Physica A 419 (2015) 64–74

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

A microscopic study of the fitness-dependent topology of the world trade network

K. Hoppe *, G.J. Rodgers

Department of Mathematical Sciences, Brunel University, Uxbridge, Middlesex UB8 3PH, United Kingdom

HIGHLIGHTS

- We study the hidden variable correlation matrix of the world trade network.
- These correlations resemble the previously found disassortativity of world trade.
- A new fitness definition that leads to more robust results is proposed.
- The impact of a changed data source on first order results is illustrated.

ARTICLE INFO

Article history:
Received 15 March 2014
Received in revised form 4 September 2014
Available online 13 October 2014

Keywords:
World trade network
Fitness model
Complex networks

ABSTRACT

Previous studies have suggested that the world-trade network belongs to the class of static hidden variable models. In this article we investigate the microscopic structure of the world trade network, that is the hidden variable correlation matrix of the network. The hidden variable is defined as a rank ordering of gross domestic products. This choice significantly reduces the noise in the statistical analysis found in previous studies. The hidden variable correlation matrix, that expresses the probability that a trade relationship between two countries of given fitness exists, suggests an attachment kernel that at least partially favours trading pairs or dissimilar fitness rather than the purely multiplicative one found previously. Additionally, we provide an in-depth look at the data source and reveal that first-order results, such as the degree distribution, exhibit significant qualitative differences depending on the data provider. Furthermore, we shed light on the intertemporal activity of international trade and point out that fluctuations occur mostly between countries with strong dissimilarities of fitness and connectivity.

1. Introduction

The world trade network (WTN) has received a considerable amount of attention in the last decade. A detailed understanding of the network structure of international trade facilitates a better comprehension of crises that spread across a country's frontiers. It is generally agreed that a static hidden variable model [1,2] best describes the topology of the WTN. The static model is most suitable because there are only a finite number of distinct countries in the world. Depending on the scope of the analysis, previous contributions understand the network as growing since the number of vertices increased from 86 in the year 1950 to 190 in the year 2000 [3]. However, since there are – depending on the definition – approximately 195 countries in the world to date, the growth of the network cannot be sustained. A dynamic perspective on the world trade network, by investigating the duration and strength of individual connections and also the relationship between the development status of a country and its ability to establish long-lasting trade relationships has been discussed in Refs. [4,5]. It

* Corresponding author. Tel.: +44 7415413056.
E-mail address: konrad.hoppe@brunel.ac.uk (K. Hoppe).

http://dx.doi.org/10.1016/j.physa.2014.09.036

has been shown in Refs. [6,7], that the static fitness model is an adequate model to describe the structure of the world trade network for several years, not just on the level of node degree distributions, but up to higher order statistics, such as the average nearest neighbour degree. In this article, we look into the microscopic structure of the WTN. The main quantity of interest is the edge density, that is the probability matrix for any two nodes of fitness x and y to share an edge. The fitness model is found to be an excellent fit for the WTN. However, the previously found multiplicative attachment kernel varies significantly from the kernel that we find by investigating the edge density.

A review of the literature on the network of international trade reveals that many results are not without controversy. Even first order results, such as the form of the degree distribution, differ between various publications. One reason for these discrepancies is the choice of data source. Every trade relationship between two countries should be reported twice, once as an export and once as an import. However, this is not always the case [8]. As a result, different data providers report different trading data. Another issue is the way these disparities are corrected. Some [9,10] simply use only import data, which is deemed more reliable than export data, others [11–13] interpolate between reports of exports and imports.

The data set we use for the present article is provided by the National Bureau of Economic Research (NBER) [8] and forms an excellent starting point for this investigation. The special feature of this data set is that one part of it has been adjusted for errors, while another part has not. The adjustments have been made by specialised macroeconomists and can therefore be deemed reliable.

The world trade network is treated as directed and binary in this article. That means that directionality of trade is taken into account, but the traded amounts are ignored. Others [7,14,12,15] have argued that directionality could be ignored because of the high reciprocity, i.e. export links are usually reciprocated with import links. However, since the world trade network is directed by nature, the analysis in the present article is laid out in terms of a directed network. A weighted representation of the WTN has been investigated for example in Refs. [12,16,13]. Curiously, the additional information of edge weight is not always of greater explanatory power than binary network analysis [10,17]. In Refs. [10,17] a network randomisation technique was used to show that by knowing the degree sequence of the binary WTN, higher order statistics like the average neighbour degree and the clustering coefficient can be obtained. However, using a weighted network approach and fixing the sequence of interaction strengths, these higher order statistics cannot be found anymore. This suggests that binary network analysis used in economics is a powerful tool, because all the necessary information is contained in the first order statistics.

In principle, the WTN is a spatially embedded network. However it has been illustrated that distances between countries do not add a significant amount of information to a binary analysis of international trade [18]. Similarly, also in the weighted regime, the importance of geographic distance for understanding international trade is declining over time [19]. The results of Ref. [19] are in contrast to the standard notion in economics, that the intensity of trade between pairs of countries is strongly related to their distance, see for instance [20]. Another aspect that will not feed into the analysis in this article is the multi-layered architecture of international trade. Every reported trade flow is an aggregation over different product categories. Refs. [10,17,9,21] investigate these different layers separately.

The investigation of the topology of the world trade is not conducted as an end in itself. Its aim is to understand how trade can affect economic welfare. Many aspects are yet to be understood. The theory of complex systems is just one of the building blocks towards a good understanding of those effects [22,23]. Network-theoretic measures have been shown to explain parts of nation's income. In Ref. [24] it is shown that an improvement of the degree centrality ranking by ten units increases the average GDP per capita by 0.27%. Others [11,6,25,3] investigate correlation structures of income, connectivity and interaction strength in weighted networks and illustrate that an involvement in international trade has a direct impact on income and vice versa. Additionally, it has been shown that network properties have good explanatory power to detect vulnerable economies in the WTN [26,27].

This paper is organised as follows. In Section 2, the static fitness model is reviewed and the central quantities for this study are derived. In Section 3, the data set is introduced and reasons for its choice are discussed. In Section 4, the definition of fitness is elucidated. In Section 5, the inter-temporal structure of the WTN is clarified and in Section 6, the static structure is investigated. Section 7 closes the article with concluding remarks.

2. The static fitness model

The empirical analysis in this article is based on the static fitness model, as it was introduced in Refs. [1,28,2]. The investigation that is presented later in the text relies on results that are reviewed in the following.

The static fitness model is a network model with N nodes and M directed edges. Each node inside the network is endowed with a fitness value x, that is drawn from a probability density function $\rho(x)$. The probability that a node with fitness x originates a link towards a node with fitness y is proportional to the attachment kernel $f(x, y)$. Although the network is static, the edge deployment process can be understood as dynamic. The probability that an edge is added to a pair of nodes $i \to j$ is $f(x_i, x_j)/\sum_{k,l} f(x_k, x_l)$. The reader should be reminded at this point that this formalism ignores the possible problem of edge duplications. Edge duplications are not an issue in sparse networks, i.e. $M \ll N^2$, however they will lead to bias in dense networks, such as the WTN. However, correcting for the impossibility of duplications leads to non-linearities in the formalism, so that closed form solutions can no longer be found.

To begin with, the probability that two nodes with fitness x_i and x_j are connected in a network with M edges and N nodes is derived. Denote this probability with $C_{M,N}(x_i, x_j)$ and further define $\overline{C}_{M,N}(x_i, x_j) = 1 - C_{M,N}(x_i, x_j)$. We will refer to $C_{M,N}$