CHARLES UNIVERSITY IN PRAGUE

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Milan Hanousek

International Trade Network

Master Thesis

Prague 2014

Author: **Bc. Milan Hanousek**
Supervisor: **PhDr. Ladislav Krištoufek, Ph.D.**
Academic year: **2013/2014**

Declaration of Authorship

I hereby proclaim that I wrote this master thesis using literature and other sources listed in the reference list. Furthermore, I declare that this thesis was not used to obtain another academic degree.

I also grant a permission to lend and reproduce the thesis for study and research purposes.

Prague, July 15, 2014

Signature

Acknowledgments

I would like to express my deep gratitude to Ladislav Krištoufek for his guidance, recommendations and comments. Furthermore, I would like to thank to Matúš Baniar for his assistance with Wolfram Mathematica. Finally, I give thanks to my family and friends for their support and encouragement.

Bibliographic Record

HANOUSEK, Milan. *International Trade Network*. Prague 2014. 66 p. Master Thesis. Charles University in Prague, Faculty of Social Sciences, Institute of Economic Studies. Supervisor: PhDr. Ladislav Krištoufek, Ph.D.

Abstract

This paper studies the topological properties of the International Trade Network (ITN) among world countries using a network analysis. We explore the distributions of the most important network statistics measuring connectivity, assortativity and clustering. We show that the topological properties of the weighted representation of the ITN are very different from those obtained by a binary network approach. In particular, we find that: (i) the majority of countries are characterized by weak trade relationships, (ii) well connected countries tend to trade with poorly connected partners and (iii) countries holding more intense trade relationships are more clustered. Finally, we display that all structural properties of the ITN have remained remarkably stable over time.

Keywords

Econophysics, international trade network, topological properties, stylized facts, binary network, weighted network, network statistics

JEL Classification

D85, F10

Bibliografický záznam

HANOUSEK, Milan. *Síť mezinárodního obchodu*. Praha 2014. 66 s. Magisterská práce. Univerzita Karlova v Praze, Fakulta sociálních věd, Institut ekonomických studií. Vedoucí práce: PhDr. Ladislav Krištoufek, Ph.D.

Abstrakt

Tato práce studuje topologické vlastnosti sítě mezinárodního obchodu (ITN) mezi světovými zeměmi pomocí síťové analýzy. Zkoumáme rozdělení nejdůležitějších síťových statistik, které měří propojenost, uspořádání a shlukování. Ukazujeme, že topologické vlastnosti vážené reprezentace ITN jsou velmi odlišné od těch, které signalizuje binární síťový přístup. Konkrétně znázorňujeme, že: (i) většina zemí je charakterizována slabými obchodními vztahy, (ii) dobře propojené země mají tendenci obchodovat se slabě propojenými partnery a (iii) země držící intenzivnější obchodní vztahy jsou více shlukovány. Na závěr demonstrujeme, že všechny strukturální vlastnosti ITN jsou velmi stabilní v průběhu času.

Klíčová slova

Ekonofyzika, síť mezinárodního obchodu, topologické vlastnosti, stylizovaná fakta, binární síť, vážená síť, síťové statistiky

JEL Klasifikace

D85, F10

Contents

List of Figures

List of Tables

Acronyms

ANND	Average Nearest-Neighbour Degree
ANNS	Average Nearest-Neighbor Strength
BCC	Binary Clustering Coefficient
CCDF	Complementary Cumulative Distribution Function
CDF	Cumulative Distribution Function
CIA	Central Intelligence Agency
DOTS	Direction of Trade Statistics
FTA	Free Trade Agreement
GM	Gravity Model
GDP	Gross Domestic Product
GDPpc	Gross Domestic Product Per Capita
GNP	Gross National Product
HHI	Herfindahl-Hirschman Index
i.i.d.	Independently and Identically Distributed
IMF	International Monetary Fund
ITN	International Trade Network
ND	Node Degree
NS	Node Strength
PWT	Penn World Tables
US	United States
USD	United States Dollar
WANND	Weighted Average Nearest-Neighbor Degree
WCC	Weighted Clustering Coefficient
WED	World Export Data

WTW World Trade Web

WWW World Wide Web

Master Thesis Proposal

Author:	Bc. Milan Hanousek
Supervisor:	PhDr. Ladislav Krištoufek, Ph.D.
Proposed topic:	International Trade Network: Beyond Gravity Models

Topic Characteristics

The researchers have been recently contributing to the modeling of economic and financial behavior by using tools and methods developed in statistical physics. The interaction between economics and physics has given rise to the field "econophysics". This new interdisciplinary field might be very helpful in modeling and analyzing various financial systems like trading, banking, stock markets etc. Econophysics uses (1) non-linearity, (2) scaling laws, (3) statistical mechanics and (4) the entire family of stable distributions to explain economic and financial behavior more robustly than traditional economic and financial tools. Most of the work is focused on understanding statistical features of financial data. It is possible that financial data viewed from a different perspective might yield new results.

The object of this thesis is to analyze the structure, function and dynamics of international trade. We will show that the international trade network (ITN) can be looked upon as the weighted network obeying the scale invariance and the universality. The standard (indeed the only) model used to examine the international trade is the gravity model. The name comes from an analogy with Newton's law of gravitation, thus the model can be seen as the older example of the interaction between economics and physics. The model acquired its great popularity, because it reproduces well observed trade flows between countries. On the other hand, the model has some serious and irreducible limitations which emerged especially after the publication of several empirical analyses showing the topology of the ITN. For that reason we will try to develop a new dynamical model in the context of the "network" approach.

Hypothesis

1. The ITN is the weighted network obeying the scale invariance and the universality.

2. The gravity model fails to replicate the empirical features of the ITN.

3. The dynamical model based on the gravity law perfectly reproduces empirical features of the ITN.

Methodology

At first, we will perform a detailed analysis of the real data of the ITN. We will look at the evolution of nodes and links, degree distribution, distribution of link weights and strength of nodes. Secondly, we will estimate the gravity model and build predictions for the properties of the ITN. The traditional approach for estimating the gravity model includes a logarithmic transformation of gravity equation, however some authors show that the log-linearized estimation method can lead to a highly misleading result. The proper method is to estimate the gravity model in its multiplicative form using a Poisson pseudo-maximum likelihood (PPML) estimator. We will examine the international trade for several last years, thus we will use various panel data models, i. e. pooled ordinary least squares (OLS) model, fixed effects (FE) model and random effects (RE) model. To find out the most appropriate model we will apply some specification tests. The statistical features of the predicted ITN will be then compared to those observed in the real ITN. Finally, we will develop a dynamical model which will have the gravity law as a starting point. After its estimation we will repeat the procedure of comparison between the reproduced ITN and the real ITN.

Outline

1. Introduction

2. Statistical properties of the international trade network

3. Gravity model

4. Dynamical model based on the gravity law

5. Conclusion

Core Bibliography

1. ABERGEL, Frederic., Hideaki AOYMA., Bikas K. CHAKRABARTI., Anirban CHAKRABORTI. and Asin GHOSH. *Econophysics of Agent-Based Models*. Springer International Publishing, 2014.

2. BHATTACHARYA, Kankan., Goutam MUKHERJEE., Jari SARAMAKI., Kimmo KASKI. and Subhrangshu S. MANNA. *The International Trade Network: weighted network analysis and modelling*. Journal of Statistical Mechanics: Theory and Experiment, 2008.

3. CHATTERJEE, Arnab. and Bikas K. CHAKRABARTI. *Econophysics of Markets and Business Networks*. Springer Milan, 2007.

4. DUENAS, Marco. and Giorgio FAGIOLO. *Modeling the International Trade-Network: a gravity approach*. Journal of Economic Interaction and Coordination, Vol. 8, No. 1, 155-178, 2013.

5. FAGIOLO, Giorgio. *The international-trade network: gravity equations and topological properties*. Journal of Economic Interaction and Coordination, Vol. 5, No. 1, 1-25, 2010.

6. MANTEGNA, Rosario N. and H. Eugene STANLEY. *An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge: Cambridge University Press, 2000.

7. SANTOS SILVA, Joao M. C. and Silvana TENREYRO. *The Log of Gravity*. CEP Discussion Paper, No. 701, 2005.

8. VOIT, Johannes. *The Statistical Mechanics of Financial Markets*. Springer Berlin Heidelberg, 2005.

9. WESTERLUND, Joakim. and Fredrik WILHELMSSON. *Estimating the gravity model without gravity using panel data*. Applied Economics, No. 43, 2011.

Author Supervisor

Chapter 1

Introduction

Two last decades have witnessed the emergence of a large body of papers utilizing methods and tools from statistical physics to explain economic behavior. Statistical physics describes the complex behavior observed in many physical systems in terms of their simple basic constituents and simple interaction laws. Complexity arises from the interaction and disorder and from the cooperation and competition of basic units. Financial markets are certainly complex systems, which is judged both by their output and structure. A growing number of physicists have therefore attempted to analyze and model financial markets and more generally economic systems. The interest of physical community in economic systems has given rise to the field of "econophysics". This new interdisciplinary field generally uses non-linearity, scaling laws, statistical mechanics and the entire family of stable distributions to explain economic behavior more robustly than traditional economic tools. The econophysics has proved to be especially fruitful in a research of the structure and function of complex economic network systems like trading, banking, stock markets and so on.

Over the last two decades, there has been also an increasing interest in the study of networks across many scientific disciplines. The study of networks has primarily flourished thanks to contributions stemming from mathematics, physics and computer science. With new powerful tools researchers have begun to explore statistical properties of biological, information and technological networks [1, 2, 3, 4]. These new methods have been naturally applied to social and economic systems [5]. The added value of using the network approach to economic problems is the possibility to investigate indirect effects arising as the combination of many pairwise interactions between economic agents. The idea that economic systems like trading, banking and stock markets might be considered as network structures have been increasingly accepted. However, much earlier studies recognized that socio-economic systems can be described as networks [6, 7]. In fact, psychologists and sociologists have employed social network analysis since the beginning of the 20th century to explore the interactions established among people or groups [8].

The network approach has been recently used in empirical studies of international trade [9, 10, 11, 12, 13, 14, 15, 16, 17]. The idea is to describe trade relations as a network, where countries play the role of nodes and the presence of an export/import relation between any two countries is described by a link. Such a network is called International Trade Network (ITN) or World Trade Web (WTW). Understanding the topological properties of the ITN allows for a better description of processes such as economic globalization and internalization [18, 19]. The standard approach to the empirics of international trade employs indicators, which characterize the profile of a country by referring only to its direct bilateral-trade relationships (direct export/import relationships). On the one hand, direct bilateral-trade linkages are known to be one of the most important ways of interaction between world countries [20]. For example, they can help to explain the extent to which economic policies affect foreign markets [21]. Alternatively, they can explain how economic shocks to any single country can be easily transmitted to countries that are relatively minor bilateral trading partners. Furthermore, they can help to stress global interdependencies that explain spreading of economic crises. On the other hand, direct bilateral-trade linkages can explain only a part of the effect that an economic shock arising in one country can have on another country that is not among its direct trade partners [22, 23]. For that reason, a complex network analysis [1, 2, 3, 4] that goes far beyond the standard indicators of international trade is required.

The earlier literature [9, 10, 11] exploring the topological properties of the ITN has employed a binary network analysis, where a link between any two countries is either present or not according to whether the trade flow that it carries is larger then a given threshold. More recent contributions [12, 13, 14, 15, 16, 17] have adopted a weighted network analysis, where each link between any two countries is weighted by some value of trade flow that it carries. The reason is that the binary approach treats all relationships equally, which might dramatically underestimate the role of heterogeneity in trade linkages. This heterogeneity might be crucial to better understand the architecture of complex networks, therefore the weighted analysis is better to grasp a more complete and truthful picture of the ITN.

In this thesis, we employ a network analysis to explore the topological properties of the ITN among large set of world countries over the period 1980-2000. Although the ITN is an excellent example of the weighted network, we perform both binary and weighted network analyses to get a more complete picture. We use international trade data to build a network of links between pair of countries in each year. This enables us to apply statistical network techniques and characterize some robust stylized facts of international trade. The empirical regularities revealed in the data allow us to understand the structure and evolution of the ITN. In other words, empirical regularities can provide a theoretical explanation of why the international trade is organized in this way. The main contribution of this thesis in a comparison with other studies is a thoroughness of the analysis. We provide a detailed theoretical background to the analysis of networks. Furthermore, we review all existing literature related to the topic. Finally, we present a comprehensive empirical analysis of the topological properties of the ITN, which includes several dimensions. First we check the directionality of the ITN to justify a directed or an undirected analysis. Second, we describe the distribution of the most important network statistics measuring connectivity, assortativity and clustering. Third, we study whether the empirical regularities of the ITN have been changing over time. Fourth, we compare the weighted network results to those obtained by a binary approach. Finally, we investigate the extent to which the topological properties of the ITN relate to country specific characteristics (GDP per capita).

The study is organized as follows. In chapter 2, we present the main concepts related to the analysis of networks. We specifically provide explanations as well as more formal definitions of the network statistics for both binary and weighted networks. Chapter 3 summarizes the relevant literature on the ITN. Chapter 4 describes the origin of data and the selection of the sample. The results are reported in Chapter 5. Finally, we conclude the thesis and discuss the future work in chapter 6.

The computations are performed in Wolfram Mathematica. All figures and tables provided in the study are also produced in this software. The document with the script called Wolfram Mathematica notebook necessary for understanding of computational processes is available upon request.

Chapter 2

Statistical Analysis

We study the topological properties of the ITN among world countries using a network analysis. This chapter provides the main concepts related to the empirical analysis of networks. We start with an introduction of basic notions. We continue with a description of the procedure to check whether a network is directed or undirected. We expect that the ITN is sufficiently symmetric, therefore we will be able to use tools for an undirected network analysis. The tools are initially summarized for binary networks and consequently extended to a weighted perspective. This statistical analysis is apart from other cited works based on [13, 14, 15, 16, 17, 24, 25, 26].

2.1 Basic Notions

A *network* is a set of *nodes* $\{1, 2, \ldots, N\}$ connected through *links*[1]. The network can be alternatively characterized by a $N \times N$ matrix $\tilde{M} = \{\tilde{m}_{ij}\}$, where the out-of-diagonal element $\tilde{m}_{ij}$ is non-zero if and only if a link from a node i to a node j is present. The diagonal elements $\tilde{m}_{ii}$ are either all different from zero or all equal to zero. It depends whether self-loops[2] are allowed or not. Networks are divided into *binary* and *weighted*. In binary networks, any two nodes are either connected by a link or not, i. e. $\tilde{m}_{ij} \in \{0, 1\}$. The matrix $\tilde{M}$ is then called an *adjacency matrix*. In weighted networks, each link is weighted by some proxy of the flow intensity that it carries. The non-zero element $\tilde{m}_{ij}$ measures the weight of the link from node i to node j and the resulting matrix $\tilde{M}$ is referred as a *weight matrix*.

[1] Standard terms of graph theory are a graph, vertices and edges, respectively.
[2] Self-loops are links, which connect a node to itself.

Both binary and weighted networks can be *directed* or *undirected*. In directed networks, all links are directed from one node to another. Directed networks are not symmetric meaning that there exists at least a pair of connected nodes in which one directed link is not reciprocated, i. e. $\exists(i,j), i \neq j : \tilde{m}_{ij} > 0$ and $\tilde{m}_{ji} = 0$. Analysing the topological properties of directed networks might be very complicated and convoluted, because one has to distinguish inward and outward links in computing network statistics. In undirected networks, all links are instead bilateral. All pairs of connected nodes mutually affect each other, i. e. $\forall(i,j), i \neq j : \tilde{m}_{ij}\tilde{m}_{ji} > 0$. Disregarding a direction of links greatly simplifies the analysis, since the tools are much more developed and understood for undirected networks.

The main issue is to empirically distinguish directed and undirected networks. If the empirical analysis considers mutual economic and social relationships (e. g. friendship, marriage, business partnership), the constructed matrix $\tilde{M}$ is symmetric and tools for undirected network analysis can be used. However, the majority of interaction relationships are notionally non-mutual. The constructed matrix $\tilde{M}$ (especially in a weighted case) is then hardly found to be symmetric. Strictly speaking such networks should be treated as directed. Since a directed network analysis is more difficult and convoluted, one should check whether the "amount of directedness" of the observed matrix justifies the use of a more complicated procedure.

A traditional way quantifying whether a network is sufficiently symmetric to justify an undirected analysis is to measure its *reciprocity* as the fraction of the number of reciprocated links $L_{\leftrightarrow}$ to the total number of directed links L_D [6, 9, 27]:

$$r = \frac{L_{\leftrightarrow}}{L_D}. \tag{2.1}$$

The reciprocity is zero for the fully-directed network, while it is one for the fully-undirected one. The value of r generally represents the average probability that a link is reciprocated. If this ratio is "reasonably" large, one can symmetrize the network and employ the appropriate tools for an undirected network. However, the above definition of the reciprocity poses various conceptual problems. First, the value of r has to be compared to the value of r^{rand} expected in a random network with the same number of nodes and links. The reason is an assessment whether mutual links occur more, less or just as often than expected by chance. Second, the definition (2.1) is heavily dependent on the *density* of the network, which is defined as the number of observed directed links L_D to the number of possible directed links $N(N-1)$ [6, 27]:

$$\rho = \frac{L_D}{N(N-1)}. \tag{2.2}$$

The value of r^{rand} is naturally larger in a network with larger density, because mutual links occur by chance more often in a network with more links. Finally, even in two networks with the same density the definition (2.1) can lead to inconsistent results if L_D contains the number of self-loops. In order to avoid aforementioned problems, Garlaschielli and Loffredo [24] propose a new definition of the reciprocity as the correlation coefficient between the entries of the adjacency matrix of a directed network. Unfortunately, there is even one more drawback. If a network is weighted, the reciprocity does not consider the effect of link weights. A bilateral link exists between node i and node j if and only if $\tilde{m}_{ij}\tilde{m}_{ji} > 0$. Of course, the sub-case where $m_{ij} >> 0$ and $m_{ji} \simeq 0$ is very different from the sub-case where $m_{ji} \simeq m_{ji} > 0$. To overcome this problem Fagiolo [25] develops a new index, which is introduced in the following section.

2.2 Checking for Symmetry

To ground a "directed vs undirected" decision, we have decided to employ an index developed by Fagiolo [25]. The index has two main properties. First, it can be applied with minor modifications to both binary and weighted networks. Second, the standardized version of the index follows a standardized normal distribution (over all possible adjacency/weight matrices).

Suppose a directed weighted network $\tilde{G} = (N, \tilde{W})$, where N is the number of nodes and $\tilde{W} = \{\tilde{w}_{ij}\}$ is the $N \times N$ matrix of the link weights. Without loss of generality, it can be assumed that $\tilde{w}_{ij} \in [0, 1]$ for all $i \neq j$ and $\tilde{w}_{ii} = \tilde{w} \in \{0, 1\}$ for all i, where $i, j = 1, \ldots, N$. A directed link from node i to node j exists if and only if $\tilde{w}_{ij} > 0$.

The index is based on a very simple idea. If the network $\tilde{G}$ is undirected (symmetric), any norm of the suitably rescaled difference between $\tilde{W}$ and $\tilde{W}^T$ (the transpose of $\tilde{W}$) should converge to zero. Without loss of generality, one can define:

$$Q = \{q_{ij}\} = \tilde{W} + (1 - \tilde{w})I_N, \tag{2.3}$$

where I_N is the $N \times N$ identity matrix. The network $G = (N, Q)$ can be then established. Notice that $q_{ij} = \tilde{w}_{ij}$ for all $i \neq j$ and $q_{ii} = 1$ for all i[3]. As a next step, consider the square of the Frobenius (or the Hilbert-Schmidt) form:

$$||Q||_F^2 = \sum_i \sum_j q_{ij}^2 = N + \sum_i \sum_{j \neq i} q_{ij}^2, \tag{2.4}$$

where all sums (also in what follows) span from 1 to N. The index used to check for the symmetry takes the following form:

$$\tilde{S}(Q) = \frac{||Q - Q^T||_F^2}{||Q||_F^2 + ||Q^T||_F^2} = \frac{||Q - Q^T||_F^2}{2||Q||_F^2} = \frac{1}{2}\Big(\frac{||Q - Q^T||_F}{||Q||_F}\Big)^2. \tag{2.5}$$

By using the symmetry $(q_{ij} - q_{ji})^2$, one can easily get:

$$\tilde{S}(Q) = 1 - \frac{\sum_i \sum_j (g_{ij} - g_{ji})^2}{2 \sum_i \sum_j q_{ij}^2}. \tag{2.6}$$

By expanding the squared term at the numerator, we obtain:

$$\tilde{S}(Q) = 1 - \frac{\sum_i \sum_j q_{ij} q_{ji}}{\sum_i \sum_j q_{ij}^2} = \frac{\sum_i \sum_{j \neq i} q_{ij}^2 - 2 \sum_i \sum_{j > i} q_{ij} q_{ji}}{N + \sum_i \sum_{j \neq i} q_{ij}^2}. \tag{2.7}$$

Since $\tilde{S}(Q) \in [0, \frac{N-1}{N+1}]$, its scaled version:

$$S(Q) = \frac{N+1}{N-1} \tilde{S}(Q) \tag{2.8}$$

ranges in $[0, 1]$ and therefore has a more straightforward interpretation. The index is zero if the observed matrix is fully-symmetric, whereas it is one if the observed matrix is fully-asymmetric.

The further step includes a standardization of the index S in order to statistically check for the symmetry of empirically-observed matrix $\tilde{W}$. The distribution of the index S depends on (i) the size of the matrix N and (ii) the underlying nature of the network (binary/weighted). Fagiolo [25] for each $N \in \{5, 10, 50, 100, 200, 500, 700, 1000\}$ generates 100,000 random matrices Q obeying the restriction that $q_{ii} = 1$ for all i. In the binary case, the elements q_{ij} are supposed to be independently and identical distributed (i.i.d.) Bernoulli random variables with probability $p(q_{ij} = 0) = p(q_{ij} = 1) = 0.5$. In the weighted case, the elements q_{ij} are instead supposed to be i.i.d. uniform ran-

[3] The inclusion of self-loops is only required to have an index, which is strictly increasing in the degree of asymmetry of the underlying network.

dom variables over $[0, 1]$. The results of simulations are summarized in the following points:

$$m_B(N) \simeq 0.50 + \exp(-1.786369 - 1.680938 \ln N), \tag{2.9}$$
$$m_W(N) \simeq 0.25 - \exp(-1.767551 - 0.937586 \ln N), \tag{2.10}$$
$$s_B(N) \simeq \exp(-0.135458 - 1.001695 \ln N), \tag{2.11}$$
$$s_W(N) \simeq \exp(-0.913297 - 0.982570 \ln N), \tag{2.12}$$

where $m_B(N)$ (respectively $m_W(N)$) is the sample mean of the index S for the binary (respectively weighted) network and $s_B(N)$ (respectively $s_W(N)$) is the sample standard deviation of the index S for the binary (respectively weighted) network. Given the approximate relations in equations 2.9-2.12, the index S is standardized as follows:

$$S_B(Q) = \frac{S(Q) - m_B(N)}{s_B(N)}, \tag{2.13}$$
$$S_W(Q) = \frac{S(Q) - m_W(N)}{s_W(N)}. \tag{2.14}$$

The standardized versions of the index S are well approximated by a $N(0, 1)$. Positive (respectively negative) values of the standardized index suggest that the network is directed (respectively undirected).

If the notionally-directed network turns out to be sufficiently undirected (symmetric), the typical procedure is to symmetrize the originally-observed matrix. For binary networks the symmetric matrix is defined as:

$$A = \{a_{ij}\} = max\{\tilde{a}_{ij}, \tilde{a}_{ji}\}, \tag{2.15}$$

whereas for weighted networks it takes the form:

$$W = \{w_{ij}\} = \frac{1}{2}(\tilde{w}_{ij} + \tilde{w}_{ji}). \tag{2.16}$$

2.3 Binary Network

The simplest type of networks is binary and undirected. This means that any two nodes are either connected by a link or not and the directions of links do not count. Such a type of networks can be characterized by a symmetric $N \times N$ adjacency matrix $A = \{a_{ij}\}$, where $a_{ij} = a_{ji} = 1$ for all $i \neq j$ if and only if a link between nodes i and j is present and zero otherwise[4]. The most important binary statistics are summarized below.

2.3.1 Node Degree

The most common statistics is the *node degree* (ND), which is simply the total number of connections that a node i holds:

$$d_i = \sum_j a_{ij} = A_{(i)}\mathbf{1}, \tag{2.17}$$

where $A_{(i)}$ is the i-th row of A and $\mathbf{1}$ is the N vector of ones. The shape of the ND distribution can provide a lot of information about the structure of a network. For instance, random networks have an unimodal ND distribution meaning that node degrees are distributed around the mean. On the contrary, real networks often indicate a right-skewed ND distributions with a majority of nodes holding few links and a minority of nodes (known as *hubs*) holding many links. Some networks are found to have the ND distribution, which approximately follows a power law. These networks known as scale-free networks have recently attracted a lot of attention for their structural and dynamical properties. Perfect examples are the world wide web (WWW), biological networks or social networks [1, 2, 3, 4]. The ND statistics is the first-order indicator, because it takes into account nodes lying one step away from the one under analysis.

2.3.2 Average Nearest-Neighbor Degree

The ND only considers nodes that are directly connected to the analysed one, however the importance of a node in the network is also determined by connections of its partners. The *average nearest-neighbor degree* (ANND) therefore measures how much the partners of a node are themselves connected in the network. The ANND is basically the average of ND of i's partners:

$$annd_i = d_i^{-1} \sum_j a_{ij} d_j = d_i^{-1} \sum_j \sum_h a_{ij} a_{jh} = \frac{A_{(i)} A\mathbf{1}}{A_{(i)}\mathbf{1}}. \tag{2.18}$$

Nodes with the largest ND and ANND usually hold the most intense interaction relationships in the network. The correlation between the ANND and ND is a

[4]Self-loops are not considered, i. e. $a_{ii} = 0$ for all i.

measure of the network assortativity. If the correlation is positive (respectively negative), the network is called assortative (respectively disassortative). The ANND statistics is the second-order indicator, because it looks at nodes lying two step away from the analysed one.

2.3.3 Binary Clustering Coefficient

The third important characteristics determining the structure of the network is clustering[5]. The *binary clustering coefficient* (BCC) measures how much the partners of a node are themselves partners. The node i's BCC is formally defined as the ratio between the number of triangles with i as one node and the maximum number of triangles that a node i could have formed given its degree:

$$bcc_i = \frac{\frac{1}{2}\sum_{j\neq i}\sum_{h\neq(i,j)} a_{ij}a_{ih}a_{jh}}{\frac{1}{2}d_i(d_i-1)} = \frac{(A^3)_{ii}}{d_i(d_i-1)}, \tag{2.19}$$

where $(A^3)_{ii}$ is the i-th element of the main diagonal of $A^3 = A \cdot A \cdot A$. Each product $a_{ij}a_{ih}a_{jh}$ is intended to count whether a triangle is present around i or not. The order of subscripts is very important, because all entries in A are symmetric. In a random graph, where links are in place independently of each other with a probability $p \in (0,1)$, the expected value of the BCC is equal to p. Node clustering is very important, because geographically structured networks are usually highly-clustered with more short-distance links. The BCC is also the second-order indicator, as it takes into account nodes lying two step away from the analysed one.

2.4 Weighted Network

A binary approach treats all links present in the network as completely homogeneous, however many researchers have recently argued that real networks exhibit a relevant heterogeneity in the capacity and intensity of their connections [28, 29]. This heterogeneity might be crucial to better understand the architecture of complex networks. Each link ij present in the weighted network (i.e. $a_{ij} = 1$) is assigned a value $w_{ij} > 0$ proportional to the flow intensity carried by that link. For example, weights can represent the amount of trade volumes

[5] Clustering is a well known concept in sociology, where terms such as "cliques" and "transitive triads" have been widely used [6, 7]. For example, friendship networks are typically highly clustered, as any two friends of a person are very probable to be friends.

exchanged between countries (as a fraction of their gross domestic product), the number of passengers traveling between airports, the traffic between two Internet nodes, the number of e-mails exchanged between pairs of individuals and so on. For that reason, we need to move from a binary perspective to a weighted network approach.

Suppose that the underlying network is now weighted and undirected. The network is characterized by a symmetric $N \times N$ weight matrix $W = \{w_{ij}\}$, where $w_{ij} = w_{ji} > 0$ for all $i \neq j$ if and only if nodes i and j are connected by a link and zero otherwise[6]. The three statistics above (ND, ANND, and BCC) can be easily extended to a weighted perspective.

2.4.1 Node Strength

Whereas the ND counts how many connections a node holds, the *node strength* (NS) measures a weighted connectivity. The NS is defined as the sum of weights associated to the links, which a node i holds:

$$s_i = \sum_j w_{ij} = W_{(i)}\mathbf{1}, \tag{2.20}$$

where $W_{(i)}$ is the i-th row of W. Nodes with the larger NS are distinguished by a higher intensity of interaction relationships. It is also evident that any two nodes with the same ND can end with very different levels of the NS. Similarly as the ND, the NS is the first-order indicator.

The NS statistics is only an aggregate measure of the intensity of trade relationships mediated by a node, however particular weights associated with the links of a node might vary a lot. A simple way to measure a dispersion (or concentration) of weights is derived from the Herfindahl-Hirschman Index (HHI) [30, 31]. After a slight adjustment, one can define the *node disparity* among i's weights as follows:

$$\tilde{h}_i = \sum_j \left(\frac{w_{ij}}{s_i}\right)^2 = \frac{1}{s_i^2}\sum_j w_{ij}^2 = \frac{W_{(i)}^2\mathbf{1}}{(W_{(i)}\mathbf{1})^2}. \tag{2.21}$$

As $\tilde{h}_i \in [\frac{1}{N-1}, 1]$, it is reasonable to define its rescaled version:

$$h_i = \frac{(N-1)\tilde{h}_i - 1}{N-2}, \tag{2.22}$$

[6] Self-loops are not considered, i. e. $w_{ii} = 0$ for all i.

which ranges in $[0,1]$. The disparity is one if a node concentrates all its relationships on one partner, whereas it converges to zero if the relationships are more differentiated. Moreover, the disparity should equal to the inverse of the ND if the weights associated with the links of a node are of the same order.

2.4.2 Weighted Average Nearest-Neighbor Degree and Average Nearest-Neighbor Strength

The intensity of trade relationships maintained by the partners of a given node is measured by either the *weighted average nearest-neighbor node degree* (WANND) or the *average nearest-neighbor node strength* (ANNS). The WANND is computed as the weighted average of ND of i' partners:

$$wannd_i = s_i^{-1} \sum_j w_{ij} d_j = s_i^{-1} \sum_j \sum_h w_{ij} a_{jh} = \frac{W_{(i)} A \mathbf{1}}{W_{(i)} \mathbf{1}}. \tag{2.23}$$

The formula implies that $wannd > annd$ if the links with the larger weights are pointing out to the neighbors with larger degree and $wannd < annd$ in the opposite case. The WANND is therefore a measure of the effective affinity to connect with high- or low-degree neighbors according to the significance of actual interactions. The correlation between the WANND and ND measures the network assortativity (if positive) or disassortativity (if negative). The ANNS is instead defined as the average of NS of i's partners:

$$anns_i = d_i^{-1} \sum_j a_{ij} s_j = d_i^{-1} \sum_j \sum_h a_{ij} w_{jh} = \frac{A_{(i)} W \mathbf{1}}{A_{(i)} \mathbf{1}}. \tag{2.24}$$

Analogically to the previous case, the correlation between the ANNS and NS measures the network assortativity (if positive) or disassortativity (if negative). Once again, any two nodes with the same ANND can be associated to very different levels of the WANND or ANNS. Both statistics are second-order indicators.

2.4.3 Weighted Clustering Coefficient

An extension of the BCC to a weighted perspective is not straightforward, because one has to take into account the weights associated to the links in the neighborhood of a given node. A motivation and comparison of selected features for different *weighted clustering coefficients* (WCC) are provided by

Saramäki et al. [32]. For example, suppose that a triangle ihj is in place. One might consider only the weights of the links ih and ij [28]. Alternatively, one might employ the weights of all the links. The total contribution of a triangle can be then defined as the geometric mean of its weights [29] or simply as the product among them [33, 34]. We focus on the extension of the BCC to a weighted perspective originally presented by Onnela et al. [29]:

$$wcc_i = \frac{\frac{1}{2}\sum_{j\neq i}\sum_{h\neq(i,j)} w_{ij}^{\frac{1}{3}} w_{ih}^{\frac{1}{3}} w_{jh}^{\frac{1}{3}}}{\frac{1}{2}d_i(d_i-1)} = \frac{(W^{[\frac{1}{3}]})^3_{ii}}{d_i(d_i-1)}, \tag{2.25}$$

where $W^{[\frac{1}{k}]} = \{w_{ij}^{\frac{1}{k}}\}$, i. e. the matrix obtained from W by taking the k-th root of each entry. The index wcc_i ranges in [0,1] an reduces to bcc_i, when weights become binary. Moreover, it considers the weights of all the links in a triangle and is invariant to the weight permutation for one triangle. In a random graph, where links are in place independently of each other with a probability $p \in (0,1)$, the expected value of the WCC is equal to $(\frac{3}{4})^3 p$. The WCC is also the second-order indicator.

Chapter 3

Literature Review

We have already stated that a lot of effort has been devoted to analysing the international trade from a network perspective in the last two decades. The goal of this chapter is to examine and summarize all relevant literature. We start with the first contributions that recognized that international trade can be described as a network, however they did not employ the statistical analysis discussed in the previous chapter. Furthermore, we move to the network approach, which has been flourished thanks to the significant contributions stemming from mathematics, physics and computer science. Finally, we review other contributions relating a network analysis.

3.1 First Contributions

The first contributions describing the international trade flows as a network have been originally presented in sociology and political science. The researchers have gradually showed that relational characteristics are more relevant than (or at least as relevant as) individual country characteristics in explaining the macroeconomic dynamics resulting from export and import. Early studies have been greatly influenced by "dependency" and "world-system" theories, which attempt to explain unequal economic relations between poor and wealthy countries. These theories are based on an international division of labor, which divide the world into three major categories: core, semiperiphery and periphery. In their extreme forms, they state that poor and undeveloped countries are exploited by wealthy ones through world economic system[1].

[1] We recognize that the dependency and world-system theories are not necessarily identical, however their distinction is not important for our purposes.

The initial paper by Snyder and Kick [35] addresses dependency and world-system theories of differential economic growth among countries. They present a block-model for four types of binary undirected networks (trade flows, military interventions, diplomatic relations and conjoint treaty membership) among 118 countries circa 1965. The results provide a strong support for a core-semiperiphery-periphery structure in the world system. They further perform a regression analysis of the effects of structural positions on a nations' economic growth (change in GNP per capita) from 1955 to 1970. The estimation shows that the location of a country in the structure can explain an economic growth, which is consistent with dependency and world-system arguments.

Nemeth and Smith [36] also use a network analysis of international interactions in an attempt to derive the structure of the world economy. They specially sort countries into structural positions in the world system according to their patterns of commodity trade. A block-model is based on five types of binary directed networks (heavy/high-technology manufactures, intermediate manufactures, raw materials, light manufactures and food products) among 86 non-centrally planned countries in 1970. The results again indicate a strong core-semiperiphery-periphery structure in the world system. In particular, the authors refer to four distinct structural positions in the world economy. These findings are confirmed by the regression analysis of the effects of structural positions on a nations' economic growth and strength, income inequality and level of social welfare.

A similar approach is followed by Breiger [37], however he expands the research upon a weighted analysis. A standard block-model procedure is performed on a four types of undirected networks (raw materials, energy resources, manufactured goods and agriculture products) among 24 highly industrialized countries in 1972. The results suggest the existence of a smaller and more predominant core in the world system than has been recognized by other studies. If one takes into account the weighted approach, a considerably more differentiated structure is shown to underlie the original findings. A main characteristic of this underlying structure is the presence of multiple competing cores.

More recent paper is provided by Smith and White [38], who improve on previous network studies of the world economy in two ways. First, they apply a more general approach for measuring positional proximity in a network. Second, they add a dynamic aspect into the analysis by comparing international trade flows in three different years (1965, 1970 and 1980). A block-model process includes an investigation of five types of binary directed networks (heavy/high technology manufactures, sophisticated extractives, simple extractives, low wage/light manufactures and food products) among 63 countries. The results support dependency and world-system formulations about the asymmetrical flows of raw materials versus processed goods. Furthermore, recent declines in core position's share of low wage/light manufactures and simple extractives are in full accordance with the "new international division of labor" argument[2].

Kim and Shin [40] employ a social network analysis to examine effects of globalization and regionalization, which are defined as specific types of linkages between countries. The analysis includes a binary directed approach for large set of commodities among 105 countries in three different snapshots (1959, 1975 and 1996). They find that the world became increasingly globalized between 1959 and 1996. Countries had significantly more trading partners in 1996 than in 1959 and the ITN became more denser. The main source of this process was a development of countries in the middle strata. They also show that the the ITN became decentralized during this period, which provides a stronger support for neoclassical theory rather than for dependency and world-system theories. They finally argue that intraregional density is higher than interregional density and intraregional ties are stronger than interregional ties across years indicating that the ITN became regionalized. These findings demonstrate that globalization and regionalization are not contradictory processes.

Kastelle et al. [41] propose a new way for measuring globalization, which is based on a complex network analysis. They study the evolution of the topological properties of the ITN using a binary directed analysis with longitudinal trade data between 1938 and 2003. The paper shows that several network characteristics have changed significantly over examined 65 years, however the basic structure of the ITN has been remarkably stable over the period. The authors proclaim that there is some globalization, but the perception that international trade is integrated into one huge market is inaccurate.

3.2 Network Approach

The study of the ITN has been recently influenced by contributions stemming from mathematics, physics and computer science. The researchers have started

[2] The term new international division of labor is associated with a structural change that forces companies to reorganize their production on a global scale. The most common pattern is a shift of manufacturing industries from advanced countries to developing ones [39].

to explore the statistical properties of the ITN with new and more powerful statistical tools. The idea is that the international trade might be viewed as a complex network. The main purpose of the revived approach is to analyse the mechanics and topological properties of the ITN by abstracting from any economic and social effects.

3.2.1 Binary Network

The initial works employed a binary network analysis, where a link is either present or not according to whether the trade flow that it carries is larger than a given lower threshold[3].

Serrano and Boguna [9] construct a binary undirected network of trade relationships among 179 countries in 2000. The network displays the typical properties of complex networks, i. e. scale-free degree distribution, “small-world” property, degree correlation between different countries and high clustering coefficient. Specifically, the degree distribution approximately follows a power law, which implies that the ITN is a scale-free network. Furthermore, the ITN appears to be a disassortative network, where highly connected countries tend to connect with poorly connected countries. Finally, the ITN is characterized by a hierarchical structure, which means that the partners of highly connected countries are less interconnected than partners of poorly connected ones. The highly connected countries form large degree centers (hubs) in the network. The authors argue that these results refer to a high similarity between the ITN and the Internet.

Garlaschelli and Loffredo [10] provide an empirical test whether the “hidden variable” model reproduces all the relevant topological properties of the ITN. According to this model the topological properties of the ITN can be well explained by a single country characteristic. The undirected analysis is based on a very detailed dataset of 191 countries in 1995. First, they explore the topological properties of the ITN. They find that power-law region is only a small part of the whole degree distribution, therefore they conclude that the ITN is not a scale-free network. This finding is in contrast with the previous work [9]. Second, they find that all studied properties (degree distribution, degree correlation and clustering) are in excellent agreement with the predictions

[3]There is no agreement on how much this threshold should be. Kim and Shin [40] set thresholds of USD 1 million and 10 million. Kastelle et al. [41] use a threshold in a way to have a connected network in each year. On the contrary, other researchers [9, 10, 11] define a link whenever non-zero trade flow occurs.

of the hidden-variable model. The single country characteristic called hidden variable (fitness) is the GDP.

While previous two contributions focused on undirected version of a single snapshot of the ITN, Garlaschelli and Loffredo [11] study the properties of the ITN as a directed and evolving network. The analysis is based on a comprehensive dataset, which reports the trade activity for all world countries over the period 1950 and 1996. They find that degree distribution does not follow a power law. Furthermore, the results confirm that the topology of the ITN shows a peculiar dependence on the gross domestic product, which is in full accordance with the hidden-variable model. Moreover, the network variables are quite stable over time, which might cast some doubts on a process of economic integration (globalization) in the second half of 20th century.

3.2.2 Weighted Network

A binary analysis only counts the mere presence or absence of an interaction between any two nodes, however the majority of economic and social relationships also involve an assessment of the intensity of the interaction between any two nodes. If one uses a binary network analysis, a lot of information might be disregarded and a role of heterogeneity in trade linkages might be significantly underestimated. Many researchers have therefore adopted a weighted network approach to the study of the ITN.

Li et al. [12] attempt to investigate the effect of dynamics on the ITN. First, they present the scale-free features of the degree distribution and link weight distribution for a weighted directed network among 188 countries in 2000. The United States (US) appears to be the biggest node in the weighted degree sense. Second, they study a synchronization of economy cycles due to its scale-free features. The real GDP data for 21 developed countries between 1975 and 2000 are analysed in terms of their correlations with the United States. The results show that 18 developed countries are significantly synchronized with the United States.

Bhattacharya et al. [13] study in detail the variations of different network quantities over the period 1948 and 2000 by using a weighted undirected approach. They demonstrate that the deviation in the size of the giant component of the ITN from the fully connected graph declines exponentially. They also show that the distribution of link weights is better approximated by the log-normal distribution. Furthermore, the size of a few rich countries that trade among themselves one half of the total world trade decreases over time. Finally, the three disparity measures using link weights as the total trade, export and import increase in similar manner.

The following paper by Bhattacharya et al. [14] is very similar to the previous one. The authors analyse international trade data over the same period and confirm the properties of the ITN. In addition, they show that many of these features are reproduced by a non-conservative dynamical model based on the gravity model of social and economic sciences.

Fagiolo et al. [15] show that the topological properties of the ITN viewed as a weighted network are significantly different from those obtained by a binary network approach. For instance, the most countries are characterized by weak trade relationships (the weighted representation of the ITN leads to a weakly connected graph). Additionally, the weighted ITN is only weakly disassortative. Finally, highly connected countries are more likely to trade with partners, which are strongly connected among themselves. The research is based on a weighted undirected network among 159 countries between 1981 and 2000.

Fagiolo et al. [16] investigate how the most important network statistics (connectivity, assortativity, clustering and centrality) have evolved over time. They construct a weighted undirected network among 159 over the period 1981 and 2000. The results show that node statistic distributions and their correlation structure have remained highly stable over time. In the contrast, the distribution of link weights is slightly changing from a log-normal distribution to a power law. They also describe the autoregressive properties of network-statistics dynamics. They find that network-statistics growth rates are well described by fat-tailed distributions (the Laplace or the asymmetric exponential power).

A more thorough analysis of the statistical properties of the ITN is presented by Fagiolo et al. [17], who expand the preceding contributions along four dimensions. First, they present a more complete description of statistical properties of the ITN by discussing several measures (indicators). Second, they explore how these properties correlate with node characteristics. Third, they study the extent to which statistical features are robust to different ways of weighting links. Finally, they evaluate whether the observed ITN is sufficiently undirected to justify a weighted undirected approach instead of weighted directed one. The results of the work are also based on the empirical analysis of international trade data among 159 countries between 1981 and 2000.

3.3 Additional Contributions

We finally review the papers whose goal is not primarily to show the topology of the ITN. They have instead focused on specific features of the structure and evolution of the ITN, on the determinants of the topological properties of the ITN, on replicating of the structure of the ITN and so on.

Garlaschelli et al. [42] report a range of empirical results and theoretical arguments showing the interplay between the dynamics of the GDP values and the evolution of the ITN. A weighted directed analysis includes increasing number of countries during the period 1950 and 2000. The authors find that the topological properties of the ITN are determined by the GDP of all world countries, which supports the presence of the hidden variable (fitness). On the other hand, they show that the topology of the ITN determines the GDP values due to the exchange between countries. These results lead to a new framework, where the hidden variable (fitness) is a dynamical variable determining and simultaneously depending on the network topology.

Kali and Reyes [43] combine international trade data with network methods to explore global trading system as an interdependent complex network. The data includes exports and imports of all commodities between 182 countries for the years 1992 and 1998. The authors outline the topology of the ITN and suggest new network based measures of international economic integration. These measures embody the structure and function of the network and might provide a more reasonable approach to globalization than current measures based on trade volumes. They find that in terms of participation in the network the global trade is hierarchical with core-periphery structure, although the integration of smaller countries into the network increased greatly over the 1990s. They further show that a country's position in the network can have significant impacts for economic growth. They finally suggest that a network approach to international economic integration has a potential for useful applications in international finance and development.

Serrano et al. [44] instead study the world network of merchandise trade imbalances and describe its overall flux organization. They develop a general procedure, which is able to filter out in a consistent and quantitative way the dominant trade channels. They build a weighted directed network of trade imbalances between independent countries in the world during the period 1948 and 2000. The obtained networks are characterized by a high density of connections and heterogeneity of the particular fluxes among countries. The analysis exhibits the presence of high-flux backbones, which are sparse subnetworks of connected trade fluxes carrying the most of the total flux in the ITN.

The evolution of trade “islands” in the ITN in which countries are linked with directed links carrying a total trade flow larger then some given thresholds is examined by Tzekina et al [45]. The term islands is used as a means to identify communities and hubs. The dataset contains bilateral merchandise trade data for 186 countries during the years 1948 and 2005. The results show that the evidence for or against globalization is mixed. On the one hand, many more countries trade significantly in 2005 than in 1950, which implies that the international trade was more spread out in 1950 than in 2005. On the other hand, many more islands (trade centers) have evolved over time. This mixed evidence for the globalization supports the previous work [41].

Fagiolo [46] explores the determinants of the statistical properties of the ITN. He employs international trade data among 159 countries between 1981 and 2000 to build an undirected weighted network. He subsequently estimates a standard gravity model to build a residual ITN. The bilateral trade flows are regressed on the country GDP, geographical distance between countries, border effects, trade agreements etc. The statistical properties of the residual ITN are then compared to those of the observed ITN. The results indicate that the residual ITN has a very different topological structure. It is characterized by power-law shaped distributions of node statistics (e. g. strength, clustering and random-walk betweenness centrality) and link weights. Whereas the observed ITN indicates a structure with a few large-sized hubs and a relatively strong connectivity among close countries, the residual ITN is organized around many small-sized but trade-oriented countries that play the role of local hubs or attract large and rich countries in complex trade-interaction patterns.

Kali and Reyes [47] provide a similar paper to their previous one [43]. They combine international trade data with network approach to map global trading system as an interdependent complex network. The sample is this time slightly longer, i.e. 182 countries for the period 1992 and 2000. Their network based measures of connectedness allow to explain stock market returns during recent financial crises. They show that a crisis is stronger if the epicenter country is better integrated into the ITN. On the other hand, target countries affected by a crisis are more able to dissipate the impact if they are well integrated into the network. Finally, they show that a network approach including the cascading and diffusion of interdependent ripples when a shock hits a specific part of the ITN provides the explanation of financial contagion.

Squartini et al [48] point out that it is still unclear whether the network approach conveys additional information in a comparison with traditional economic approaches that characterize the international trade only in terms local (first-order) properties. They employ a recently proposed randomization method to evaluate in detail the role, which local properties have in forming higher-order patterns of the ITN. They use yearly bilateral data on exports and imports from 1992 to 2000 to explore all possible representation of the ITN (binary/weighted, directed/undirected, aggregated/disaggregated). The results show that the properties of all binary projections of the ITN can be completely explained by the degree sequence. In other words, the degrees of world countries are maximally informative about the ITN as a whole. The implication of this contribution is that explaining the observed degree sequence of the ITN should become one of the main focuses of models of trade. Furthermore, Squartini et al [49] stress that current economic models of the ITN generally aim at explaining local weighted properties, not local binary ones. They analyse the binary projections of the ITN by considering its weighted representations. The results show that all possible weighted representations of the ITN (directed/undirected, aggregated/disaggregated) cannot be traced back to local country-specific characteristics. These two papers suggest that the traditional macroeconomic approaches fail to capture the main properties of the ITN. In the binary case, they do not take an interest in the degree sequence and therefore cannot characterize or reproduce higher-order properties. In the weighted case, they in generally focus on the strength sequence, however it is not enough in order to understand or replicate indirect effects.

Another paper also examines whether a particular model can explain the statistical properties of the ITN. Duenas and Fagiolo [50] specially inspect the relation between the topology of the ITN and the gravity model of trade. The authors employ the international trade data for all available countries between 1970 and 2000. They predict international trade flows using alternative estimation methods and build the predictions for the topological properties of the ITN. The properties of the predicted ITN are then compared to those observed in the original ITN. The first finding is that the gravity model is able to partially replicate the weighted network structure only if the binary architecture is fixed to the observed one. Second, the gravity model fails to explain higher-order statistics, which require the knowledge of triadic link-weight topological patterns even if the binary structure perfectly replicates the observed one. Finally, the gravity model works very badly in predicting the presence of a link or the level of trade flow it carries whenever the binary structure has to be simultaneously estimated.

The last reviewed paper is proposed by Mastrandrea et al. [51], who stress that the strength (total value of relationships) of a given node has always an important economic meaning. Null models of networks capturing the observed strengths of all nodes are crucial in order to either detect interesting deviations of an empirical network from economically meaningful indicators or reconstruct the most probable structure of an economic network. However, several works have showed that real economic networks are topologically very different from configurations implied only from node strengths. The authors compare the ITN to an enhanced model, which they propose in order to simultaneously replicate the degree and strength of each node. Their comprehensive panel includes 162 countries between the years 1992 and 2002. Moreover, the study employs several different layers (commodity classes). They suggest that the observed properties of the ITN are well reproduced by their model. This allows them to introduce the concept of extensive and intensive bias, which is defined as a measurable tendency of the network to prefer either the formation of new links or the reinforcement of existing ones.

Chapter 4

Data

We use international trade data provided by Gleditsch [52] to build a sequence of binary and weighted directed networks. The original dataset includes GDP per capita and population of independent states (1950-2000) and trade flows between independent states (1948-2000). The GDP per capita and population come from the Penn World Tables (PWT) produced by the Center for International Comparisons at the University of Pennsylvania. To address the case of countries not included, two additional sets of estimates have been generated. First, a set of GDP per capita and population estimates is based on the figures from World Factbook reported by Central Intelligence Agency (CIA). Second, missing lead or tail parts of series have been filled in by estimates based on the first/last nonmissing observations. These estimates suppose that the real GDP per capita remains the same for the lead/tail parts. The values have been deflated to current prices using a US GDP deflator. The origin of the GDP per capita and population observations is shown in table 4.1.

Trade data are based on the Direction of Trade Statistics (DOTS) produced by the International Monetary Fund (IMF), however this database contains only about 40% of all export and import figures between 1948 and 1996. The coverage is particularly poor for developing and socialist countries. Missing data are replaced with additional estimates through several different procedures. First, the World Export Data (WED) database has been used to compile data to fill in some of the gaps for export figures. Second, missing data have been substituted with the reserve flows whenever available. Third, in the absence of other information i' exports to j is probably a reasonable estimate for i' imports from j and vice versa. Fourth, missing data within time series have been estimated by a linear interpolation. Finally, many time series have

Table 4.1: GDP and population data categories

Data origin	Share
Observed data from PWT	76.16
Estimate based on figures from the World Factbook	12.70
Lags and leads based on first nonmissing observations	11.14

Source: PWT [53], World Factbook [54]

spells of missing data at the beginning or end. These data have been filled in by estimates based on the first/last nonmissing observation deflated to current international prices. The origin of the export and import observations is presented in table 4.2 and 4.3.

The original dataset includes 196 countries for which trade data are available from 1948 to 2000. The sample used in this study has been chosen according to following considerations. First, the sample size must be as large as possible to achieve statistical significance. Second, trade data contain many missing figures for small countries before 1970. The data availability basically causes an increasing number of countries over the years. More specifically, the number of nodes has increased from 82 in 1948 to 190 in 2000 (figure 4.1, left). This might be a problem if we want to study the dynamics of the topological properties of the ITN. For that reason we need to fix the number of countries in the sample period. Finally, our analysis requires to synchronize trade data with real GDP. The reason is that we want to correlate network statistics with country-specific variables (GDP per capita). In summary, the choice of countries has been driven by following conditions: (1) the country sample size and time horizon as large as possible and (2) no missing values in trade data and GDP. By applying these two conditions we get 83 countries for the period 1950-2000, 112 countries for the period 1960-2000, 138 countries for the period 1970-2000, 161 countries for the period 1980-2000 and 168 countries for the period 1990-2000. We have decided to choose a sample of 161 countries for the period 1980-2000. Several authors [16, 17] remove countries that have total exports equal to zero in some years, however we keep them in the sample. Our balanced panel finally refers to 161 countries (table A.1) and 21 years (1980-2000).

Trade data are expressed in current US dollars, while GDP data are provided in current US dollars as well as in 1996 US dollars. The application of a standard reference money unit cancels out the effects of inflation and allows for meaningful across-years comparisons. We therefore deflate trade data using

Table 4.2: Export from country i to country j data categories

Data origin	Share
Observed data from DOTS	40.13
Observed data from WED	11.08
Estimate based on imports of j from i (from DOTS)	9.29
Estimate based on imports of i from j	6.35
Interpolated estimate	5.19
Lags and leads based on first nonmissing observations	2.85
Pairs of countries with no observed data assumed to be 0	25.11

Source: DOTS [55], WED [56]

Table 4.3: Import of country i from country j data categories

Data origin	Share
Observed data from DOTS	43.90
Estimate based on exports from j to i (from DOTS or WED)	17.60
Estimate based on exports from i to j	6.35
Interpolated estimate	5.19
Lags and leads based on first nonmissing observations	2.85
Pairs of countries with no observed data assumed to be 0	25.11

Source: DOTS [55], WED [56]

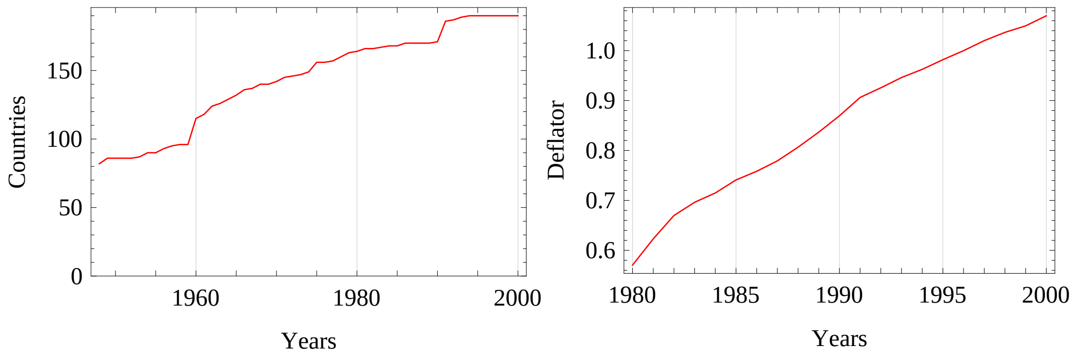
the ratio of the value of current US dollars to the value of 1996 US dollars. The ratio is specifically defined as the US GDP deflator, i.e. the current US GDP to the 1996 US GDP:

$$f^t = \frac{USD^t}{USD^{1996}} = \frac{GDP^t_{USA}}{GDP^{1996}_{USA}} \tag{4.1}$$

Figure 4.1 (right) shows the evolution of the US GDP deflator over time. In what follows, both GDP and trade data will be expressed in millions of 1996 US dollars.

Annual trade between two countries i and j is described by four different measures $\tilde{x}^t_{ij}$, $\tilde{x}^t_{ji}$, $\tilde{m}^t_{ij}$ and $\tilde{m}^t_{ji}$. The figures $\tilde{x}^t_{ij}$ and $\tilde{m}^t_{ji}$ should be generally the same, however they have been quoted differently, because exports from country i to country j and imports of country j from country i are reported as different flows in the DOTS data. The magnitudes of these measures are approximately the same, but they vary in many instances due to different reporting procedures followed and different rates of duties applied in different countries. We denote

Figure 4.1: Selection of the sample: (Left) number of countries and (Right) US GDP deflator



Note: (Left) the scale set to possible numbers of countries (from 0 to N)

the amount of exports from country i to country j by x_{ij}^t and the amount of imports of country i from country j by m_{ij}^t. They are subsequently defined as:

$$x_{ij}^t = \frac{1}{2}(\tilde{x}_{ij}^t + \tilde{m}_{ji}^t) \tag{4.2}$$

$$m_{ij}^t = \frac{1}{2}(\tilde{x}_{ji}^t + \tilde{m}_{ij}^t). \tag{4.3}$$

In order to build adjacency and weight matrices, we follow the flow of goods. This means that rows stand for exporting countries, while columns represent importing countries.

We define a "trade relationship" by setting the generic entry of the adjacency matrix $a_{ij}^t = 1$ if and only if exports from country i to country j (labeled by x_{ij}^t) are strictly positive in year t. Following papers [12, 13, 14, 16, 17], the weight of a link from country i to country j in year t is defined as:

$$\tilde{w}_{ij}^t = x_{ij}^t, \tag{4.4}$$

where x_{ij}^t are the deflated exports from country i to country j in year t. Therefore, we get a sequence of $N \times N$ adjacency and weight matrices $(\tilde{A}^t, \tilde{W}^t)$ with $t = 1, \ldots, N$, which fully describe the dynamics of the ITN from a binary and weighted directed perspective.

Chapter 5

Results

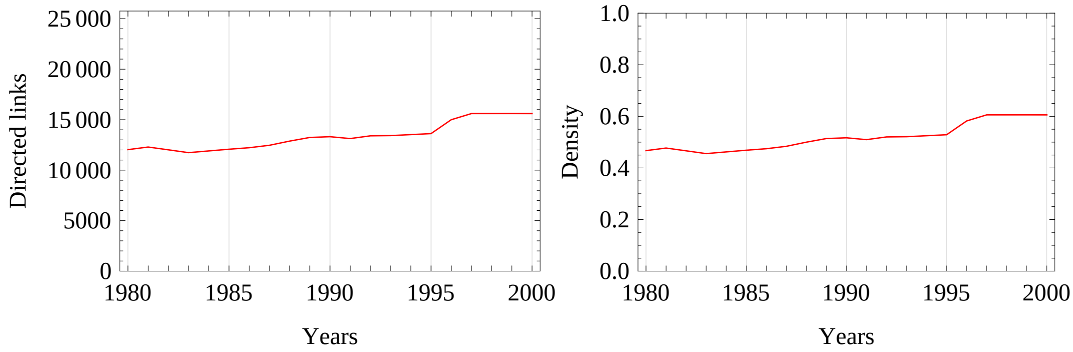
We study the topological properties of the ITN among 161 countries over the period 1980-2000 using a network analysis. We begin with a quick overview of the main global properties of the ITN. This overview especially checks whether the symmetry of the ITN is so strong to justify an undirected analysis. The main part of the analysis presents network statistics, which allow us to address the study of node characteristics in terms of three dimensions: connectivity (ND and NS), assortativity (ANND and WANND/ANNS) and clustering (BCC and WCC). We further explore in more detail the stability of the distributions of the network statistics. The weighted network results are then compared to those obtained by a binary network approach. We finally focus on the correlation of the network indicators and country-specific characteristics (GDP per capita).

5.1 Global Properties

The ITN is in general a directed network with two opposite flows along a link, however we have observed that few links have only one flow. The number of directed links has increased almost systematically over the years. More specifically, the number of links has increased from 12,513 in 1980 to 15,603 in 2000 (figure 5.1, left).

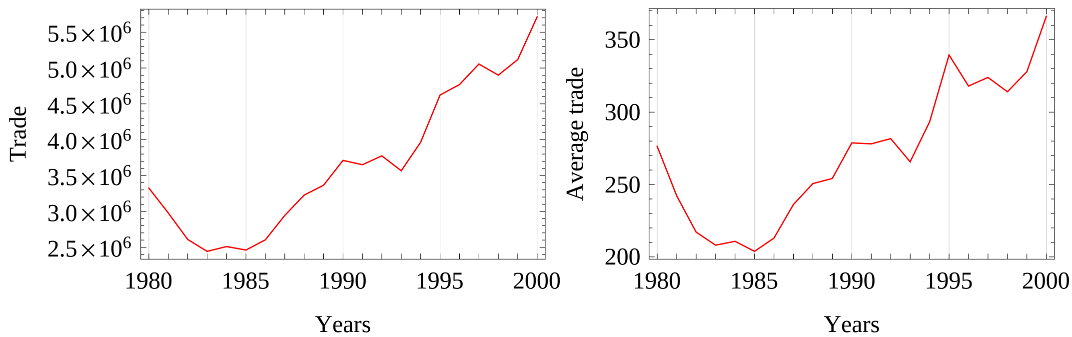
The number of links that are actually in place determines the density of the network. Since we have fixed the number of countries, the maximum number of possible links is always the same, i.e. $N(N-1)$. This implies that the graph of density is equivalent to the graph of the number of links. The average density of the ITN over the years is 0.52. A relatively high value indicates that the ITN is a highly connected network. Moreover, a slightly increasing trend of

Figure 5.1: Directed links: (Left) number of directed links and (Right) density of the network



Note: (Left) the scale set to possible numbers of directed links (from 0 to $N(N-1)$) and (Right) the scale set to possible values of the network density (from 0 to 1)

Figure 5.2: Volume of world trade: (Left) total volume of world trade and (Right) average total trade per link

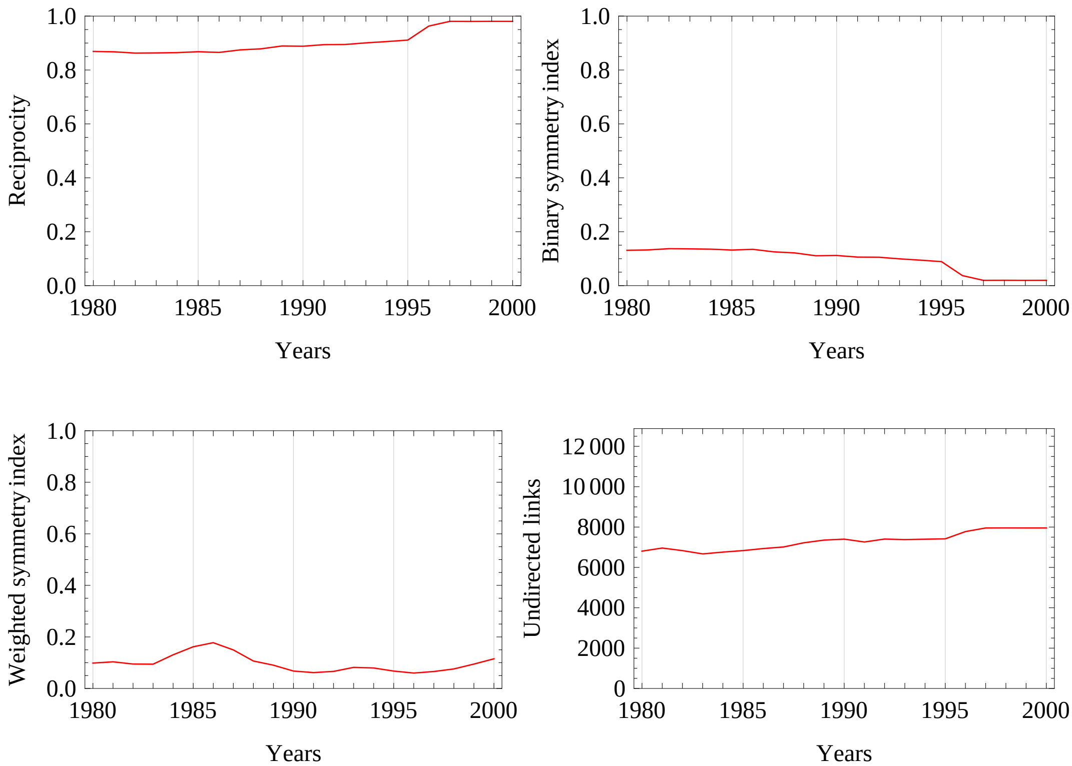


the density witnesses for the increasing participation of countries in the world trade over the last 20 years of the 20th century (figure 5.1, right).

We observe two quantities illustrating the volume of international trade over the sample period. They have also grown almost systematically over the years. The total volume of world trade has increased from 3.3×10^6 USD million in 1980 to 5.7×10^6 USD million in 2000 (figure 5.2, left). Accordingly, the average total trade per link has increased from 276 USD million in 1980 to 366 USD million in 2000 (figure 5.2, right).

If the observed network is sufficiently symmetric, we can ignore the direction of links and define an undirected link between an arbitrary pair of nodes. The first measurement checking the symmetry of network is the reciprocity. On

Figure 5.3: Symmetry of the network: (Top-left) reciprocity of the network, (Top-right) symmetry index for the binary network, (Bottom-left) symmetry index for the weighted network and (Bottom-right) number of undirected links



Note: (Top-left) the scale set to possible values of the network reciprocity (from 0 to 1), (Top-right) the scale set to possible values of the binary symmetry index (from 0 to 1), (Bottom-left) the scale set to possible values of the weighted symmetry index (from 0 to 1) and (Bottom-right) the scale set to possible numbers of undirected links (from 0 to $N(N-1)/2$)

average about 90% of links are reciprocated in each given year (figure 5.3, top-left). This means that if country i exports to country j, then country j almost always exports to country i. We can also see one of the drawbacks of the reciprocity discussed in section 2.1. The reciprocity is heavily dependent on the density of the network. The mutual links simply occur more often with the increasing density (number of links).

The symmetric pattern of the ITN is also confirmed by the symmetry index S for both adjacency and weight matrices $(\tilde{A}^t, \tilde{W}^t)$. In the binary case, the index ranges in the sample period between 0.02 and 0.14 (figure 5.3, top-right). In the weighted case, the index ranges in the sample period between 0.06 and 0.18 (figure 5.3, bottom-left). The corresponding standardized versions of the index are equal to values at least 20 standard deviations below zero. These results signalize a strong and stable symmetry of both adjacency and weight matrices, therefore we will explore the statistical properties of the symmetrized version of the ITN. We define the entry a_{ij}^t of the new adjacency matrix A^t as:

$$a_{ij}^t = max\{\tilde{a}_{ij}^t, \tilde{a}_{ji}^t\}. \quad (5.1)$$

Accordingly, the entry w_{ij}^t of the new adjacency matrix W^t is defined as:

$$w_{ij}^t = \frac{1}{2}(\tilde{w}_{ij}^t + \tilde{w}_{ji}^t). \quad (5.2)$$

In order to have well behaved weights, we finally divide all entries in W^t by their maximum value:

$$w_*^t = max\{w_{ij}^t\}. \quad (5.3)$$

This does not cause any bias in our analysis and ensures that $w_{ij}^t \in [0,1]$ for all (i,j) and t [29].

Since an undirected analysis is justified, we finally take a look at undirected links. The number of undirected links is approximately half the number of directed links. More specifically, the number of undirected links has increased from 6,086 in 1980 to 7,955 in 2000 (figure 5.3, bottom right).

5.2 Connectivity

We start by studying the behavior of the node degree (ND) and node strength (NS) distributions. First, we explore how the shape of the ND and NS distributions look like. More specifically, we investigate the extent to which countries are more or less connected in both terms of number of partners (ND) and interaction intensity (NS). Second, we inspect how the shape of the ND and NS distributions have changed over time.

The ND distribution $P(d)$ is one of the most important topological properties of a network. This quantity measures the probability of a randomly chosen node to have d connections to other nodes. Real networks often indicate a highly asymmetric (right-skewed) ND distribution, which means that most of the nodes have low NDs, while a small fraction of nodes have an extraordinarily high NDs. Nodes with remarkably high NDs are called hubs. The ND distribution $P(d)$ plotted as a function of the degree d therefore displays a long tail, which is much fatter than the tail of a normal (Gaussian) or exponential

Figure 5.4: ND distribution in 1980, 1990 an 2000: (Left) Kernel density estimation and (Right) Complementary cumulative distribution function

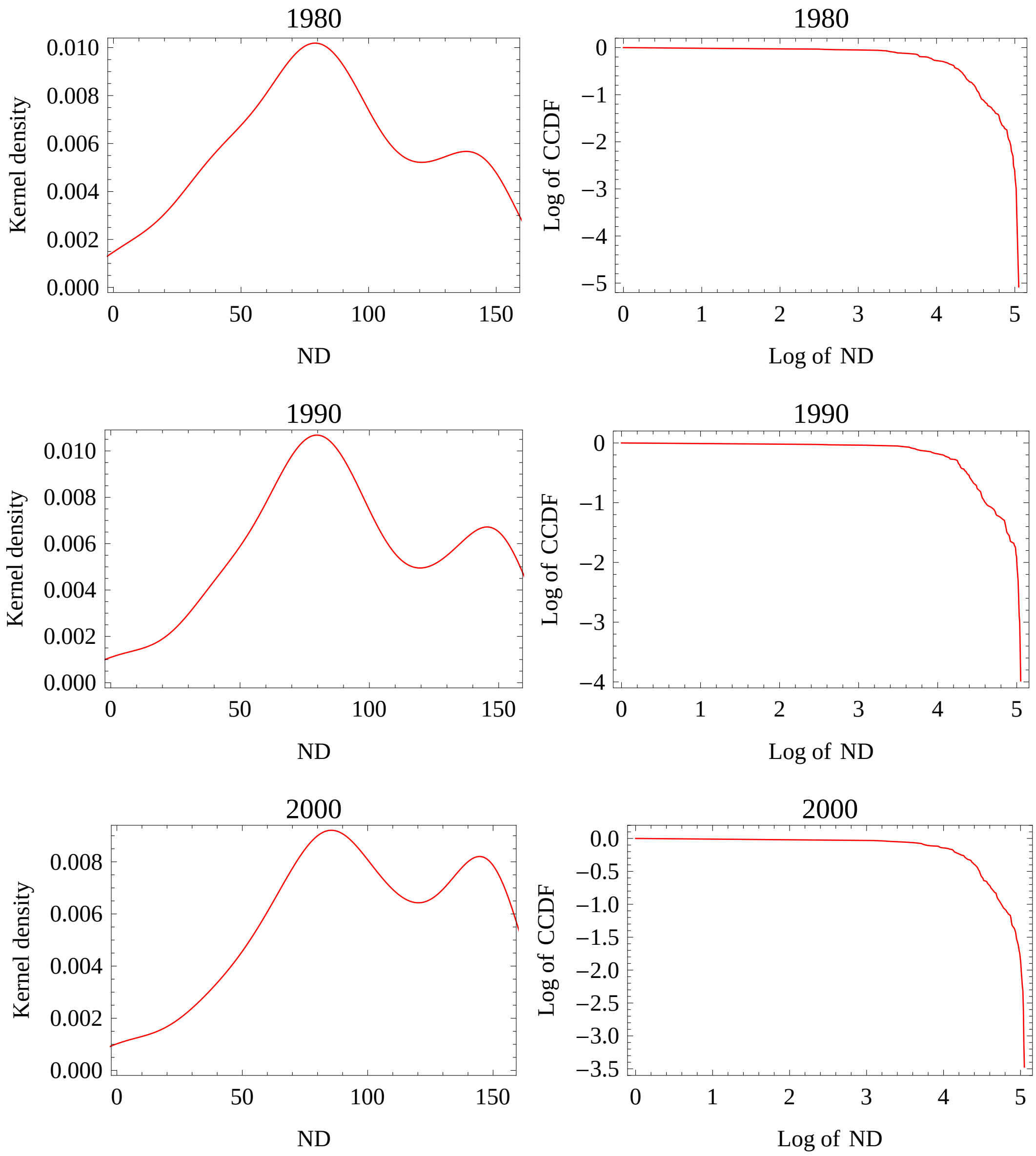


Table 5.1: P-values for ND normality tests

Test	1980	1990	2000
Shapiro-Wilk test	0.0000***	0.0000***	0.0000***
Jarque-Bera test	0.0000***	0.0000***	0.0000***

Note: (*) the null hypothesis rejected at 10%, (**) the null hypothesis rejected at 5% and (***) the null hypothesis rejected at 1%

distribution. The most popular long tail probability distribution is the power law. Networks whose the ND distribution follows a power law are called scale-free networks. Serrano and Boguna [9] just show that the ND distribution of the ITN approximately follows a power law, which implies that the ITN is a scale-free network. Li et al. [12] also present the scale-free features of the ND distribution of the ITN. On the contrary, other authors [10, 11, 13, 17] find out that the power-law region is only a small part of the whole ND distribution, therefore the ITN is not a scale-free network.

We compare the ND distribution of the ITN to power law and log-normal densities. The kernel density estimation shows that the ND distribution can be hardly proxied by these densities. The ND distribution does not appear to be as skewed as expected. Moreover, the ND distribution exhibits some bimodality. Besides a modal value around 80, there exists a second peak around 150. This means that there is a large group of countries, which trade with almost everyone else in the sample. The bimodality is more obvious at the end of the period (figure 5.4, left).

To check quantitatively whether the distribution comes from pre-defined densities, we apply two additional procedures. We start with the power law, which states that the probability $P(d)$ of having a node with d neighbors is defined as:

$$P(d) = \alpha d^{-\gamma}, \tag{5.4}$$

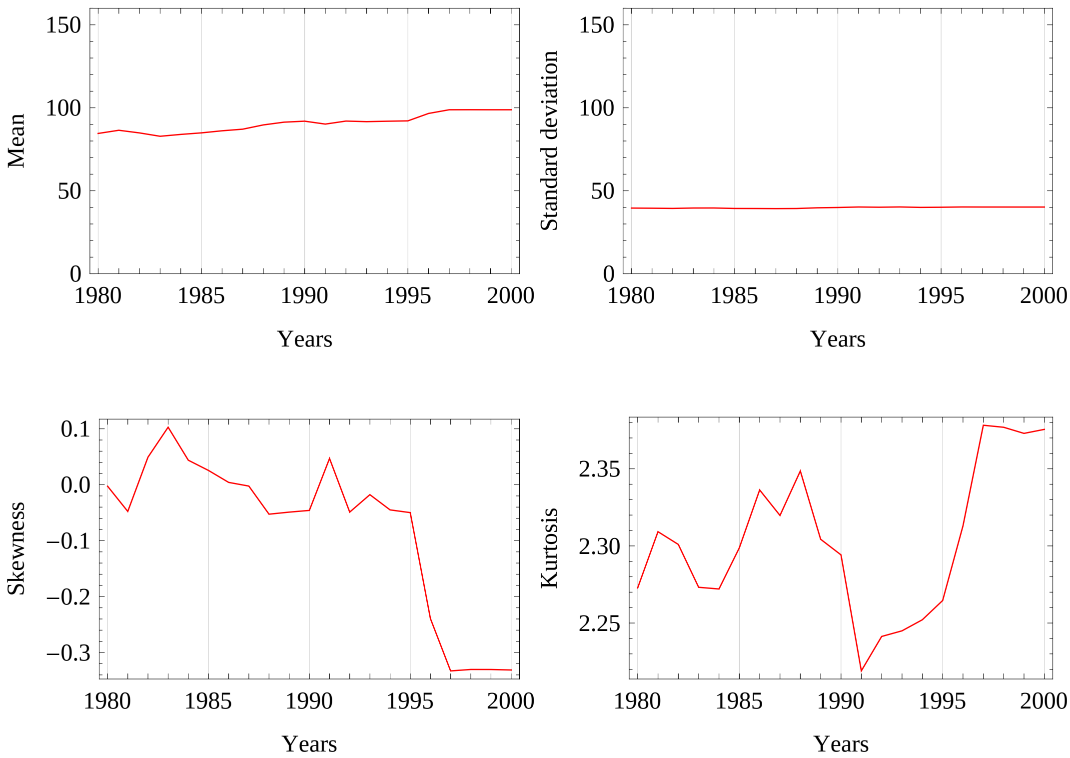
where α is a constant and γ is the exponent of the power law. Values in the range $2 \leq \gamma \leq 3$ are typical for the ND distribution of networks. Taking the logarithm of both sides of equation 5.4, we get:

$$\log P(d) = \log \alpha - \gamma \log d. \tag{5.5}$$

If we depict $P(d)$ as a function of d on log-log plot[1], we should observe a

[1] A log-log plot is a two-dimensional graph, which uses logarithmic scales on both axes.

Figure 5.5: Sample moments of the ND distribution: (Top-left) mean, (Top-right) standard deviation, (Bottom-left) skewness and (Bottom-right) kurtosis



Note: (Top-left) the scale set to possible values of the mean (from 0 to $N-1$) and (Top-right) the scale set to possible values of the standard deviation (from 0 to $N-1$)

straight line of slope $-\alpha$. An alternative (and a more convenient) method to visualize and detect a power law behavior is to draw the complementary cumulative distribution function (CCDF) on a log-log plot. The CCDF gives the probability that a random variable X with a given probability distribution is higher than or equal to x:

$$\tilde{F}(x) = P(X > x) = 1 - F(x), \tag{5.6}$$

where $F(x)$ is the cumulative distribution function (CDF). For our purpose the CCDF $\tilde{F}(d)$ describes the fraction of nodes, which has degree equal to or greater then d:

$$\tilde{F}(d) = \sum_{x=d}^{\infty} P(x). \tag{5.7}$$

Figure 5.6: NS distribution in 1980, 1990 an 2000: (Left) Kernel density estimation and (Right) Complementary cumulative distribution function

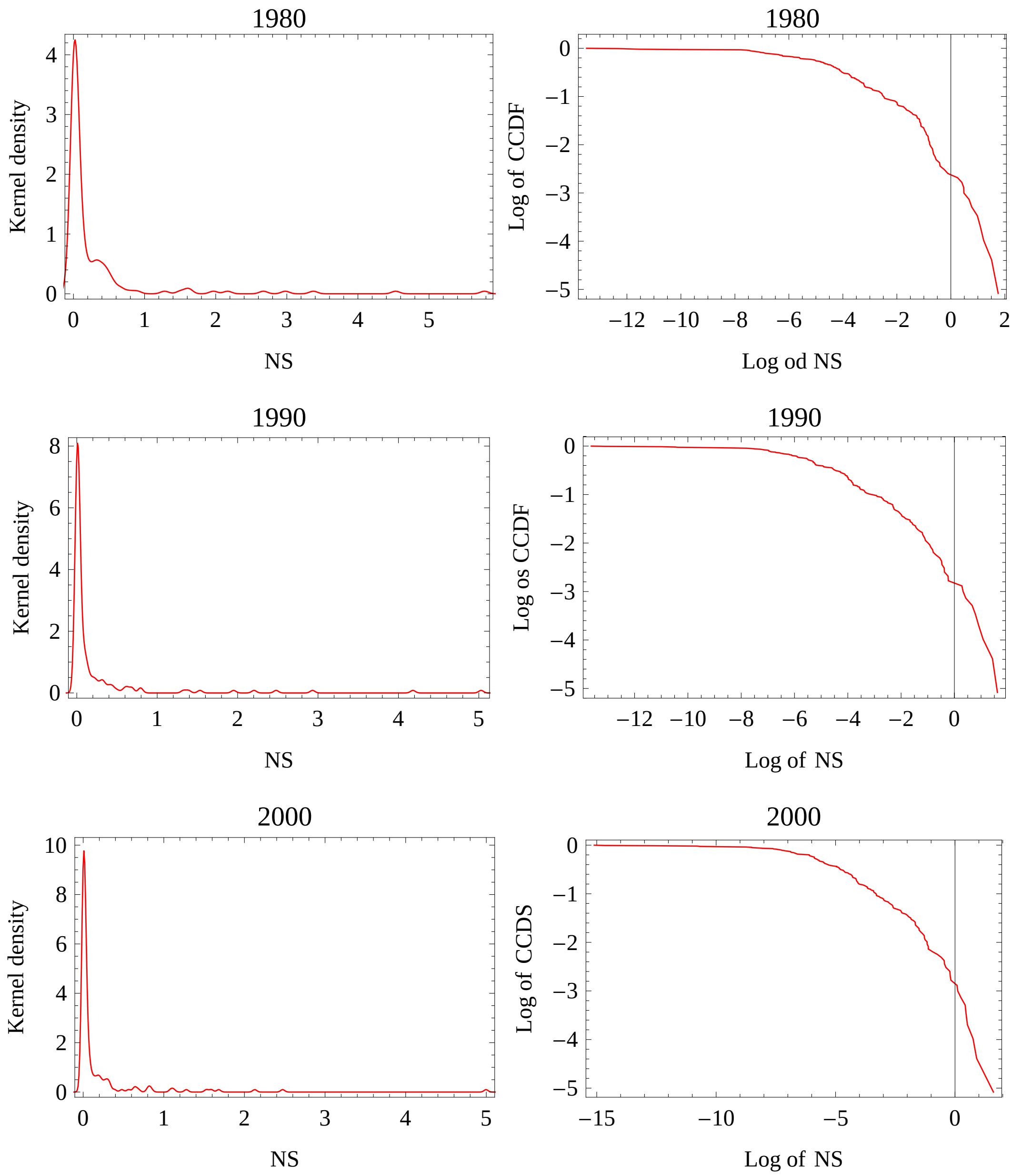


Table 5.2: P-values for NS normality tests

Test	1980	1990	2000
Shapiro-Wilk test	0.0000***	0.0023***	0.0002***
Jarque-Bera test	0.0007***	0.0044***	0.0005***

Note: (*) the null hypothesis rejected at 10%, (**) the null hypothesis rejected at 5% and (***) the null hypothesis rejected at 1%

Given relation 5.4 and $\gamma > 1$, the CCDF $\tilde{F}(d)$ takes the form:

$$\tilde{F}(d) = \sum_{x=d}^{\infty} P(x) = \alpha \sum_{x=d}^{\infty} x^{-\gamma} \simeq \alpha \int_{d}^{\infty} x^{-\gamma}\,\mathrm{d}x = \frac{\alpha}{\gamma - 1} d^{-(\gamma-1)}. \tag{5.8}$$

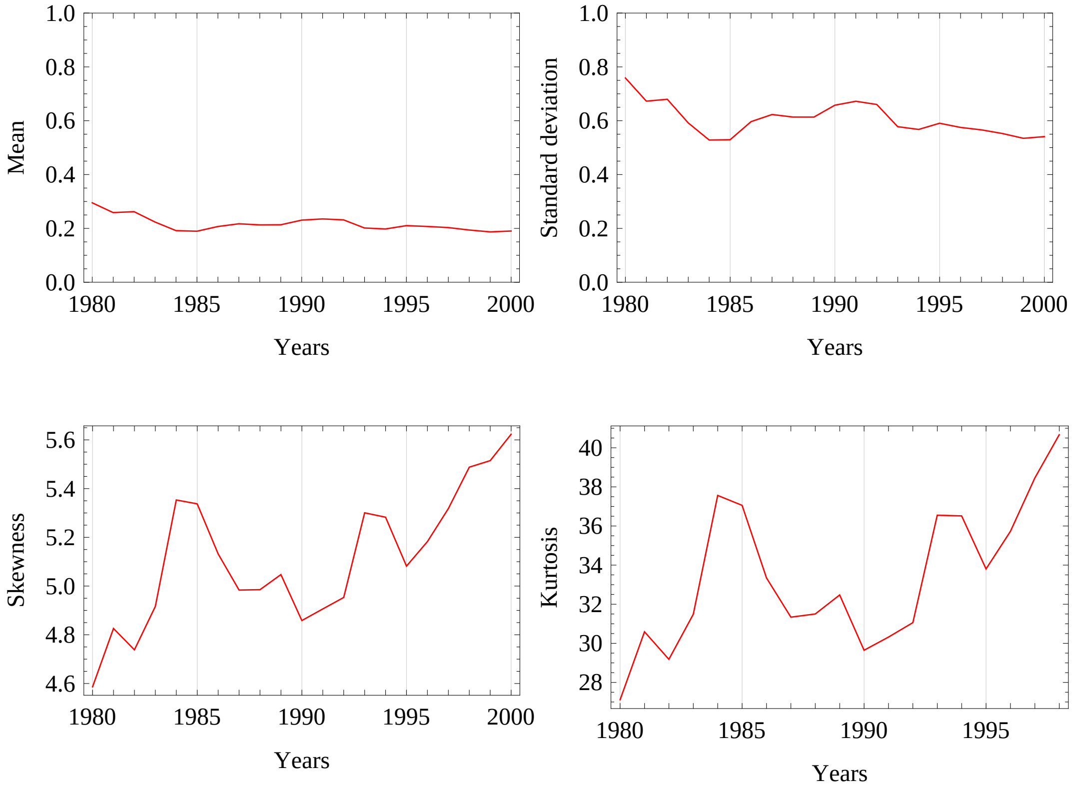
If a ND distribution follows a power law, then the CCDF of the distribution follows a power law. The only difference is that a given exponent is one less than the original exponent. The CCDF of a power law plotted on a log-log scale should also appear as a straight line. The curve of the CCDF shows a clear downward curvature, therefore the ND distribution does not match the power law (figure 5.4, right). We have also preferred to display the shape of the distribution using the CCDF instead of standard (two-tailed) CDF, because the existing literature on the ITN has taken an interest in the upper-tail behavior of node statistics[2].

To investigate whether the ND distribution follows a log-normal density we run normality tests on the log of the ND statistics. We employ the Shapiro-Wilk test [57] and the Jarque-Bera test [58, 59]. The hull hypothesis is that a random variable underlying the data is normally distributed. More specifically, the null hypothesis is that the log of the ND is normally distributed. Table 5.1 shows that the ND distribution is never log-normal, i. e. the log of ND is never normal. We can conclude that the ND distribution is approximated by neither power law or log-normal densities.

We further discuss in more detail the evolution of the first four moments of the ND distribution. Figure 5.5 shows that the ND distribution seems to be stable over time. As already noted in section 5.1, the binary ITN is characterized by an extremely high density. Each country holds on average 91 trade

[2] Analogical method is to draw a rank size plot, which is the transformation of a standard cumulative distribution function (CDF). Suppose that $\{x_1, \ldots, x_N\}$ are the available observations of a random variable X. Afterwards, sort the N observations to obtain $\{x_{(1)}, \ldots, x_{(N)}\}$, where $x_{(1)} \geq x_{(2)} \geq \cdots \geq x_{(N)}$. A rank size-plot depicts $\log r$ against $\log x_{(r)}$, where r is the rank. Since $r/N = 1 - F(x_{(r)})$, then $\log r = \log[1 - F(x_{(r)})] + \log N$.

Figure 5.7: Sample moments of the NS distribution: (Top-left) mean, (Top-right) standard deviation, (Bottom-left) skewness and (Bottom-right) kurtosis

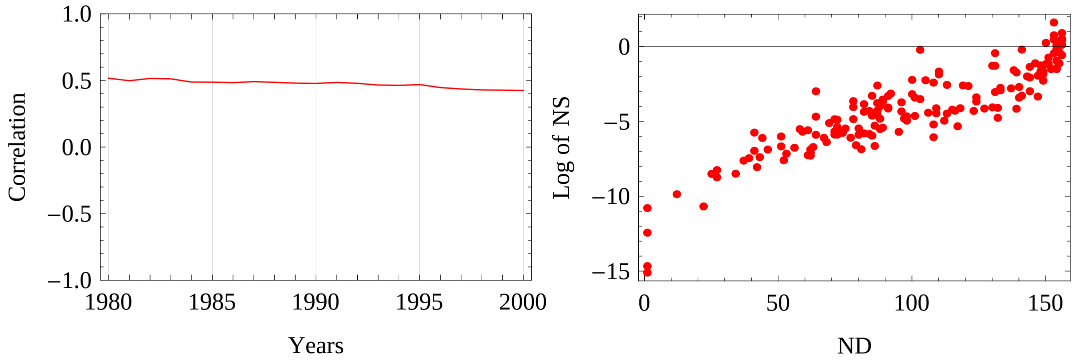


Note: (Top-left) the scale set from 0 to 1 and (Top-right) the scale set from 0 to 1

partners (over a maximum of 160). Moreover, the average ND has slightly increased over the years, which means that the number of trade relationships has been weakly growing during the time. The standard deviation displays a strong stability in the sample period, which implies that the integration has increased without any rise in the heterogeneity of the number of trade relationships. This conclusion is in contrast with the evolution of the skewness and kurtosis in the last few years of the sample. The ND distribution has become less symmetric (more left-skewed) and signalized fatter tails. We can therefore state that more countries now display extreme ND values.

The results changes significantly if one measures the connectivity in the weighted representation of the ITN. The researchers [15, 16, 17] show that the NS distribution of the ITN approximately follows a log-normal density. We compare the NS distribution to power law and log-normal densities. We can observe that the NS distribution is definitely right-skewed, which means that

Figure 5.8: NS-ND correlation patterns: (Left) correlation coefficient and (Right) scatter plot in 2000



Note: (Left) the scale set to possible values of the correlation (from −1 to 1)

the majority of countries characterized by weak trade relationships coexist with a small number of countries characterized by very intense trade relationships (figure 5.6, left). To verify whether the NS distribution is proxied by pre-defined densities, we apply the same procedures as in the case of the ND distribution. The CCDF displays a clear downward curvature (does not display a straight line), thus the NS distribution does not follow a power law (figure 5.6, right). Furthermore, table 5.2 indicates that the NS statistics is not even log-normal. These findings are in contrast with the previous papers. We can see that the NS statistics seems to be more log-normal than the ND one, however the results are still very clear. The explanation of this contrast might be that we have used a slightly different dataset. More specifically, our sample includes more distant observations. For example, we have not removed countries that have total exports equal to zero in some years. We can conclude that the NS distribution is proxied by neither power law or log-normal densities.

Although the first four moments of the NS distribution are more volatile than the moments of the ND distribution, the shape of the NS distribution seems to be quite stable over time (figure 5.7). The average NS is relatively low (at least in a [0,1] scale) in a comparison to the high average ND. Moreover, the weak increase in the ND is not matched by a similar behavior in the NS. The average NS rather seems to decline in the sample period. The recent wave of globalization resulted in an increased number connections, but they are characterized by a lower intensity. The decreasing trend of standard deviation indicates a decrease in the heterogeneity of the intensity of trade rela-

Figure 5.9: Node disparity distribution in 1980, 1990 an 2000: (Left) Kernel density estimation and (Right) Complementary cumulative distribution function

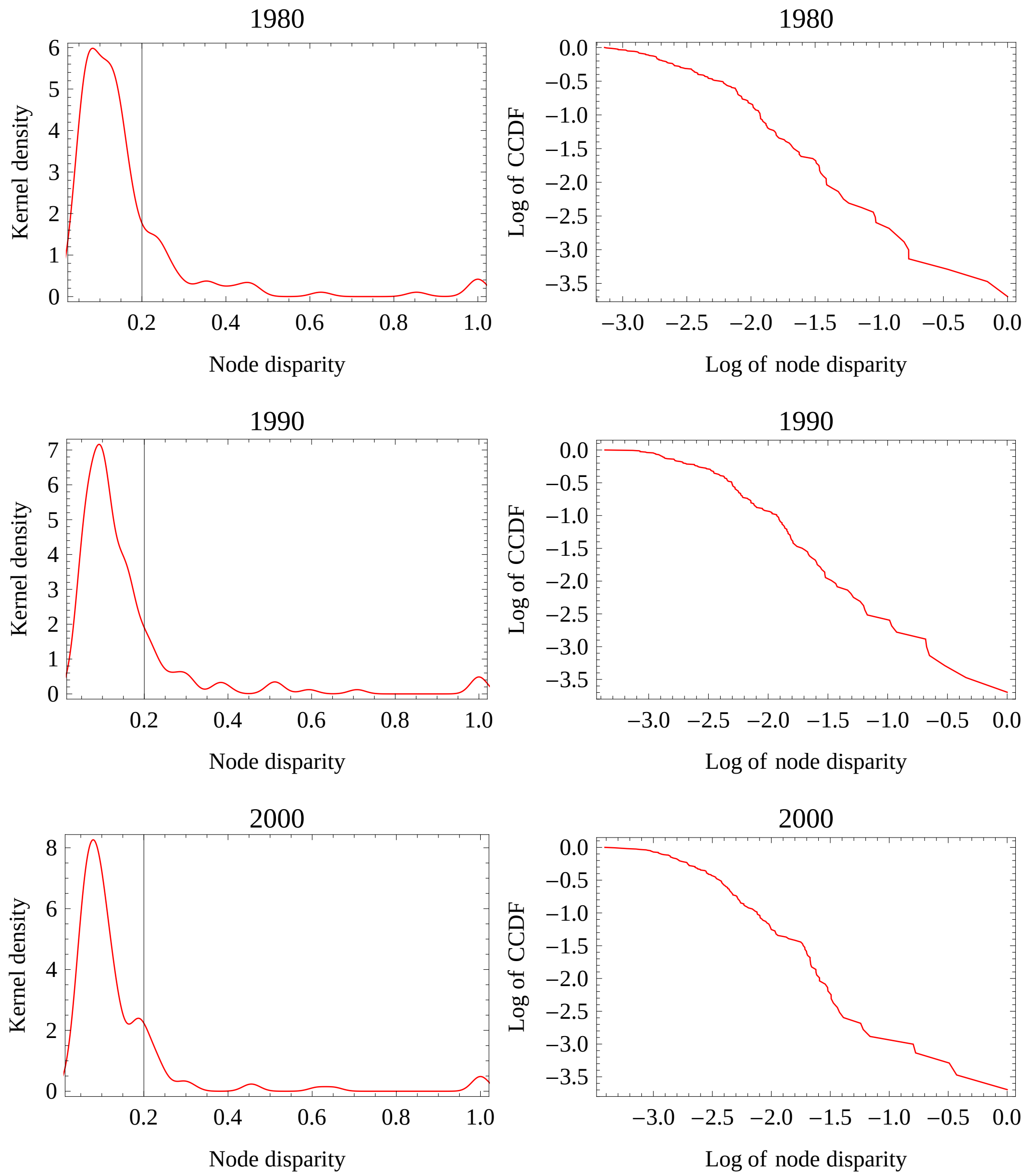


Table 5.3: P-values for node disparity normality tests

Test	1980	1990	2000
Shapiro-Wilk test	0.0000***	0.0000***	0.0000***
Jarque-Bera test	0.0009***	0.0005***	0.0000***

Note: (*) the null hypothesis rejected at 10%, (**) the null hypothesis rejected at 5% and (***) the null hypothesis rejected at 1%

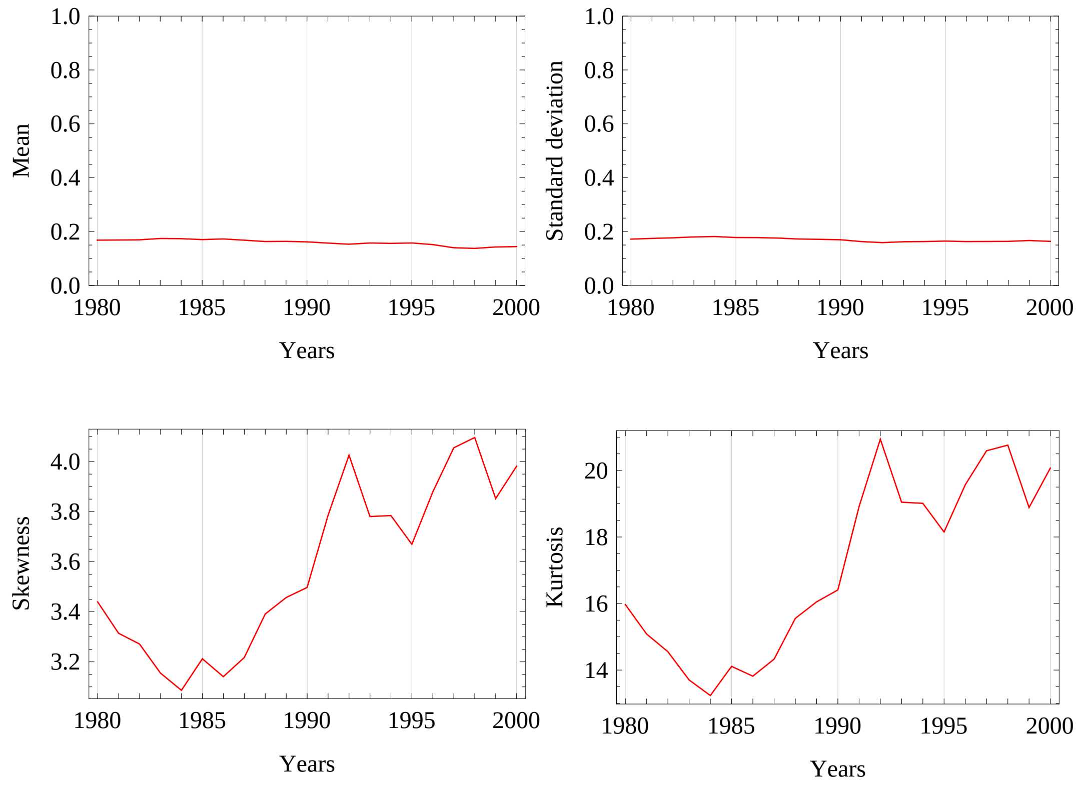
tionships. This conclusion is undermined by the evolution of the skewness and kurtosis. The NS distribution has become less symmetric (more right-skewed) and displayed the fatter tails. We cannot therefore say whether fewer or more countries now display extreme NS values.

The difference between the ND and NS distributions can be further examined by computing the correlation between the NS and ND. Figure 5.8 (left) shows that this correlation is quite stable around 0.47. The countries having many trade partners therefore tend to hold more intense trade relationships. However, a high ND does not automatically implies a high NS. This can be better appreciated by looking at NS-ND scatter plot (figure 5.8, right). We can see that the NS variability for any given ND value is relatively high in 2000. There are definitely countries having a low ND and relatively high NS. The conclusion is that only a subset of countries holding many trade partners actually have a very high NS.

Since the NS is only an aggregate measure of the interaction intensity mediated by a node, we finally measure the extent to which a node holds a dispersed (or concentrated) weight profile. The right-skewness of the NS distribution implies the right-skewed node disparity distribution. The majority of countries having a portfolio of very dispersed trade relationships therefore coexist with a fraction of countries concentrating almost all their trade relationships on a small number of partners. The CCDF does not still indicate a straight line, therefore the node disparity distribution does not follow a power law (figure 5.9, right). It does not also seem to be approximated by a log-normal density (table 5.3). This suggests that the node disparity distribution matches neither power law or log-normal densities.

The first four moments of the node disparity distribution have remained quite stable over the years, which hints to a relatively strong stability of the shape of the node disparity distribution. The average node disparity is relatively low (around 0.16), which means that on average countries hold more

Figure 5.10: Sample moments of the node disparity distribution: (Top-left) mean, (Top-right) standard deviation, (Bottom-left) skewness and (Bottom-right) kurtosis

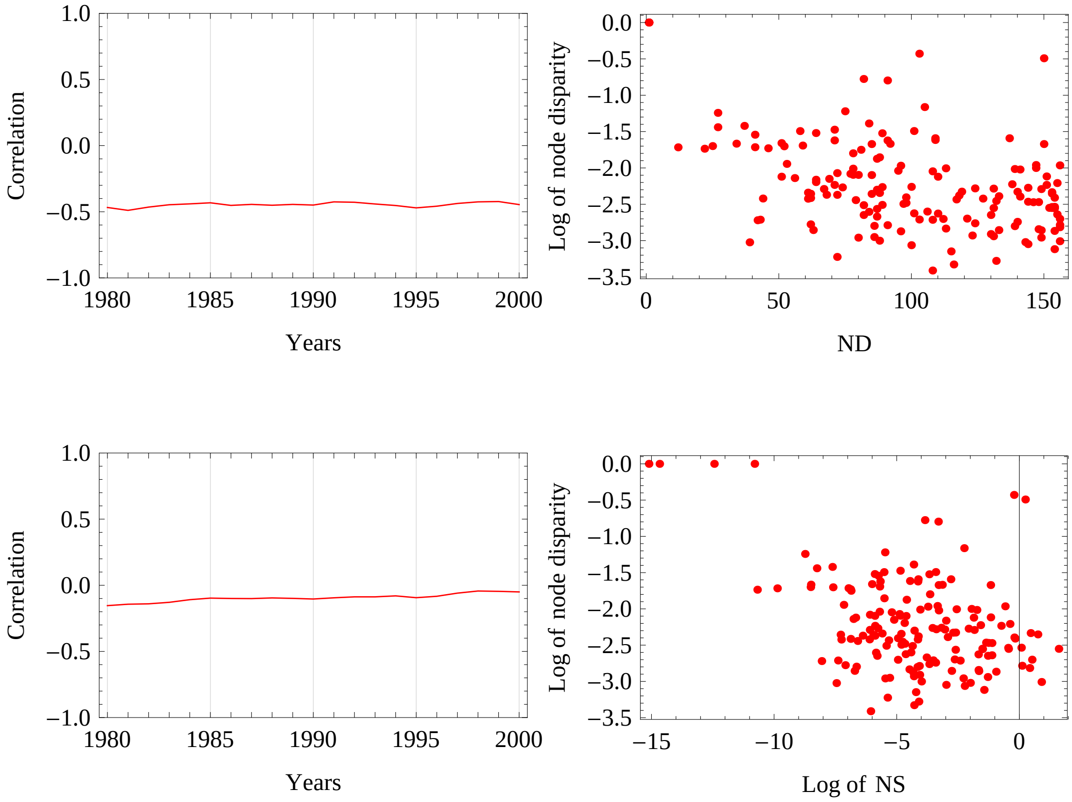


Note: (Top-left) the scale set to possible values of the mean (from 0 to 1) and (Top-right) the scale set to possible values of the standard deviation (from 0 to 1)

dispersed trade relationships. The increase of the skewness and kurtosis of the NS is matched by a similar behavior for the skewness and kurtosis of the node disparity (figure 5.10). More interestingly, the node disparity is negatively correlated with both ND (on average -0.45) and NS (on average -0.10). These correlation patterns are also visible on the scatter plots in year 2000 (figure 5.11). We can conclude that if country holds more trade partners and more intense trade relationships, then its trade portfolio is more differentiated. This conclusion can be partially expected, because in the case of equally-distributed weights node disparity should equal to the inverse of the ND.

The aforementioned results show that the binary representation of the ITN leads to a highly connected network. On the other hand, the picture changes substantially if the ITN is analysed from a weighted perspective. In that case,

Figure 5.11: Node disparity correlation patterns: (Left) correlation coefficient and (Right) scatter plot in 2000

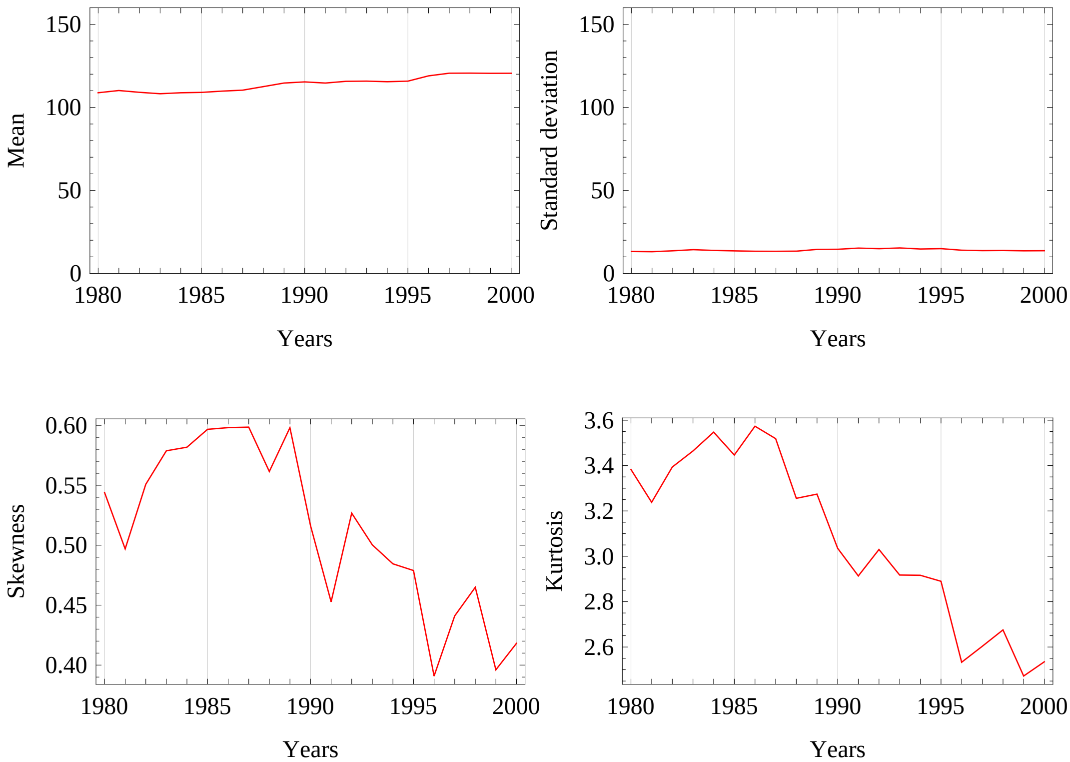


Note: (Top-left) the scale set to possible values of the correlation (from −1 to 1) and (Bottom-left) the scale set to possible values of the correlation (from −1 to 1)

a majority of trade flows are weak and coexist with a few very intense trade partnerships. This findings reflects the difference between intensive and extensive interpretations stated in the microeconomic trade literature. The export intensity is found to be much more important than the number of exporting firms in explaining aggregate export performances.

Based on the first part of analysis, we can make an important general conclusion. A binary approach cannot fully extract the information about the intensity of trade relationships. If a network is analysed from a binary perspective, a lot of information might be disregarded and a role of heterogeneity in trade relationships might be dramatically underestimated. A weighted network analysis can instead provide a more complete description of the underlying topological structure of the ITN.

Figure 5.12: Sample moments of the ANND distribution: (Top-left) mean, (Top-right) standard deviation, (Bottom-left) skewness and (Bottom-right) kurtosis



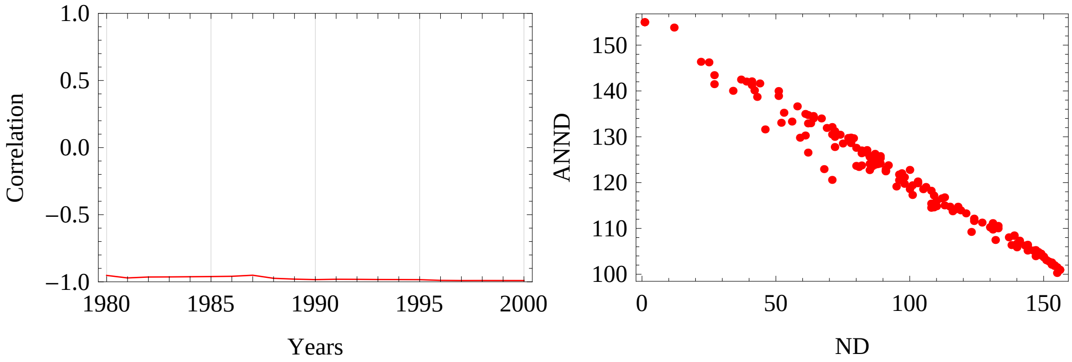
Note: (Top-left) the scale set to possible values of the mean (from 0 to $N-1$) and (Top-right) the scale set to possible values of the standard deviation (from 0 to $N-1$)

5.3 Assortativity

The ND and NS statistics are only first-order indicators, because they take into account the information about nodes lying one step away from the original one. In other words, they do not provide any information about the greater structure of the network. In fact, countries holding many trade relationships can only trade with poorly connected countries. Such a network is called "disassortative". On the contrary, countries holding many linkages can tend to trade with other highly connected countries. In that case, the network is referred as "assortative". The statistics describing assortativity are second-order indicators, as they look at nodes that are two step away from the analysed one.

We begin to explore the assortativity in the ITN from a binary perspective. More specifically, we explore the behavior of the average nearest-neighbor de-

Figure 5.13: ANND-ND correlation patterns: (Left) correlation coefficient and (Right) scatter plot in 2000

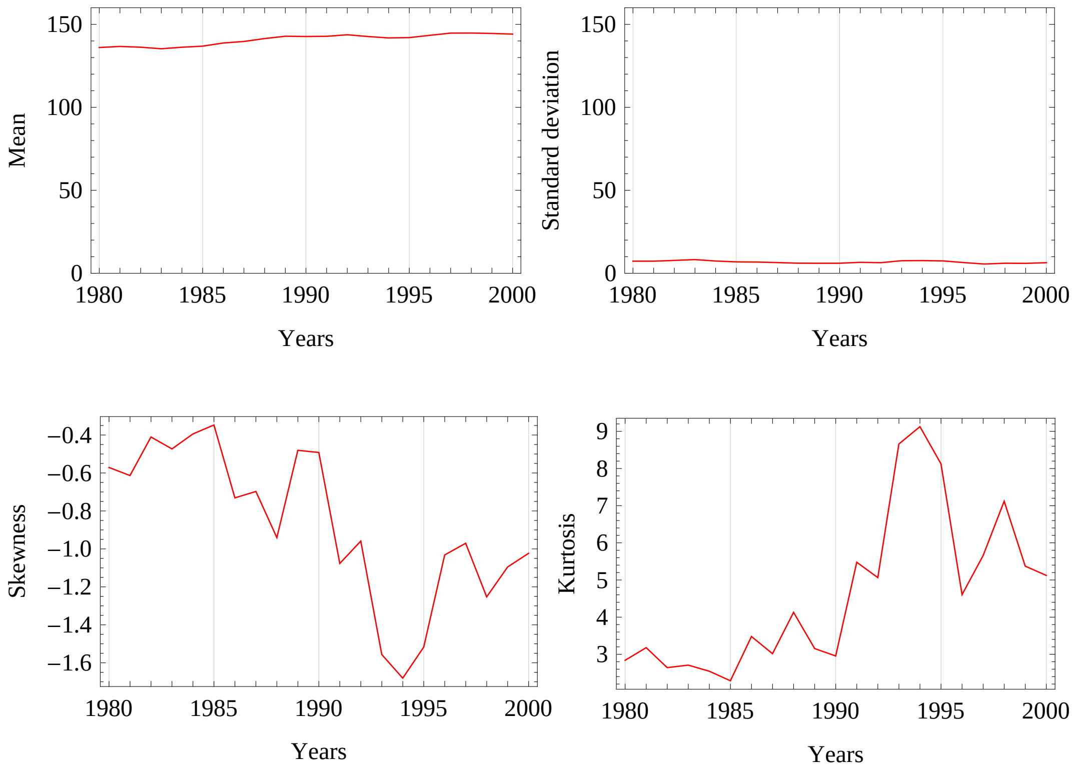


Note: (Left) the scale set to possible values of the correlation (from −1 to 1)

gree (ANND) and its correlation with other network statistics. The previous papers [9, 10, 11, 15, 16, 17] present a strongly disassortative network. We basically confirm their results. The first four moments of the ANND distribution display a relatively stable pattern in the sample period, which implies a stability of the shape of the ANND distribution. The average ANND stays always above the average ND, which means that the partner of a country holds on average more trade relationships than a given country. Specifically, each partner holds on average 114 trade relationships (over maximum of 160). The reduction of the skewness and kurtosis in the second half of the sample signifies that fewer countries now display extreme ANND values (figure 5.12). The correlation between the ANND and ND clearly signalizes a strongly disassortative network, as the figures stays very close to −1 in the entire period. Moreover, the correlation is characterized by a very limited variability. The scatter plot shows a linear dependence between the ANND and ND (figure 5.13). In the ITN viewed as a binary network, countries with many relationships definitely trade with countries holding few partnerships.

If the ITN is studied from a weighted perspective, the disassortative nature of the ITN remains evident, but the results are much weaker [15, 16, 17]. We measure the intensity of trade relationships carried by partners of a node by the weighted average nearest-neighbor node degree (WANND) and the average nearest-neighbor node strength (ANNS). Figure 5.14 shows the evolution of the first four moments of the WANND over the years. The stable moments display a stability of the shape of the WANND distribution. The average WANND

Figure 5.14: Sample moments of the WANND distribution: (Top-left) mean, (Top-right) standard deviation, (Bottom-left) skewness and (Bottom-right) kurtosis

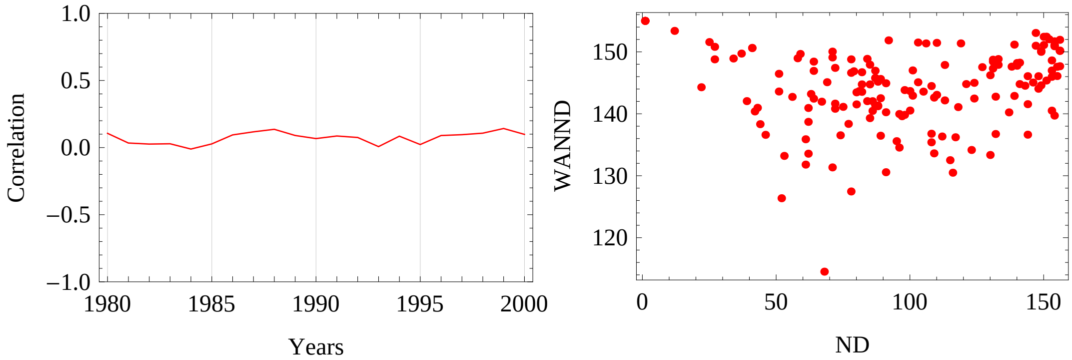


Note: (Top-left) the scale set to possible values of the mean (from 0 to $N-1$) and (Top-right) the scale set to possible values of the standard deviation (from 0 to $N-1$)

remains always above the average ANND, which means that links with larger weights point out to the neighbors with the larger ND. The decline of the skewness and the increase of the kurtosis indicate that more countries now have extreme WANND values. The correlation between WANND and ND is on average 0.07 across the years. Furthermore, the corresponding scatter plot is characterized by a much more dispersed cloud of points (figure 5.15).

Similarly, the stable first four moments of the ANNS distribution hint to a stability of the shape of the ANNS distributions. The average ANNS is always above the average NS, which implies that the partner of a country holds on average more intense trade relationships than a given country. The evolution of the skewness and kurtosis basically shows a constant trend (figure 5.16). The correlation between the ANNS and NS is still negative but weaker in the

Figure 5.15: WANND-ND correlation patterns: (Left) correlation coefficient and (Right) scatter plot in 2000

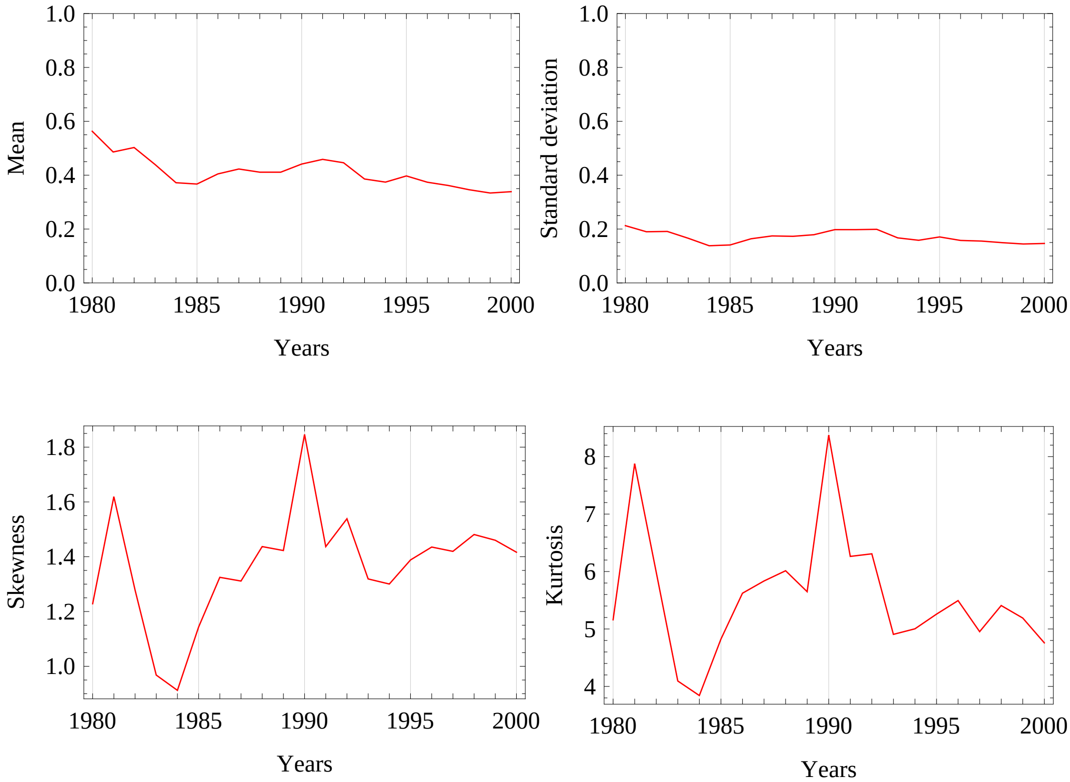


Note: (Left) the scale set to possible values of the correlation (from -1 to 1)

magnitude in all years (on average -0.37). Moreover, the ANNS-NS scatter plot shows a higher variability of the correlation (figure 5.17). If the ITN is analysed from a weighted perspective, the disassortative pattern does not seem to hold that robustly. The study of the ITN from a weighted perspective just provides a different (and more insightful) picture.

The disassortative nature of the ITN indicates that countries holding many and more intense relationships tend to trade with less and more weakly connected countries. This evidence suggests that the ITN has a core-periphery structure, which is common pattern in many social and economic networks. For example, Hojman and Szeidl [60] propose a model of network formation, where the unique equilibrium network architecture is the "periphery-sponsored star". In this equilibrium, there is only one country as a center and all other countries maintain one link to that center. Alternatively, Rombach et al. [61] develop a new method to investigate the core-periphery structure, which can identify multiple cores in a network and takes into account different possible cores. The core-periphery structure often implies that peripheral countries suffer from a sort of marginalization. Fagiolo et al. [17] specifically state that such a polarized structure is not necessarily the most efficient outcome and that a more balanced structure of trade relationships would allow both developing and industrialized countries to better exploit the gains from trade. Moreover, we have already shown that the WANND-ND and ANNS-NS scatter plots are characterized by a much more dispersed cloud of points. This means that there exists a number of countries holding relatively many partners or rel-

Figure 5.16: Sample moments of the ANNS distribution: (Top-left) mean, (Top-right) standard deviation, (Bottom-left) skewness and (Bottom-right) kurtosis



Note: (Top-left) the scale set from 0 to 1 and (Top-right) the scale set from 0 to 1

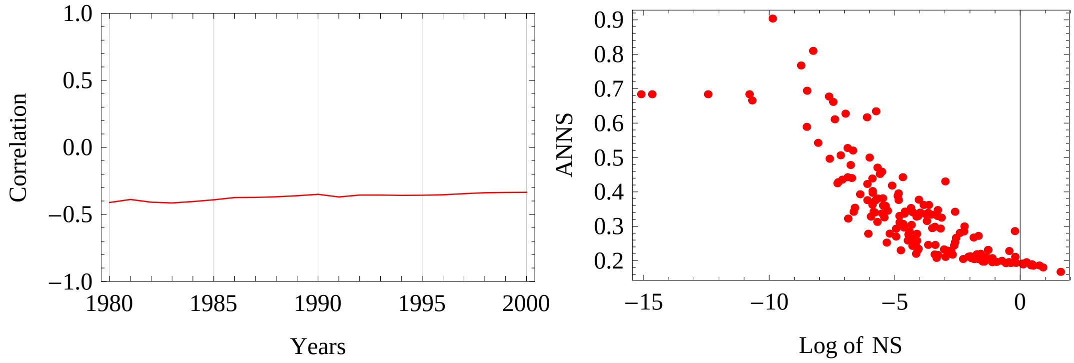
atively intense relationships, which tend to trade with well connected countries. In other words, there exist an intermediate periphery within the core-periphery structure of the ITN.

5.4 Clustering

We now turn to investigating clustering patterns and how they correlate with the connectivity. The main question is whether more and better connected countries tend to establish trade relationships with countries that also trade with each other. The statistics measuring clustering are also second-order indicators, because they take into account nodes lying two step away from the one under analysis.

We initially address the issue of clustering in terms of binary network. According to papers [9, 10, 11, 15, 16, 17], the average binary clustering coefficient

Figure 5.17: ANNS-NS correlation patterns: (Left) correlation coefficient and (Right) scatter plot in 2000

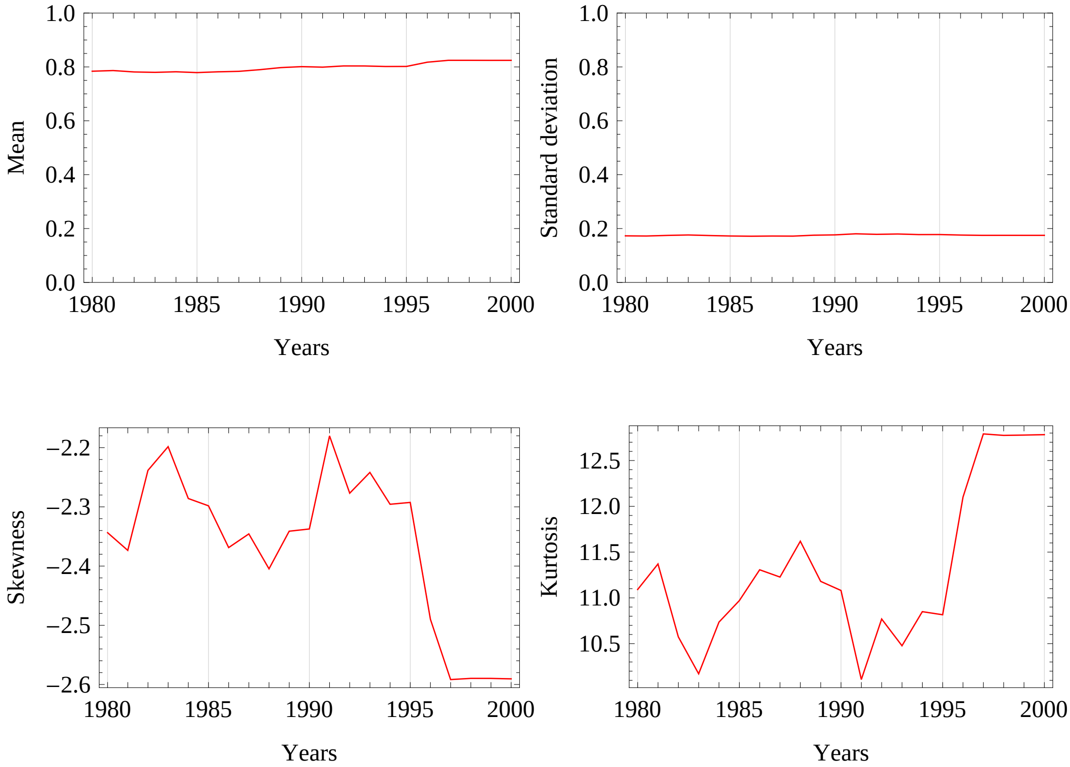


Note: (Left) the scale set to possible values of the correlation (from −1 to 1)

(BCC) is very high and countries holding more trade partners are usually associated to lower clustering coefficients. We confirm these results for our data. Figure 5.18 shows the behavior of the BCC over the years. The first four moments have remained quite stable, which implies the stability of the shape of the BCC distribution. The average BCC has a slightly increasing trend around 0.80. Moreover, the average BCC stays always above the density of the network (figure 5.1, right). In a random network, where each link is in place with probability $p \in (0, 1)$, the expected value of the BCC is equal to p. The probability of the placement of each link is equivalent to the density of the network, therefore in a random network the expected value of the BCC is equal to its density. This means that the binary ITN is statistically more clustered then its random counterpart. Therefore, the countries tend to form on average trade relationships with countries that also trade with each other. In other words, the link between any two partners of a given node is very likely to be present. This conclusion implies that local (regional) links still play a very important role. The localism in this sense might not have automatically a geographic meaning, but it rather represents a tendency to interact with traditional partners. Such an interpretation is also confirmed by the fact that geographically structured networks are usually highly clustered with more short-distance links. The traditional members can be members of regional group, countries with similar level of development or simply historically close partners.

Figure 5.19 displays the correlation patterns across the years from the binary perspective. The correlation between the BCC and ND is quite strong and

Figure 5.18: Sample moments of the BCC distribution: (Top-left) mean, (Top-right) standard deviation, (Bottom-left) skewness and (Bottom-right) kurtosis



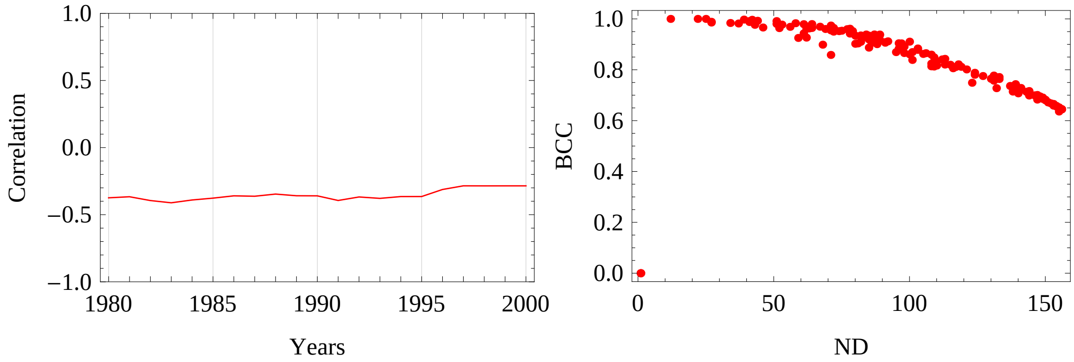
Note: (Top-left) the scale set to possible values of the mean (from 0 to 1) and (Top-right) the scale set to possible values of the standard deviation (from 0 to 1)

negative (around −0.35). Countries holding more trade parters are therefore less clustered than countries holding few partners. Alternatively, partners of well connected countries are less interconnected than partners of poorly connected countries. This property is sometimes referred as "hierarchy". The negative correlation between the BCC and ND is also reflected in the scatter plot. The correlation patterns support the hypothesis that the ITN has a core-periphery structure. Countries holding few trade partnerships simply do not trade with each other, but they are connected to the hubs.

The conclusion changes again if one take into account the intensity of trade relationships. The researchers [15, 16, 17] indicate that the average weighted clustering coefficient (WCC) is actually very low. Furthermore, they display that the correlation between the WCC and NS is very strong and positive[3].

[3]To be more accurate, the strength of the correlation depends on a weighting scheme

Figure 5.19: BCC-ND correlation patterns: (Left) correlation coefficient and (Right) scatter plot in 2000



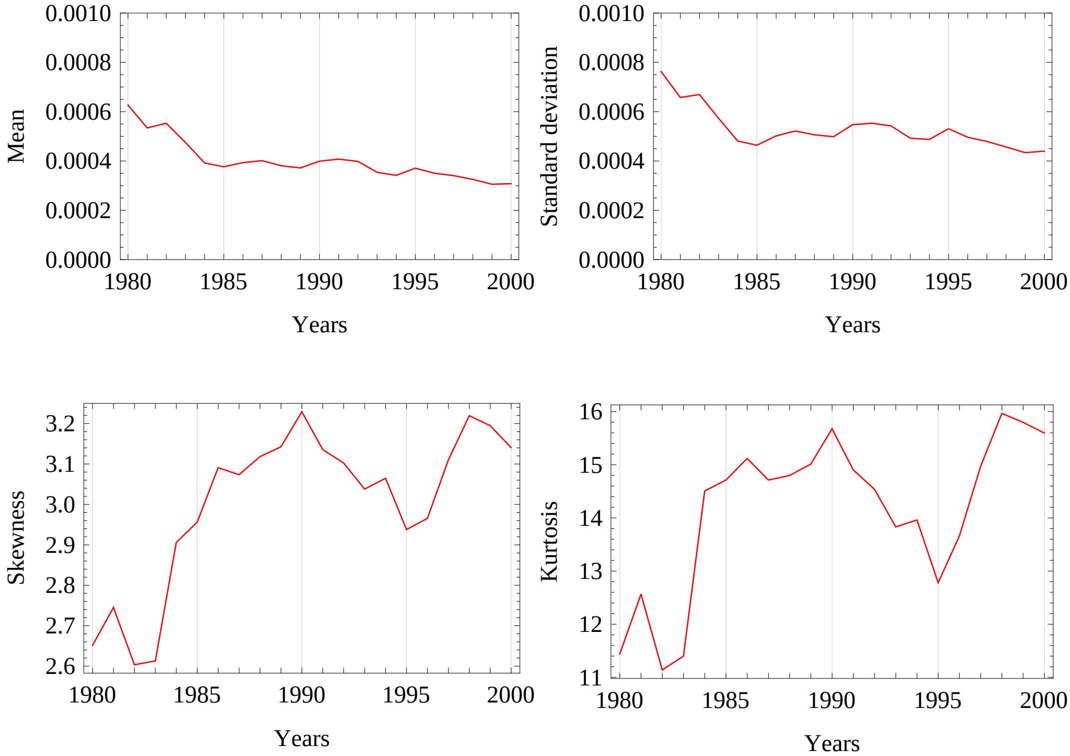
Note: (Left) the scale set to possible values of the correlation (from -1 to 1)

Our weighted analysis also gives the opposite results. The first four moments seem to be still stable over the years, which implies the stability of the shape of the WCC distribution (figure 5.20). The average WCC is however significantly smaller than its expected value in the random network. The expected value of the WCC in a random network is equal to $(\frac{3}{4})^3 p$. The probability p is still equivalent to the density of the network. The average WCC ranges from 0.0003 in 1999 to 0.0006 in 1980, whereas the expected value of the WCC moves from 0.1971 to 0.2556 in the corresponding years. More interestingly, the WCC distribution lies to the left of random-network expected values in every year. This means that no country is ever characterized by a WCC, which is above the expected value.

The correlation between the WCC and NS is now very strong and positive (around 0.92). The scatter plot also confirms this evidence (figure 5.21). We can conclude that countries holding more intense trade relationships are more likely to establish highly connected trade triangles. This feature reminds the so-called "rich club phenomenon". The rich club represents a small group of rich countries, which control a large part of the world trade. Bhattacharya et al. [13] show that a few top rich countries control one half of the world's total trade volume. Similarly, Fagiolo et al. [16] present that the 10 richest countries in terms of the NS are responsible for about 40% of the total trade flows. The presence of the "rich club phenomenon" is also in accordance with the paper by Furusawa and Konishi [62], who propose a new model to examine the formation

used. Fagiolo et al. [16] show the WCC-NS correlation to be very close to 1 in all years. Other authors [15, 17] suggest a sharply increasing WCC-NS correlation across time.

Figure 5.20: Sample moments of the WCC distribution: (Top-left) mean, (Top-right) standard deviation, (Bottom-left) skewness and (Bottom-right) kurtosis



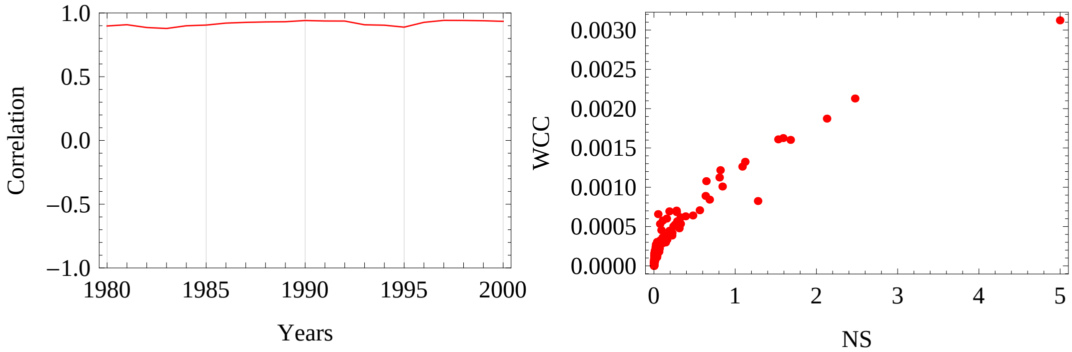
Note: (Top-left) the scale set from 0 to 0.0010 and (Top-right) the scale set from 0 to 0.0010

of free trade agreements (FTAs). They show that under several assumptions countries sign a trade agreement only if their industrialization levels are not very different.

5.5 Stability of Probability Distributions

We have already discussed the stable patterns for the sample moments of the distributions of all node statistics. If the moments seemed to be quite stable throughout the whole period (figure 5.5, 5.7, 5.12, 5.14, 5.16, 5.18 and 5.20), we have concluded that the distribution of a node statistic is stable. Let us now examine this evidence more quantitatively. We compute the time average of the absolute value of 1-year growth rates of the first four moments of all node statistics (ND, NS, ANND, WANND/ANNS, BCC and WCC). The time average of the absolute value of 1-year growth rates of the k-th moment of the

Figure 5.21: WCC-NS correlation patterns: (Left) correlation coefficient and (Right) scatter plot in 2000



Note: (Left) the scale set to possible values of the correlation (from -1 to 1)

statistic is defined as:

$$gr_t = \frac{1}{T-1} \sum_{t=2}^{T} \left| \frac{M^k(X_i^t) - M^k(X_i^{t-1})}{M^k(X_i^{t-1})} \right|, \tag{5.9}$$

where X_i^t is the value of the node statistic X at time t for country i and $M^k(\cdot)$ is the moment operator that for $k = 1, 2, 3, 4$ gives respectively the mean, standard deviation, skewness and kurtosis.

The results are summarized in table 5.4. The average absolute growth rates are generally lower for the moments of the binary statistics, which hints to a higher stability of the binary representation of the ITN. The average absolute growth rates of the first two moments range in our sample between 0.0034 and 0.0640, which indicates a relatively high stability of them. The average absolute growth rates of the skewness for the WANND and ANNS display somewhat larger values (0.2911 respectively 0.1152). Similarly, the average absolute growth rates of the kurtosis for the WANND and ANNS are larger (0.2407 respectively 0.1544). These findings confirm that the weighted version of the ITN displays a lower stability. The biggest issue relates the skewness of the ND. The average absolute growth rate is very high (2.7667), however the figure 5.5 shows that the skewness seems to be quite stable over time. The skewness of the ND is almost zero in several years, therefore the average absolute growth rate reaches such a high value. We can conclude that the moments of all indicators have remained quite stable over the years, which implies that structural properties of the ITN viewed either as a binary

Table 5.4: Average absolute growth rates

Statistics	Mean	Std. dev.	Skewness	Kurtosis
ND	0.0145	0.0034	2.7667	0.0104
ANND	0.0079	0.0264	0.0746	0.0388
BCC	0.0043	0.0078	0.0254	0.0363
NS	0.0528	0.0477	0.0275	0.0649
WANND	0.0058	0.0605	0.2910	0.2407
ANNS	0.0595	0.0640	0.1152	0.1544
WCC	0.0611	0.0596	0.0556	0.0059

or as a weighted network have not been much influenced by the process of globalization. This conclusion is consistent with the paper by Garlaschelli and Loffredo [11].

5.6 Binary vs Weighted Network

The recent papers have shown that the ITN [15, 16, 17] is a great example of the weighted networks. In order to get a more complete picture, we have decided to investigate both binary and weighted versions of the ITN. We compare connectivity (ND and NS), assortativity (ANND and WANND/ANNS) and clustering (BCC and WCC). We have shown that the topological properties of the ITN viewed as weighted network are significantly different from those obtained by a binary network approach. The comparison of both approaches is summarized in table 5.5.

Our findings suggests that accounting for heterogeneity in the capacity and intensity of the trade relationships is crucial to better understand the architecture of complex networks. The binary representation of the ITN leads to a highly connected network, where all links have the same effect on the presented statistics. Such a highly connected graph almost automatically implies very large values of the ANND and BCC for the majority of nodes. This might lead to the biased computation of correlation patterns. In fact, the links are characterized by a very different interaction intensity in the ITN. The weighted representation of the ITN suggests that the majority of links show very low export/import flows. Furthermore, the network seems to be only weakly disassortative. Finally, countries holding more intense trade relationships tend to form strongly connected trade triangles, The only statistical feature, which is common for both approaches, is the constancy of the network properties over

Table 5.5: Binary vs weighted network

	Connectivity	Assortativity	Clustering
Binary network	Highly connected + Bimodal ND distribution	Strongly disassortative	Highly clustered + Negative BCC-ND correlation
Weighted network	Weakly connected + Skewed NS distribution	Weakly disassortative	Weakly clustered + Positive WCC-NS correlation

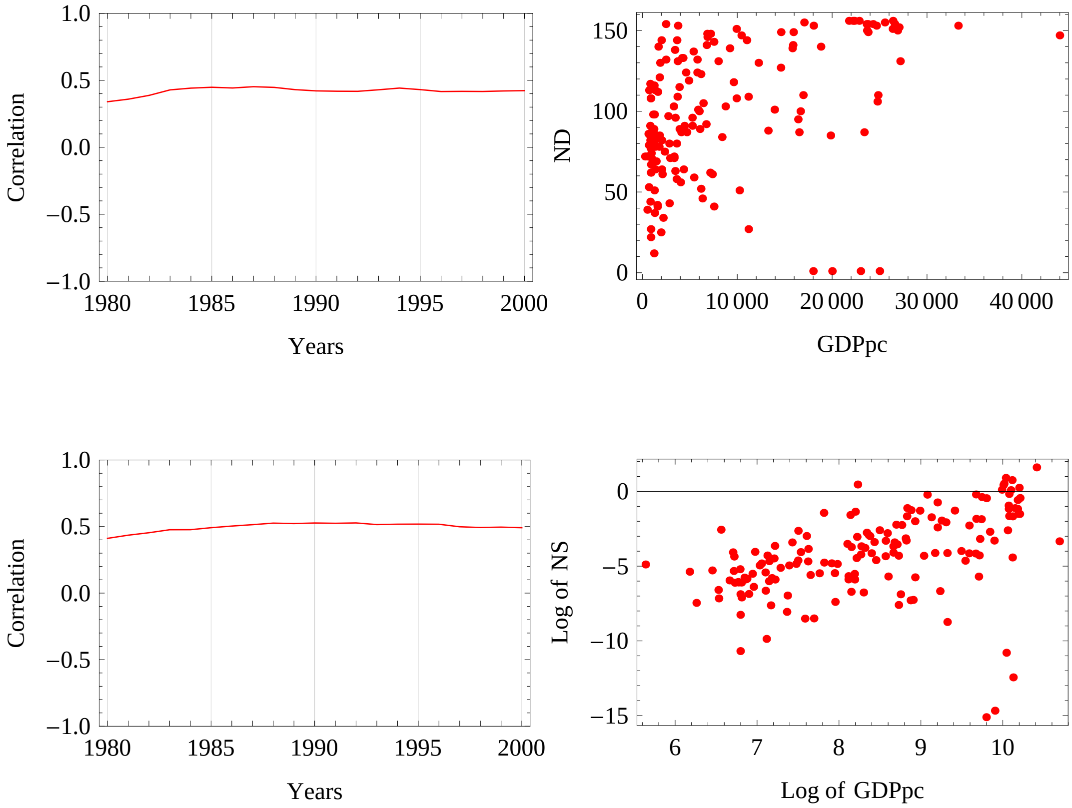
time. We can therefore conclude that taking into account the intensity of the trade relationships allow us to better appreciate the topological properties of the analysed network.

5.7 Country-specific Characteristics

We further explore a relation between network properties and country-specific characteristics. Baier and Bergstrand [63] show that country characteristics determine the formation of free trade agreements. Alternatively, Kali and Reyes [43] outline that a country's position in the network can have substantial implications for economic growth. Garlaschelli and Loffredo [11] confirm the interplay between topology of the ITN and the dynamic of the GDP. They show that at each time-step the GDP distribution is determined by the ITN as a weighted network and at the same time the ITN topology is determined by the GDP values. We specifically explore the correlation patterns between network statistics and the GDP per capita (GDPpc) in order to see whether countries with higher income are more or less integrated into world trade. For the purpose of the analysis, we employ the GDPpc expressed in 1996 US dollars.

We start with the correlation patterns between connectivity levels and the GDPpc. Figure 5.22 (left) signifies that the correlation appears to be relatively strong and positive both in terms of the number of trade partners (ND) and the intensity of trade relationships (NS). High-income countries therefore tend to hold more partners and more intense relationships. We can further see that

Figure 5.22: Connectivity-GDPpc correlation patterns: (Left) correlation coefficient and (Right) scatter plot in 2000



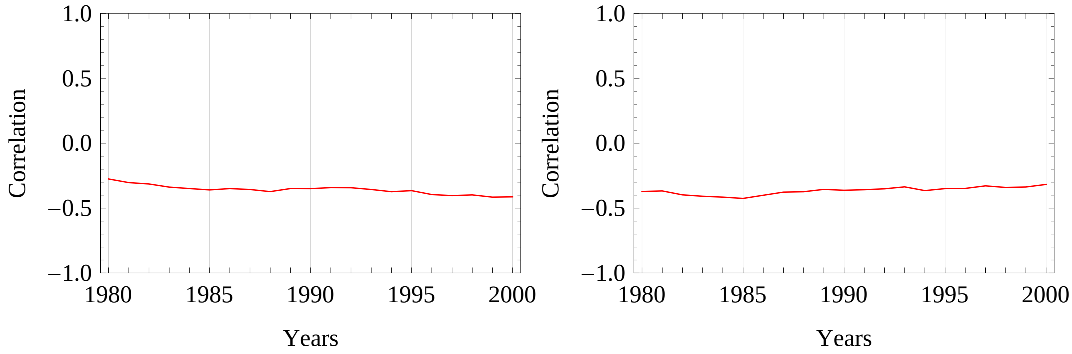
Note: (Left) the scale set to possible values of the correlation (from −1 to 1)

the NS-GDPpc correlation (on average 0.50) is slightly stronger than the ND-GDPpc one (on average 0.42). This pattern is also confirmed by the scatter-plots in year 2000 (figure 5.22, right). Whereas there seems to be a linear dependence between the ND and GDPpc, a log-log relation is observed between the NS and GDPpc. This implies that the GDPpc has a larger effect on the NS rather than the ND.

The outcome is also very clear as far as the assortativity is concerned. Figure 5.23 shows that the correlation between assortativity levels and the GDPpc is relatively strong and negative both in terms of the ANND (on average −0.36) and ANNS (on average −0.37). The conclusion is now that high-income countries tend to trade with less and more weakly connected countries.

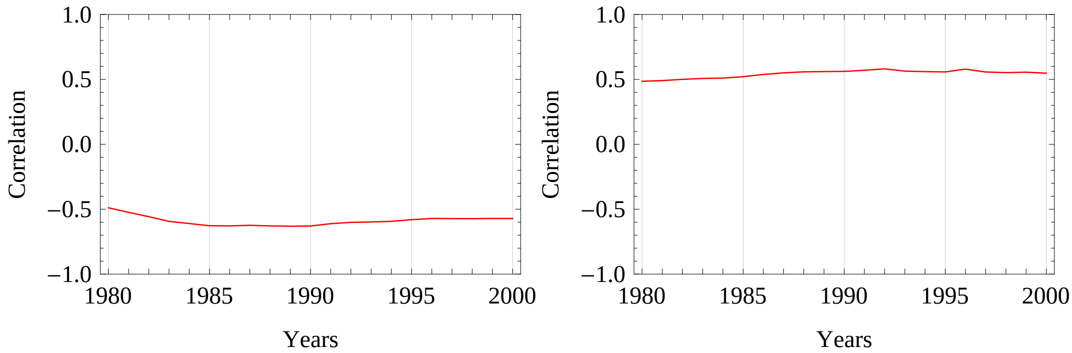
We finally turn to investigate the correlation patterns between the clustering and GDPpc. Figure 5.24 suggests that the results for correlation between clustering and GDPpc mimic those obtained for the correlation between the

Figure 5.23: Assortativity-GDPpc correlation patterns: (Left) ANND-GDPpc correlation coefficient and (Right) ANNS-GDPpc correlation coefficient



Note: (Left) the scale set to possible values of the correlation (from −1 to 1) and (Right) the scale set to possible values of the correlation (from −1 to 1)

Figure 5.24: Clustering-GDPpc correlation patterns: (Left) binary CC-GDPpc correlation coefficient and (Right) weighted CC-GDPpc correlation coefficient



Note: (Left) the scale set to possible values of the correlation (from −1 to 1) and (Right) the scale set to possible values of the correlation (from −1 to 1)

clustering and connectivity. The BCC-GDPpc correlation seems to be relatively strong and negative (on average −0.59), whereas the WCC-GDPpc correlation is relatively strong and positive (on average 0.54). Considering the weighted version of the ITN, we can conclude that high income countries tend to form highly connected trade triangles. This result supports the "rich club phenomenon", which has been already discussed above.

Chapter 6

Conclusion

In this thesis, we have employed a network analysis to explore the topological properties of the international trade network (ITN). More specifically, we have described trade relations as a network, where countries play the role of nodes and trade flows represent links between nodes. This allows us to better explain the degree of international economic integration. We have studied export and import flows among 161 world countries over the period 1980-2000. We have build on other studies and we have provided a more thorough analysis of the ITN. Beside a full description of the theoretical background and a review of all relevant literature, we have presented a comprehensive empirical analysis of the topological properties of the ITN.

We have explored the topological properties of the ITN from a purely descriptive perspective. Indeed, we have tried to characterize some robust stylized facts pertaining to the evolution of the ITN. A network analysis allows us to investigate not only first-order indicators associated to direct bilateral-trade relationships of a given country but also second-order and higher-order empirical facts. For example, we have been able to study the extent to which well connected countries tend to hold trade relationships with partners that are themselves well connected, the probability that partners of well connected countries are themselves partners, the importance of countries in the network and so on. The empirical regularities displayed by the data can be used as a starting point to explain and replicate the structure of the ITN. In other words, empirical regularities can provide some guidance for theoretical models that attempt to explain the evolution of trade relationships.

We have employed both binary and weighted network analyses to provide a complete picture of the topological properties of the ITN. A binary analysis only accounts for the mere presence of trade relationships between any two nodes. In a weighted approach, each link is weighted by some value of trade flows that it carries. Our analysis displays that the topological structure of the ITN viewed from a weighted perspective is substantially different from that obtained by using a binary approach. In fact, some findings obtained by considering only the number of trade relationships are absolutely reversed if we take into account the intensity of trade linkages. Based on these results, we can make an important methodological point. If the study of the ITN is performed from a binary perspective, one can get a misleading picture of the underlying structure of the ITN. A weighted network analysis instead allows us to better appreciate the relational patterns.

Our results show that the ITN is a highly symmetric network, which means that almost all trade relationships are reciprocal. We have therefore employed an undirected analysis to study the ITN. While the binary representation of the ITN leads to a highly connected network, the average intensity of trade relationships is rather low. In particular, the majority of countries holding weak relationships coexist with a small fraction of countries holding very intense relationships. The ITN further exhibits a disassortative pattern, which implies that countries holding many and more intense relationships tend to trade with less and weakly connected countries. We also show that the binary version of the ITN is highly clustered. Moreover, countries with more trade partners are less clustered than those with few partners. The clustering patterns are completely different if we take into account the intensity of trade relationships. The ITN viewed as a weighted network indicates very low clustering level and countries with more intense trade relationships tend to establish highly connected trade triangles. This implies that there exists a small group of countries playing a dominant role in the international trade, which reminds the "rich club phenomenon".

We can also state one important conclusion from a policy point of view. The aforementioned results hint to a core-periphery structure of the ITN. Moreover, we have suggested that several relatively well connected countries tend to trade with countries that are themselves well connected. This supports the hypothesis that there exists an intermediate periphery within the core-periphery structure of the ITN. We have already state that the polarized structure of the ITN is not the most efficient one. Since the main benefit from the trade lies in the flows of capital and the market size, it is easy to see that the growth and development of peripheral countries might be limited.

Furthermore, we have investigated the evolution of the topological properties of the ITN over time. The average number of trade relationships has slightly increased through time, however all other network statistics have remained remarkably stable across time. The stability of structural properties of the ITN implies that the international economic integration has not increased significantly over the last 20 years. A possible explanation might be that a peak of the integration had been achieved before the sample period. We can conclude that the recent wave of globalization has not influenced significantly the structure of the ITN.

Finally, we have investigated the relation between network properties and country-specific characteristics (GDP per capita). We have found that high-income countries tend to hold more partners and more intense relationships, to trade with less and more weakly connected countries and to be more clustered. This also suggests the presence of the “rich club phenomenon”.

This study can be considered as a initial step towards a better understanding of the structure and evolution of the ITN, therefore many extensions can be conceived. First, one may try to explore the robustness of the results. The procedure would include an experiment with other economically-meaningful weighting systems. This might be an important point, because a weighted analysis might be sensible to a particular choice of weighting procedure. For example, the papers [16, 17] present that their results seem to be quite robust to all employed alternatives. Second, one would like to examine in more details the in-sample dynamics and out-of-sample evolution of the topological properties of the ITN. Third, one can be interested in whether network statistics measuring connectivity, assortativity and clustering can be employed as explanatory variables for the macroeconomic dynamics of growth and development. Finally, one may attempt to develop economic-meaningful models that are able to explain and reproduce the topological properties of the ITN.

References

[1] ALBERT, Réka. and Albert-László BARABÁSI. *Statistical mechanics of complex networks*. Reviews of modern physics, Vol. 74, 47-97, 2002.

[2] DOROGOVTSEV, Sergey N. and Jose F. F. MENDES. *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford: Oxford University Press, 2003.

[3] NEWMAN, Mark E. J. *The Structure and Function of Complex Networks*. SIAM Review, Vol. 45, 167-256, 2003.

[4] PASTOR-SATORRAS, Romualdo. and Alessandro VESPIGNANI. *Evolution and Structure of the Internet*. Cambridge: Cambridge University Press, 2004.

[5] WATTS, Duncan J. *Small worlds: the dynamics of networks between order and randomness*. Princeton university press, 1999.

[6] WASSERMAN, Stanley and Katherine FAUST. *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press, 1994.

[7] SCOTT, Stanley. *Social network analysis: A Handbook*. London: Sage, 2000.

[8] FREEMAN, Linction C. *Some Antecedents of Social Network Analysis*. Connections, Vol. 19, No. 1, 39-42, 1996.

[9] SERRANO, M. Angeles. and Marian BOGUNA. *Topology of the world trade web*. Physical Review E, Vol. 68, No. 1, 2003.

[10] GARLASCHELLI, Diego. and Maria I. LOFFREDO. *Fitness-dependent topological properties of the World Trade Web*. Phisical Review Letters, Vol. 93, No. 18, 2004.

[11] GARLASCHELLI, Diego. and Maria I. LOFFREDO. *Structure and evolution of the world trade network*. Physica A: Statistical Mechanics and its Applications, Vol. 355, No. 1, 138-134, 2005.

[12] LI, Xiang., Yu YING JIN. and Guanrong CHEN. *Complexity and synchronization of the world trade web*. Physica A: Statistical Mechanics and its Applications, Vol. 328, No. 1, 287-296, 2003.

[13] BHATTACHARYA, Kunal., G. MUKHERJEE. and Subhrangshu S. MANNA. *The international trade network*. In CHATTERJEE, Arnab. and Bikas K. CHAKRABARTI. *Econophysics of Markets and Business Networks*. Springer Milan, 2007.

[14] BHATTACHARYA, Kunal., G. MUKHERJEE, Jari SARAMAKI., Kimmo KASKI. and Subhrangshu S. MANNA. *The international trade network: weighted network analysis and modelling*. Journal of Statistical Mechanics: Theory and Experiment, 2008.

[15] FAGIOLO, Giorgio., Javier REYES. and Stefano SCHIAVO. *On the topological properties of the world trade web: A weighted network analysis*. Physica A: Statistical Mechanics and its Applications, Vol. 387, No. 15, 2008.

[16] FAGIOLO, Giorgio., Javier REYES. and Stefano SCHIAVO. *World-trade web: Topological properties, dynamics, and evolution*. Physical Review E, Vol. 79, No. 3, 2009.

[17] FAGIOLO, Giorgio., Javier REYES. and Stefano SCHIAVO. *The evolution of the world trade web: a weighted-network analysis*. Journal of Evolutionary Economics, Vol. 20, No. 4, 2010.

[18] STIGLITZ, Joseph E. *Globalization and its Discontents*. New York: WW Norton & Company, 2002.

[19] DREHER, Axel., Noel GASTON. and Pim MARTENS. *Measuring globalisation: Gauging its consequences*. Berlin: Springer, 2008.

[20] KRUGMAN, Paul R. *Growing World Trade: Causes and Consequences*. Brookings papers on economic activity, Vol. 26, No. 1, 327-377, 1995.

[21] HELLIWELL, John F. and Tim PADMORE. *Empirical studies of macroeconomic interdependence*. In JONES, Ronald W. and Peter KENEN. *Handbook of International Economics*. Elsevier, Vol. 2, 1985.

[22] FORBES, Kristin J. *Are Trade Linkages Important Determinants of Country Vulnerability to Crises?*. In SEBASTIAN, Edwards. and A. Frankel JEFFREY. *Preventing Currency Crises in Emerging Markets*. Chicago: University of Chicago Press, 2002.

[23] ABEYSINGHE, Tilak. and Kristin J. FORBES. *Trade Linkages and Output-Multiplier Effects: a Structural VAR Approach with a Focus on Asia*. Review of International Economics, Vol. 13, No. 2, 356-375, 2005.

[24] GARLASCHELLI, Diego. and Maria I. LOFFREDO. *Patterns of link reciprocity in directed networks*. Physical Review Letters, Vol. 93, No. 26, 2004.

[25] FAGIOLO, Giorgio. *Directed or undirected? A new index to check for directionality of relations in socio-economic networks*. Economics Bulletin, Vol. 3, No. 34, 1-12, 2006.

[26] FAGIOLO, Giorgio. *Clustering in complex directed networks*. Physical Review E, Vol. 76, No. 2, 2007.

[27] NEWMAN, Mark E. J. *Networks: An Introduction*. Oxford: Oxford University Press, 2010.

[28] BARRAT, Alain., Marc BARTHELÉMY., Romualdo PASTOR-SATORRAS. and Alessandro VESPIGNANI. *The architecture of complex weighted networks*. Proceedings of the National Academy of Sciences of the United States of America, Vol. 101, 3747-3752, 2004.

[29] ONNELA, Jukka-Pekka., Jari SARAMÄKI., János KERTÉSZ. and Kimmo KASKI. *Intensity and coherence of mofits in weighted complex networks*. Physical Review E, Vol. 71, No. 6, 2005.

[30] HERFINDAHL, Orris Clemens. *Copper costs and prices: 1870-1957*. Baltimore: Published for Resources for the Future, inc. by the Johns Hopkins Press, 1959.

[31] HIRSCHMAN, Albert O. *The paternity of an index*. The American Economic Review, Vol. 54, No. 5, 761-762, 1964.

[32] SARAMÄKI, Jari., Mikko KIVELÄ., Jukka-Pekka ONNELA., Kimmo KASKI. and János KERTÉSZ. *Generalizations of the clustering coefficient to weighted complex networks*. Physical Review E, Vol. 75, No. 2, 2007.

[33] HOLME, Petter., Sung Min PARK., Beom Jun KIM. and Christopher R. ENDLING. *Korean university life in a network perspective: Dynamics of a large affiliation network*. Physica A: Statistical Mechanics and its Applications, Vol. 373, 821-830, 2007.

[34] ZHANG, Bin. and Steve HORVATH. *A general Framework for Weighted Gene Co-Expression Network Analysis*. Statistical applications in genetics and molecular biology, Vol. 4, No. 1, 2005.

[35] SNYDER, David. and Edward L. KICK. *Structural Position in the World System and Economic Growth, 1955-1970: A Multiple Network Analysis of Transnational Interactions*. American Journal of Sociology, Vol. 85, No. 5, 1096-1126, 2010.

[36] NEMETH, Roger J. and David A. SMITH. *International Trade and World-System Structure: A Multiple Network Analysis*. Review (Fernand Braudel Center), Vol. 8, No. 4, 517-560, 1985.

[37] BREIGER, Roland. *Structures of economic interdependence among nations*. In BLAU, Peter M. and Robert K. MERTON. *Continuities in structural inquiry*. Newbury Park, CA: Sage, 353-380, 1981.

[38] SMITH, David A. and Douglas R. WHITE. *Structure and Dynamics of the Global Economy: Network Analysis if International Trade 1965-1980*. Social Forces, Vol. 70, No. 4, 857-893, 1992.

[39] FRÖBEL, Folker., Jürgen HEINRICHS. and Otto KREYE. *The new international division of labour*. Social Science Information, Vol. 17, No.1, 123-142, 1978.

[40] KIM, Sangmoon. and Eui-Hang SHIN. *A Longitudinal Analysis of Globalisation and Regionalisation in International Trade: A Social Network Approach*. Social Forces, Vol. 81, No. 2, 445-468, 2002.

[41] KASTELLE, Tim., John STEEN. and Peter LIESCH. *Measuring Globalisation: An Evolutionary Economic Approach to Tracking the Evolution of*

International Trade. Paper to be presented at the DRUID Summer Conference on Knowledge, Innovation and Competitivness: Dynamics of Firms, Networks, Regions and Institutions, Copenhagen, 2005.

[42] GARLASCHELLI, Diego., Tiziana DI MATTEO., Tomaso ASTE., Guido CALDARELLI. and Maria I. LOFFREDO. *Interplay between topology and dynamics in the World Trade Web*. The European Physical Journal B, Vol. 57, No. 2, 159-164, 2007.

[43] KALI, Raja. and Javier REYES. *The architecture of globalization: a network approach to international economic integration*. Journal of International Business Studies, Vol. 38, No. 4, 595-620, 2007.

[44] SERRANO, M. Angeles., Marian BOGUNA. and Alessandro VERSPIGNANI. *Patterns of dominant flows in the world trade web*. Journal of Economic Interaction and Coordination, Vol. 2, No. 2, 11-124, 2007.

[45] TZEKINA, Irena., Karan DANTHI. and Daniel N. ROCKMORE. *Evolution of community structure in the world trade web*. The European Physical Journal B, Vol. 63, No. 4, 541-545, 2008.

[46] FAGIOLO, Giorgio. *The international-trade network: gravity equations and topological properties*. Journal of Economic Interaction and Coordination, Vol. 5, No. 1, 1-25, 2010.

[47] KALI, Raja. and Javier REYES. *Financial contagion on the international trade network*. Economic Inquiry, Vol. 48, No. 4, 1072-1101, 2010.

[48] SQUARTINI, Tiziano., Giorgio FAGIOLO. and Diego GARLASCHELLI. *Randomizing world trade. I. A binary network analysis*. Physical Review E, Vol. 84, 2011.

[49] SQUARTINI, Tiziano., Giorgio FAGIOLO. and Diego GARLASCHELLI. *Randomizing world trade. II. A weighted network analysis*. Physical Review E, Vol. 84, 2011.

[50] DUENAS, Marco. and Giorgio FAGIOLO. *Modeling the International-Trade Network: a gravity approach*. Journal of Economic Interaction and Coordination, Vol. 8, No. 1, 1-24, 2013.

[51] MASTRANDREA, Rossana., Tiziano SQUARTINI, Giorgio FAGIOLO, Diego GARLASCHELLI. *Intensive and extensive biases in economic networks: reconstructing the world trade multiplex.* [online]. 2014 [cit. 2014-05-13]. Available at: `http://arxiv.org/abs/1402.4171`.

[52] GLEDITSCH, Kristian. *Expanded Trade and GDP data.* Journal of Conflictict Resolution, Vol. 46, No. 5, 712-724, 2002.

[53] SUMMERS, Robert. and Alan HESTON. *The Penn World Table (Mark 5): An expanded set of international comparisons, 1950–1988.* The Quarterly Journal of Economics, Vol. 106 No. 2, 327-368, 1991.

[54] CENTRAL INTELLIGENCE AGENCY. *World Factbook* [online]. 1998. Available at: `https://www.cia.gov/library/publications/the-world-factbook/`

[55] INTERNATIONAL MONETARY FUND. *Direction of Trade Statistics: Codebook* (ICPSR no. 7628). Washington DC: International Monetary Fund, 1997.

[56] FABER, Jan. and Tom NIEROP. *World Export Data (WED), 1948-1983* (ICPSR no. 9116). Amsterdam: University of Amsterdam (producers); Ann Arbor, MI: Inter-university Consortium for Political and Social Research (distributors).

[57] SHAPIRO, Samuel Sanford. and Martin B. WILK. *An analysis of variance test for normality (complete samples).* Biometrika, Vol. 52, No. 3 and 4, 591-611, 1965.

[58] JARQUE, Carlos M. and Anil K. BERA. *Efficient tests for normality, homoscedasticity and serial independence of regression residuals.* Economics Letters, Vol. 6, No. 3, 255-259, 1980.

[59] JARQUE, Carlos M. and Anil K. BERA. *Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence.* Economics Letters, Vol. 7, No. 4, 313-318, 1981.

[60] HOJMAN, Daniel A. and Adam SZEIDL. *Core and periphery in networks.* Journal of Economic Theory, Vol. 139, No. 1, 2008.

[61] ROMBACH, M. Puck., Mason A. PORTER., James H. FOWLER. and Peter J. MUCHA. *Core-periphery structure in networks*. SIAM Journal on Applied mathematics, Vol. 74, No. 1, 2014.

[62] FURUSAWA, Taiji. and Hideo KONISHI. *Free Trade Networks*. Journal of International Economics, Vol. 72, No. 2, 310-335, 2007.

[63] BAIER, Scott L. and Jeffrey H. BERGSTRAND. *Economic determinants of free trade agreements*. Journal of International Economics, Vol. 64, No. 1, 29-63, 2004).

Appendix A

Tables

Table A.1: Countries in balanced panel

ID	Acro	Name	ID	Acro	Name
2	USA	United States	110	GUY	Guyana
20	CAN	Canada	115	SUR	Surinam
31	BHM	Bahamas	130	ECU	Ecuador
40	CUB	Cuba	135	PER	Peru
41	HAI	Haiti	140	BRA	Brazil
42	DOM	Dominican Rep.	145	BOL	Bolivia
51	JAM	Jamaica	150	PAR	Paraguay
52	TRI	Trinidad/Tobago	155	CHL	Chile
53	BAR	Barbados	160	ARG	Argentina
54	DMA	Dominica	165	URU	Uruguay
55	GRN	Grenada	200	UKG	United Kingdom
56	SLU	Saint Lucia	205	IRE	Ireland
57	SVG	St. Vincent	210	NTH	Netherlands
70	MEX	Mexico	211	BEL	Belgium
90	GUA	Guatemala	212	LUX	Luxembourg
91	HON	Honduras	220	FRN	France
92	SAL	El Salvador	221	MNC	Monaco
93	NIC	Nicaragua	223	LIE	Liechtenstein
94	COS	Costa Rica	225	SWZ	Switzerland
95	PAN	Panama	230	SPN	Spain
100	COL	Colombia	232	AND	Andorra
101	VEN	Venezuela	235	POR	Portugal

ID	Acro	Name	ID	Acro	Name
260	GFR	Germany	461	TOG	Togo
290	POL	Poland	471	CAO	Cameroon
305	AUS	Austria	475	NIG	Nigeria
310	HUN	Hungary	481	GAB	Gabon
325	ITA	Italy	482	CEN	Centr African Rep.
331	SNM	San Marino	483	CHA	Chad
338	MLT	Malta	484	CON	Congo
339	ALB	Albania	490	DRC	Congo (Zaire)
345	YUG	Yugoslavia	500	UGA	Uganda
350	GRC	Greece	501	KEN	Kenya
352	CYP	Cyprus	510	TAZ	Tanzania
355	BUL	Bulgaria	516	BUI	Burundi
360	RUM	Rumania	517	RWA	Rwanda
365	RUS	Russia	520	SOM	Somalia
375	FIN	Finland	522	DJI	Djibouti
380	SWD	Sweden	530	ETH	Ethiopia
385	NOR	Norway	540	ANG	Angola
390	DEN	Denmark	541	MZM	Mozambique
395	ICE	Iceland	551	ZAM	Zambia
402	CAP	Cape Verde	552	ZIM	Zimbabwe
403	STP	Sao Tome	553	MAW	Malawi
404	GNB	Guinea-Bissau	560	SAF	South Africa
411	EQG	Eq. Guinea	570	LES	Lesotho
420	GAM	Gambia	571	BOT	Botswana
432	MLI	Mali	572	SWA	Swaziland
433	SEN	Senegal	580	MAG	Madagascar
434	BEN	Benin	581	COM	Comoros
435	MAA	Mauritania	590	MAS	Mauritius
436	NIR	Niger	591	SEY	Seychelles
437	CDI	Cote Divore	600	MOR	Morocco
438	GUI	Guinea	615	ALG	Algeria
439	BFO	Burkina Faso	616	TUN	Tunisia
450	LBR	Liberia	620	LIB	Libya
451	SIE	Sierra leone	625	SUD	Sudan
452	GHA	Ghana	630	IRN	Iran

ID	Acro	Name	ID	Acro	Name
640	TUR	Turkey	771	BNG	Bangladesh
645	IRQ	Iraq	775	MYA	Myanmar
651	EGY	Egypt	780	SRI	Sri Lanka
652	SYR	Syria	781	MAD	Maldives
660	LEB	Lebanon	790	NEP	Nepal
663	JOR	Jordan	800	THI	Thailand
666	ISR	Israel	811	CAM	Cambodia
670	SAU	Saudi Arabia	812	LAO	Laos
678	YEM	Yemen	816	DRV	Vietnam
690	KUW	Kuwait	820	MAL	Malaysia
692	BAH	Bahrain	830	SIN	Singapore
694	QAT	Qatar	840	PHI	Philippines
696	UAE	Arab Emirates	850	INS	Indonesia
698	OMA	Oman	900	AUL	Australia
700	AFG	Afghanistan	910	PNG	Papua
710	CHN	China	920	NEW	New Zealand
712	MON	Mongolia	935	VAN	Vanuatu
713	TAW	Taiwan	940	SOL	Solomon's
731	PRK	North Korea	950	FJI	Fiji
732	ROK	South Korea	970	KBI	Kiribati
740	JPN	Japan	971	NAU	Nauru
750	IND	India	972	TON	Tonga
760	BHU	Bhutan	973	TUV	Tuvalu
770	PAK	Pakistan			