; moreover, the Czech import connects to the Hungarian one only. At the same time, the depth of the Czech import is much deeper, while the Hungarian has only three levels, and it is shortest in the region. This suggests that the Hungarian car part import has no 'history'; most imports happen at the same time and the producers only assemble the product. This may also refer to the use of the just-in-time system; however, this is less likely if it is not applied in the neighbouring countries with the same profile.

[78] Differences in import volume is adjusted by standardisation.

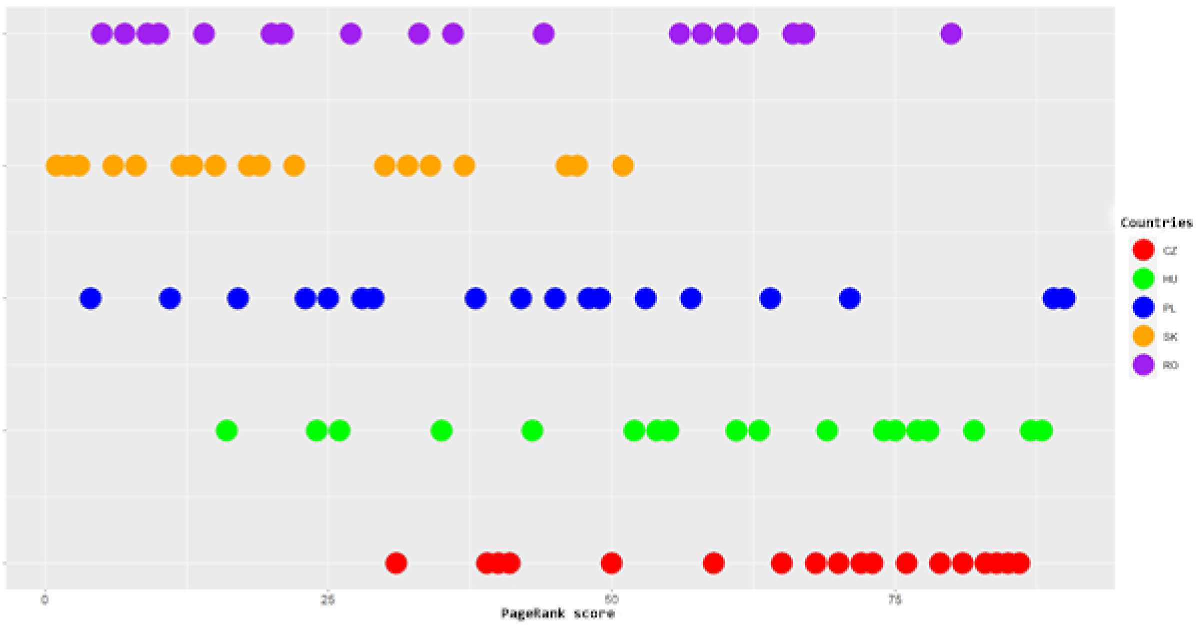
43. Figure: The HVG graph of the regional automotive industry by products and countries



Source: own calculation (R14b)

The PageRank method based on the data visualised in Figure 43 puts Slovakia to the bottom. Out of 90 identified sequences, half of the first 20 belong to Slovakia. There are 10 Czech and seven Hungarian products in the last sequence. Figure 44 illustrates that the Slovakian import is at the beginning of the sequences, while the Czech is at the end. The Hungarian import varies in the whole interval, which suggests that the production is less specialised.

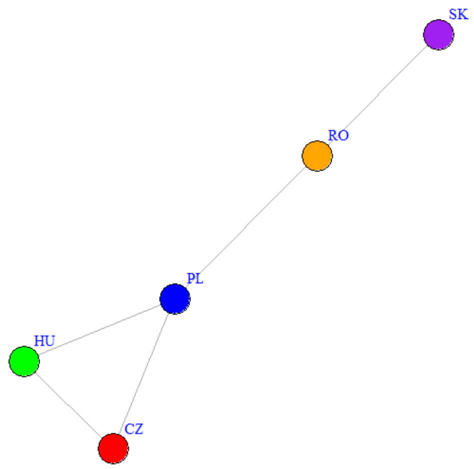
44. Figure: PageRanks scores of graph presented on 43. Figure



Source: own estimation (R15)

When the aggregation is done through the countries, the result will be a simplified version of the product network. Results show that the Hungarian import usually takes place after the Slovakian and the Romanian on average, followed by the Polish and the Czech imports (see figure 45.). This order is not general, however it calls one's attention to the fact that the Slovakian import advances the region on average, while the Hungarian producers join the process a bit later.

45. Figure: The HVG graph at country level of regional automotive import



Source: own calculation (R16)

22. Table: The PageRank scores of the aggregated HVG graph at country level of the regional automotive import

Country	PageRank score
Slovakia	0.09
Romania	0.16
Hungary	0.18
Poland	0.23
Czech Republic	0.34

Source: own estimation (R16)

7.6 Summary, conclusion

On the basis of the map of sequences of five countries from the CEE region (Czech Republic, Hungary, Poland, Romania, and Slovakia), it can be asserted that there are small shifts in car parts import that can be measured in days or weeks. However, the supply cycles are very different and ranging from three to six years depending on the product. The most frequent cycle was observed to be three months, and the shift within this domain is only in days.

Algorithms revealing the sequences confirmed that products of the same categories (like electrical parts) are generally imported at the same time or right after each other. There are also parts that are positioned to the end of the sequences and are seemingly independent of other products (like seats).

The regional sequence strongly resembles the Czech one, which shows a deep network in which the imports arrive in a regular order. As compared to that, the Hungarian import of car parts followed some irregularity, and no definite order could be detected. One can observe similar tendencies in the other four EU member states; however, the import is more layered in all of them.

Car bodies got special attention because these parts constitute the base of all vehicles. Both spectral and sequential analyses positioned the purchase of these parts to the middle of the whole supply process. In country-wide aggregation, the order was somewhat altered; however, the variance was more or less the same.

The aim of the sequential analysis was to contribute to those indicators that are base index numbers. There are many measures in the world of GVC analysis in which the value-added is indexed by gross exports, while the latter also includes the value-added produced by all other firms in the value chain. This bias cannot be adjusted by the estimated sequences; however, they can be utilised as supplementary data.

These analyses showed that the first sequences of the supply and production process are dominated by Slovakian firms, that is, they import a bit earlier than the other investigated countries. The sequence after Slovakia is Romania, Hungary, Poland, and the Czech Republic. Hungary has positioned at the middle in the supply chain of the region and the network is not deep. Thus, the car parts arrive in the country more or less at the same time, which suggests that the producers generally do not make significant transformations on the imported products. In case of Slovakia, the network is deeper, and thus, Slovakian producers are more likely to be integrated deeper than the Hungarians. The first place in sequences suggests that there is some specialisation in Slovakia. This could have many forms from manufacturing to logistics.
These results reinforce the findings of Vakhal (2018b, 2018c), which suggested that the position of Hungarian carmakers in the regional automotive network is rather weak. Despite the relative strong integration, the links to other foreign producers are weak, especially compared to Slovakia. The conclusion is that for Hungary, the trade in value-added statistics indexed by the gross export is likely to bias upwards as the less deep and the sequential shift the value of the denominator is probably higher than for Slovakia and Romania. The import of Czech companies happens sequentially later, but because of the deeper integration, the domestic value-added is likely to be larger, which compensates the bias because of the larger gross export.

In the future, it could be worth analysing the bilateral links in terms of sequences[79], and the graph could be managed as directed. However, this definitely requires a global approach of which the computation capacity demand is enormous. The visualisation of the networks in case of such a high volume of nodes and edges is very complex. This chapter contributes to the field with the application of HVG algorithm that is extended to map hidden sequences. The PageRank method to determine the final order is robust (Avrachenkov & Litvak, 2006); however, the literature offers several other methods that are potentially suitable for sequence detection.

[79] This analysis requires a large volume of data, which constrains the feasibility of the investigation (Vakhal, 2018c).

8. Hungarian firm in the global value chains[80]

The role of Central and Eastern European region countries is peculiar. Generally, one cannot define an exact time interval for the development and spread of GVCs. This region, however, provides a good example to observe the integration into the world trade, as all the countries had begun the transition to market economy around the 1990s. In the beginning, there was no real alternative to this process, as the large companies founded by the FDI inflow in the region had no local supplier background that could produce competitive inputs. Subsequently, new local firms gradually emerged and caught up to the international competitive standards; however, most of them did not survive (Palócz & Vakhal, 2018b). This could have several endogenous and exogenous causes, although the underlying reasons are usually the moderation of productivity increase and the lack of innovation and corporate renewal capacity. Studies investigating the competitiveness of the Hungarian economy often conclude that some exogenous factors are just as important in entrepreneurial survival as the endogenous factors (Szerb et al., 2014). Perényi and Losoncz (2018) provide a systematic literature review of the internationalisation of Hungarian firms and the economic environment between 1998 and 2008. Nowadays, the duality of the Hungarian economy is common in which large (often foreign-owned) companies with above the average productivity operate besides the small- and medium-size (mostly owned by residents) that have below the average productivity (Czakó et al., 2016). The latter definitely has lower export capacity; however, as a supplier of a direct export exporter they can also be indirect exporters (Éltető & Udvari, 2018).

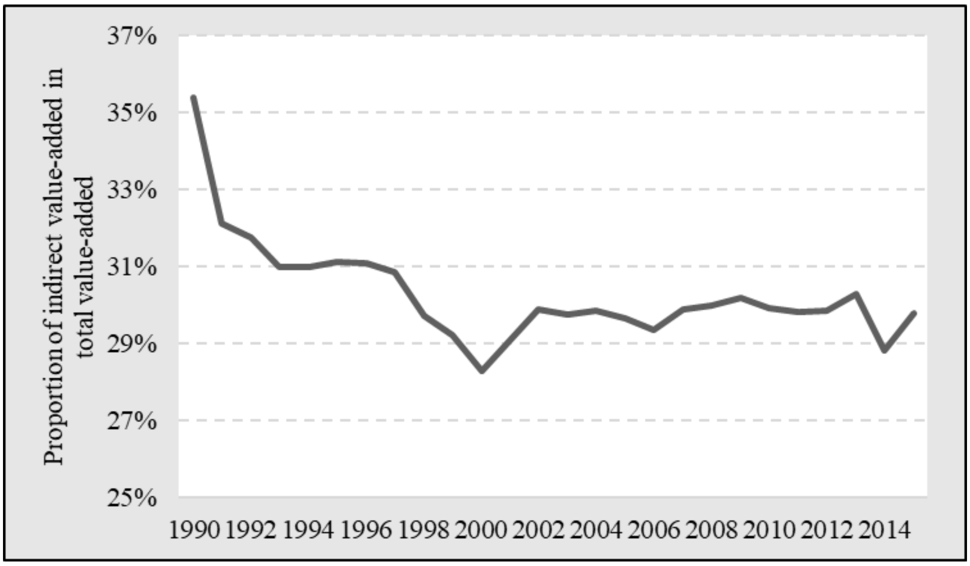
The indicator of indirect domestic value-added in gross export was already introduced in Chapter 4. This measure shows how much exported value-added is produced indirectly by those companies that are not engaged in international trade. It also shows how strong can the exporters rely on the background domestic economic network.

Figure 46 presents the evolution of this indicator in Hungary since 1990. It depicts well all shocks and milestones of the Hungarian economy that influenced this indicator. The EU accession rounds began right after the recovery period between 1990 and 1995. In the frame of that, Hungary had joined European Communities, and the customs in block trade were gradually dismantled until 2000. The FDI inflow began in parallel, and eventually,

[80] This chapter strongly relies on the work of Vakhal (2020).

the foreign affiliates could rely more and more on the resident supplier network[81]. Despite that, there was no crucial development in the economy, and the local suppliers could not significantly decrease the import dependency.

46. Figure: The proportion of indirect value-added in total value-added produced by Hungarian producers between 1990 and 2015

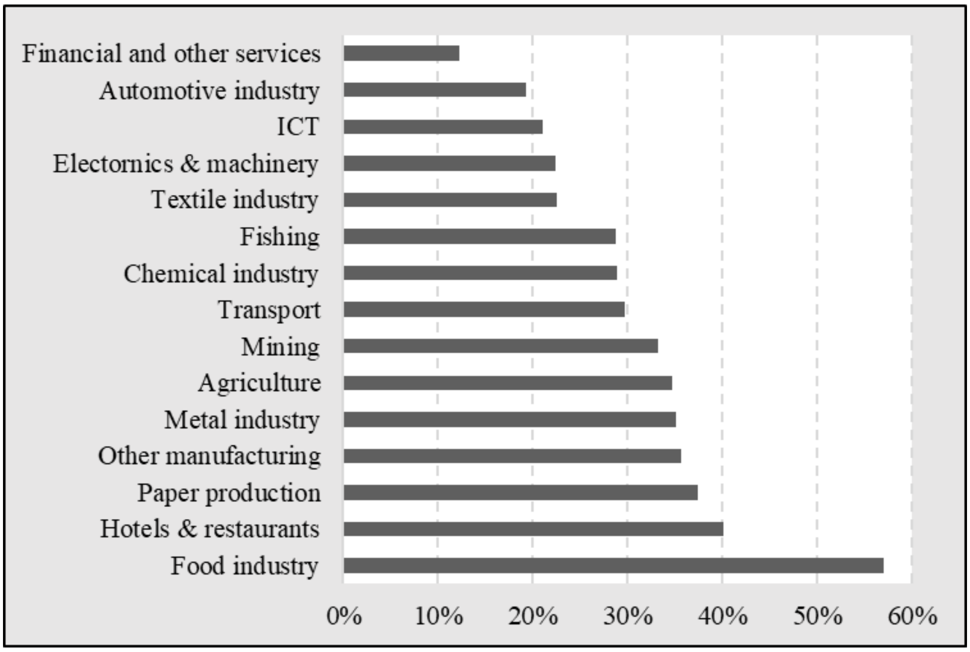


Source: own calculation based on Eora data

Today, almost 30% of domestic Hungarian value-added flowing in GVCs is indirect, and thus, almost a third of the exported value-added is produced by firms that are not engaged in foreign trade[82]. The industrial distribution of indirect value-added export is illustrated in Figure 47.

[81] It must be noted that not all companies in that resident supplier network were owned 100% by Hungarian residents.
[82] Note that a company can have both direct and indirect exports.

47. Figure: Indirect value-added on some industries[83] in proportion of the total exported value-added

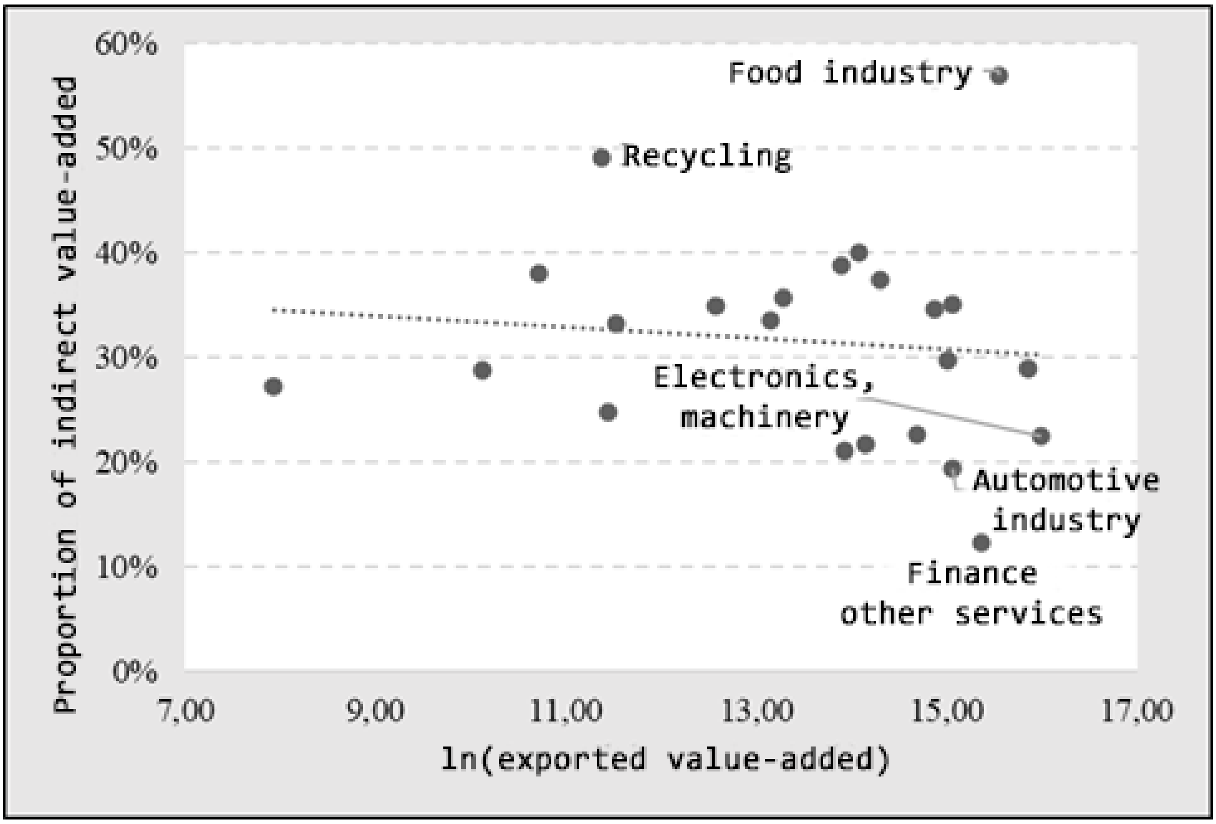


Source: own calculation based on Eora data

Figure 47 shows that the industries form heterogenous groups, and there is no association between the volume of direct and indirect value-added export. In general terms, it is true that industries that traditionally produce export (like the vehicle, electronics, and chemical industries as well as the machinery sector) rely less on the domestic suppliers. Thus, in these industries, the share of those companies that supply domestic value-added to a direct exporter is below the average. In contrast to that, there are many companies in the food industry that indirectly export, even more, this is the only industry that exports more indirectly than directly (see Figure 48).

[83] For the sake of simplicity those industries were omitted in which the volume of gross export is generally low (like education or health).

48. Figure: The relationship between the indirect value-added and exported value-added in Hungary in 2015



Source: own calculation based on Eora data

8.1 Data and methodology

This chapter aims to find how large is the Hungarian small and medium enterprises (SME) segment, which participates in global production either directly or indirectly. For this, the symmetric IO table of 2015 of the Central Statistical Office (CSO), the company balance sheet, and income statement database of the National Tax and Customs Authority (NAV) are utilised. A connection can be drawn between the aggregated (the IO table is aggregated at national and industrial levels) and the corporate level. Owing to the fact that IO tables do not provide data according to company size, this data shall be estimated[84].

[84] A similar analysis was published by Boda et al. (2019); however, the aim of the research was different.

The phenomenon of blank IO matrices is not uncommon in the literature. Both scholars and decision makers require a more detailed database. The most common process is the disaggregation of regional tables, which provides an insight into regional interdependencies, and one can also estimate the regional multipliers[85]. The theoretical background of IO disaggregation according to company size has been sparsely investigated both in the Hungarian and the international literature[86]. The methods utilised during regional disaggregation can also be applied for disaggregation of any dimension if the necessary data are in hand. This requirement is fulfilled by the NAV database and the foreign trade database of the CSO.

The edges of the IO matrix are already known because of the availability of the national data: the variables of industrial output, import use, value-added, export, and final use are known, and the sum of all these is the total output. On the basis of the company balance sheet and foreign trade by industry data, the elements of the matrix can be estimated as follows:

The total value-added, export, and output can be calculated by company size from the corporate balance sheet and income statement database; however, owing to the adjustment method applied by the statistical office, the results may be different from the GDP[87]. However, in case of gross value-added and export, there is no reason to assume that the real ratios would be significantly different from what one can calculate from the company database. Nevertheless, the estimation of the volume of import by company size is immensely taxing. The CSO publishes data concerning foreign trade according to firm size, which is only available for merchandise trade. Thus, the import volume of services is uncertain[88]. To overcome this, it is assumed that the distribution of import services by company size does not noticeably differ from the distribution of total import including both goods and services[89].

---

[85] About the estimation methodology of regional IO tables, see Szabó (2015).

[86] This topic was first studied by Nakajo (1995), and the next article partly dealing with the topic was published about 20 years after that (Grassi, 2016).

[87] In the compilation of the industrial statistics, the CSO applies sampling from the population of firms employing less than 50 persons. These data are used to estimate the quarterly GDP, because the financial reports of the companies are available afterwards.

[88] Companies must report merchandise import over 170 million HUF, while there no such criteria for services, and thus, the CSO estimates the volume by sampling.

[89] Unfortunately, this assumption cannot be confirmed for Hungary owing to lack of data. Neither the Hungarian nor the international literature contains any reliable information that can be applied in the CEE region. The only partial exemption is the statistical office of the UK, which publishes annual data concerning imports by company size. Although there are no data about the volumes (just the number of companies), the distribution supports the initial assumption (https://www.ons.gov.uk/businessindustryandtrade/business/businessservices/datasets/annualbusinesssurveyimportersandexporters)

In case of exports, there is no such problem, because income statements register all export revenues, which also includes services. Total revenue was used as a proxy of total output, which is in line with practices of statistical offices. Values in the vector of final used were distributed between large companies and SMEs according to their share in total output. Despite the fact that the official IO database includes 65 industries, the estimations were done for the three main sectors only (agriculture, industry, and services). The reason behind this is that the number of elements in the transaction matrix that should have been estimated was 4,225 in contrast to those nine that must be calculated in case of the smaller matrix. The number of required iterations significantly decreased, and together with the use side of the IO tables, only 36 elements were estimated.

The descriptions of the prorated estimations can be found in Table 23, which also provides information about data sources.

23. Table: Introductin to the data sources

Data type	Data description	Data source
Industrial output, import use, value-added, export, final use	The production components that can be found on the edge of the transaction matrix of the IO table. Values are at 2008 prices.	Central Statistical Office, table PP1109
Merchandise trade according to firm size	Foreign trade by corporate self-declaration according to firm size. In case of export the threshold is 100 million HUF, for the import it is 170 million HUF in case of trade within the European Union. Data collection in case of trade with third countries is fully covered.	Central Statistical Office, table 3.5.27.
Corporate export and financial data	The export data serve as a control variable to check the data of the CSO. Financial data are for creating distributions of output, value-added, and export according to firm size.	NAV database

Source: own collection

The edges of the IO matrix are equal to the edges of the national IO matrix. The distributions are presented in Table 24.

24. Table: The distribution of output according to company size (SME+Large companies =100%)

	Gross value added	Output	Import	Export	Domestic final use
SME	45,8%	43,7%	32,2%	20,4%	58,4%
Large companies	54,2%	56,3%	67,8%	79,6%	41,6%

Source: own estimation based on Hungarian Statistical Office data

The intra-industrial trade of intermediate goods (marked by a question mark in Annex III.) was estimated by the so-called RAS[90] method, which is a balancing algorithm (Miller & Blair, 2009). RAS is an iterative algorithm with an objective function that aims to get an estimated matrix in which the edges are equal to the predefined values. The estimation is based on a previous full matrix, and thus, the inner structure is preserved. In particular, the base matrix was the IO table containing the national intermediate-use data, and thus, one could take advantage that the objective function was such that the sum data by firm size must be equal to the sum of data at national level. The algorithm is extremely sensitive to the initial values, and therefore, the national data were evenly distributed in the cells.

The stop criterion for the algorithm was either to stop within 1% error margin or reaching 100 iterations.

8.2 Results

Although the algorithm runs until 100 iterations, no significant improvement has been observed after the 14th iteration. After that, another optimisation algorithm[91] was run to make the aggregated values of current intermediate consumption (sum of rows) equal to the real national values. As a result, slight differences occurred on the edges of the matrix, which are not common when the RAS method is applied. At the same time, the objective of this research was to estimate the inner structure of the matrix, and thus, small deviances

[90] RAS is not an abbreviation but the actual name of the process. Its name comes from the original article in which the authors calculated with three matrices that were denoted by R and S, while A refers to the transaction matrix.

[91] The generalised reduced non-linear gradient method (Excel solver) was applied here. The objective function was such that the squared error between the base matrix of the national intermediate use and the estimated matrix shall be minimal. The variables were the elements of the estimated matrix, and only the equality of row sums (intermediate consumption) was set as a constraint. As there were no significant differences between the base and the estimated matrices, the algorithm could optimise the values.

from on the edges were considered acceptable. The results have less than 2% error margin; on the side of out, the error is 1.5%, while at intermediate consumption the difference is 2.4%. The aggregated base and estimated matrices are presented in Table 25, while the estimated disaggregated matrix is tabulated in annex III.

25. Table: True and estimated technological coefficients in 2015

	True IO values		
	Agriculture	**Manufacturing**	**Services**
Agriculture	0.186	0.031	0.006
Manufacturing	0.124	0.119	0.065
Services	0.101	0.083	0.213

	Estimated IO values		
	Agriculture	**Manufacturing**	**Services**
Agriculture	0.187	0.031	0.006
Manufacturing	0.126	0.123	0.067
Services	0.102	0.086	0.218

Source: Hungarian Statistical Office, own estimations

The values of the estimated matrix are very informative. The total output of the large companies in agriculture is only 19% of the total agricultural output. This is in congruence with the official statistical data in which only 9% of the total net revenue is generated by large companies. The difference could be in the taxes and transfers, but these were not estimated. By looking at the intermediate consumption of large companies in the industry, one could spot that it is smaller than in case of SMEs (only 47% of intermediate consumption can be linked to large companies). At the same time, the export volume of large companies is six times larger than the volume of the SME sector, and 76% of the total output is done by the large firms (according to the CSO, 66% of the total net revenue in the industrial sector is generated by firms with more than 250 employees).

Regarding the supplier network, local SMEs use the output of other local SMEs (in 38%), or they import goods and services (in 39%). Domestic large companies use only 23% of the total SME output. In case of large companies, 34% of their total intermediate use comes from other local large companies, while the import is 60%.

The main users of SME output (42%) are the final users (households, governments, etc), and thus, almost half of the output is final good. Other SMEs and large companies use their output (intermediate goods) in 20–20% and only 18% is exported. There is no significant demand for the intermediate products of the large companies. The SMEs use 9%, and other large companies use 8%. The use of the final goods of the large companies is also not significant (28%) as compared to the export, which is 58%.

As per the following figure, it is clear that from SMEs, mainly those who have a chance to participate in the international values chains directly that operate in the industrial sector. In that segment, the share of direct export is 41% (70% in case of the large companies). Thus, only a small fraction of the companies can directly join one of the value chains owing to the insufficient return to scale; however, to be part of the GVC, no direct export is required. SMEs with above average productivity can export through a large direct exporter in a way that their goods and services are built into the exported good. In this manner, they also capitalise on the international labour share. Direct risks could also be lower, especially if they use the local currency in the contracts. However, undoubtedly, the competition is more intense in the supplier's network, and smaller companies are more flexible and readily adapt to changes.

26. Table: Domestic value added in gross export according to the source industry

			SME			Large companies		
			Agriculture	Manufacturing	Services	Agriculture	Manufacturing	Services
SME	Agriculture	billion forints	139.9	48.5	14.8	48.5	335,3	11,4
	Manufacturing		7.7	969.6	43.0	14.7	480,5	33,3
	Services		14.1	106.2	1281.6	27.0	748,6	184,4
Large com-panies	Agriculture		1.0	2.0	0.6	204.0	13,7	0,4
	Manufacturing		5.5	46.0	29.5	9.0	4057,5	20,0
	Services		11.9	90.4	198.0	20.1	556,8	2148,5
Exported value added / export			0,66	0,43	0.71	0.76	0.35	0.79
Exported value added / output			0,08	0,18	0.08	0.56	0.27	0.16

Source: own estimation based on official data

The value-added content of the export can be interpreted in two ways (Wang et al., 2013). Vertical summation measures the origin of the value-added (industry, firm size), while the horizontal summation indicates the use of the value-added.

Hungarian SMEs are integrated into GVCs in a similar degree as the large companies. The difference is that most of the value-added is indirectly exported, that is, it is serves as an input for direct exporters. Despite that, the domestic value-add share in the gross export is only one-third of the same value of the large firms. The share of value-added in the gross export of SMEs operating in the industrial sector is 43%, which exceeds the large companies' value (35%). That is, the direct and indirect export of SMEs have a higher share of domestic value-added in their gross export than large companies. This does not contradict the official data. According to the company balance sheet and income statement data, the value-added created by the SMEs is twice as large as their gross export[92]. In case of large firms, the same value is only 0.6, because their export volume is much larger.

One can filter the bias caused by different export volumes if the value-added is indexed by the total output (Y) instead of the gross export. This clearly shows the lag of SMEs behind the large firms. For example, in the agriculture sector, the measure is only 8%,

[92] Most likely because their primary source of revenue is from domestic partners.

while it is 56% for the large companies. In the industrial sector, the exported value-added in the output is 27% in case of the large firms and only 18% in case of SMEs.

Table 27 compares the results of the analysis with the indicators from international IO tables.

27. Table: Hungarian domestic value added in proportion of gross exports, comparison of estimation results with other data sources for year 2015

	Estimation	OECD	Eora
Agriculture	72%	71%	69%
Manufacturing	36%	44%	34%
Services	76%	77%	72%

Source: own estimation, OECD, Eora

On the basis of Table 27, one can conclude that the estimated model is in line with control values from international databases. A considerable, however relatively not significant difference can be observed in case of the industrial sector (the difference is not significant in case of the Eora database). The reason behind this could the differences in the estimation methods. It must be noted that data for small countries tend to be larger in the OECD database as compared to other data providers (see Chapter 4). Despite that, there is no reason for doubt that the estimation is not in congruence with international statistics.

8.3 Summary, conclusion

This chapter analysed the integration of Hungarian firms by size into GVCs. Estimations based on international IO databases (domestic value-added content of export, indirect value-added) suggest that there is a considerable layer in the Hungarian economy, which is hidden from the official foreign trade statistics (because they do not export directly); however, they supply inputs to direct exporters, and to GVC participants, too.

The analysis based on time-series revealed that 30% of exported domestic value-added in Hungary is supplied by indirect exporters, and this value has remained stable over the past decade. In other words, the domestic supplier network has been unable to adequately extend since the transition to market economy since 1990, despite the fact that resident exporters definitely need to rely more on that group. In comparison, the share of indirect export in the gross export is 45% in Poland, 36% in Romania, 35% in the Czech Republic, and 26% in Slovakia (however, it is 31% in Austria).

To reveal the causes behind this, one must disaggregate the indices based on firm size by the utilisation of the corporate balance sheet data of NAV and the national IO table. The two datasets were linked using two optimisation algorithms. Estimations were made concerning the distribution of exported value-added within the main sectors (agriculture, industry, and services). Results show that large companies (above 250 employees) rather export directly, while SMEs export indirectly into GVCs. in case of the latter, the volume of exported value-added is very low as compared to the share of companies in the economy.

Consequently, the layer of SMEs that are able to become a supplier of a large company is weak. This could be because the suppliers also become large companies themselves; however, this theory is not supported by the panel analysis (Palócz & Vakhal, 2018b). In the past 15 years, the layer of 100% Hungarian-owned companies has been eroded. As a result of the inappropriate resident supplier network, large companies fall back on imports.

The lack of competitiveness of the resident SME sector has several diversified aspects, one of which, that is, economic policy is crucial. Besides that, studies evaluating endogenous and exogenous factors pointed that corporate development is strongly dependent on assets (resources) and human factors (Ábel & Czakó, 2013). The latter suggests the importance of other factors of competitiveness such as education, financial culture, and strategic thinking (Chikán, 2008).

This analysis is not comprehensive – the disaggregation could be continued to get a more detailed insight, like more industries or groups by ownership, because these factors also impact corporate performance. However, data are not available for the public, only the statistical offices are in the possession of them; thus, this must constitute the base of further research.

9. Summary, conclusions, and further research

The aim of this dissertation was to connect the disciplines of graph theory and economics to map the position Hungarian industries and companies in GVCs. With the help of methodological support of the two fields together with the self-developed tools, the author was able to find answers to his research questions as well as position Hungary in GVCs. This research successfully revealed those paths that link Hungary to the regional, European, and global trade. Besides that, the dynamic factors were identified that describe the operation of the network.

The research questions and hypotheses of this dissertation can be found in Chapter 1, while the remainder chapters contain the methodological documentation and the analyses that propose answers to the research questions and confirm or decline the hypotheses.

Mapping the accurate position of the Hungarian economy in GVCs, studying the relevant supplier network, and revealing the dynamics of the network is anticipated to contribute to further research on this topic. Furthermore, it supports all fields pertinent to economic policies either at micro- or macro-level. Chapter 2 introduces these challenges and the framework of GVCs.

The actuality and relevance of the research are given by the fact that GVC accounting of statistical offices is mediocre, and thus, the main data sources for GVC analysis are secondary. Owing to globalisation, the system of global trade and capital transaction has been rapidly changing, which can bias the main macroeconomic indicators of the economy. This topic is analysed in Chapter 3. It was proven that because of inadequate and ambiguous definition of resident companies and the hardly traceable transaction in the complex corporate networks, severe biases can emerge in macroeconomic statistics (in particular through import price indices). Other studies also proved that bilateral trade statistics do not reflect the real economic relations between the partners. On the basis of that, the first hypothesis was confirmed.

Besides the case studies, international IO tables are the first source of information regarding GVCs. The structure of IO tables, which also serves as the framework of GVC analysis, was introduced in Chapter 4. From the available international IO tables, the one provided by the Eora database was selected for this research because it covers 189 countries, and to provide the most accurate position of Hungary in the GVCs, the most extended geographical coverage is required. The said chapter also introduced those standard indicators that serve as a measurement of value chains. All indices were analysed and interpreted with regard to the Visegrad countries (the Czech Republic, Hungary, Poland, and Slovakia) between 1995 and 2015. The relative unfavourable position was first revealed in this chapter. The most concerning problem is that those resident firms that are directly participating in the GVCs do not have a domestic supplier background and must thus rely on imports. The volume of indirect exports is the second lowest in Hungary after Slovakia.

With the help of MRIO tables, the GVC analyses could overcome the analysis of bilateral trade relations, because one can map the whole network of suppliers in the dimension of countries and industries. The representation of IO tables is equivalent to the adjacency matrices utilised in graph theory. At the same time, in contrast to the standard community networks visualised by network science, GVC networks are complete, that is, everyone is in connection with everyone.

As a result, most scholars begin the graph analysis of GVCs by reducing the density of the networks. However, the pruning algorithms are usually calibrated to eliminate those nodes and edges that have low weights. Classical cluster analysis is also unable to map the partitions in the network because the elements of low weights are simply grouped at the edge of the segments like satellites around the node with the highest weight.

These problems are analysed in the consecutive chapters, which propose three, partially self-developed approaches to map the structure of the Hungarian network of suppliers, the flow of value-added exports, and regional sequences of the regional import. As there are about 25 million data points at the industrial level and more than 8 million at the product level, in some cases the analysis was narrowed to the relation of the largest export partner (Germany) and to the most important value chain (automotive parts).

Chapter 5 sought the proper definition and visualisation of value chain networks. A self-developed algorithm could find those partners that are equally crucial for the country in focus and other members of the network. The results showed that Hungary is chiefly linked to the network of the CEE region. However, at later stages, this network is also extended by the USA, Japan, and China. The latter three countries are less integrated into the Hungarian network, as it is clearly dominated by the Austrian, German, Polish, Czech, and Slovakian partners. The value-added network of Hungary is very similar to the networks of other Visegrad countries. Thus, the second hypothesis was confirmed, most partners of Hungary are from Europe.

Chapter 6 introduced the methodology of value-added path analysis, which mapped how far the Hungarian value-added 'gets' in the value chains. It was a surprising result that it is not the automotive industry that 'lives the longest' in the system but the chemical industry. The share of value-added in the partner's export of the latter is much higher than the volume of the vehicle industry. Analysis covering one to two rounds of value-added flow revealed that the Hungarian value-added chiefly circulates within Europe; however, it can get also to the USA, Australia, Singapore, and China, although the volume of the feedback from these countries to Europe is not significant. Consequently, the third hypothesis could also be accepted.

Another novel approach was proposed in Chapter 7, which analysed the dynamics of networks in the dimension of the automotive industry. It was concluded that Hungary is situated in the middle of the sequences after Slovakia but before the Czech Republic and Poland. The analysis of dynamism was critical because the actual position in the production sequence can bias GVC indicators. Results suggest that the domestic value-added/gross export indicator is likely to be downward-biased because the companies enter into the production process a bit later. The delay is not more than one to two months.

The analyses confirmed the fourth hypothesis, and the proposed methods are suitable to map the possible order of production. However, the time shifts were very small, and the variances were quite large.

In Chapter 8, the research analysed the integration of Hungarian enterprise into GVCs and studied the direct and indirect export of value-added in the dimension of company size. It was concluded that there is a thin layer of enterprises in Hungary that are indirectly embedded into the value chains. Within this layer, the ratio of SMEs is considerable; however, their absolute number as compared to the total SME sector is low, in particular in the field of manufacturing. This reiterated the findings of Chapter 4 and confirmed the fifth hypothesis.

For the sake of clarity, Table 28 summarises the research questions, followed by the hypotheses and the decision over them:

28. Table: Summary of the research questions, hypothesises and the results

Research question	Hypothesis	Result
How does the ever-globalising world trade, production in value chains and the new forms of trade affect official statistical data collection?	The current statistical accounting of transactions between companies in the value chain can significantly bias macroeconomic statistics.	Confirmed
Where can Hungary, Hungarian industries and companies be placed in global value chains, taking into account bilateral relations? How to display important connections for the country in the production network?	Hungary is strongly linked to countries in the region in global value chains, with only weak economic ties with non-European states.	Confirmed
In the short term, how far will Hungarian added value reach in global value chains, which are the most significant routes and hubs for Hungarian value added exports?	Hungarian added value circulates mainly in Europe, less reaches outside the continent.	Confirmed
How do sequential differences in supplies affect the value of different GVC indicators? Is it relevant at what stage of production countries with similar production profiles join value chains?	Hungary join production later than V4 countries with similar production profile, which biases the value of the GVC indicators downwards.	Confirmed
How much are domestic companies involved in the flow of value added in global value chains?	The value added supplied by Hungarian companies indirectly connected to the value chains may be greater than that exported by directly related companies, and this is mainly generated by the small and medium-sized enterprises.	Confirmed

Source: own edition

This dissertation presented two self-developed models (Chapters 5 and 8) and introduced improvements in a previous one (Chapter 7). In their current form, these models have not been discussed yet in the literature. The algorithm in Chapter 5 was the first one that proposed an accurate position of Hungary in GVCs by narrowing the complete global graph, while it keeps the network of Hungary in focus. By utilising the methodology described in Chapter 6, one could track the path of direct and indirect value-added flows in the world, even through multiple mediators.

The method discussed in Chapter 7 is a sequence-mapping algorithm that is suitable to cluster time series and reveal the dynamism in the networks. This topic was sparsely visited in the literature. The methodology in Chapter 8 satisfied a longstanding demand when it revealed a new dimension of IO tables with the help of firm financial data.

As other investigations, this also has its flaws. National IO tables are highly aggregated, and thus, they conceal the differences across different companies. These distinctions strongly affect the volume of value-added that a firm can produce as compared to its competitors alike. Such factors are the structure of ownership, location, and differences in productivity or employment (blue- or white-collar workers). These data would contribute to more accurate estimations. Another constraint is that all values in international IO tables are at current prices. However, the indexation method applied in this dissertation can filter the effect of price changes and would extend the base of comparisons if one could also analyse the changes in volumes. Therefore, this dissertation had to neglect time-series analysis; however, the utilised Eora database provides data since 1990.

All analyses in this dissertation were conducted with the intention that they can be improved and continued in the future. The further analysis of those neuralgic points discovered in the field of GVCs would add valuable contributions to the literature. Out of these, two need to be emphasised: the identification of growth sources and the estimation of risks in the networks. The first point would reveal the multiplier links between the industries of different countries, and thus, one could estimate the growth contributions because the nominal GDP positively correlates with the positions in the GVCs (Dorrucci et al., 2019). The COVID-19 pandemic in 2019 and 2020 proved that a disturbance in a chain in the GVC can give rise to shocks at other units as well. Most likely, this is not related to the actual size of the economies but rather to the market structure of product, which leads to the second point – the analysis of the interdependent relations and risks in the system.

References

Ábel, I., & Czakó, E. (2013). *Az exportsiker nyomában*. Alinea.

Adamic, L. A., Huberman, B. A., Barabási, A.-L., Albert, R., Jeong, H., & Bianconi, G. (2000). Power-Law Distribution of the World Wide Web. *Science*, *287*(5461), 2115. https://doi.org/10.1126/science.287.5461.2115a

Ágoston, K. C., Burka, D., Kovács, E., Vaskövi, Á., & Vékás, P. (2019). Klaszterelemzési eljárások halandósági adatokra. *Statisztikai Szemle*, *97*(7), 629–655. https://doi.org/10.20311/stat2019.7.hu0629

Aguiar de Medeiros, C., & Trebat, N. (2017). Inequality and Income Distribution in Global Value Chains. *Journal of Economic Issues*, *51*(2), 401–408. https://doi.org/10.1080/00213624.2017.1320916

Ahn, K. J., Guha, S., & McGregor, A. (2012). Graph sketches: Sparsification, spanners, and subgraphs. *Proceedings of the 31st ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, 5–14. https://doi.org/10.1145/2213556.2213560

Albert, R., Jeong, H., & Barabási, A.-L. (1999). Diameter of the World-Wide Web. *Nature*, *401*(6749), 130–131. https://doi.org/10.1038/43601

Alves, L. G. A., Mangioni, G., Cingolani, I., Rodrigues, F. A., Panzarasa, P., & Moreno, Y. (2019). The nested structural organization of the worldwide trade multi-layer network. *Scientific Reports*, *9*(1), 2866. https://doi.org/10.1038/s41598-019-39340-w

Alves, L. G. A., Mangioni, G., Rodrigues, F. A., Panzarasa, P., & Moreno, Y. (2018). Unfolding the Complexity of the Global Value Chain: Strength and Entropy in the Single-Layer, Multiplex, and Multi-Layer International Trade Networks. *Entropy*, *20*(12), 909. https://doi.org/10.3390/e20120909

Amador, J., & Cabral, S. (2017). Networks of Value-added Trade. *The World Economy*, *40*(7), 1291–1313. https://doi.org/10.1111/twec.12469

Amighini, A. A. (2012). China and India in the international fragmentation of automobile production. *China Economic Review*, *23*(2), 325–341. https://doi.org/10.1016/j.chieco.2012.01.002

Ancarani, A., Di Mauro, C., & Mascali, F. (2019). Backshoring strategy and the adoption of Industry 4.0: Evidence from Europe. *Journal of World Business*, *54*(4), 360–371. https://doi.org/10.1016/j.jwb.2019.04.003

Anderson, J. E., & van Wincoop, E. (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review*, *93*(1), 170–192. https://doi.org/10.1257/000282803321455214

Antràs, P. (2020). *De-Globalisation? Global Value Chains in the Post-COVID-19 Age* (Working Paper No. 28115; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w28115

Arellano, M., & Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*, *58*(2), 277–297. https://doi.org/10.2307/2297968

Arto, I., Rueda-Cantuche, J. M., & Peters, G. P. (2014). Comparing the Gtap-Mrio and Wiod Databases for Carbon Footprint Analysis. *Economic Systems Research*, *26*(3), 327–353. https://doi.org/10.1080/09535314.2014.939949

Avrachenkov, K., & Litvak, N. (2006). The Effect of New Links on Google Pagerank. *Stochastic Models*, *22*(2), 319–331. https://doi.org/10.1080/15326340600649052

Backer, K. D., DeStefano, T., Menon, C., & Suh, J. R. (2018). *Industrial robotics and the global organisation of production* (No. 2018/03; OECD Science, Technology and Industry Working Papers). OECD. https://www.oecd-ilibrary.org/content/paper/dd98ff58-en

Baldwin, R. (2012). *Global Supply Chains: Why They Emerged, Why They Matter, and Where They are Going* (SSRN Scholarly Paper ID 2153484). Social Science Research Network. https://papers.ssrn.com/abstract=2153484

Baldwin, R. (2014). *Keynote speech at Global Value-Chain Training and Research Workshop*. Global Value-Chain Training and Research Workshop, Bejing.

Baldwin, R., & Lopez-Gonzalez, J. (2015). Supply-chain Trade: A Portrait of Global Patterns and Several Testable Hypotheses. *The World Economy*, *38*(11), 1682–1721. https://doi.org/10.1111/twec.12189

Baldwin, R., & Tomiura, E. (2020). Thinking ahead about the trade impact of COVID-19. In R. Baldwin & W. di Mauro (Eds.), *Economics in the Time of COVID-19* (pp. 59–72). Centre for Economic Policy Research.

Baldwin, R., & Weder di Mauro, B. (2020). *Mitigating the COVID Economic Crisis: Act Fast and Do Whatever It Takes | VOX, CEPR Policy Portal*. VoxEU CEPR. https://voxeu.org/content/mitigating-covid-economic-crisis-act-fast-and-do-whatever-it-takes

Banga, R. (2013). Measuring value in global value chains. *UNCTAD Background Paper*, *21*. https://doi.org/10.18356/dd847f8a-en

Barabási, A.-L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, *286*(5439), 509–512. https://doi.org/10.1126/science.286.5439.509

Barabási, A.-L., & Pósfai, M. (2016). *Network science*. Cambridge University Press.

Barigozzi, M., Fagiolo, G., & Mangioni, G. (2011). Identifying the community structure of the international-trade multi-network. *Physica A: Statistical Mechanics and Its Applications*, *390*(11), 2051–2066. https://doi.org/10.1016/j.physa.2011.02.004

Bell, H. E. (1965). Gershgorin's Theorem and the Zeros of Polynomials. *The American Mathematical Monthly*, *72*(3), 292–295. https://doi.org/10.2307/2313703

Bella K., & Kazimir I. (2020). A multinacionális nagyvállalatok stratégiai döntéseinek hatása a termelés oldali GDP alakulására. *Statisztikai Szemle*, *98*(3), 212–241.

Bems, R., Johnson, R. C., & Yi, K.-M. (2010). Demand Spillovers and the Collapse of Trade in the Global Recession. *IMF Economic Review*, *58*(2), 295–326. https://doi.org/10.1057/imfer.2010.15

Bergstrand, J. H. (1985). The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence. *The Review of Economics and Statistics*, *67*(3), 474–481. https://doi.org/10.2307/1925976

Bod, P. Á. (2015). Átmeneti ütemvesztés vagy a 'közepes jövedelem csapdája' – kommentár a magyar gazdaságfejlesztési teendőkhöz. *Gazdaság És Pénzügy*, *2*(1), 2–17.

Boda, G., Révész, T., Losonci, D., & Fülöp, Z. (2019). A növekedési ütem és a foglalkoztatás növelésének lehetőségeiről. *Közgazdasági Szemle*, *66*(4), 376–417. https://doi.org/10.18414/KSZ.2019.4.376

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *The Journal of Mathematical Sociology*, *2*(1), 113–120. https://doi.org/10.1080/0022250X.1972.9989806

Borgatti, S. P., & Everett, M. G. (2006). A Graph-theoretic perspective on centrality. *Social Networks*, *28*(4), 466–484. https://doi.org/10.1016/j.socnet.2005.11.005

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, *30*(1), 107–117. https://doi.org/10.1016/S0169-7552(98)00110-X

Bródka, P., Skibicki, K., Kazienko, P., & Musiał, K. (2011). A degree centrality in multilayered social network. *2011 International Conference on Computational Aspects of Social Networks (CASoN)*, 237–242. https://doi.org/10.1109/CASON.2011.6085951

Bruner, J., Rassier, D. G., & Ruhl, K. J. (2018). *Multinational Profit Shifting and Measures throughout Economic Accounts* (Working Paper No. 24915; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w24915

Caldarelli, G., Cristelli, M., Gabrielli, A., Pietronero, L., Scala, A., & Tacchella, A. (2012). A Network Analysis of Countries' Export Flows: Firm Grounds for the Building Blocks of the Economy. *PLOS ONE*, *7*(10), e47278. https://doi.org/10.1371/journal.pone.0047278

Campbell, D. L. (2010). *History, Culture, and Trade: A Dynamic Gravity Approach* (Working Paper No. 26/2010). EERI Research Paper Series. https://www.econstor.eu/handle/10419/142588

Cappariello, R., Franco-Bedoya, S., Gunnella, V., & Ottaviano, G. I. P. (2020). *Rising Protectionism and Global Value Chains: Quantifying the General Equilibrium Effects* (SSRN Scholarly Paper ID 3612910). Social Science Research Network. https://doi.org/10.2139/ssrn.3612910

Cerina, F., Zhu, Z., Chessa, A., & Riccaboni, M. (2015). World Input-Output Network. *PLOS ONE*, *10*(7), e0134025. https://doi.org/10.1371/journal.pone.0134025

Chikán, A. (2008). *Vállalatgazdaságtan*. Aula.

Chikán, A., Kovács, E., Matyusz, Z., Sass, M., & Vakhal, P. (2018). *Inventories in National Economies: A Cross-Country Analysis of Macroeconomic Data* (1st ed. 2018). Springer London : Imprint: Springer. https://doi.org/10.1007/978-1-4471-7371-7

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, *70*(6), 066111. https://doi.org/10.1103/PhysRevE.70.066111

Clelland, D. A. (2014). The Core of the Apple: Degrees of Monopoly and Dark Value in Global Commodity Chains. *Journal of World-Systems Research*, 82–111. https://doi.org/10.5195/jwsr.2014.564

Cobb, C., & Douglas, P. (1928). A Theory of Production. *The American Economic Review*, *18*(1), 139–165.

Consul, P. C., & Jain, G. C. (1973). A Generalization of the Poisson Distribution. *Technometrics*, *15*(4), 791–799. https://doi.org/10.2307/1267389

Criscuolo, C., & Timmis, J. (2017). The relationship between global value chains and productivity. In A. Sharpe & G. Nicoletti (Eds.), *International Productivity Monitor* (pp. 61–83). Centre for the Study of Living Standards.

Criscuolo, C., & Timmis, J. (2018). *GVCS and centrality: Mapping key hubs, spokes and the periphery*. https://doi.org/10.1787/d4a9bd6f-en

Csontos, L., & Ray, S. C. (1992). The Leontief Production Function as a Limiting Case of the CES. *Indian Economic Review, 27*(2), 235–237.

Czakó E., Juhász P., & Reszegi L. (2016). Versenyképesség és export – vállalati szintű kvalitatív és kvantitatív kutatási eredmények összevetése. *Vezetéstudomány - Budapest Management Review, 47*(8), 3–14. https://doi.org/10.14267/VEZTUD.2016.08.01

Czakó, E., Reszegi, L., Bartók, I., & Vállalatgazdasági Tudományos Egyesület. (2010). *Nemzetközi vállalatgazdaságtan*. Alinea : Vállalatgazdasági Tudományos Egyesület.

Czakó, E., & Vakhal, P. (2020). Hungary in Global Value Chains. In C. Xin (Ed.), *Toward Center or Periphery in Global Value Chains*. China-CEE Institute.

Dachs, B., Kinkel, S., & Jäger, A. (2019). Bringing it all back home? Backshoring of manufacturing activities and the adoption of Industry 4.0 technologies. *Journal of World Business, 54*(6), 101017. https://doi.org/10.1016/j.jwb.2019.101017

Daudin, J.-J., Picard, F., & Robin, S. (2008). A mixture model for random graphs. *Statistics and Computing*, *18*(2), 173–183. https://doi.org/10.1007/s11222-007-9046-7

Degain, C., Meng, B., & Wang, Zhi. (2017). Recent trends in global trade and global value chains. In *Global value chain development report 2017: Measuring and analyzing the impact of GVCs on economic development* (pp. 37–68). World Bank.

Demsetz, H. (1997). *The economics of the business firm: Seven critical commentaries* (1. paperback ed). Cambridge Univ. Press.

Dhillon, I. S., Guan, Y., & Kulis, B. (2007). Weighted Graph Cuts without Eigenvectors A Multilevel Approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 29*(11), 1944–1957. https://doi.org/10.1109/TPAMI.2007.1115

Diakantoni, A., Escaith, H., Roberts, M., & Verbeet, T. (2017). *Accumulating Trade Costs and Competitiveness in Global Value Chains* (SSRN Scholarly Paper ID 2906866). Social Science Research Network. https://doi.org/10.2139/ssrn.2906866

Diewert, E., & Lawrence, D. (1999). Measuring New Zealand's Productivity. In *Treasury Working Paper Series* (No. 99/05; Treasury Working Paper Series). New Zealand Treasury. https://ideas.repec.org/p/nzt/nztwps/99-05.html

Diewert, W. E. (1971). An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function. *Journal of Political Economy, 79*(3), 481–507.

Doherty, M. (2015). Reflecting Factoryless Goods Production in the U.S. Statistical System. In S. N. Houseman & M. Mandel (Eds.), *Measuring Globalization: Better Trade Statistics for Better Policy* (pp. 13–44). W.E. Upjohn Institute. https://doi.org/10.17848/9780880994903.vol2ch2

Donato, M., Ahsan, K., & Shee, H. (2015). Resource dependency and collaboration in construction supply chain: Literature review and development of a conceptual framework. *International Journal of Procurement Management, 8*(3), 344–364. https://doi.org/10.1504/IJPM.2015.069157

Dorrucci, E., Gunnella, V., Al-Haschimi, A., Benkovskis, K., Chiacchio, F., de Soyres, F., Di Lupidio, B., Fidora, M., Franco-Bedoya, S., Frohm, E., Gradeva, K., López-García, P., Koester, G., Nickel, C., Osbat, C., Pavlova, E., Schmitz, M., Schroth, J., Skudelny, F., … Wörz, J. (2019). *The impact of global value chains onthe euro area economy* (Research Report No. 221). ECB Occasional Paper. https://doi.org/10.2866/870210

Dridi, J., & Zieschang, K. (2004). Export and Import Price Indices. *IMF Staff Papers*, *51*(1), 157–194. https://doi.org/10.2307/30035869

Eichengreen, B., Park, D., & Shin, K. (2013). *Growth Slowdowns Redux: New Evidence on the Middle-Income Trap* (Working Paper No. 18673; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w18673

Éltető, A., & Udvari, B. (2018). Nemzetköziesedés a válság után – a magyar kis- és középvállalatok exportjára ható tényezők. *Közgazdasági Szemle*, *65*(4), 402–425.

Emmert-Streib, F., Dehmer, M., & Shi, Y. (2016). Fifty years of graph matching, network alignment and network comparison. *Information Sciences*, *346–347*, 180–197. https://doi.org/10.1016/j.ins.2016.01.074

Engel, J., & Taglioni, D. (2017). The middle-income trap and upgrading along global value chains. In *Global Value Chain Report 2017: Measureing and analyzing the impact of GVCs on economic development* (pp. 119–139). World Bank.

Erdős, P., & Rényi, A. (1960). On the Evolution of Random Graphs. *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, *5*(1), 17–61.

EUROSTAT-OECD. (2015). *Eurostat-OECD Compilation guide on land estimations*. Eurostat. https://doi.org/10.1787/9789264235175-en

Fernandes, N. (2020). *Economic Effects of Coronavirus Outbreak (COVID-19) on the World Economy* (SSRN Scholarly Paper ID 3557504). Social Science Research Network. https://doi.org/10.2139/ssrn.3557504

Ferrarini, B. (2013). Vertical Trade Maps. *Asian Economic Journal*, *27*(2), 105–123. https://doi.org/10.1111/asej.12005

Fiedler, M. (1973). Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, *23*(2), 298–305. https://doi.org/10.21136/CMJ.1973.101168

Fortunato, S., Boguñá, M., Flammini, A., & Menczer, F. (2008). Approximating PageRank from In-Degree. In W. Aiello, A. Broder, J. Janssen, & E. Milios (Eds.), *Algorithms and Models for the Web-Graph* (pp. 59–71). Springer. https://doi.org/10.1007/978-3-540-78808-9_6

Garlaschelli, D., & Loffredo, M. I. (2005). Structure and evolution of the world trade network. *Physica A: Statistical Mechanics and Its Applications*, *355*(1), 138–144. https://doi.org/10.1016/j.physa.2005.02.075

Gáspár, T., & Koppány, K. (2020). A globális értékláncok mérése nemzetközi ÁKM-ek alapján. *Statisztikai Szemle*, *98*(9), 1035–1065. https://doi.org/10.20311/stat2020.9.hu1035

Gereffi, G. (1994). *The Organization of Buyer-Driven Global Commodity Chains: How U.S. Retailers Shape Overseas Production Networks* (pp. 95–122).

Gereffi, G. (2020). What does the COVID-19 pandemic teach us about global value chains? The case of medical supplies. *Journal of International Business Policy*, *3*(3), 287–301. https://doi.org/10.1057/s42214-020-00062-w

Gereffi, G., & Fernandez-Stark, K. (2011). *Global Value Chain Analysis: A Primer*.

Gereffi, G., Humphrey, J., & Sturgeon, T. (2005). The governance of global value chains. *Review of International Political Economy*, *12*(1), 78–104. https://doi.org/10.1080/09692290500049805

Gereffi, G., & Luo, X. (2014). *Risks and Opportunities of Participation in Global Value Chains* (No. WPS6847; Policy Research Working Paper). World Bank. https://documents1.worldbank.org/curated/en/914141468325443509/pdf/WPS6847.pdf

Giljum, S., Wieland, H., Lutter, S., Eisenmenger, N., Schandl, H., & Owen, A. (2019). The impacts of data deviations between MRIO models on material footprints: A comparison of EXIOBASE, Eora, and ICIO. *Journal of Industrial Ecology*, *23*(4), 946–958. https://doi.org/10.1111/jiec.12833

Goldstein, B., & Newell, J. P. (2020). How to track corporations across space and time. *Ecological Economics*, *169*, 106492. https://doi.org/10.1016/j.ecolecon.2019.106492

Grassi, B. (2016). *Io in io: Size, industrial organization, and the input-output network make a firm structurally important*. Bocconi University.

Gullickson, W., & Harper, M. J. (1987). Multifactor Productivity in U.S. Manufacturing, 1949-83. *Monthly Labor Review*, *110*, 18–28.

Győrffy, D. (2021). Felzárkózási pályák Kelet-Közép-Európában két válság között. *Közgazdasági Szemle*, *68*(1), 47–75. https://doi.org/10.18414/KSZ.2021.1.47

Hajnal, P. (2003). *Gráfelmélet*. Polygon Kiadó.

Hambÿe, C., Hertveldt, B., & Michel, B. (2018). Does consistency with detailed national data matter for calculating carbon footprints with global multi-regional input–output tables? A comparative analysis for Belgium based on a structural decomposition. *Journal of Economic Structures*, *7*(1), 11. https://doi.org/10.1186/s40008-018-0110-6

Hayter, R., & Watts, H. D. (1983). The Geography of Enterprise: A Reappraisal. *Progress in Human Geography*, *7*(2), 157–181. https://doi.org/10.1177/030913258300700201

Heathfield, D. F., & Wibe, S. (1987). *An Introduction to Cost and Production Functions*. Macmillan Education UK. https://doi.org/10.1007/978-1-349-18721-8

Heintz, J. (2006). Low-wage manufacturing and global commodity chains: A model in the unequal exchange tradition. *Cambridge Journal of Economics*, *30*(4), 507–520. https://doi.org/10.1093/cje/bei095

Helpman, E., & Krugman, P. R. (1985). *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy*. MIT Press.

Heuser, C., & Mattoo, A. (2017). *Services trade and global value chains* (No. 8126; Policy Research Working Paper). World Bank.

IMF. (2017). *Globalization and Global Value Chains in External Sector Statistics: Measurement and Challenge* (BOPCOM—17/04). IMF. https://www.imf.org/external/pubs/ft/bop/2017/pdf/17-04.pdf

Inklaar, R., Timmer, M. P., & Ark, B. van. (2007). Mind the Gap! International Comparisons of Productivity in Services and Goods Production. *German Economic Review*, *8*(2), 281–307. https://doi.org/10.1111/j.1468-0475.2007.00408.x

Jiang, X. (2013). *Trade and employment in a vertically specialized world* [ILO Working Paper]. International Labour Organization. https://econpapers.repec.org/paper/iloilowps/994855113402676.htm

Johanson, J., & Vahlne, J. (1990). The Mechanism of Internationalisation. *International Marketing Review*, *7*(4). https://doi.org/10.1108/02651339010137414

Johnson, R. C., & Noguera, G. (2012). Accounting for intermediates: Production sharing and trade in value added. *Journal of International Economics*, *86*(2), 224–236. https://doi.org/10.1016/j.jinteco.2011.10.003

Johnson, R. C., & Noguera, G. (2017). A Portrait of Trade in Value-Added over Four Decades. *The Review of Economics and Statistics*, *99*(5), 896–911. https://doi.org/10.1162/REST_a_00665

Jones, C. I. (2003). Growth, capital shares, and a new perspective on production functions. *Proceedings*, *Nov*. https://ideas.repec.org/a/fip/fedfpr/y2003inovx2.html

Kaiser, C., Festl, A., Pucher, G., Fellmann, M., & Stocker, A. (2019). *The Vehicle Data Value Chain as a Lightweight Model to Describe Digital Vehicle Services*. 68–79. https://doi.org/10.5220/0008113200680079

Kaplinsky, R. (2000). Globalisation and Unequalisation: What Can Be Learned from Value Chain Analysis? *The Journal of Development Studies*, *37*(2), 117–146. https://doi.org/10.1080/713600071

Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika*, *18*(1), 39–43. https://doi.org/10.1007/BF02289026

Knight, G. A., & Cavusgil, S. T. (2004). Innovation, Organizational Capabilities, and the Born-Global Firm. *Journal of International Business Studies*, *35*(2), 124–141.

Knorringa, P., & Pegler, L. (2006). Globalisation, Firm Upgrading and Impacts on Labour. *Tijdschrift Voor Economische En Sociale Geografie*, *97*(5), 470–479. https://doi.org/10.1111/j.1467-9663.2006.00357.x

Kolaczyk, E. D., & Csárdi, G. (2014). *Statistical Analysis of Network Data with R*. Springer-Verlag. https://doi.org/10.1007/978-1-4939-0983-4

Koopman, R., Powers, W., Wang, Z., & Wei, S.-J. (2010). *Give Credit Where Credit Is Due: Tracing Value Added in Global Production Chains* (Working Paper No. 16426; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w16426

Koopman, R., Wang, Z., & Wei, S.-J. (2014). Tracing Value-Added and Double Counting in Gross Exports. *American Economic Review*, *104*(2), 459–494. https://doi.org/10.1257/aer.104.2.459

Koopmans, L. H. (1995). *The spectral analysis of time series*. http://site.ebrary.com/id/10244424

Koppány, K. (2017). A növekedés lehetőségei és kockázatai. Magyarország feldolgozóipari exportteljesítményének és ágazati szerkezetének vizsgálata, 2010–2014. *Közgazdasági Szemle*, *64*(1), 17–53.

Koroknai, P., & Lénárt-Odorán, R. (2011). A speciális célú vállalatok szerepe a hazai gazdaságban és a statisztikákban. *MNB Szemle*, *6*(3), 51–60.

Kortum, S. S. (1997). Research, Patenting, and Technological Change. *Econometrica*, *65*(6), 1389–1419. https://doi.org/10.2307/2171741

Kovács, E. (1989). Idősorok kointegrációja. *Statisztikai Szemle*, *67*(5), 599–619.

Kovács, E. (2014). *Többváltozós adatelemzés*. Typotex.

Kovács, O. (2017). Az ipar 4.0 komplexitása – I. *Közgazdasági Szemle*, *64*(7–8), 823–854.

Lacasa, L., Luque, B., Ballesteros, F., Luque, J., & Nuño, J. C. (2008). From time series to complex networks: The visibility graph. *Proceedings of the National Academy of Sciences*, *105*(13), 4972–4975. https://doi.org/10.1073/pnas.0709247105

Lacasa, L., Nicosia, V., & Latora, V. (2015). Network structure of multivariate time series. *Scientific Reports*, *5*(1), 15508. https://doi.org/10.1038/srep15508

Lancichinetti, A., & Fortunato, S. (2009). Community detection algorithms: A comparative analysis. *Physical Review E*, *80*(5), 056117. https://doi.org/10.1103/PhysRevE.80.056117

Lee, J., & Gereffi, G. (2015). Global value chains, rising power firms and economic and social upgrading. *Critical Perspectives on International Business*, *11*(3/4), 319–339. https://doi.org/10.1108/cpoib-03-2014-0018

Lemmers, O., & Wong, K. F. (2019). Distinguishing Between Imports for Domestic Use and for Re-Exports: A Novel Method Illustrated for the Netherlands. *National Institute Economic Review*, *249*(1), R59–R67. https://doi.org/10.1177/002795011924900115

Lenzen, M., Moran, D., Kanemoto, K., & Geschke, A. (2013). Building Eora: A Global Multi-Region Input–Output Database at High Country and Sector Resolution. *Economic Systems Research*, *25*(1), 20–49. https://doi.org/10.1080/09535314.2013.769938

Leslie, D., & Reimer, S. (1999). Spatializing commodity chains. *Progress in Human Geography*, *23*(3), 401–420. https://doi.org/10.1177/030913259902300304

Li, X., Meng, B., & Wang, Zhi. (2019). Recent patterns of global production and GVC participation. In D. Dollar, E. Ganne, V. Stolzenburg, & Wang, Zhi (Eds.), *Global Value Chain Development Report 2019* (pp. 9–44). World Trade Organization. http://ea.makowave.net/docs/20190517013703293004.pdf#page=19

Lipsey, R. E. (2006). *Measuring International Trade in Services* (Working Paper No. 12271; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w12271

Lorenz, E. N. (1963). Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, *20*(2), 130–141. https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2

Los, B., Timmer, M. P., & Vries, G. J. de. (2015). How Global Are Global Value Chains? A New Approach to Measure International Fragmentation. *Journal of Regional Science*, *55*(1), 66–92. https://doi.org/10.1111/jors.12121

Lund, S., Manyika, J., Woetzel, J., Bughin, J., Krishnan, M., Seong, J., & Muir, M. (2019). *Globalization in transition: The future of trade and value chains* (No. 144). McKinsey Global Institute.

Luque, B., Lacasa, L., Ballesteros, F., & Luque, J. (2009). Horizontal visibility graphs: Exact results for random time series. *Physical Review E*, *80*(4), 046103. https://doi.org/10.1103/PhysRevE.80.046103

Lütkepohl, H. (2006). Structural vector autoregressive analysis for cointegrated variables. *Allgemeines Statistisches Archiv*, *90*(1), 75–88. https://doi.org/10.1007/s10182-006-0222-4

Maher, M., Narodytska, N., Quimper, C.-G., & Walsh, T. (2008). Flow-Based Propagators for the SEQUENCE and Related Global Constraints. In P. J. Stuckey (Ed.), *Principles and Practice of Constraint Programming* (pp. 159–174). Springer. https://doi.org/10.1007/978-3-540-85958-1_11

Manders, J. H. M., Caniëls, M. C. J., & Ghijsen, P. W. Th. (2016). Exploring supply chain flexibility in a FMCG food supply chain. *Journal of Purchasing and Supply Management*, *22*(3), 181–195. https://doi.org/10.1016/j.pursup.2016.06.001

McGrath, S. (2013). Fuelling global production networks with slave labour?: Migrant sugar cane workers in the Brazilian ethanol GPN. *Geoforum*, *44*, 32–43. https://doi.org/10.1016/j.geoforum.2012.06.011

Mead, D. (2014). Analyzing alternatives to export price indexes. *Beyond the Numbers: Global Economy*, *3*(27). https://www.bls.gov/opub/btn/volume-3/analyzing-alternatives-to-export-price-indexes.htm?view_full

Meghanathan, N. (2014). Spectral Radius as a Measure of Variation in Node Degree for Complex Network Graphs. *2014 7th International Conference on U- and e- Service, Science and Technology*, 30–33. https://doi.org/10.1109/UNESST.2014.8

Meng, B., Ye, M., & Wei, SJ. (2017). *Value-added Gains and Job Opportunities in Global Value Chains* (No. 668; IDE Discussion Papers).

Milberg, W., & Winkler, D. E. (2010). *Trade Crisis and Recovery: Restructuring of Global Value Chains* (SSRN Scholarly Paper ID 1601769). Social Science Research Network. https://papers.ssrn.com/abstract=1601769

Miller, R. E., & Blair, P. D. (2009). *Input–Output Analysis: Foundations and Extensions* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511626982

Miroudot, S., & Cadestin, C. (2017). *Services In Global Value Chains: From Inputs to Value-Creating Activities* (OECD Trade Policy Papers No. 197; OECD Trade Policy Papers, Vol. 197). https://doi.org/10.1787/465f0d8b-en

Miroudot, S., & Nordström, H. S. (2019). *Made in the World Revisited* (SSRN Scholarly Paper ID 3489137). Social Science Research Network. https://doi.org/10.2139/ssrn.3489137

Mudambi, R. (2008). Location, control and innovation in knowledge-intensive industries. *Journal of Economic Geography*, *8*(5), 699–725. https://doi.org/10.1093/jeg/lbn024

Muradov, K. (2016). *Counting Borders in Global Value Chains* (SSRN Scholarly Paper ID 2808130). Social Science Research Network. https://doi.org/10.2139/ssrn.2808130

Nádudvari, Z. (2013). A globális vállalatcsoport gazdaságstatisztikai mutatói. *Statisztikai Szemle*, *91*(7), 763–773.

Nakajo, A. (1995). Analysis of Firm Size Effect on R&D Activities in Japan. *Journal of Applied Input-Output Analysis*, *2*(2), 80–93.

Nakamura, A., Diewert, E., Greenless, J., Nakamura, L., & Reinsdorf, M. (2015). Sourcing substitution and related price index biases. In S. N. Houseman & M. Mandel (Eds.), *Measuring Globalization: Better Trade Statistics for Better Policy* (pp. 21–88). W.E. Upjohn Institute.

Newman, M. E. J. (2018). *Networks* (Second edition). Oxford University Press.

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, *69*(2), 026113. https://doi.org/10.1103/PhysRevE.69.026113

Ngai, P., & Chan, J. (2012). Global Capital, the State, and Chinese Workers: The Foxconn Experience. *Modern China*, *38*(4), 383–410. https://doi.org/10.1177/0097700412447164

Nicita, A., Ognivtsev, V., & Shirotori, M. (2013). Global Supply Chains: Trade And Economic Policies For Developing Countries. In *UNCTAD Blue Series Papers* (No. 55; UNCTAD Blue Series Papers). United Nations Conference on Trade and Development. https://ideas.repec.org/p/unc/blupap/55.html

OECD. (2001). *Measuring Productivity - OECD Manual: Measurement of Aggregate and Industry-level Productivity Growth*. OECD. https://doi.org/10.1787/9789264194519-en

OECD. (2008). *Multinational enterprises in the global economy*. OECD. https://www.oecd.org/industry/ind/MNEs-in-the-global-economy-policy-note.pdf

OECD. (2016). *Irish GDP up by 26.3% in 2015?* OECD. https://www.oecd.org/sdd/na/Irish-GDP-up-in-2015-OECD.pdf

OECD (Ed.). (2017). *OECD Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations*. OECD.

Owen, A., Steen-Olsen, K., Barrett, J., Wiedmann, T., & Lenzen, M. (2014). A Structural Decomposition Approach to Comparing Mrio Databases. *Economic Systems Research*, *26*(3), 262–283. https://doi.org/10.1080/09535314.2014.935299

Palócz É., & Vakhal P. (2018a). Fél pohár víz—Avagy hogyan értelmezhető a magyar versenyképesség az objektív és szubjektív mutatók szerint. In Kolosi T. & Tóth I. G. (Eds.), *Társadalmi Riport 2018* (pp. 217–232). TÁRKI.

Palócz É., & Vakhal P. (2018b). Mi lett velük?: Egy kiterjesztett esettanulmány tanulságai a középvállalati réteg sorsának alakulásáról 2000-2016 között. In Kolosi T. & Tóth I. G. (Eds.), *Társadalmi Riport 2018* (pp. 203–216). TÁRKI. http://publikaciotar.repozitorium.uni-bge.hu/1127/

Pemmaraju, S., & Skiena, S. (2003). *Computational Discrete Mathematics: Combinatorics and Graph Theory with Mathematica* ®. Cambridge University Press. https://doi.org/10.1017/CBO9781139164849

Perényi, Á., & Losoncz, M. (2018). A Systematic Review of International Entrepreneurship Special Issue Articles. *Sustainability*, *10*(10), 3476. https://doi.org/10.3390/su10103476

Porter, M. E. (1998). *Competitive advantage: Creating and sustaining superior performance: with a new introduction* (1st Free Press ed). Free Press.

Pratono, A. H. (2019). Cross-cultural collaboration for inclusive global value chain: A case study of rattan industry. *International Journal of Emerging Markets*, *15*(1), 149–170. https://doi.org/10.1108/IJOEM-01-2017-0028

Rassier, D. G. (2017). Improving the SNA Treatment of Multinational Enterprises. *Review of Income and Wealth*, *63*(s2), S287–S320. https://doi.org/10.1111/roiw.12323

Rodrik, D. (2012). *The Globalization Paradox: Democracy and the Future of the World Economy* (Reprint edition). W. W. Norton & Company.

Schreyer, P., & Pilat, D. (2001). Measuring productivity. *OECD Economic Studies*, *33*(2), 127–170.

Serrano, M. Á., & Boguñá, M. (2003). Topology of the world trade web. *Physical Review E*, *68*(1), 015101. https://doi.org/10.1103/PhysRevE.68.015101

Shin, N., Kraemer, K. L., & Dedrick, J. (2012). Value Capture in the Global Electronics Industry: Empirical Evidence for the “Smiling Curve” Concept. *Industry and Innovation*, *19*(2), 89–107. https://doi.org/10.1080/13662716.2012.650883

Soundarajan, S., Eliassi-Rad, T., & Gallagher, B. (2014). A Guide to Selecting a Network Similarity Method. *Proceedings of the 2014 SIAM International Conference on Data Mining*, 1037–1045. https://doi.org/10.1137/1.9781611973440.118

Stöllinger, R. (2019). *Testing the Smile Curve: Functional Specialisation in GVCs and Value Creation* (Working Paper No. 163). wiiw Working Paper. https://www.econstor.eu/handle/10419/204028

Strange, R. (2020). The 2020 Covid-19 pandemic and global value chains. *Journal of Industrial and Business Economics, 47*(3), 455–465. https://doi.org/10.1007/s40812-020-00162-x

Stringer, C., & Michailova, S. (2018). Why modern slavery thrives in multinational corporations' global value chains. *Multinational Business Review, 26*(3), 194–206. https://doi.org/10.1108/MBR-04-2018-0032

Stubnya B. (2020, November 18). *A kínai autóvásárlók elkezdték kihúzni a gödörből a magyar járműgyártást | G7—Gazdasági sztorik érthetően*. https://g7.hu/vallalat/20201118/a-kinai-autovasarlok-elkezdtek-kihuzni-a-godorbol-a-magyar-jarmugyartast/

Sturgeon, T. (2001). How Do We Define Value Chains and Production Networks? *IDS Bulletin, 32*(3), 9–18. https://doi.org/10.1111/j.1759-5436.2001.mp32003002.x

Sturgeon, T. (2015, May 20). Trade in value added indicators: What they are, what they aren't, and where they're headed. *VoxEU.Org*. https://voxeu.org/article/trade-value-added-indicators-caveat-emptor

Sturgeon, T., Nielsen, P., Linden, G., Gereffi, G., & Brown, C. (2013). Direct Measurement of Global Value Chains: Collecting Product- and Firm-Level Statistics on Value Added and Business Function Outsourcing and Offshoring. In A. Mattoo, Z. Wang, & S.-J. Wei (Eds.), *Trade in value added—Developing new measures of corss-border trade*. World Bank.

Sturgeon, T., Van Biesebroeck, J., & Gereffi, G. (2008). Value chains, networks and clusters: Reframing the global automotive industry. *Journal of Economic Geography, 8*(3), 297–321. https://doi.org/10.1093/jeg/lbn007

Szabó, N. (2015). A regionális input-output táblák becslési módszerei. *Területi Statisztika, 55*(01), 3–27.

Szalavetz, A. (2012). Az immateriális beruházások és a nem közvetlenül a termelésben foglalkoztatottak szerepe a gazdasági felzárkózásban. *Közgazdasági Szemle, LIX*(11), 1187–1206.

Szalavetz, A. (2013). Feljebb lépés a multinacionális vállalatok globális értékláncain belül – a hazai leányvállalatok tapasztalatai. *Külgazdaság, 58*(11–12), 53–75.

Szerb, L., Márkus, G., & Csapi, V. (2014). Versenyképesség és nemzetköziesedés a magyar kisvállalatok körében a 2010-es években. *Külgazdaság, 58*(11–12), 65–82.

Timmer, M., Los, B., Stehrer, R., & de Vries, G. (2016). *An Anatomy of the Global Trade Slowdown based on the WIOD 2016 Release* (GGDC Research Memorandum GD-162). Groningen Growth and Development Centre, University of Groningen. https://econpapers.repec.org/paper/grorugggd/gd-162.htm

Timmer, M. P., Erumban, A. A., Los, B., Stehrer, R., & de Vries, G. J. (2014). Slicing Up Global Value Chains. *Journal of Economic Perspectives, 28*(2), 99–118. https://doi.org/10.1257/jep.28.2.99

Timmer, M. P., O'Mahony, M., & Ark, B. van. (2007). EU KLEMS Growth and Productivity Accounts: An Overview. *International Productivity Monitor, 14*, 71–85.

Tröster, B. (2020). *Blockchain technologies for commodity value chains: The solution for more sustainability?* (Research Report No. 27). ÖFSE Briefing Paper. https://www.econstor.eu/handle/10419/224986

UNECE. (2011). *The impact of globalization on national accounts*.

United Nations (Ed.). (1999). *Handbook of input-output table compilation and analysis*. United Nations.

UNSTAT. (2015). *Central Product Classification* (Statistical Papers ST/ESA/STAT/SER.M/77/Ver.2.1). United Nations, Department of Economic and Social Affairs , Statistical Division. https://unstats.un.org/unsd/classifications/unsdclassifications/cpcv21.pdf

Vakhal, P. (2016a). A GVC-k felülnézetből, avagy mit lát a hivatalos statisztika a globális értékláncokból? In E. Czakó, *A globális értékláncok-elméleti alapok és számbavételi lehetőségek.: Vol. Fejezetek a nemzetközi üzleti gazdaságtanból* (pp. 60–71). Budapesti Corvinus Egyetem. http://edok.lib.uni-corvinus.hu/481/1/Nkzi_163.pdf

Vakhal, P. (2016b). *A hozzáadott-érték kereskedelem tendenciái az OECD-országokban* (No. 50; Kopint-Tárki Műhelytanulmányok). Kopint-Tárki.

Vakhal P. (2017). *Magyarország elhelyezkedése a globális érték- és termelési láncban*. Kopint-Tárki. https://www.parlament.hu/documents/126660/1249496/Magyarorszag+elhelyezkedese+a+globalis+erteklancban.pdf

Vakhal P. (2018a). A termelési tényezők szerepe az európai járműipari értékláncban. *Külgazdaság, 62*(11–12), 32–65.

Vakhal, P. (2018b, November 14). *Hol helyezkedik el Magyarország a regionális beszállítói láncban?* A Magyar Logisztikai, Beszerzési és Készletezési Társaság 26. kongresszusa, Siófok.

Vakhal, P. (2018c, November 26). Szekvencia, motívum, dinamika feltáró algoritmusok. *Adatelemzés a Gyakorlatban*. BCE Adatelemző Központ workshop, Budapest.

Vakhal, P. (2020). Magyar kis- és középvállalkozások a globális értékláncokban. *Külgazdaság, 64*(5–6), 30–59. https://doi.org/10.47630/KULG.2020.64.5-6.30

Wall, R. S., & van der Knaap, G. A. (2011). Sectoral Differentiation and Network Structure Within Contemporary Worldwide Corporate Networks. *Economic Geography, 87*(3), 267–308. https://doi.org/10.1111/j.1944-8287.2011.01122.x

Wang, Z., Wei, S.-J., Yu, X., & Zhu, K. (2017a). *Characterizing Global Value Chains: Production Length and Upstreamness* (No. w23261; p. w23261). National Bureau of Economic Research. https://doi.org/10.3386/w23261

Wang, Z., Wei, S.-J., Yu, X., & Zhu, K. (2017b). *Measures of Participation in Global Value Chains and Global Business Cycles* (Working Paper No. 23222; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w23222

Wang, Z., Wei, S.-J., & Zhu, K. (2013). *Quantifying International Production Sharing at the Bilateral and Sector Levels* (Working Paper No. 19677; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w19677

Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics, 15*(2), 70–73. https://doi.org/10.1109/TAU.1967.1161901

Wetherell, C., & Shannon, A. (1979). Tidy Drawings of Trees. *IEEE Transactions on Software Engineering*, *SE-5*(5), 514–520. https://doi.org/10.1109/TSE.1979.234212

Xiao, H., Meng, B., Ye, J., & Li, S. (2020). Are global value chains truly global? *Economic Systems Research*, *32*(4), 540–564. https://doi.org/10.1080/09535314.2020.1783643

Annex I.

Country code	Country
ABW	Aruba
AFG	Afghanistan
AGO	Angola
ALB	Albania
AND	Andorra
ANT	Netherlands Antilles
ARE	UAE
ARG	Argentina
ARM	Armenia
ATG	Antigua
AUS	Australia
AUT	Austria
AZE	Azerbaijan
BDI	Burundi
BEL	Belgium
BEN	Benin
BFA	Burkina Faso
BGD	Bangladesh
BGR	Bulgaria
BHR	Bahrain
BHS	Bahamas
BIH	Bosnia and Herzegovina
BLR	Belarus
BLZ	Belize
BMU	Bermuda
BOL	Bolivia
BRA	Brazil
BRB	Barbados
BRN	Brunei
BTN	Bhutan
BWA	Botswana
CAF	Central African Republic
CAN	Canada
CHE	Switzerland
CHL	Chile
CHN	China
CIV	Cote dIvoire
CMR	Cameroon
COD	DR Congo
COG	Congo
COL	Colombia
CPV	Cape Verde
CRI	Costa Rica
CUB	Cuba
CYM	Cayman Islands

Country code	Country
DJI	Djibouti
DNK	Denmark
DOM	Dominican Republic
DZA	Algeria
ECU	Ecuador
EGY	Egypt
ERI	Eritrea
ESP	Spain
EST	Estonia
ETH	Ethiopia
FIN	Finland
FJI	Fiji
FRA	France
GAB	Gabon
GBR	United Kingdom
GEO	Georgia
GHA	Ghana
GIN	Guinea
GMB	Gambia
GRC	Greece
GRL	Greenland
GTM	Guatemala
GUY	Guyana
HKG	Hong Kong
HND	Honduras
HRV	Croatia
HTI	Haiti
HUN	Hungary
IDN	Indonesia
IND	India
IRL	Ireland
IRN	Iran
IRQ	Iraq
ISL	Iceland
ISR	Israel
ITA	Italy
JAM	Jamaica
JOR	Jordan
JPN	Japan
KAZ	Kazakhstan
KEN	Kenya
KGZ	Kyrgyzstan
KHM	Cambodia
KOR	South Korea
KWT	Kuwait

Annex I.

Country code	Country
CYP	Cyprus
CZE	Czech Republic
DEU	Germany
LBY	Libya
LIE	Liechtenstein
LKA	Sri Lanka
LSO	Lesotho
LTU	Lithuania
LUX	Luxembourg
LVA	Latvia
MAC	Macao SAR
MAR	Morocco
MCO	Monaco
MDA	Moldova
MDG	Madagascar
MDV	Maldives
MEX	Mexico
MKD	North-Macedonia
MLI	Mali
MLT	Malta
MMR	Myanmar
MNE	Montenegro
MNG	Mongolia
MOZ	Mozambique
MRT	Mauritania
MUS	Mauritius
MWI	Malawi
MYS	Malaysia
NAM	Namibia
NCL	New Caledonia
NER	Niger
NGA	Nigeria
NIC	Nicaragua
NLD	Netherlands
NOR	Norway
NPL	Nepal
NZL	New Zealand
OMN	Oman
PAK	Pakistan
PAN	Panama
PER	Peru
PHL	Philippines
PNG	Papua New Guinea
POL	Poland
PRK	North Korea
LAO	Laos
LBN	Lebanon
LBR	Liberia
ROU	Romania
ROW	Rest Of the World
RUS	Russia
RWA	Rwanda
SAU	Saudi Arabia
SDS	South Sudan
SEN	Senegal
SGP	Singapore
SLE	Sierra Leone
SLV	El Salvador
SMR	San Marino
SOM	Somalia
SRB	Serbia
STP	Sao Tome and Principe
SUD	Sudan
SUR	Suriname
SVK	Slovakia
SVN	Slovenia
SWE	Sweden
SWZ	Swaziland
SYC	Seychelles
SYR	Syria
TCD	Chad
TGO	Togo
THA	Thailand
TJK	Tajikistan
TKM	Turkmenistan
TTO	Trinidad and Tobago
TUN	Tunisia
TUR	Turkey
TWN	Taiwan
TZA	Tanzania
UGA	Uganda
UKR	Ukraine
URY	Uruguay
USA	USA
USR	Former USSR
UZB	Uzbekistan
VEN	Venezuela
VGB	British Virgin Islands
VNM	Viet Nam
VUT	Vanuatu

Annex I.

Country code	Country
PRT	Portugal
PRY	Paraguay
PSE	Gaza Strip
PYF	French Polynesia
QAT	Qatar

Country code	Country
WSM	Samoa
YEM	Yemen
ZAF	South Africa
ZMB	Zambia
ZWE	Zimbabwe

Annex II.

		Indus-try	Region R			Region S		Region T		Region U		
			1	**2**	**3**	**1**	**2**	**1**	**2**	**1**	**2**	**3**
T=	**Region R**	**1**	150	500	50	25	75	0	0	40	80	50
		2	200	100	400	200	100	0	0	60	30	90
		3	300	500	50	60	40	0	0	70	65	75
	Region S	**1**	75	100	60	200	250	300	320	90	100	180
		2	50	25	25	150	100	350	390	90	85	155
	Region T	**1**	0	0	0	60	100	350	200	75	55	60
		2	0	0	0	150	200	400	550	60	75	95
	Region U	**1**	60	100	50	100	70	200	150	400	300	330
		2	40	60	90	60	100	100	160	250	220	200
		3	60	80	50	75	50	150	90	350	280	390
	Intermediate comsumption		935	1465	775	1080	1085	1850	1860	1485	1290	1625
	Value added		225	775	415	655	835	600	440	600	400	650

	RR				**SR**			**TR**			**UR**		
A=	0.13	0.22	0.04		0.01	0.04		0.00	0.00		0.02	0.05	0.02
	0.17	0.04	0.34		0.12	0.05		0.00	0.00		0.03	0.02	0.04
	0.26	0.22	0.04		0.03	0.02		0.00	0.00		0.03	0.04	0.03
	RS				**SS**			**TS**			**US**		
	0.06	0.04	0.05		0.12	0.13		0.12	0.14		0.04	0.06	0.08
	0.04	0.01	0.02		0.09	0.05		0.14	0.17		0.04	0.05	0.07
	RT				**ST**			**TT**			**UT**		
	0.00	0.00	0.00		0.03	0.05		0.14	0.09		0.04	0.03	0.03
	0.00	0.00	0.00		0.09	0.10		0.16	0.24		0.03	0.04	0.04
	RU				**SU**			**TU**			**UU**		
	0.05	0.04	0.04		0.06	0.04		0.08	0.07		0.19	0.18	0.15
	0.03	0.03	0.08		0.03	0.05		0.04	0.07		0.12	0.13	0.09
	0.05	0.04	0.04		0.04	0.03		0.06	0.04		0.17	0.17	0.17
VA	0.19	0.35	0.35		0.38	0.43		0.24	0.19		0.29	0.24	0.29

L=			R 1	R 2	R 3	S 1	S 2	T 1	T 2	U 1	U 2	U 3
	R	1	1.33	0.39	0.23	0.13	0.13	0.09	0.10	0.13	0.17	0.13
	R	2	0.48	1.32	0.54	0.27	0.19	0.14	0.15	0.19	0.20	0.20
	R	3	0.51	0.44	1.26	0.17	0.14	0.11	0.11	0.17	0.19	0.17
	S	1	0.25	0.20	0.20	1.29	0.30	0.36	0.40	0.23	0.26	0.27
	S	2	0.17	0.12	0.12	0.22	1.19	0.35	0.39	0.20	0.22	0.22
	T	1	0.06	0.04	0.05	0.11	0.13	1.27	0.22	0.13	0.13	0.11
	T	2	0.10	0.08	0.08	0.23	0.25	0.40	1.50	0.18	0.21	0.19
	U	1	0.27	0.22	0.23	0.24	0.21	0.32	0.31	1.46	0.46	0.40
	U	2	0.20	0.17	0.21	0.17	0.18	0.22	0.26	0.31	1.33	0.27
	U	3	0.25	0.20	0.21	0.20	0.18	0.27	0.25	0.42	0.43	1.41

Region R	Export	S	U	
	1	100	170	
	2	300	180	
	3	100	210	
Region S	Export	R	T	U
	1	235	620	370
	2	100	740	330
Region T	Expo0rt	S	U	
	1	160	190	
	2	350	230	
Region U	Export	R	S	T
	1	210	170	350
	2	190	160	260
	3	190	125	240

	Direct and indirect export of S to R					
<VA>	0.38	0.00		**L**	1.29	0.30
	0.00	0.43			0.22	1.19
<VA>L	0.49	0.11		**E**	235	
	0.10	0.52			100	
DVA_e	125.57					
	74.53					

	Direct and indirect export of T to R via S					
<VA>	0.245	0		**L**	0.1	0.13
	0	0.2			0.2	0.25
<VA>L	0.027	0		**E**	235	
	0.044	0			100	
DVA_e	9.628					
	15.22					

	Direct and indirect export of U to R via S						
<VA>	0.29	0.00	0.00		**L**	0.24	0.21
	0.00	0.24	0.00			0.17	0.18
	0.00	0.00	0.29			0.20	0.18
<VA>L	0.07	0.06			**E**	235	
	0.04	0.04				100	
	0.06	0.05					
DVA_e	21.99						
	13.68						
	18.69						

	Re-import of R via S						
<VA>	0.19	0.00	0.00		**L**	0.13	0.13
	0.00	0.35	0.00			0.27	0.19
	0.00	0.00	0.35			0.17	0.14
<VA>L	0.02	0.03			**E**	235	
	0.09	0.07				100	
	0.06	0.05					
DVA_e	8.41						
	28.49						
	18.79						

	Direct and indirect export of S to R via U							
<VA>	0.38	0.00			**L**	0.23	0.26	0.27
	0.00	0.43				0.20	0.22	0.22
<VA>L	0.09	0.10	0.10		**E**	210		
	0.09	0.09	0.10			190		
						190		
DVA_e	56.85							
	54.19							

Direct and indirect export of T to R via U								
<VA>	0.24	0			**L**	0.13	0.13	0.11
	0	0.19				0.18	0.21	0.19
<VA>L	0.03	0.03	0.03		**E**	210		
	0.03	0.04	0.04			190		
						190		
DVA_e	17.69							
	21.59							

	Direct and indirect export of U to R							
<VA>	0.29	0.00	0.00		**L**	1.46	0.46	0.40
	0.00	0.24	0.00			0.31	1.33	0.27
	0.00	0.00	0.29			0.42	0.43	1.41
<VA>L	0.42	0.13	0.11		**E**	210		
	0.07	0.32	0.06			190		
	0.12	0.12	0.40			190		
DVA_e	135.01							
	87.34							
	124.60							

Re-import of R via U								
<VA>	0.19	0.00	0.00		**L**	0.13	0.17	0.13
	0.00	0.35	0.00			0.19	0.20	0.20
	0.00	0.00	0.35			0.17	0.19	0.17
<VA>L	0.02	0.03	0.03		**E**	210		
	0.07	0.07	0.07			190		
	0.06	0.07	0.06			190		
DVA_e	16.33							
	40.08							
	36.32							

Source: own edition

Annex III.

The edges of the Hungarian symmetric input-output table, 2015 (according to methodology of 2018) (billion forint)

		SME			Large companies			Domestic final use	Export	Output / final use
		Intermediate consumption								
		Agri. culture	Manu- facturing	Services	Agri. culture	Manu- facturing	Services			
SME	Agriculture	?	?	?	?	?	?	473.8	271.9	2 393,8
	Manufacturing	?	?	?	?	?	?	749.5	2 932.2	7 087,6
	Services	?	?	?	?	?	?	11 462.6	2 208.2	20 444,5
Large compa-	Agriculture	?	?	?	?	?	?	115.0	426.5	581,0
	Manufacturing	?	?	?	?	?	?	2 432.6	17 720.5	23 004,0
	Services	?	?	?	?	?	?	8 430.4	3 028.7	15 036,3
Import		388,4	3231.4	2 837.6	24.45	11 675.3	1 889.9			
Value-added		1 010,6	2187.0	9 631.8	275.7	4 910.2	10 008.7			
Output		2 393,8	7087.6	20 444.5	581.0	23 004.0	15 036.3			

Notes:
Agriculture: Industry A
Manufacturing: Industries B+C+D+E+F
Services: Industries from G to S
Estimation of output by company size and sector: $x_{i,j} = \frac{\pi_{i,j}}{\sum_{i=1}^{n}\sum_{j=1}^{k}\pi_{i,j}} \times output_{GDP,}$ where π is the revenue, *i* is the industry, *j* is the category of size, $output_{GDP}$ is output in the GDP.

Estimation of value added by company size and sector: $VA_{i,j} = \frac{VA_{i,j}}{\sum_{i=1}^{n}\sum_{j=1}^{k}VA_{i,j}} \times VA_{GDP}$, where *VA* =Output–Intermediate use,[93] VA_{GDP} value-added in the GDP.

Estimation of domestic consumption by size and sector: $x_{i,j} = \frac{\pi_{i,j}}{\sum_{i=1}^{n}\sum_{j=1}^{k}\pi_{i,j}} \times F_{GDP}$, where F_{GDP} is domestic use in the GDP.

Source: own calculations based on CSO and company data, (sum ont he edges equal to the national data).

[93] For more see *Máténé és Ritzlné* (2020).

The estimated disaggregated symmetric input-output table in 2015 (billion forint)

		SME			Large companies								
		Agri cul-ture	Manufac-turing	Services	Agri. cul-ture	Manufac-turing	Services	Interme-diate use	Diference	Domestic fi-nal use	Export	Output	Difference
SME	Agriculture	413.2	192.2	145.8	124.1	718.8	63.1	1 657.1	1.7%	473.8	271.9	2 402.9	0,4%
	Manufacturing	144.3	395.0	846.3	43.2	1 615.1	360.2	3 404.1	3.0%	749.5	2 932.2	7 085.9	0,0%
	Services	146.2	346.7	3 504.4	44.3	1 342.0	1 452.9	6 836.5	2.4%	11 462.6	2 208.2	20 507.3	0,3%
Large comp.	Agriculture	15.0	7.2	5.8	3.5	26.0	1.8	59.3		115.0	426.5	600.8	3,3%
	Manufacturing	151.4	414.3	859.2	35.9	1 285.2	296.9	3 043.0		2 432.6	17 720.5	23 196.1	0,8%
	Services	90.2	216.2	2 056.9	21.6	675.9	711.3	3 772.0		8 430.4	3 028.7	15 231.0	1,3%
Import	RoW	388.4	3 231.5	2 837.7	24.5	11 675.3	1 889.9						
Value-added		1 010,6	2 187.0	9 631.8	275.7	4 910.2	10 008.7						
Output		2 359,2	6 990.0	19 887.8	572.7	22 248.5	14 784.9						
Difference		1,5%	1.4%	2.8%	1.4%	3.4%	1.7%						

Source: own estimations

Source codes in R (comments are in Hungarian)

R1 Barabási-Albert-féle véletlen hálózatok generálása

```
require(igraph)

g<-barabasi.game(10, m=2, directed=FALSE) # 10 csúcs generálása
plot(g, vertex.size=10, vertex.label=NA, layout=layout.circle) # hálózat rajzolása

g<-barabasi.game(100, m=2, directed=FALSE) # 100 csúcs generálása
plot(g, vertex.size=10, vertex.label=NA, layout=layout.circle) # hálózat rajzolása

g<-barabasi.game(200, m=2, directed=FALSE) # 100 csúcs generálása
plot(g, vertex.size=10, vertex.label=NA, layout=layout.circle) # hálózat rajzolása
```

R2 Generált idősort látható gráfja (n=10)

```
set.seed(1000)
require(statcomp)
require(igraph)

x<-(arima.sim(model=list(ar = 0.3), n = 10)) # idősor generálása
x<-x+abs(min(x)) # idősor pozitív tartományba való eltoltás
plot(x, main="Generált idősor (n=10)", xlab="t", ylab="x", type="b") #idősor
megjelenítése
axis(side = 1, c(1:10)) # x-tengely feliratok korrekciója
g<-HVG(x, meth = "HVG", maxL = 10^9, rho = NA) # VG számítások elvégzése

plot(graph.adjacency(g$A, mode="undirected"), layout=layout.grid) # gráf
megjelenítése
plot(graph.adjacency(g$A, mode="undirected"), layout=layout.reingold.tilford) #
gráf megjelenítése
```

R3 Generált idősort látható gráfja (n=100)

```
set.seed(1000)
require(statcomp)
require(igraph)

x<-(arima.sim(model=list(ar = 0.3), n = 100)) # idősor generálása
x<-x+abs(min(x)) # idősor pozitív tartományba való eltoltás
plot(x, main="Generált idősor (n=100)", xlab="t", ylab="x", type="b") #idősor
megjelenítése
g<-HVG(x, meth = "HVG", maxL = 10^9, rho = NA) # VG számítások elvégzése

plot(graph.adjacency(g$A, mode="undirected"), layout=layout.grid) # gráf
megjelenítése
plot(graph.adjacency(g$A, mode="undirected"), layout=layout.reingold.tilford) #
gráf megjelenítése
```

R4 Három generált idősor képe, HVG hőtértképe és gráfja (n=3x10)

```
set.seed(1000)
require(statcomp)
require(igraph)

x<-arima.sim(model=list(ar=0.3), n=12) # idősor generálása
x1<-ts(window(x,1, 10)+abs(min(x))) # pozitív tartományba tolás

# idősor eltolása tetszőleges értékkel
x2<-ts(window(x+abs(min(x)), 2, 11)) #+abs(rnorm(1)))
x3<-ts(window(x+abs(min(x)), 3, 12)) #+abs(rnorm(1)))

# idősorok ábrázolása
par(mfrow=c(3,1))
plot(x1, main="X1")
plot(x2, main="X2")
plot(x3, main="X3")
dev.off()

# HVG számítása
hvg<-HVG(cbind(x1,x2,x3), meth = "HVG", maxL = 10^9, rho = NA)
g<-graph.adjacency(hvg$A, mode = "undirected") # gráf objektum készítése

# hőtérkép
heatmap(h, sym=T, Colv = NA, Rowv = NA, scale = "none", revC = TRUE,
col=c("lightblue", "orangered"))

# gráf színezése, címkézése
V(g)[1:10]$color<-"red"
V(g)[11:20]$color<-"green"
V(g)[21:30]$color<-"blue"
V(g)$name<-c(seq(1:10), seq(1:10), seq(1:10))

plot(g, layout=layout.reingold.tilford) # gráf rajzolása

# idősorok egy ábrán jelmagyarázattal
ts.plot(cbind(x1,x2,x3), gpars = list(col=c("red", "green", "blue")), lwd=3)
legend(4,1, legend=c("x1", "x2", "x3"), fill=c("red", "green", "blue"), lwd=3)
```

R5 Generált idősorok módosított gráfja

```
require(igraph)
# szomszédsági mátrix
x<-matrix(c(0,1,0,0,0,1,0,0,0), nrow=3, ncol=3, byrow = TRUE)
g_pr<-graph.adjacency(x) # igraph objektum létrehozása

# gráf színezése, rajzolása
V(g_pr)[1]$color<-"red"
V(g_pr)[2]$color<-"green"
V(g_pr)[3]$color<-"blue"
plot(g_pr)

page.rank(graph.adjacency(x)) # PageRank centralitási értékek
```

R6 Aggregált termékcsoportok hőtértéképe

```
require(ggplot)
# adatok átalakítása
heat_prod<-comex %>% group_by(product, year, month) %>% summarise(value=sum(value))
heat_prod<-unstack(form=value~product, x=as.data.frame(heat_prod)) %>% cor() %>%
melt()

# hőtértkép készítése
ggplot(data = heat_prod, aes(x=Var1, y=Var2, fill=value)) + geom_tile(color="black")
+
  scale_fill_gradient2(low = "blue", high = "red", mid="white", midpoint=0,
limit=c(-1,1),
                       space = "Lab") + theme_minimal() + coord_fixed() +
  theme(axis.text.x = element_text(angle = 90))+
  labs(x="Termékek", y="Termékek")
```

R7 Keresztkorrelációk

```
require(ggplot)
comb<-combn(18,2) # kombinációk segédmátrix
cross<-comex %>% group_by(product, year, month) %>% summarise(value=sum(value)) #
aggregálás
cross<-unstack(form=value~product, x=as.data.frame(cross)) # unpivot
cross_m<-matrix(nrow=73, ncol = ncol(comb)) # betároló mátrix
cross_m_names<-c() # neveket tároló vektor

for (i in 1:ncol(comb)) { # keresztkorrelációk számítása
  cross_m[,i]<-ccf(as.numeric(ts(cross[comb[1,i]])),                       as.nu-
meric(ts(cross[comb[2,i]])),
                   lag.max=36, plot=FALSE)$acf
  cross_m_names[i]<-
paste(colnames(cross)[comb[1,i]],"_",colnames(cross)[comb[2,i]],
                              sep="")
}
colnames(cross_m)<-cross_m_names # oszlopok elnevezése
cross_m<-as.data.frame(cross_m)
cross_m2<-cbind(melt(cross_m), rep(c(-36:36),153)) # lag információk létrehozása
colnames(cross_m2)[3]<-"lag"

# ábra
ggplot(cross_m2, aes(x=value, y=variable)) + geom_point(aes(color=lag), alpha=0.5)
+
  theme(axis.text.y = element_blank()) + labs(x="Keresztkorrláció", y="Változók") +
  geom_vline(aes(xintercept=0)) + scale_y_discrete(breaks=NULL) +
  scale_color_gradient2(low = "blue", high = "red", mid="green")
```

R8 Keresztkorrelációk kibontása a 784.21-es termékkódra

```
require(tidyr)
rownames(cross_m)<-as.character(c(-36:36))
cbind(sapply(c(1:153),                                                function(x)
cbind(which(abs(cross_m[,x])==max(abs(cross_m[,x])))-37,

cross_m[which(abs(cross_m[,x])==max(abs(cross_m[,x]))),x])) %>%
  t(), colnames(cross_m))->cross_results

cross_results[which(grepl("X78421", cross_results[,3])=="TRUE"),] # alvázak
```

R9 Idősorok spektrálanalízise

```
cross_ts<-ts(cross, start = c(2004,1), frequency = 12) # idősorok kibontása
s<-stats::spectrum(diff(cross_ts)) # idősorok spektrálanalízise

# eredmények ábrázolása
plot(s, main="", xlab="Frekvencia (év)", ylab="Spektrális sűrűség")
points(x=s$freq, y=apply(s$spec, 1, mean), type = "p", pch=19, cex=1)
points(x=s$freq, y=apply(s$spec, 1, median), type = "p", pch=19, cex=1, col="red")
legend(x="topleft", legend = c("Átlag", "Medián"), pch=19, col = c("black", "red"))
```

R10 Az azonos frekvenciatartományba tartozó fáziseltolások sűrűsége

```
ph<-matrix(nrow=ncol(comb), ncol=6) # tároló mátrix létrehozása

for (i in 1:ncol(comb)) {
  c<-which((s$spec[,comb[1,i]]+s$spec[,comb[2,i]])==
             max(s$spec[,comb[1,i]]+s$spec[,comb[2,i]])) # legmagasabb spektrális
sűrűségű közös frekvenciatartomány keresése
  ph[i,1]<-s$freq[c]
  ph[i,2]<-stats::spectrum(ts(cbind(cross[,comb[1,i]],
cross[,comb[2,i]])))$phase[c] # tartományhoz tartozó fáziseltolás megkeresése
  ph[i,3]<-prods[comb[1,i]] # termékcsoport nevek párosítása
  ph[i,4]<-prods[comb[2,i]]
  ph[i,5]<-comb[1,i] # sorrend a kombinációban
  ph[i,6]<-comb[2,i]
}

colnames(ph)<-c("Length", "Shift", "prod1", "prod2", "c1", "c2")
ph<-as.data.frame(ph)
ph$Shift_month<-ph$Shift*12 # éves fáziseltolások havi szintre konvertálása

plot(density(ph$Shift_month), main="", xlab="Késleltetés hossza (hónap)",
     ylab="Sűrűség") # ábra
```

R11 A 784.21-es csoport alkatrészekkel vett fáziseltolásai (táblázat)

```
vaz<-matrix(nrow = nrow(s$phase), ncol=18) # tárolómátrix létrehozása
colnames(vaz)<-gsub("X", replacement = "", colnames(cross)) # oszlopnevek
korrekciója

for (i in c(1:ncol(cross))) { # páros fáziseltolások számítása
  i<<-i
  x<-stats::spectrum(diff(ts(cbind(cross[,i], cross[,10])),
                          start=c(2004,1), frequency=12))
  vaz[,i]<-x$phase
}

vaz<-melt(vaz[54,]) # 0.25-ös frekvenciatartomány kiválasztása
vaz
```

R12 A 784.21-es csoport alkatrészekkel vett fáziseltolásai (boxplot)

```
x<-subset(ph, prod1=="78421" | prod2=="78421") # szűkítés
c<-rownames(subset(ph, prod1=="78421" | prod2=="78421")) %>% as.numeric # áru ki-
választása
vaz<-matrix(nrow = nrow(s$phase), ncol=length(c)) # tárolómátrix létrehozása
vaz_n<-c() # tárolóvektor neveknek

for (i in 1:length(c)) { # kiválasztott fáziseltolások szűrése
  vaz[,i]<-s$phase[,i]
}

for (i in 1:length(c)) { # oszlopnevek gyűjtése
  if (x[i,3]!="78421") {
    vaz_n[i]<-x[i,3]
  } else {
    vaz_n[i]<-x[i,4]
    vaz[,i]<-vaz[,i]*-1 } # x-beli sorrend kiigazítása
}

colnames(vaz)<-vaz_n # oszlopnevek csatolása
vaz<-melt(vaz)
vaz$value<-vaz$value*12 # évek konvertálása hónapokká

# ábra
ggplot(vaz, aes(x=as.factor(Var2), y=value)) + geom_boxplot() +
  labs(x="Termék", y="Fáziseltolás (hónap)") +
  theme(axis.title=element_text(size = 14), axis.text = element_text(size=12)) +
  geom_hline(aes(yintercept=0), linetype="dashed", color="red")
```

R13 Többdimenziós idősorok HVG gráfja

```
# szekvencia többdimenziós idősor mátrixból
network_sequence<-function(series) {
  require(statcomp)
  series<-apply(series, 2, scale)
  series<-series+min(series)
  g_network<-HVG(ts(series))$A # szomszédsági mátrix az input adatokból
  g_network[upper.tri(g_network)]<-0 # alsó háromszög mátrix
  g_network<<-g_network # HVG eredmény szomszédszági mátrix
  vg<-seq(1, c(nrow(series)*ncol(series)), nrow(series)) # idősor kezdő időpontok
  dom<-cbind(vg, c(vg+nrow(series)-1), colnames(series)) %>% as.data.frame # idősor
tartományok
  dom$vg<-as.numeric(dom$vg)
  dom$V2<-as.numeric(dom$V2)
  colnames(dom)<-c("start", "end", "name")

  # tárolómátrix létrehozása
  g_seq<-matrix(nrow = ncol(series), ncol = ncol(series), data=0)
  rownames(g_seq)<-colnames(series)
  colnames(g_seq)<-colnames(series)

  # Szekvenciák keresése
  for (i in 1:length(vg)) {
    #i<<-i
    m<-g_network[vg[i]:c(vg[i]+nrow(series)-1),  1:c(nrow(series)*ncol(series))]  #
blokkmátrix
    m[,vg[i]:(vg[i]+nrow(series)-1)]<-0 # saját kapcsolatok törlése
    d<-which(as.matrix(m)!=0, arr.ind = TRUE) %>% as.data.frame()
    if (nrow(d)==0) { # ha nincs kapcsolat egyetlen partnerrel sem
      i<-i+1
      if (i>length(vg)) { break}
      m<-g_network[vg[i]:c(nrow(series)-1),   1:c(nrow(series)*ncol(series))]   #
blokkmátrix
      m[,vg[i]:(vg[i]+nrow(series)-1)]<-0 # saját kapcsolatok törlése
      d<-which(as.matrix(m)!=0, arr.ind = TRUE) %>% as.data.frame()
    }
    # a helyes index a partner idősorában
    d$col_adj<-sapply(c(1:nrow(d)),        function(x)        rep(1:nrow(series),
length(vg))[d$col[x]])
    d$base<-rep(i, nrow(d)) # bázis idősor

    x<-c()
    for (k in 1:nrow(d)) { # partner hozzárendelése
      #k<<-k
      x[k]<-which(abs(dom$end-d$col[k])==min(abs(dom$end-d$col[k])))
    }
    d$partner<-x
    rm(x)
    d$diff<-d$row-d$col_adj # differenciák számítása

    x<-rownames(as.matrix(table(d$partner))) %>% as.numeric() # hány partner van,
és kik ők
    for (k in 1:length(x)) {
      m<-subset(d, d$partner==x[k])
      if (mean(m$diff)<0) {
        g_seq[mean(m$base), mean(m$partner)]<-c(-1)
      }
      else {
```

```
        g_seq[mean(m$base), mean(m$partner)]<-1 }
    }
  }
  g_network2<-g_seq
  g_network2[g_network2!=0]<-1
  g_network2<<-g_network2 # csoportok szerinti szomszédsági mátrix
  sort(page.rank(graph.adjacency(g_network2))$vector)
}

# ábra

network_sequence(cross) #

g<-graph.adjacency(g_network, mode = "undirected")
x<-c()
for (i in 1:ncol(cross)) {
  c<-colors()[runif(1, 1, 657)]
  while (grepl("gray", c) | grepl("grey", c)) {
    c<-colors()[runif(1, 1, 657)]
  }
  x[i]<-c
}

for (i in 1:ncol(cross)) {
  V(g)$color[vg[i]:(vg[i]+201)]<-x[i]
}
x1<-which(degree(g)==max(degree(g)))
plot(g, layout=layout_as_tree(g, root = x1), vertex.label=NA, vertex.size=5, mar-
gin=-0.35,
     vertex.color=adjustcolor(V(g)$color, alpha.f = 0.4))
legend("bottomright", legend = colnames(cross), fill = x, title = "Termékek")
```

R14a A módosított termék szerinti HVG ábra

```
network_sequence(cross)

g<-graph.adjacency(g_network2, mode = "undirected")
x<-c()
for (i in 1:ncol(cross)) {
  c<-colors()[runif(1, 1, 657)]
  while (grepl("gray", c) | grepl("grey", c)) {
    c<-colors()[runif(1, 1, 657)]
  }
  x[i]<-c
}

V(g)$color<-x

x1<-which(degree(g)==max(degree(g)))
plot(g,   layout=layout_as_tree(g,   root   =   x1),   margin=-0.35,   vertex.la-
bel.color="blue",
     vertex.label.cex=0.8, vertex.label.dist=2,
     vertex.color=adjustcolor(V(g)$color, alpha.f = 1))
```

R14b Regionális járműipari import országok és termékek szerint

```
v5_prods<-sapply(c(1:90), function(x) cbind(get(var_names[x]))) # termék_ország mátrix létrehozása
colnames(v5_prods)<-var_names

network_sequence(v5_prods) # szekvenciák keresése
g<-graph.adjacency(g_network2, mode = "undirected") # gráf kialakítása
x<-c("red", "green", "blue", "orange", "purple") # színek meghatározása

x1<-seq(1,90,18) # színezés
for (i in 1:length(x1)){
  V(g)$color[x1[i]:(x1[i]+17)]<-x[i]
}

# ábra
x1<-which(degree(g)==max(degree(g))) # azok a csúcsok legyen a gyökérben, amelyeknek a fokszáma a legnagyobb
plot(g, layout=layout_as_tree(g, root = x1), margin=-0.35, vertex.size=5, vertex.label.cex=0.1,
     vertex.color=adjustcolor(V(g)$color, alpha.f = 1))
legend("bottomright", legend = c("CZ", "HU", "PL", "RO", "SK"), fill = x, title = "Ország")
```

R15 A regionális járműipari import országok és termékek szerinti HVG hálózatának PageRank pontszáma (ábra)

```
v5_prod_pr<-sapply(c(1:5), function(x) which(grepl(countries[x],
                                       sort(page.rank(graph.adjacency(g_network2))$vector)
                                       %>% names)==TRUE))
colnames(v5_prod_pr)<-countries

melt(v5_prod_pr) %>%
  ggplot(aes(x=value, y=(as.numeric(sapply(c(1:5), function(x) rep(x, 18)))),
             color=Var2)) +
  geom_point(size=12) +
  scale_color_manual(values=c("red", "green", "blue", "orange", "purple")) +
  xlab("PageRank pontszám") + ylab("") + labs(color="Országok") +
  theme(axis.text = element_text(size=14), axis.title = element_text(size=16),
        legend.text = element_text(size=12), legend.title = element_text(size=14),
        axis.text.y = element_blank())
```

R16 Az országszinten aggregált gépjárműipari import HVG gráfja

```
# input mátrix előállítása
v5<-comex %>% filter(flow=="import") %>% group_by(year, month, reporter) %>%
  summarise(value=sum(value)) %>% spread(reporter, value)
v5<-v5[,-(1:2)]
v5<-ts(as.data.frame(v5), start = c(2004,1), frequency = 12)

network_sequence(v5) # HVG gráf kibontása
g<-graph.adjacency(g_network2, mode = "undirected") # gráf objektum felvétele

x<-c("red", "green", "blue", "orange", "purple") # színek
V(g)$color<-x

x1<-which(degree(g)==max(degree(g)))
plot(g, margin=-0.35, vertex.label.color="blue",
     vertex.label.cex=1, vertex.label.dist=2,
     vertex.color=adjustcolor(V(g)$color, alpha.f = 1))

sort(page.rank(graph.adjacency(g_network2))$vector) # PageRank pontszám
```

R17 Hozzáadott értékek tranzakciók hőtérképe

```
# adatok szerkezetének átalakítása
dva_int_plot<-dva_int[,-c(1:2)]
dva_int_plot[is.na(dva_int_plot)]<-0
dva_int_plot<-apply(dva_int_plot, 2, as.numeric)
dva_int_plot<-as.data.frame(dva_int_plot)
dva_int_plot$V1<-as.factor(dva_int[,1])
dva_int_plot<-dva_int_plot[-which(dva_int[,1]=="ROW"),]
dva_int_plot<-dva_int_plot %>% group_by(V1) %>% summarise_each(funs(sum))
dva_int_plot$V1<-NULL
dva_int_plot<-as.data.frame(dva_int_plot)
rownames(dva_int_plot)<-colnames(dva_int_plot)

# oszlopok szerint sorba rendezés (beszállítás alapján)
dva_int_plot<-as.data.frame(t(dva_int_plot))
dva_int_plot$sum<-colSums(dva_int_plot)
dva_int_plot<-dva_int_plot[order(dva_int_plot$sum),]
dva_int_plot$sum<-NULL
dva_int_plot<-as.data.frame(t(dva_int_plot))

# ábra
dva_int_plot[dva_int_plot==0]<-1

heatmap(log(as.matrix(dva_int_plot)), Colv = NA, Rowv = NA, scale = "none",
        col=topo.colors(5), revC=TRUE)
legend(x="topright", legend=c("min", "Q1", "medián", "Q2", "max"), fill=topo.col-
ors(5))
```

R18 Modularitás szegmentáció

```
# adatok szerkezetének átalakítása
g<-dva_int[,-c(1:2)]
g[is.na(g)]<-0
g<-apply(g, 2, as.numeric)
g<-as.data.frame(g)
g$V1<-as.factor(dva_int[,1])
g<-g[-which(dva_int[,1]=="ROW"),]

g<-g %>% group_by(V1) %>% summarise_each(funs(sum))
g$V1<-NULL
g<-as.data.frame(g)
rownames(g)<-colnames(g)

# gráf objektummá alakítás és klaszterezés
g_hc<-graph.adjacency(as.matrix(g), weighted = TRUE, mode="undirected", diag=FALSE)
kg<-fastgreedy.community(g_hc)
```

R19 A hierarchikus klaszterek gráfai

```
g<-dva_int[,-c(1:2)]
g[is.na(g)]<-0
g<-apply(g, 2, as.numeric)
g<-as.data.frame(g)
g$V1<-as.factor(dva_int[,1])
g<-g[-which(dva_int[,1]=="ROW"),]

g<-g %>% group_by(V1) %>% summarise_each(funs(sum))
g$V1<-NULL
g<-as.data.frame(g)
rownames(g)<-colnames(g)

g_hc<-graph.adjacency(as.matrix(g), weighted = TRUE, mode="undirected", diag=FALSE)
kg<-fastgreedy.community(g_hc)

x<-cbind(melt(g), rep(rownames(g), nrow(g)))
colnames(x)<-c("reporter", "value", "partner")
y<-as.data.frame(as.matrix(membership(kg)))
y$country<-rownames(y)
y<-as.data.frame(cbind(y$country, y$V1))
x<-merge(x, y, by.x = "reporter", by.y = "V1")
x<-merge(x, y, by.x = "partner", by.y = "V1")

# 1.klaszter
x1<-subset(x, V2.x==1 & V2.y==1)
x1<-x1[-c(which(x1$partner=="USR" | x1$reporter=="USR")),]
x1<-x1[-c(which(x1$partner==x1$reporter)),]
x1<-x1[-c(which(log(x1$value)<quantile(log(x1$value))[3])),]
g<-graph.edgelist(as.matrix(x1[,1:2]))
g<-set_edge_attr(g, "weight", value=x1$value)
plot(g,  edge.arrow.size=0.3,  vertex.size=log(strength(g,  mode  =  "all",  loops  =
FALSE)),
     layout=layout.kamada.kawai, edge.color="gray80", margin=-0.4)
sort(page.rank(g)$vector)
```

```
# 2.klaszter
x2<-subset(x, V2.x==2 & V2.y==2)
x2<-x2[-c(which(x2$partner==x2$reporter)),]
x2<-x2[-c(which(log(x2$value)<quantile(log(x2$value))[3])),]
g<-graph.edgelist(as.matrix(x2[,1:2]))
g<-set_edge_attr(g, "weight", value=x2$value)
plot(g, edge.arrow.size=0.3, vertex.size=log(strength(g, mode = "all", loops = FALSE)),
     layout=layout.kamada.kawai, edge.color="gray80", margin=-0.4, vertex.color="#ed7d31",
     vertex.label.color="black")
sort(page.rank(g)$vector)

# 3.klaszter
x3<-subset(x, V2.x==3 & V2.y==3)
x3<-x3[-c(which(x3$partner==x3$reporter)),]
x3<-x3[-c(which(log(x3$value)<quantile(log(x3$value))[3])),]
g<-graph.edgelist(as.matrix(x3[,1:2]))
g<-set_edge_attr(g, "weight", value=x3$value)
plot(g, edge.arrow.size=0.3, vertex.size=log(strength(g, mode = "all", loops = FALSE)),
     layout=layout.kamada.kawai, edge.color="gray80", margin=-0.4, vertex.color="#548235",
     vertex.label.color="white")
sort(page.rank(g)$vector)

# 4.klaszter
x4<-subset(x, V2.x==4 & V2.y==4)
x4<-x4[-c(which(x4$partner==x4$reporter)),]
x4<-x4[-c(which(log(x4$value)<quantile(log(x4$value))[3])),]
g<-graph.edgelist(as.matrix(x4[,1:2]))
g<-set_edge_attr(g, "weight", value=x4$value)
plot(g, edge.arrow.size=0.3, vertex.size=log(strength(g, mode = "all", loops = FALSE)),
     layout=layout.kamada.kawai, edge.color="gray80", margin=-0.4, vertex.color="#ffc000",
     vertex.label.color="black")
sort(page.rank(g)$vector)
```

R20 Harmadik klaszter PageRank centralitásai

```
sort(page.rank(g)$vector, decreasing = TRUE)[1:20] %>% barplot
```

R21 Hálózatok mélységi és szélességi feltárása

```
function(country, ratio1, simplify=TRUE) {
  require(igraph)
  require(DescTools)

  #Regional direct partners
  c1<-min(which(labels_T$Country2==country))
  c2<-max(which(labels_T$Country2==country))
  country_total<-sum(apply(dva_total2[c1:c2,], 2, sum))

  if (cumsum(sort(apply(dva_total2[c1:c2,], 2, sum), decreasing = TRUE))[1]/country_total>ratio1) {
    stop("Ratio1 is too low, regional network is higly concentrated. Try higher ratio1 value.") }
```

```
  regional_1_names<-names(which(cumsum(sort(apply(dva_total2[c1:c2,], 2, sum), de-
creasing = TRUE))/country_total<ratio1))
  regional_1_names<-append(country, regional_1_names)
  regional_1_values<<-sort(apply(dva_total2[c1:c2,],     2,     sum),     decreas-
ing=TRUE)[1:length(regional_1_names)]

  potential_names<-setdiff(country_names, regional_1_names)
  potential_names<-potential_names[-which(potential_names=="ROW")]
  potential_values<<-matrix(nrow=length(potential_names),          ncol=length(re-
gional_1_names), data=0)
  rownames(potential_values)<<-potential_names
  colnames(potential_values)<<-regional_1_names


  for (i in 1:length(potential_names)) {
    for (k in 1:length(regional_1_names)) {
      c1<-min(which(labels_T$Country2==regional_1_names[k]))
      c2<-max(which(labels_T$Country2==regional_1_names[k]))
      c3<-which(colnames(dva_total2)==potential_names[i])
      potential_values[i,k]<<-sum(dva_total2[c1:c2,c3]) }}

  potential_values<<-potential_values[-as.numeric(which(apply(potential_values,  1,
max)<min(regional_1_values))),]

  indirect_names<-names(which(apply(potential_values,      1,      Gini)<=Gini(re-
gional_1_values)))

  network_names<-append(regional_1_names, indirect_names) #oszlop: beszállító, sor:
partner
  print(regional_1_names)
  print(indirect_names)

  sub_network<<-matrix(nrow=length(network_names),     ncol=length(network_names),
data=NA)
  colnames(sub_network)<<-network_names
  rownames(sub_network)<<-network_names

  for (i in 1:length(network_names)) {
    for (k in 1:length(network_names)) {
      c1<-min(which(labels_T$Country2==network_names[i]))
      c2<-max(which(labels_T$Country2==network_names[i]))
      c3<-which(colnames(dva_total2)==network_names[k])
      sub_network[i,k]<<-sum(dva_total2[c1:c2,c3]) }}

  cat("Regional coverage: ", sum(regional_1_values)/country_total, "\nGlobal cover-
age: ", sum(sub_network[1,])/sum(dva_total2))

  if (simplify==TRUE) {
    sub_network[sub_network<min(regional_1_values)]<<-0 }

  sub_network[upper.tri(sub_network)]<<-0

  delete<-as.numeric(which(rowSums(sub_network)==0 & colSums(sub_network)==0))

  if (length(delete)>0) {
    sub_network<<-sub_network[-delete,-delete] }

  g<<-graph.adjacency(sub_network, mode="directed", weighted = TRUE)
```

```
  E(g)$width<<-E(g)$weight*5/max(E(g)$weight)
  V(g)$size<<-(rowSums(sub_network)+colSums(sub_network))
  V(g)$size<<-V(g)$size*(20/max(V(g)$size))+10
  #V(g)$size<<-log(rowSums(sub_network)+colSums(sub_network), exp(1))/15
  #V(g)$size<<-V(g)$size/mean(V(g)$size)

  plot(g, edge.arrow.size=0.3, layout=layout.auto)
}
barplot(page.rank(g)$vector %>% sort(decreasing = TRUE), main="") # PageRank értékek
ábrázolása
```

R22 Magyarországra beáramló HÉ-k

```
subnetwork("HUN", 0.9, simplify = FALSE)

# Magyarországra áramló BHÉ
g_hun_to<-delete.edges(g, setdiff(seq(1, gsize(g), 1), E(g)[to("HUN")] %>% as.nu-
meric))
g_hun_to<-delete.vertices(g_hun_to, "BGR")
plot(as.undirected(g_hun_to), layout=layout.auto, margin=-0.4)
```

R23 Magyorszság fókuszú mélységi és szélességi feltárás evolúciós dendrogramja

```
# A távolsági mátrix megszerkesztése Excel-ben történt (dendro2.csv)
dendo3<-as.data.frame(dendo2)
for (i in 1:ncol(dendo2)) {
  for (k in 1:ncol(dendo2)) {
    dendo3[i,k]<-dendo3[k,i]
  }
}
dendo3[is.na(dendo3)]<-7

hclust(dist(dendo3), method="complete") %>% plot(main="", xlab="",
                                                 ylab="Magasság       (normalizált
hasonlóság)")
```

R24 Mélységi és szélességi feltárás a régióban (egy ország példáján keresztül)

```
# Magyarország gráfja
subnetwork("HUN", ratio1 = 0.9, simplify = TRUE)
g_hun<-g

# Lengyelország
subnetwork("POL", ratio1 = 0.9, simplify = TRUE)
plot(g, margin=-0.3, edge.arrow.size=0.3)
g_pl<-g
x<-setdiff(rownames(as.matrix(V(g_pl))), rownames(as.matrix(V(g_hun))))
y<-page.rank(g_pl)$vector
sapply(c(1:length(x)), function(k) y[which(names(y)==x[k])]) %>% sort %>% round(dig-
its = 3)
```

R25 Hozzáadott érték export folyamok feltárása

```
flow<-function(origin_country, destination_country, round=1, industry="all") {
  library(dplyr)
 # országindexek meghatározása
  c1<-min(which(labels_T$Country2==origin_country))
  c2<-max(which(labels_T$Country2==origin_country))
  c3<-min(which(labels_T$Country2==destination_country))
  c4<-max(which(labels_T$Country2==destination_country))

  # közvetlenül beszállított hozzáadott érték
  if (industry=="all") { # minden iparágra
    export_direct<-T[c1:c2,c3:c4] %>% rowSums
  } else { # csak egy iparágra
    export_direct<-T[c1:c2,c3:c4] %>% rowSums
    z<-which(labels_T$Industry==industry)[1]
    x<-export_direct
    export_direct<-export_direct*0
    export_direct[z]<-x[z]
  }

  dva_direct<<-diag(VAsh[c1:c2])%*%B[c1:c2,c1:c2]%*%export_direct

  # közvetve beszállított hozzáadott érték első kör
  dva_indirect_1st<-matrix(nrow=26, ncol=length(country_names)) #26 ágazat, összes
ország-ROW
  rownames(dva_indirect_1st)<-labels_T$Industry[1:26]
  colnames(dva_indirect_1st)<-country_names

  for (i in 1:length(country_names)) {
    if (i==which(country_names=="ROW")) { # ROW-val nem foglalkozunk
      i<-i+1
    }
    # közvetett partnerország megkeresése
    c5<-min(which(labels_T$Country2==country_names[i]))
    c6<-max(which(labels_T$Country2==country_names[i]))

    if (industry=="all"){
      export_indirect<-T[c5:c6,c3:c4] %>% rowSums
    } else {
      z<-which(labels_T$Industry==industry)[1]
      export_indirect<-T[c5:c6,c3:c4] %>% rowSums
      x<-export_indirect
      export_indirect<-export_indirect*0
      export_indirect[z]<-x[z]
    }

    dva_indirect_1st[,i]<-diag(VAsh[c1:c2])%*%B[c1:c2,c5:c6]%*%export_indirect
  }
  dva_indirect_1st<<-dva_indirect_1st[,-which(country_names=="ROW")]

  if (round>1) { # ha egynél több körre van szükség
    dva_indirect_list<-list(dva_indirect_1st)
    combi<-combn(189, 2) #kombinációk létrehozása
    combi<-combi[,-which(combi==145,   arr.ind   =   TRUE)[,2]]   #   ROW   kiszűrési
(ROW=145.elem)

    # origin és destination kiszűrése
```

```
    combi<-combi[,-which(combi==which(country_names==origin_country),   arr.ind   =
TRUE)[,2]]
    combi<-combi[,-which(combi==which(country_names==destination_country),
                         arr.ind = TRUE)[,2]]

    # Korea kiszűrése
    combi<-combi[,-which(combi==which(country_names=="KOR"), arr.ind = TRUE)[,2]]

    for (i in 1:c(ncol(combi))) {
      # két közbenső partner
      c5<-min(which(labels_T$Country2==country_names[combi[1,i]])) #S
      c6<-max(which(labels_T$Country2==country_names[combi[1,i]]))
      c7<-min(which(labels_T$Country2==country_names[combi[2,i]])) #T
      c8<-max(which(labels_T$Country2==country_names[combi[2,i]]))

      # U-origin, R-destination
      if (industry=="all") { # ha minden iparág szükséges
        export_us<-T[c1:c2,c5:c6] %>% rowSums
        dva_indirect_ust<-diag(VAsh[c1:c2])%*%B[c7:c8,c1:c2]%*%export_us
      } else { # ha csak egy iparág szükséges
        export_us<-T[c1:c2,c5:c6] %>% rowSums
        z<-which(labels_T$Industry==industry)[1]
        x<-export_us
        export_us<-export_us*0
        export_us[z]<-x[z]
        dva_indirect_ust<-diag(VAsh[c1:c2])%*%B[c7:c8,c1:c2]%*%export_us
      }

      # megnézzük, hogy a VA hogyan aránylik a bruttó exporthoz
      dvax<-dva_indirect_ust/export_us
      dvax[is.infinite(dvax)]<-0
      dvax<-dvax*(T[c7:c8,c5:c6] %>% rowSums)

      # ezzel az aránnyal korrigáljuk a kéz közbenső ország exportját
      dva_indirect_trs<-diag(VAsh[c5:c6])%*%B[c5:c6,c7:c8]%*%dvax

      # eredeti, nem korrigált verzió, csak VA-val felszorozva
      # erősebben lokalizált eredményt ad
      #dva_indirect_trs<-diag(VAsh[c5:c6])%*%B[c5:c6,c7:c8]%*%dva_indirect_ust
      name_path<-paste(origin_country,country_names[combi[1,i]],coun-
try_names[combi[2,i]],
              destination_country, sep="_")
      dva_indirect_list[[i]]<-dva_indirect_trs
      names(dva_indirect_list)[i]<-name_path
    }
    # Másik irányból is
    combi2<-combi
    combi2[1,]<-combi[2,]
    combi2[2,]<-combi[1,]

    for (i in 1:c(ncol(combi2))) {
      # két közbenső partner
      c5<-min(which(labels_T$Country2==country_names[combi2[1,i]])) #S
      c6<-max(which(labels_T$Country2==country_names[combi2[1,i]]))
      c7<-min(which(labels_T$Country2==country_names[combi2[2,i]])) #T
      c8<-max(which(labels_T$Country2==country_names[combi2[2,i]]))

      # U-origin, R-destination
      if (industry=="all") {
```

```
        export_us<-T[c1:c2,c5:c6] %>% rowSums
      } else {
        export_us<-T[c1:c2,c5:c6] %>% rowSums
        z<-which(labels_T$Industry==industry)[1]
        x<-export_us
        export_us<-export_us*0
        export_us[z]<-x[z]
      }

      dva_indirect_ust<-diag(VAsh[c1:c2])%*%B[c7:c8,c1:c2]%*%export_us
      dvax<-(dva_indirect_ust/export_us)
      dvax[is.infinite(dvax)]<-0
      dvax<-dvax*(T[c7:c8,c5:c6] %>% rowSums)
      dva_indirect_trs<-diag(VAsh[c5:c6])%*%B[c5:c6,c7:c8]%*%dvax

      #dva_indirect_trs<-diag(VAsh[c5:c6])%*%B[c5:c6,c7:c8]%*%dva_indirect_ust
      name_path<-paste(origin_country,country_names[combi2[1,i]],coun-
try_names[combi2[2,i]],
                       destination_country, sep="_")
      dva_indirect_list[[i]]<-dva_indirect_trs
      names(dva_indirect_list)[i]<-name_path
    }
  }
  if (round>1) {
    dva_indirect_list<<-dva_indirect_list
  }
}

flow("HUN", "DEU", 2, industry = labels_T$Industry[10])
sort(colSums(dva_indirect_1st), decreasing = TRUE)[1:20]
x<-as.data.frame(dva_indirect_list)
sort(colSums(x), decreasing = TRUE)[1:20]

### Térkép -------------------------------------------------------------
library(maps)
library(ggplot2)
midpoints<-read.csv("g:/country_midpoints.csv", sep=";") # ország középpontok
midpoints$latitude<-midpoints$latitude %>% as.numeric
midpoints$longitude<-midpoints$longitude %>% as.numeric

# javítások
midpoints$latitude[midpoints$Code2=="COL"]<-4.570868
midpoints$longitude[midpoints$Code2=="COL"]<--74.297333
midpoints$latitude[midpoints$Code2=="HRV"]<-45.1
midpoints$longitude[midpoints$Code2=="HRV"]<-15.2
midpoints$latitude[midpoints$Code2=="CZE"]<-49.817492
midpoints$longitude[midpoints$Code2=="CZE"]<-15.472962
midpoints$latitude[midpoints$Code2=="KOR"]<-35.907757
midpoints$longitude[midpoints$Code2=="KOR"]<-127.766922
midpoints$latitude[midpoints$Code2=="USA"]<-37.09024
midpoints$longitude[midpoints$Code2=="USA"]<--95.712891

### ...HUN-DEU_all_1st_round ---------------------------------------------
o<-"HUN"
d<-"DEU"
flow(o, d, round = 1)
x<-sort(colSums(dva_indirect_1st), decreasing = TRUE)[3:150]

points<-matrix(nrow = 1, ncol=7)
```

```
colnames(points)<-c("origin", "origin_coordinates_x", "origin_coordinates_y",
                    "destination", "destination_coordinatates_x",
                    "destination_coordinatates_y",
                    "weight")

points[1,]<-cbind(o,
                  midpoints$latitude[which(midpoints$Code2==o)],
                  midpoints$longitude[which(midpoints$Code2==o)],
                  d,
                  midpoints$latitude[which(midpoints$Code2==d)],
                  midpoints$longitude[which(midpoints$Code2==d)],
                  sum(dva_direct))

for (i in 2:length(x)){
  points2<-cbind(o,
                    midpoints$latitude[which(midpoints$Code2==o)],
                    midpoints$longitude[which(midpoints$Code2==o)],
                    names(x[i]),
                    midpoints$latitude[which(midpoints$Code2==names(x[i]))],
                    midpoints$longitude[which(midpoints$Code2==names(x[i]))],
                    as.numeric(x[i]))
  points<-rbind(points, points2)
  rm(points2)
}
points<-as.data.frame(points)
points$weight<-as.numeric(points$weight)
points$norm_weight<-(points$weight-min(points$weight))/(max(points$weight)-
min(points$weight))

# Térkép

worldmap<-borders("world", colour = "#FFD127", fill="#FFD127") # alaptérkép

ggplot() + worldmap +
  geom_curve(data = points, aes(x=as.numeric(origin_coordinates_y),
                                y=as.numeric(origin_coordinates_x),
                                xend=as.numeric(destination_coordinatates_y),
                                yend=as.numeric(destination_coordinatates_x),
                                alpha=log(as.numeric(weight))),
             col="black", size=1.5, curvature = 0.2) +
  scale_alpha_continuous(c(0.1,1))
```

R26 Hozzáadott érték folyamok feltárása Magyarország esetében második kör

```
o<-"HUN"
d<-"DEU"
flow(o, d, round = 2)
x<-as.data.frame(dva_indirect_list)
x<-sort(colSums(x), decreasing = TRUE)[1:20]

points<-matrix(nrow=1, ncol=13)

x_split<-str_split(names(x[1]), '_')[[1]]
points[1,]<-cbind(x_split[1],
                  midpoints$latitude[which(midpoints$Code2==x_split[1])],
                  midpoints$longitude[which(midpoints$Code2==x_split[1])],
                  x_split[2],
                  midpoints$latitude[which(midpoints$Code2==x_split[2])],
                  midpoints$longitude[which(midpoints$Code2==x_split[2])],
```

```
                  x_split[3],
                  midpoints$latitude[which(midpoints$Code2==x_split[3])],
                  midpoints$longitude[which(midpoints$Code2==x_split[3])],
                  x_split[4],
                  midpoints$latitude[which(midpoints$Code2==x_split[4])],
                  midpoints$longitude[which(midpoints$Code2==x_split[4])],
                  as.numeric(x[1]))

for (i in 2:length(x)) {
  x_split<-str_split(names(x[i]), '_')[[1]]
  points2<-cbind(x_split[1],
                 midpoints$latitude[which(midpoints$Code2==x_split[1])],
                 midpoints$longitude[which(midpoints$Code2==x_split[1])],
                 x_split[2],
                 midpoints$latitude[which(midpoints$Code2==x_split[2])],
                 midpoints$longitude[which(midpoints$Code2==x_split[2])],
                 x_split[3],
                 midpoints$latitude[which(midpoints$Code2==x_split[3])],
                 midpoints$longitude[which(midpoints$Code2==x_split[3])],
                 x_split[4],
                 midpoints$latitude[which(midpoints$Code2==x_split[4])],
                 midpoints$longitude[which(midpoints$Code2==x_split[4])],
                 as.numeric(x[i]))
  points<-rbind(points, points2)
  rm(points2)
}

# Térkép
colnames(points)<-c("origin", "o_y", "o_x", "p1", "p1_y", "p1_x", "p2", "p2_y",
"p2_x",
                    "destination", "d_y", "d_x", "weight")
points<-as.data.frame(points)
points$weight<-as.numeric(points$weight)
points$norm_weight<-(points$weight-min(points$weight))/
  (max(points$weight)-min(points$weight))

points_o1<-points[,1:6] %>% as.matrix
points_o2<-points[,4:9] %>% as.matrix
points_o3<-points[,7:12] %>% as.matrix
points_m<-rbind(points_o1, points_o2, points_o3)
points_m<-points_m[-which(duplicated(points_m)),]
points_m<-as.data.frame(points_m)
points_m$c<-"Export első kör"
points_m$c[points_m$p1==d]<-"Export második kör"

#points_m<-points_m[-which(points_m$p1==d),] # Final dest kivétele

worldmap<-borders("world", colour = "#FFD127", fill="#FFD127") # alaptérkép

ggplot() + worldmap +
  geom_curve(data=points_m, aes(x=as.numeric(o_x), y=as.numeric(o_y), xend=as.nu-
meric(p1_x),
                                yend=as.numeric(p1_y), color=as.factor(c)),
             size=1, curvature = 0.1, arrow = arrow(type = "closed", length=unit(2,
"mm")))+
  labs(color="HÉ iránya")
```