Summer June, 2018

Distance-varying assortativity and clustering of the international trade network

Angela Abbate, *Deutsche Bundesbank*
Luca De Benedictis, *University of Macerata*
Giorgio Fagiolo, *Sant'Anna School of Advanced Studies - Pisa*
Lucia Tajoli, *Politecnico di Milano*

doi:10.1017/nws.2018.7

Distance-varying assortativity and clustering of the international trade network

ANGELA ABBATE
Economic Affairs Division, Swiss National Bank, Bern/Zürich, Switzerland
(*e-mail:* angela.m.abbate@gmail.com)

LUCA DE BENEDICTIS
DED, University of Macerata, Via Crescimbeni 20, Macerata 62100, Italy
(*e-mail:* luca.debenedictis@unimc.it)

GIORGIO FAGIOLO
Institute of Economics, Scuola Superiore Sant'Anna, Piazza Martiri della Libertà 33, Pisa 56127, Italy
(*e-mail:* giorgio.fagiolo@sssup.it)

LUCIA TAJOLI
Dipartimento di Ingegneria Gestionale, Politecnico di Milano, Piazza Leonardo da Vinci 32, Milano 20133, Italy
(*e-mail:* lucia.tajoli@polimi.it)

Abstract

In this paper, we study how the topology of the International Trade Network (ITN) changes in geographical space, and along time. We employ geographical distance between countries in the world to filter the links in the ITN, building a sequence of subnetworks, each one featuring trade links occurring at similar distance. We then test if the assortativity and clustering of ITN subnetworks changes as distance increases, and we find that this is indeed the case: distance strongly impacts, in a non-linear way, the topology of the ITN. We show that the ITN is disassortative at long distances, while it is assortative at short ones. Similarly, the main determinant of the overall high-ITN clustering level are triangular trade triples between geographically close countries. This means that trade partnership choices and trade patterns are highly differentiated over different distance ranges, even after controlling for the economic size and income per capita of trading partners, and it is persistent over time. This evidence has relevant implications for the non-linear evolution of globalization.

Keywords: *international trade, network analysis, distance*

1 Introduction

Overcoming a relatively long neglect, geography and distance have re-gained momentum in international trade analysis. After a strand of literature in the 1990s showing how the geographical position of countries could affect specialization, agglomeration of production, and the gains from economic integration (Krugman, 1998; Fujita and Krugman, 2004), the recent international trade models displaying heterogeneous productivities (Eaton and Kortum, 2002; Melitz, 2003) assign a fundamental role to distance. In both types of literature, distance matters, especially because of the existence of non-negligible international trade costs (Anderson and

Van Wincoop, 2004; Krugman and Venables, 1995). The characteristics of these trade costs are therefore crucial in determining the models' results. Also, in the empirical analysis of trade flows, the overwhelming use of the gravity equation[1] as the preferred functional form in the prediction of the volume of bilateral trade flows is emphasizing once again the persistent role of distance in international exchanges.[2]

Nevertheless, especially in applied analyses, the concept of distance is still essentially considered as an absolute, bilateral concept (i.e., as an individual characteristic of each single country i with respect to any other country j). Even if the inclusion of concepts such as *market potential* (Head and Mayer, 2010), *remoteness* (Wei, 1996), or *multilateral resistance* (Anderson and van Wincoop, 2003) in international trade analysis encouraged researchers to move from the two-country-world of international trade theory and the bilateral focus of trade empirics to a more global analysis, where the structure of the links among *all* countries in influencing countries' dyadic relations is taken into account,[3] the practical treatment of the issue through country fixed-effects in panel data (Feenstra, 2002; Redding and Venables, 2004) makes the truly structural dimension of multilateral resistance still largely under-explored.

In this paper, we want to explicitly address the issue of distance in a multidimensional system, so to better understand how geographical distance and the trade costs connected to it can influence the choice and the number of trade partners of each country, shaping the evolution of globalization and generating a possible hierarchical structure of trade relations. Even if certainly trade costs are not only determined by geographical distance, this is still a major barrier to trade according to recent estimates (Arvis et al., 2016). Therefore, we will use a network approach to study international trade flows between countries, taking geography explicitly into account. Network analysis of international trade (see Fagiolo, 2015, for a recent review of the literature) has started to uncover the properties of the web of trade flows by focusing precisely on the structural, multilateral aspects of international trade relations (Fagiolo et al., 2009; De Benedictis and Tajoli, 2011; De Benedictis et al., 2014). This approach describes international trade flows using graph-theoretic structures where countries (nodes) are linked by arcs or edges (possibly weighted) representing import–export relationships between countries. The literature hints to the existence of interesting statistical regularities in the topology of trade networks, such as the *disassortative* nature of the network, (i.e., the fact that highly-connected countries tend to be linked to countries which are not well-connected), or the relatively high level of *clustering* (i.e., a high probability that any two trade partners of a country are themselves trade partners).

There are a number of implications related to these findings. First of all, the uneven structure of the network of international trade and the strongly non-normal

[1] See Disdier and Head (2008) for a metaanalysis of the empirical literature on the gravity model in international trade and the role of distance, Santos Silva and Tenreyro (2006) for the possible implication of the inclusion of distance in a log form, and De Benedictis and Taglioni (2011); Anderson (2011) and Head and Mayer (2014) for recent overviews of the gravity model.

[2] The economic relevance of space, i.e., the cost associated to physical distance, has also influenced macroeconomists in singling it out as a fundamental explanation of discrepancies between canonical models and empirical evidence (Obstfeld and Rogoff, 2001).

[3] The literature on "third country effects" is still very sparse. See Baltagi et al. (2007) for an application to Foreign Direct Investments, and Egger and Larch (2008) and Chen and Joshi (2010) for an application to Regional Trade Agreements (RTAs).

distribution of countries in terms of connectivity to the world trading system point to the fact that countries are very heterogeneous in their choice of trade partners, implying that in trade models the concept of an "average country" presents severe limits (De Benedictis and Tajoli, 2011). Furthermore, the network structure affects directly the distribution of the gains from trade by affecting countries' market power (Jackson, 2010), the transmission of economic shocks and economic growth (Kali and Reyes, 2007; Kali et al., 2007), the efficiency and costs of the transport system, the heterogeneous effect of Preferential Trade Agreements (Eicher and Henn, 2011), and the stability of the entire trade structure.

These findings apply to the entire world trading system, without assigning a role to geography, as the network of trade flows has never been embedded in a geographical space. No attempt has been made to link the topological distance among countries to its geographical counterpart.[4] In this paper, we begin to fill this gap by including geographical space in a network analysis of international trade flows, because, as suggested by a large body of literature, geography can play a relevant role in globalization patterns. To do so, we employ the geographical distance between countries to filter the international trade network (ITN) and build subnetworks of countries located within a given geographical range. We then test if the topological properties of the ITN vary as the distance range increases.

The short answer we get is: *yes*. We find that the effect of distance on trade networks is strongly non-linear. Many of the properties observed at the aggregate level (i.e., without considering distance) are not robust to a geographical breakdown. For example, focusing on *assortativity* (the tendency of the likes to stick together) and *clustering* (the tendency of one's partners to be partners among themselves), the network is disassortative at long distances, while it is assortative at short ones, the switch occurring at a distance of approximately 9,000 km. A similar finding applies to clustering: short-distance countries-triples are the major contributors to the strong level of overall international clustering.

The main implication of such a change in the topological properties is that the drivers of trade at long and short distance can be different, and we should expect different international trade patterns between neighboring countries and countries that are far apart, as it is suggested by Eaton and Kortum (2002) and Hillberry and Hummels (2008). The difference in the assortativity mix at various distances suggests that distance-related trade costs indeed play a role, and the non-linearity of effects can be interpreted as a change in the relative weight of the fixed (e.g., sunk costs à la Melitz, 2013) and variable costs (e.g., "iceberg" type trade costs) component. By repeating this exercise over time, we show that this evidence is persistent from the year 1970 to the year 2000, and that it remains consistent even after explicitly conditioning the analysis to the economic size (e.g., GDP) and GDP per capita of countries involved in trade. This confirms an important insight for the analysis of trade networks and international trade in general: *international trade flows are not isomorphic along distance.*

[4] This is also the case for network analysis in general. There are however notable exceptions in the area of geography, urban systems, and transport analysis. See Barthelemy (2011) for a review of the issue, and Wilson (2000) and Haggett and Chorley (1969).

The rest of the paper is organized as follows. In Section 2, we review the literature dealing with geographical distance and trade models. Section 3 discusses the main findings of network-related literature on trade. In Section 4, we describe the data used in our analysis and the methodology used. Section 5 presents our main results. Finally, Section 6 concludes.

2 Distance, trade costs, and the extent of globalization

Distance matters in international trade. This is so whenever the movements of goods and services across space implies some costs. Econometric estimates of the constant elasticity of trade to geographical distance provide a measure of the relevance of those costs and their persistence over time (Brun et al., 2005), ranging within an interval of −0.7−1.2 (Disdier and Head, 2008). This means that on average, all else equal, countries twice apart show a bilateral volume of trade that is approximately half the one of neighboring countries. There is a large theoretical and empirical evidence showing that geographical distance is a proxy also for a number of other important factors that hamper international trade (Anderson and Van Wincoop, 2004), often related to geography. Differences in language (Melitz and Toubal, 2014), trade policy (Egger and Larch, 2008), history (Eichengreen and Irwin, 1995), and institutions enhance the cost of trade across geographical and cultural distance, especially by increasing uncertainty and informational issues (see, for example, De Groot et al., 2004, and the very recent work by Lendle et al., 2016).

In line with this evidence, the most recent trade models embody the idea that the geographical distance between the country where goods and services are produced and the country where they are sold and consumed has a negative effect on trade flows between the two countries. There may be several reasons why this is the case. Taking an extreme view, on the one hand, trade costs can operate proportionately to distance, increasing variable costs or reducing the quantity of the exported goods that reaches the foreign market, as in Samuelson's iceberg trade-cost formulations. On the other hand, trade costs can follow a Bernoulli process associated to the acquisition of an export status by firms extending their activities from a purely domestic context to an international one. In this case, distance acts as a fixed sunk cost, that can be specific to the each foreign market or payed once-and-for-all when the firm, having payed the fixed sunk trade cost, has acquired the knowledge necessary to be active in any foreign market.[5]

The implications of such alternative views in terms of trade partnership and trade flows among countries are clearly not the same. If trade costs are directly proportional to distance, trade partnerships would expand radially around countries and trade volumes would smoothly decrease with distance, conditional on the size of the foreign market. Conversely, if trade costs are fixed entry costs, trade partnership would be a more discontinuous process, and the selection of partners and trade volumes would be somehow independent on the distance between the home and foreign market.

[5] See Redding (2011) for a review of the recent theoretical literature on heterogeneous firms and trade and on the role of trade costs in this stream of research. The fundamental paper by Anderson and Van Wincoop (2004) gives a comprehensive overview of the issue.

Trade economists have dealt with this issue in terms of *extensive margin* and *intensive margin* of trade (i.e., new trade links vs. increasing strength of pre-existing trade links).[6] The consensual piece of evidence is that distance operates mainly through the extensive margin, as shown by Lawless (2010) and Bernard et al. (2007) and the negative effect of distance strongly emerges only at the extensive margin.

The theoretical paper by Chaney (2008), along the lines of the literature on heterogeneous firms that focuses on the micro foundation of the gravity equation (see also Anderson and van Wincoop (2003), Melitz (2003), Eaton and Kortum (2002), and Helpman et al. (2008)), reinforces the empirical evidence, showing that fixed costs affect only the extensive margin of trade. In facts, in Lawless (2010), the time-invariant variables that may influence bilateral trade, capturing the role played by fixed cost (i.e., language, internal orography, infrastructure, and import barriers) work through the extensive margin.

This evidence, distinguishing the role of distance between its effect on the extensive and intensive margin, has an implicit echo in the different analyses of the ITN. The notions of extensive and intensive margins, at the country level, can be easily applied to network analysis. Generally speaking, at the country level the extensive margin of trade is the change in the number of trade partners of a country over time, either importers or exporters or both, which coincides with the change in the number of the links of a country in the trade network. Conversely, the intensive margin—conditional to a fixed number of partners—is associated with the change of a country's pre-existing trade flows, e.g., the sum of the pre-existing weights of the links of a country.

The second issue that is worth discussing in the context of the gravity equation is the way distance is included as a covariate of bilateral trade flows. As we mentioned already, the cost of distance is in general interpreted as a fixed (sunk) cost or as a variable cost, with different implications.

In principle, given the different nature of trade costs, there is no reason to believe that distance should be related to trade in a (log)linear manner. As pointed out by Bernard et al. (2007) and Hummels (2007) among others, transportation costs can induce a selection among the goods that are sold in distant markets, and the average value of exports could increase in distance precisely to compensate the increase in trade costs. Even for a given distance, trade costs are much dependent on the characteristics of specific goods, such as fragility, perishability, size, or weight. In aggregate terms, trade costs would depend on trade composition and would be country-specific, affected by country's remoteness and sectoral specialization. Such non-linearities are usually addressed empirically using a log-log specification.[7] This procedure can be however costly, de-facto removing all zero flows and generating a selection in the data. The role played by fixed cost is identified only on those

[6] The dimensions through which aggregate trade is typically split in margins are many. Just to fix ideas, the extensive margin can be considered at the country level (new export markets or countries from which imports are coming from); at the sectoral level (new product lines get activated in export or import data, at the sectoral level); at the firm level (new firms enlarge their reference market beyond national boundaries); or at the product level (multiproduct firms start selling new product varieties abroad).

[7] For an analysis of the implications of estimating the gravity model in a log-linear form with heteroskedastic errors, see Santos Silva and Tenreyro (2006) and De Benedictis and Taglioni (2011) for a survey of non-linear estimators in the context of the gravity model.

flows that are able to cover those fixed cost and not on those that are not able to afford it.

In some cases, the assumption of a (log)linear cost of space is substituted with some ad hoc functions. In Eaton and Kortum (2002), the distance effect is associated to six non-overlapping distance intervals: [0, 375); [375, 750); [750, 1,500); [1,500, 3,000); [3,000, 6,000); and [6,000, maximum], measured in miles. Anderson and Yotov (2012) uses the same specification decomposing the distance effect into four different elasticities corresponding to the four non-overlapping distance intervals: [0, 3,000); [3,000, 7000); [7,000, 10,000); [10,000, maximum], measured in kilometers. It is amazing that neither one of the two papers explains why this stepwise functional forms has been adopted. Hillberry and Hummels (2008) bring some illuminating evidence on why the Eaton and Kortum (2002) specification might be ad hoc *but* meaningful: there is an extraordinary difference between short and long-distance trade. Using highly disaggregated data at the spatial and sectoral level on manufacturers' shipments within the United States, they find that the pattern of shipments is strongly localized. Shipments within 5-digit zip codes, with a median radius of only 4 miles, are 3 times larger than shipments outside the zip code. The analysis shows that distance reduces aggregate trade values primarily by reducing the number of commodities shipped and the number of establishments shipping. " ... Extensive margins are particularly important over very short distances."

The analysis that will follow adds a complementary motive to the micro evidence put forward by Hillberry and Hummels (2008), and brings support to the distinction between short and long-distance trade; thus, giving a further support to the Eaton and Kortum (2002) specification of the effect of distance on trade. More specifically, we embed a multilateral perspective to international trade, based on network theory, in an explicit geographic dimension, in order to ask whether the structure of international trade is isomorphic across distance.

3 Complex networks and international trade

In the last years, there was an increasing surge of interest in applying a complex-network approach to the study of international trade, as documentd in Fagiolo (2015).[8]

The ITN, aka World-Trade Web (WTW) or World Trade Network (WTN), is defined as the graph of import/export relationships between countries in a given year t. The resulting graph, $\mathscr{G}_t = (\mathscr{V}_t, \mathscr{L}_t)$, where $n_t = \mid \mathscr{V}_t \mid$ is the number of countries constituting the vertices (or nodes) of the graph, and $m_t = \mid \mathscr{L}_t \mid$ is the number of existing directed trade links (or arcs), gives rise to network $\mathscr{N}_t = (\mathscr{V}_t, \mathscr{L}_t, \mathscr{P}_t, \mathscr{W}_t)$—where $\mathscr{P}_t$ is the *vertex value function* including the exogenous or endogenous properties of vertices, and $\mathscr{W}_t$ is the *line value function* including the exogenous or endogenous weights of links—which could be a binary (unweighted) network, $\exists \mathscr{W}_t = \mathscr{A}_t$, where $\mathscr{A}_t$ is, for every year t, a $n \times n$ 0-1 matrix, if only the

[8] See, for example, Li et al. (2003), Serrano and Boguñá (2003), Garlaschelli and Loffredo (2004), Garlaschelli and Loffredo (2005), Reichardt and White (2007), Serrano et al. (2007), Bhattacharya et al. (2008), Bhattacharya et al. (2007), Garlaschelli et al. (2007), Tzekina et al. (2008), Fagiolo et al. (2008), Reyes et al. (2008), Fagiolo et al. (2009), De Benedictis and Tajoli (2011), and Chaney (2016).

presence/absence of a positive trade flow is considered, or could be a weighted network, $\exists \mathscr{W}_t \neq \mathscr{A}_t$, if links have different intensities, according to, e.g., the value of the bilateral trade flows.

A lot of effort has been recently put forward in uncovering the topological properties of the ITN architecture, both at the aggregate and at the product-specific level (see Fagiolo et al., 2010; Barigozzi et al., 2010, for a review of the main results). As mentioned, understanding the topological properties of the ITN from a complex-network perspective (Albert and Barabási, 2002; Dorogovtsev and Mendes, 2003), and their evolution over time, is fundamentally important to study issues such as economic globalization, the spreading of international crises, and the transmission of economic shocks (Helliwell and Padmore, 1985; Artis et al., 2007; Forbes, 2002).

A network approach, by focusing on direct as well as *indirect relationships* between countries, is able to single out the role of each countries in the complex web of world trade interactions. This is a remarkable change with respect to traditional analyses. Indeed, the standard approach to international-trade empirics employs statistics that fully characterize the profile of a country in the system by referring mainly to its bilateral-trade direct linkages. Whereas direct bilateral trade linkages are known to be one of the most important channels of interaction between world countries (Krugman, 1995), recent studies show that they can only explain a small fraction of the impact that an economic shock originating in a given country can have on another one, which is not among its direct-trade partners (Abeysinghe and Forbes, 2005). Along similar lines, Squartini et al. (2011a,b) show that knowledge of country-specific indicators such a the number of trade partners, total imports or total exports (which only take into account direct bilateral links in the ITN) is not enough in the weighted ITN to characterize *higher-order moments* of the distribution of trade relationships, involving, for example, the trade behavior of the partners of a given country, the likelihood that any two trade partners of a vertex are themselves partners, etc.

In order to fully account for system-wide phenomena such as globalization and crises diffusion, a more detailed knowledge of the local and global topological properties of the network is therefore required. This means, in other words, acquiring a better understanding of the presence and importance of trade paths connecting any pair of non-direct trade partners and, more generally, of topological indicators proxying the likelihood that economic shocks might be transmitted between any two countries (Kali and Reyes, 2007). This, in turn, has been shown to help explaining patterns of macroeconomic dynamics related to, e.g., growth and development (Kali et al., 2007; Reyes et al., 2008).

The issue of assortativity is central in empirical studies of real-world networks. More generally, one asks whether there exists any assortative mixing between nodes, i.e., similar nodes are linked or not. Looking at nodes' similarity in terms of connectivity, assortative networks feature well-connected nodes joining other well-connected nodes, whereas in disassortative networks strongly connected nodes are linked to weakly connected ones. Newman (2002, 2003) showed that nodes' connectivity in many social networks tends to be positively correlated. McPherson et al. (2001) cite over one hundred studies that have observed homophily in some form or another. Examples range from company director networks, co-authorship and collaboration networks, or the network of email address books. On the contrary, most biological networks (protein-protein interaction network in the yeast cell, metabolic networks in bacteria, food webs) or technological networks (the Internet at the Autonomous System level, the network of hyperlinks between pages in the World Wide Web, etc.) appear to be disassortative. Networks in economic contexts may have features of both technological and social relationships (Jackson, 2010).

The evidence produced in the above mentioned studies on the topological properties of the ITN has generally neglected the issue of geographical distance and space (for a review on spatial networks, see Barthelemy, 2011). Instead, this paper explicitly takes on board geographical space in the way ITN graphs are defined. More precisely, we build trade graphs by filtering international trade flows so as to build trade-network structures, where the presence of any bilateral link is conditioned on the geographical distance between its two end nodes (countries). We will, therefore, analyze the topological properties discussed above conditional on countries' distance.

4 Data and methodology

In our analysis of the ITN, we employ the dataset made available by Subramanian and Wei (2007), which includes aggregate bilateral imports in constant US dollar for all countries, from 1970 to 2000.[9]

More formally, let n_t be the number of countries present in the database in year t, where a country i is said to enter the database if there is at least a positive import or export flow associated to it. Define, as in Section 3, $\mathscr{W}_t$ as the $n_t \times n_t$ weight matrix of the corresponding weighted directed ITN, where the generic element of $\mathscr{W}_t$, labeled as $w_t(i,j)$ represents the logarithmic transformation of positive-valued export flows from country i to j in year t (and zero if the corresponding trade flow is zero).[10] We also define the time-t binary matrix $\mathscr{A}_t$ as the binary $n_t \times n_t$ matrix whose generic element $a_t(i,j) = 1$ if and only if $w_t(i,j) > 0$, and zero otherwise. Therefore, we constructed both a weighted directed and a binary directed representation of the ITN. The binary directed representation gives us information about the presence or absence of trade partnerships, whereas the weighted directed representation adds to the binary structure information about the heterogeneity of export flows carried by each link.

The properties of the binary directed representation of the ITN built with this dataset were examined by De Benedictis and Tajoli (2011) and Duenas and Fagiolo (2013). Table 1 presents some descriptive statistics.

Over time, the ITN displays an increasing number of participating countries. Entry of new countries in the database is due to the presence of at least a new positive trade flow involving the entrant, and may be possibly caused either by the availability of new data or by the actual entry of the country in the international

[9] This dataset is built from the IMF Direction of Trade Statistics, reporting the importing country bilateral trade values at current prices, then deflated by US CPI at 1982–83 prices. The dataset is downloadable from the website `http://www.nber.org/~wei/data.html`

[10] We use a logarithmic scale instead of a linear one in order to make easier any comparison of our results with those from standard log-log gravity-equation formulation. To allow for meaningful comparisons across years, we also re-scale link weights by logs of yearly total world exports so as to account for the overall increase in total trade over time.

Table 1. Summary statistics.

	1970	1980	1990	2000
Countries (No.)	130	143	145	157
Trade flows (No.)	6593	8180	10289	11938
Density	0.393	0.403	0.493	0.487
Countries making up to 50% of trade	7	9	7	11
Flows making up to 50% of trade	71	88	67	77
Countries making up to 90% of trade	65	75	77	78
Flows making up to 90% of trade	793	893	747	854
Average trade flow	50.72	57.20	70.96	76.04
Median trade flow	54.50	60.00	64.00	70.00

Source: IMF, Directions of trade statistics.

trade market. New trade links, however, seem to increase more than quadratically with the number of participating countries, as shown by the rising density of the network over time.[11] Note also that the number of countries making up a high percentage (50% or 90%, respectively) of total trade tends to increase over the years, hinting to an increasing engagement in trade by a larger number of countries. However, given that also the total number of reported trading countries in our sample is also becoming larger, this figure does not necessarily imply a significant decline in the concentration of world trade.

In order to consider the effect of geographical distance on trade in the ITN defined above, we begin considering the original binary ($\mathscr{A}_t$) and weight ($\mathscr{W}_t$) matrix (from now on, we remove the time label to simplify the notation). To any given link ij between countries i and j (i.e., to each ordered pair (i, j) such that $a(i, j) = 1$), we associate the geographical distance $d(i, j)$, computed using the great-circle distance measure between the capital cities of the two countries. This allows us to build a geographical-distance matrix $\mathscr{D} = \{d(j, i)\}$, whose generic element is equal to the geographical distance between countries i and j whenever the correspondent trade link exists, and zero otherwise. Note that this matrix is by construction symmetric and possibly changes through time as trade links are created or removed.

Figure 1 plots a kernel-density estimation for the distribution of the logs of geographical distances between world countries in year 2000 (i.e., of the positive upper-diagonal entries of the matrix $\mathscr{D}$). It is easy to see that the distribution of the logs of distance is skewed to the left and presents a peak at large distances.

Table 2 breaks down the values of total trade (in percent) according to the different deciles of the distribution of distance among countries. The table suggests that most international trade occurs at relatively short (less than 2,000 km) and intermediate distances, but the distribution is very uneven.

[11] The density in a directed network is computed as the ratio between existing links and the maximum number of possible links, i.e., $\gamma_t = \frac{m_t}{m_{\max}} = \frac{m_t}{n_t(n_t-1)}$. For example, in 1970 one gets: $\gamma_{1970} = \frac{6593}{130\times129} = 0.393$.

Table 2. Share of total trade (in percentage) of each distance decile, by decade.

Decile	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Interval (km)	0–1128	1129–1978	1979–2811	2812–3556	3557–4453	4454–5183	5184–5914	5915–6879	6880–8500	8501–12351
1970	31.35	16.32	3.72	5.60	4.52	12.14	6.35	9.75	4.71	5.53
1980	30.99	13.12	5.84	6.97	3.96	11.31	7.85	10.45	5.63	3.88
1990	37.14	14.64	4.01	3.50	4.66	9.01	7.55	12.46	3.17	3.85
2000	33.24	19.27	4.44	3.61	4.50	8.46	6.35	10.42	5.43	4.28

Source: IMF, Directions of trade statistics.

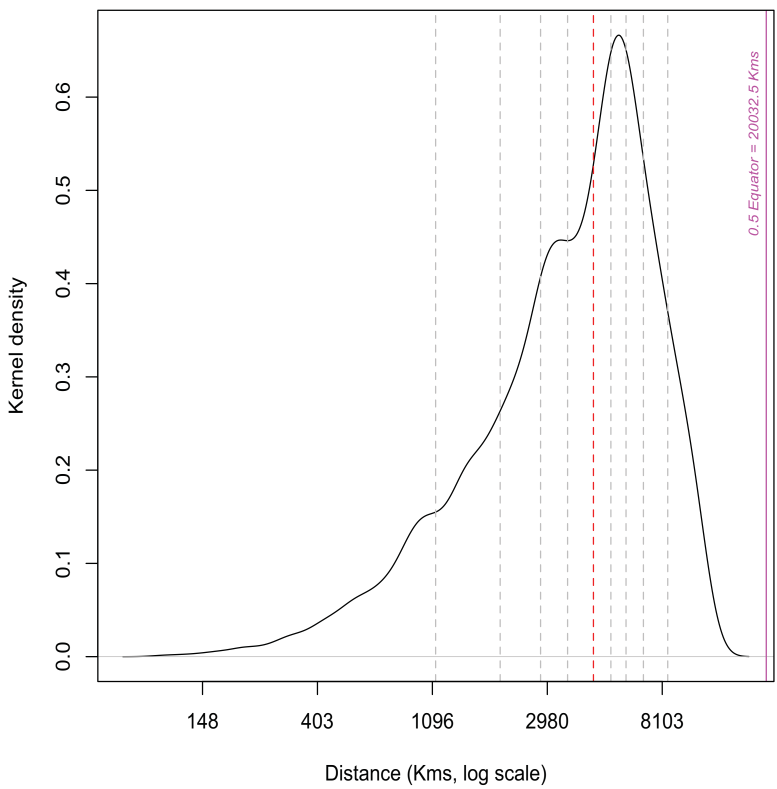

Fig. 1. The distribution of geographical distance between country pairs. Kernel density estimate of the distribution of distance (logarithmic scale). The middle dashed vertical line indicates the median value of the distribution, the other dashed lines indicate the deciles of the distribution, while the continuous vertical line, bounding the distribution to the right, indicates half of the equator distance. (color online)

We employ the deciles of the distance distribution to filter the (weighted and unweighted) ITN matrices. More specifically, we build two sets of 10 subnetworks to obtain two families of distance-conditioned ITNs. The first family, which we call the *cumulated distance-conditioned ITN*, is obtained by keeping in the networks, for each decile of the distribution of distance, only the arcs associated to geographical distances smaller than the upper-limit of that decile. For example, the fourth cumulated distance-conditioned ITN is obtained by keeping in the binary and weighted matrix only the arcs associated to pairs (i,j) of countries that are located at a distance below that of the upper-limit of the fourth decile of the distribution of distance (i.e., 3,556 km).

The second family of matrices, called simply *distance-conditioned ITN*, is obtained by keeping in each one of the ten networks, only the arcs associated to geographical distances belonging to that decile. For instance, the fourth *distance-conditioned ITN* is obtained by keeping in the binary and weighted matrices only the arcs associated to pairs (i,j) of countries that are located at a distance lower than that of the upper-limit of the fourth decile and above the upper-limit of the third decile of the distance distribution (i.e., between 2,811 km and 3,556 km).[12]

More formally, let us label with $\delta_{(1)},\ldots,\delta_{(10)}$ the deciles (in km) of the distribution distance, with $\delta_{(0)} = 0$. In each year, we build 10 *cumulated* distance-conditioned weight matrices, $\mathscr{W}_k^C = \{w_k^C(i,j)\}$, $k = 1,\ldots,10$, according to the following rule:

$$\begin{cases} w_k^C(i,j) = w(i,j) & \text{if } d(i,j) \leqslant \delta_{(k)} \\ w_k^C(i,j) = 0 & \text{otherwise} \end{cases}$$

and 10 *distance-conditioned* matrices $\mathscr{W}_k^{NC} = \{w_k^{NC}(i,j)\}$, for $k = 1,\ldots,10$:

$$\begin{cases} w_k^{NC}(i,j) = w(i,j) & \text{if } \delta_{(k-1)} \leqslant d(i,j) \leqslant \delta_{(k)} \\ w_k^{NC}(i,j) = 0 & \text{otherwise} \end{cases}$$

Obviously, the two network families are complementary. From the definitions above, one can indeed start from $\mathscr{W}_k^C$ and remove links in $\mathscr{W}_{k-1}^C$ to get $\mathscr{W}_k^{NC}$. Similarly, one can build $\mathscr{W}_k^C$ starting from $\mathscr{W}_h^{NC}$, $h = 1,\ldots,k$. Whereas cumulated networks give us a picture of the ITN for all trade relationships between countries that are distant less than a given threshold, distance-conditioned networks (non-cumulated) tell us what trade relationships can be imputed to trade relationships between pair of countries whose distance is within a certain range.

Cumulative and simple distance-conditioned networks are then analyzed both as binary (unweighted) networks and as weighted networks. In the binary case, one simply builds the corresponding binary matrices $\mathscr{A}_k^C = \{a_k^C(i,j)\}$ and $\mathscr{A}_k^{NC} = \{a_k^{NC}(i,j)\}$ by adding an arc whenever the correspondent entry in the weighted matrix is positive. It must be noticed that weighted and binary matrices provide complementary information on the role of different types of costs in international trade. In fact, the presence of fixed costs to access foreign markets should affect the characteristics of the binary matrix, determining the number of links that each country has, whereas trade variable costs (e.g., of the "iceberg" type) should affect links' weights, conditional on a link to be already in place.

In what follows, we start by presenting some baseline results for year 2000. The same exercise is repeated for each decade, from 1970to 2000, in order to test whether the impact of distance has changed over the years.

[12] More generally, one can use any quantile-based breakdown of the original distribution range, e.g., quintiles or percentiles. The choice of deciles has been made in order to efficiently trade off the need for a sufficiently larger number of distance classes and a sufficiently large number of observations in each quantile class. Note also that in principle one could have employed distance classes delimited by absolute km values, independently on the distribution. We preferred to use the quantile-based breakdown because in so doing we are sure that in each class there will be the same, fixed number of links. This implies that each subnetwork displays the same density. This avoids comparing subnetworks characterized by different densities.

5 Results

As customary in this literature (see, for example, Fagiolo et al., 2009, 2010; De Benedictis and Tajoli, 2011), we begin by analyzing the main topological characteristics of the network. The crucial difference with previous work is that we do so across the different subnetworks obtained above by conditioning to distance deciles. In particular, we focus on network connectivity (i.e., whether any two nodes can be connected in the network by a chain of links), the distributions of total node degree (i.e., a country's number of partners) and total node strength (i.e., a country's total imports plus exports), average nearest-neighbors degree and strength (i.e., average total degree or strength of the partners of a country), and clustering (i.e., the probability, possibly weighted by link weights, that any pair of partners of a node are themselves partners). Furthermore, for each decile, we examine the distribution of link weights, the correlation structure among node statistics, and the correlation structure between node statistics and some country macroeconomic characteristics (e.g., GDP and GDP per capita).

The analysis of subnetworks created for different distance deciles shows that considering the geographical distance between the nodes of the network indeed matters. Distance-conditioned trade subnetworks display topological properties that greatly change with distance deciles, as discussed in more details below.

5.1 Connectivity in the ITN

To begin with, we explore connectivity of the ITN as distance changes. The overall ITN is evidently connected, i.e., any country in the world can be reached from anywhere else in the network through undirected trade links.[13] When splitting the network based on distance, two interesting connectivity statistics are the number of connected components and the size of the giant component (i.e., the number of nodes making up the largest subset of connected nodes in the network). Of course if only one connected component is observed then the giant component has the same size of the overall network (as happens for the ITN as a whole).

Figure 2 shows how connectivity statistics change with distance in cumulated ITNs for year 2000. As we keep adding longer distance trade relationships in the network, the number of connected components sharply decreases, following a power-law shape. Initially, when only small-distance trades are taken into account, a very large number of small components emerge. This means that at small distances the ITN is extremely disconnected. Indeed, the hubs of the network are typically joined with countries that lie far apart. Accordingly, the relative size of the giant component is very small (about 7%) and quickly increases toward 1 as we consider longer-distance relationships. Complete connectivity only emerges when we start taking into the picture trade relationships occurring at 4,000 km or more. This allows one

[13] We employ throughout the concept of weak connectivity, which considers any two nodes connected if there is any link between them, irrespective of its directionality. Strong connectivity instead requires that pairs of nodes can be reachable via a directed path.

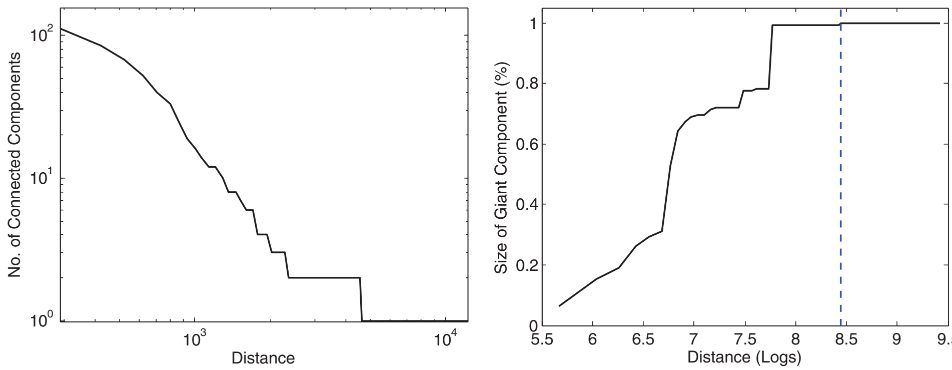

Fig. 2. Connectivity in distance-conditioned cumulated ITNs in year 2000. Left: Number of connected components vs. distance (log-log scale). Right: Size of the giant component (as a share of number of nodes) vs. distance. (color online)

to have in the network also links between the hubs and peripheral countries that are only connected to the hubs.[14]

5.2 Link weights and strength

Connectivity of the ITN over different distance intervals may be also studied in terms of link weights, and node degree and strength. Figure 3 shows that as we consider longer-distanced trade relationships in the ITN, average link weight initially decreases very rapidly and linearly, as predicted by the gravity equation. At higher distances, the effect of distance becomes strongly non-linear with respect to the logs of link weights, and in fact the average weight value increases for intermediate distances and then decreases again. All that maps into a smoother pattern for cumulated distance-conditioned networks: the steepness of the relationship between average link-weight decreases and distance decreases as we reach the median of the distribution.

Note also that the variance of link weights (not shown) follows a U-shaped pattern with respect to geographical distance , with the highest variance displayed at very low and very high distances. Taken together, this evidence confirms the very well-known negative relationship between trade flows and distance stressed in the empirical gravity literature, but highlights very marked non-linearities in the way in which distance affects both the average of (logs of) trade-flows and their conditional variance, especially for long-distance trade relationships.

We now turn to study how the correlation between node-specific network properties changes with geographical distance. Figure 4 shows the linear correlation coefficient computed between node out-degree (the number of countries to which a country exports) and out-strength (i.e., total exports), with 95% confidence intervals, in year 2000 for distance-conditioned networks.

[14] This result is consistent over the years. Indeed, when looking at simple distance-conditioned ITNs, one typically observes a stable number of connected components across deciles (between 15 and 23) and a few of isolate nodes (between 4 and 13).

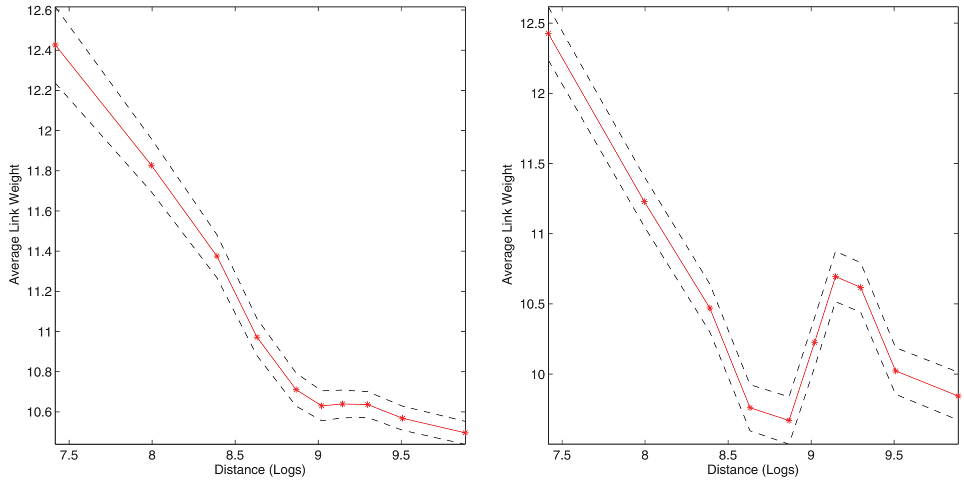

Fig. 3. Average link weights (solid lines) for simple (right) and cumulated (left) distance-conditioned networks. Dotted lines: 95% confidence bands. Year: 2000. (color online)

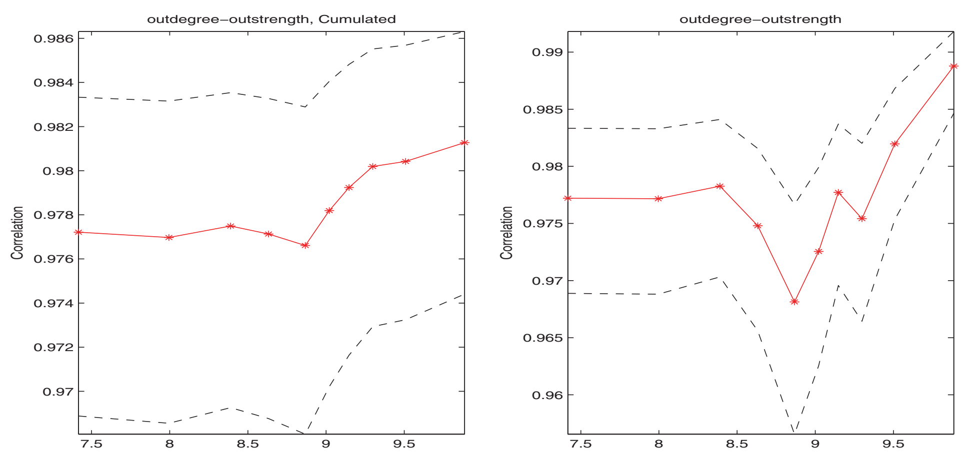

Fig. 4. Correlation coefficients (solid lines) between node degrees and node strengths for simple distance-conditioned networks. Dotted lines: 95% confidence bands. Year: 2000. (color online)

As expected, correlation coefficients are very high at all geographical distances: in general terms, countries that trade more also hold more trade partners. This evidence confirms that the geographical extensive margin and the intensive margins of trade are highly correlated. What is less expected is that distance does not affect significantly this high correlation: if any, a weak positive impact of distance on degree-strength correlation is detected, implying that the relation between number of trade partners and trade flows is marginally stronger for long-distance trade pairs. This result is in line with the suggestion by Chaney (2013). Furthermore, this evidence suggests that the relative relevance of fixed and variable costs in affecting overall trade costs—assuming that these are the variables affecting export volumes and the number of partners— is fairly constant across distance.

5.3 Assortativity and clustering

A number of previous analyses found a marked disassortativity pattern characterizing the ITN, both at the binary and weighted level (Garlaschelli and Loffredo, 2004, 2005; Serrano and Boguñá, 2003; Fagiolo et al., 2008, 2009). A disassortative (assortative) network is a graph where there is a negative (positive) correlation between node degrees/strengths and average nearest-neighbor degrees/strengths. More generally, a disassortative network is one where more (and more strongly) connected nodes are typically connected with less (and less strongly) connected nodes, i.e., countries tend to be connected with partners that are different in terms of connectivity. But do disassortativity patterns in the ITN depend on geographical distance? For example, do countries who trade more at a given distance tend to trade with partners that trade less at the same distance scale?

To answer these questions, we analyze the existing correlation structure between a country's out-degree (or out-strength) and its average nearest-neighbor out-degree (out-strength) in the distance-conditioned subnetworks obtained from our ITN. More precisely, in the binary network, we correlate the number of countries which a country exports to (its out-degree), with the average number of countries which those partners exports to (its degree out-out). The same intuition applies for the weighted network, once node degree is replaced by node strength. Figure 5 summarizes our results in year 2000, for cumulated vs. simple distance-conditioned networks, and for binary vs. weighted descriptions.

First of all, our analysis confirms that the aggregate ITN is found to be disassortative at the aggregate level: the correlation between node degree (strength) and $ANND$ ($ANNS$) is negative in the cumulated network including all distances (left panels of Figure 5). However, if one conditions the correlation structure to geographical distance, it is easy to see that short-distanced networks exhibit a very *assortative* pattern. The correlation coefficient is positive and quite high for all the short-distance networks displayed in the figure, and the pattern is very similar in the binary and weighted cases. As we add to the ITN links associated to higher distances, correlation coefficients decrease smoothly and non-linearly toward a disassortative pattern. More specifically, as Figure 5 shows, the ITN displays an assortative pattern for small distances, it becomes weakly disassortative for long-distance partnerships, while at intermediate distance no clear correlation patterns emerge. This means that, when only short-distance trade partnerships are considered, countries with many partners tend to trade with countries holding many connections. Conversely, at high distances, very connected countries typically trade with poorly connected partners.[15]

Taken together, this evidence implies that patterns of assortativity or disassortativity in the ITN are strongly dependent on distance, and the marked overall disassortativity of the ITN is mainly driven by high-distance trade relationships. A number of factors could produce this result. On the one hand, the existence of many preferential RTAs fostering trade between similar and often neighboring countries can explain the assortativity found at short distances. RTAs, if they integrate markets in a given region and reduce the cost of reaching countries within the agreement, will

[15] Note that we find a very similar pattern also when disaggregating assortativity with respect to import market distance (as mentioned, the results for the analysis on imports are not reported in this paper), suggesting that the directionality of trade flows is not a crucial factor in determining this result.

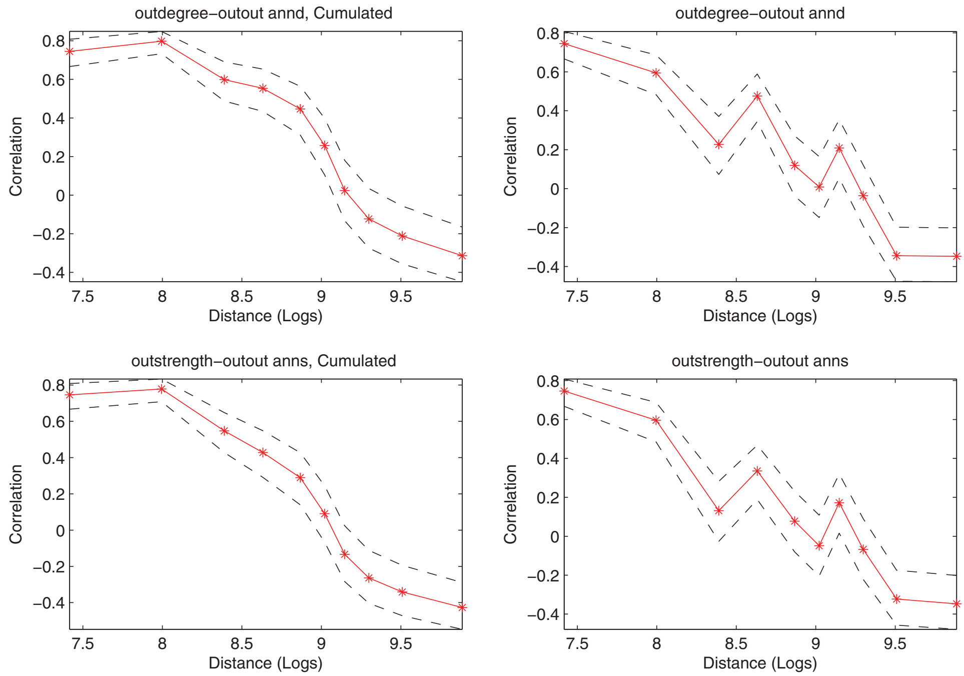

Fig. 5. Left: Assortativity in cumulated distance-conditioned directed networks. Top: The binary ITN case; bottom: The weighted ITN case. Right: Assortativity in simple distance-conditioned directed networks. Top: The binary ITN case; bottom: The weighted ITN case. For the binary case, the correlation coefficient is between node out-degree and directed average nearest-neighbor degree ($ANND$) measures. For the weighted case, the correlation coefficient is between node out-strength and directed average nearest-neighbor strength ($ANNS$) measures. Dotted lines: 95% confidence bands. Year: 2000. (color online)

typically generate a high number of strong trade links for all the member countries inside the region, producing the positive correlation result. Even, without trade agreements, barriers among countries belonging to a given geographical region tend to be lower than average (WTO, 2011) because historically a number of ties might have developed (for example, through migration or production links, or because of a common cultural and institutional heritage). On the other hand, countries that are geographically isolated from the rest of the world, when choosing an export destination for their goods might tend to prefer a well-connected country that works as a hub to connect them to rest of the system, as suggested by agglomeration economies (Fujita and Krugman, 2004) and by the presence of export platforms in the world trading system (Ekholm et al., 2007). This can in fact help to reduce distribution costs or to be closer to the relevant suppliers of the needed inputs or business services.

This preference will give rise to a disassortativity pattern at long distances. In the thirty-year period considered by our analysis, the assortativity patterns discussed so far seem to be relatively robust. If any, one notices an increasing disassortativity over time at higher distances. In other words, countries with a high out-degree/out-strength, i.e., important exporters, have been experiencing a significantly higher probability of being connected with countries that in turn export relatively little to a limited number of markets. By year 2000, the largest exporters had reached nearly every market in the world, even the countries that appear geographically isolated. Therefore, the marginal cost of accessing additional markets seem to have become very small, especially for big exporters.

The evidence of the simultaneity of disassortativity and assortativity patterns for different distances suggests a structure of the world trading system where regionalization and globalization forces co-exist. The global trading system is held together by a relatively small number of highly connected countries, which are linked to nearly every other node of the system, and play the role of hubs. At the same time, we observe some tightly connected regions formed by nearby countries. Therefore, the analysis of the trade structure shows globalization in terms of an increase in the number of trading countries and trade flows (extensive and intensive margin) and the presence of strong long-distance links connecting countries far apart, together with an elevated regionalization, that is a high and rising regional clustering, revealing a marked propensity to trade with close and similar trade partners. This confirms the suggestion that globalization and regionalization need not to be juxtaposed (as observed in Iapadre and Tajoli (2014) and Piccardi and Tajoli (2015)). The mix of fixed and variable costs changes along distance, coherently with the observed evidence of a switching from an assortative system to a disassortative one.

We turn now to explore the observed patterns of distance-conditioned clustering coefficients. The clustering coefficient of a node in the ITN measures the likelihood that a country forms intensive-trade triangles with its trade partners (Fagiolo, 2007). Previous studies show that the binary version of the ITN (at all distances) is highly clustered, whereas the weighted ITN displays a relatively weaker clustering, due to the presence of many low-trade interactions that weaken the ex-post intensity of the many triangular trade relations existing in the ITN (Fagiolo et al., 2009).

Figure 6 shows that distance plays a crucial role also in explaining evidence on clustering. Indeed, both binary (BCC) and weighted (WCC) clustering coefficient in cumulated distance-conditioned ITNs (left panels of the figure) non-linearly increase as we add to the network longer distance trade partnerships. More specifically, BCC and WCC reach a maximum when we consider trade relations close to 3,800 km, and then decrease as we approach 6,000 km, and then increase again toward the absolute maximum of average clustering. This suggests that most of the contribution to maximum clustering by binary and weighted triangular trade interactions comes from smaller-distance trade flows. This is confirmed by looking at the plot of the BCC and WCC for simple distance-conditioned ITNs (right panels). Average clustering non-linearly decreases toward zero as distance increases, meaning that triples of countries that are very distant to each other almost never engage in triangular trade relationships (with the exceptions of trade flows occurring between countries whose distance is about 9,900 km). Hence, only short-distanced triples of countries contribute to the large value of the overall BCC and WCC found in the ITN. Notice again that strong non-linear effects affect the link between logs of distance and network statistics.

Overall, also the emergence of a sharp decreasing relationship between distance and clustering can be due to the impact of RTA effects: a RTA indeed favors clustering as it creates and enforces the establishments of cliques (and thus triangular trade relationships) among countries located relatively close to each other. Instead,

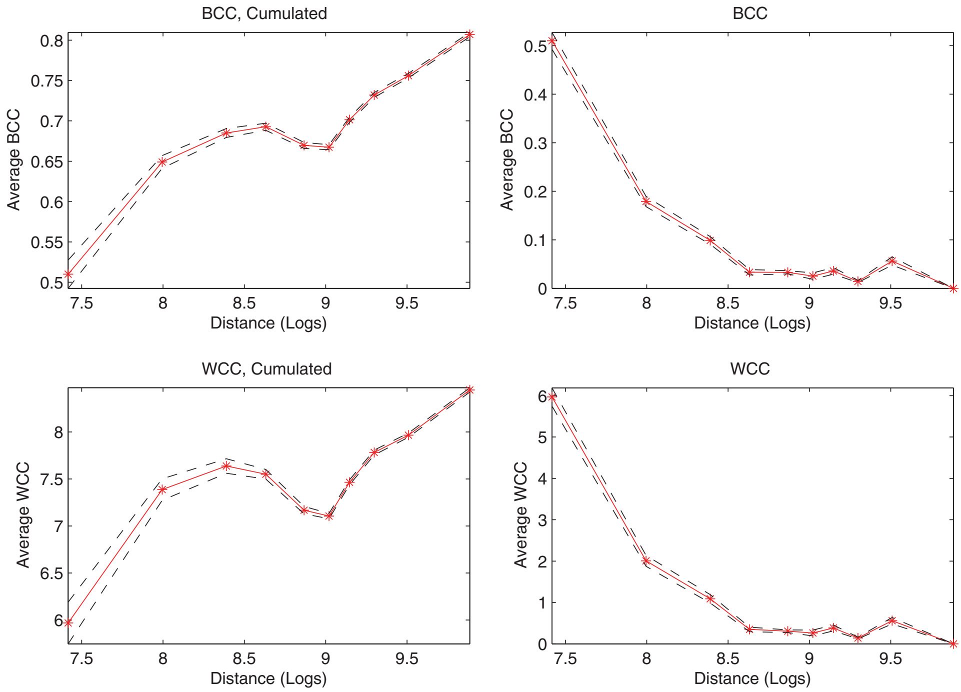

Fig. 6. Left: Clustering in distance-conditioned directed networks. Right: Clustering in simple distance-conditioned directed networks. Top: the binary ITN case; bottom: the weighted ITN case. For the binary case, the binary clustering coefficient (BCC) computes for each node the percentage of closed triangles in the node's neighborhood. In the weighted case, the weighted clustering coefficient (WCC) computes the intensity of such closed triangles, where triangle edges are weighted by link weights. Dotted lines: 95% confidence bands. Year: 2000. (color online)

the low clustering coefficient found for networks including only countries that are geographically far apart reinforces the idea that some countries may play the role of hubs of the system.

5.4 The role of country size and income

Gravity models emphasize the role played by economic size and income, in addition to geographical distance, in shaping bilateral trade flows. To address this issue in the present context, we begin by examining correlation patterns between network-based statistics and country size and income (as measured by GDP and GDP per capita) over different distance deciles. It is in fact well-known that—everything else being equal—large countries tend to be also large traders. Even if large countries tend to be less open than small ones for a variety of reasons, and therefore the trade-to-GDP ratio tends to fall with GDP, it is still true that total trade values tend to grow with GDP.

The four panels of Figure 7 show for year 2000 the plots of linear correlation coefficients between node degree/strength and node GDP/per capita GDP, conditioned on distance deciles.

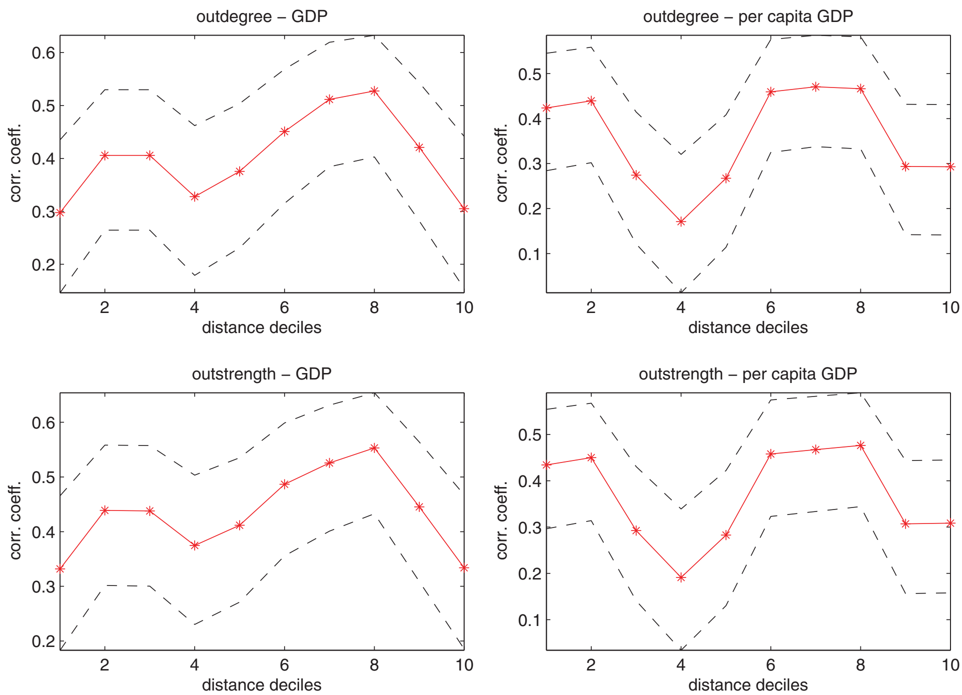

Fig. 7. Correlation between node degree/strength and GDP/per capita GDP vs. distance in year 2000. Dotted lines: 95% confidence bands. (color online)

We note that the correlation between GDP and total degree and strength is generally positive, even if it varies a lot across distance deciles. The correlation is, especially high for a subset of trade flows involving pairs of countries that lie very far apart (between 6,000 km and 7,000 km), but do not belong to the last two deciles. This is true also for per-capita GDP, and it means that at large distances economic size and income of countries heavily and positively influence their trade values. Note also that correlation coefficients first increase and then decrease when country GDP is considered. While distance certainly affects trade flows, the observed pattern indicates the extent of this influence depends also on countries' economic size. These results suggest that there could be a sort of optimal distance value that maximizes the correlation between country size and country connectivity in terms of trade partners and total trade. On the contrary, at some distances, the correlation between node connectivity and *per-capita* GDP could be very weak, or not-significantly different from zero, as shown in the figure for a distance around 4,000 km.

The foregoing size/income correlation analysis, however, does not take into account the fact that export flows bilaterally depend on both origin and destination market sizes, as gravity model estimates of trade always confirm. In order to allow for such a dependence in our data, we have investigated what happens when one re-scales link weights by the expected bilateral flow in a frictionless world. In this setup, the ij link weight is simply defined as exports from i to j (in levels) divided by the product between the GDPs of country i and country j. In this way, the impact of origin and destination market sizes are washed away and we can understand how distance affect the properties of the ITN regardless of any size effect.

As we did above, we used the logs of re-scaled link weights, namely, the log of exports minus the sum of the logs of i and j GDPs. In order to avoid negative weights, we have translated the entire distribution by a minimum threshold, in such a way to preserve existing density. We have re-computed all (weighted) topological properties of this GDP-re-scaled ITN to see how they change across the deciles of the logs of distance distribution. Of course, the binary version of the ITN is almost unaffected by this change, as the rescaling only influences positive original weights.

Results strongly confirm the main insights coming from the foregoing analysis. For example, GDP re-scaling preserves the negative relationship between link weights and distance in the distance-conditioned networks. What is more, non-linearities still emerge in the log-log relation between weights and distance. The correlation between node degree and node strength remains positive and very high, and relatively less sensitive to distance. Weighted disassortativity patterns are instead unaffected by the rescaling. This means that this result is not driven by country economic sizes. Finally, the increasing relation between average weighted clustering and distance in cumulated distance-conditioned networks is still present, even if slightly less important: when we account for country GDPs, distance seems to impact a little less on clustering. This difference wades away in simple distance-conditioned networks, where we still observe a significant and negative impact of distance on average weighted clustering.

This analysis can be of course extended by more strongly pursuing the idea of filtering away gravity-based influences on export flows. Following Fagiolo (2010), one could think to fit a gravity model to bilateral trade flows and to employ the residuals to build a trade network where now flows are net of any effect coming from size, borders, trade unions, etc. (excluding distance), and to explore the properties of such network as distance changes.

5.5 Network structure and distance over time

So far, we have explored the connection between geographical distance and network structure by focusing on year 2000. But what happens to this connection over the years?

The distribution of average link weights over distance shown in Figures 8 and 9 displays some similarities over time, but also some changes: the average link weight presents a downward trend over distance in all years, but the slope and the kinks are different. In particular, in 1970 and in 1980, the weight of trade links moves irregularly over a wide range of intermediate distances. In the more recent decades, some of these swings smooth out, but the increase in the link weight at middle-high distances becomes more evident. Overall, both from a simple and cumulated perspective, the relation between trade flows and distance was very non-linearly shaped also in the past, with only small-distanced and large-distanced trade flows markedly decreasing with geographical distance.

Correlations between node degree/strength and country size/income across distance follows a very similar pattern from 1970 to 2000, basically reproducing what we observe in Figure 7. The persistence in the role of distance that we pick up in

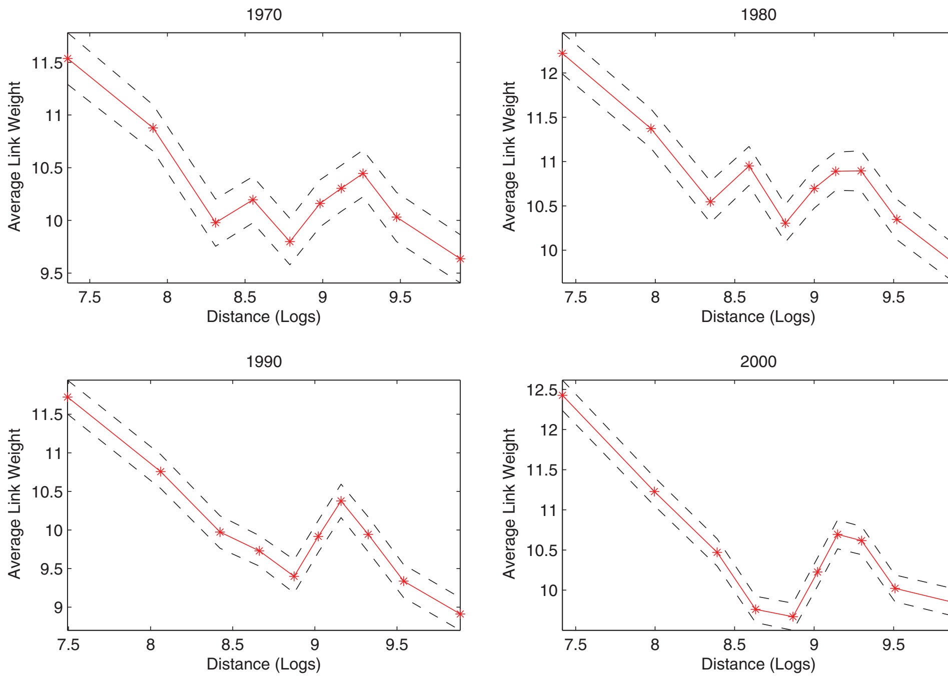

Fig. 8. Average link weights (solid lines) for simple distance-conditioned networks. Dotted lines: 95% confidence bands. Years: 1970, 1980, 1990, 2000. (color online)

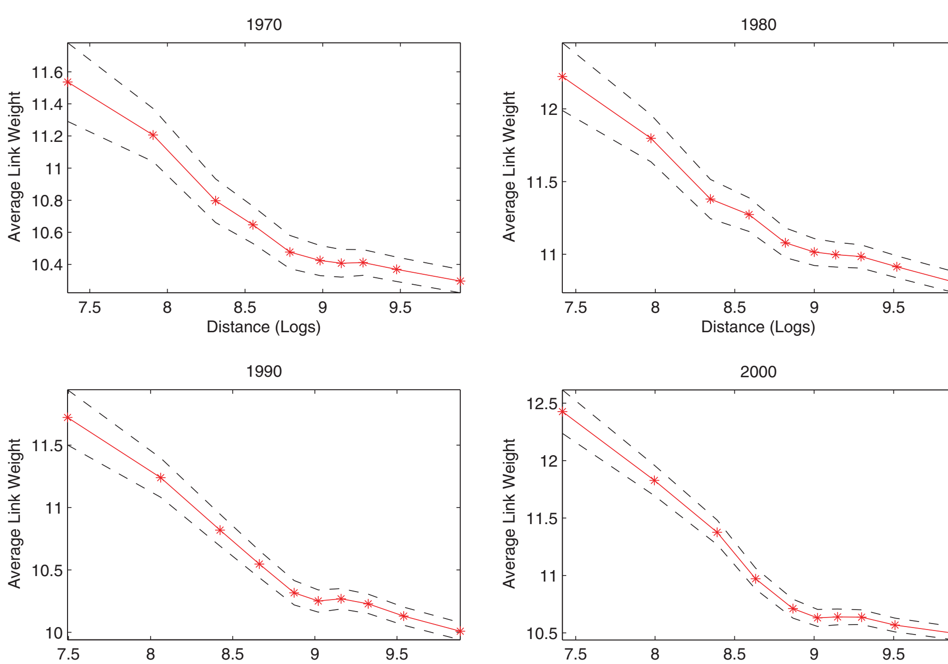

Fig. 9. Average link weights (solid lines) for cumulative distance-conditioned networks. Dotted lines: 95% confidence bands. Years: 1970, 1980, 1990, 2000. (color online)

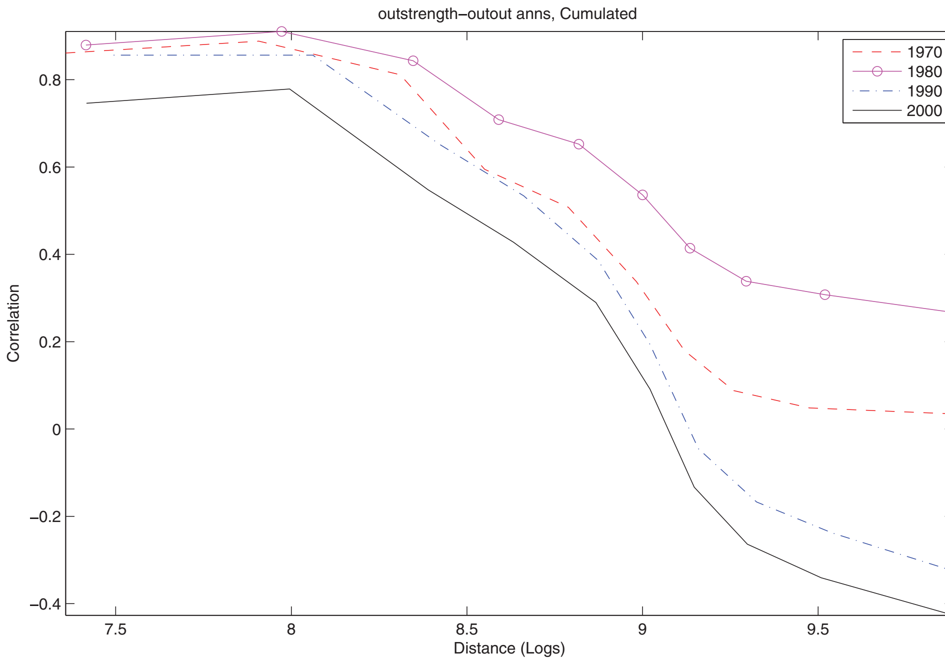

Fig. 10. Assortativity and disassortativity. Years: 1970, 1980, 1990, 2000. (color online)

time is in line with the results found elsewhere in the literature using the gravity model of trade (Disdier and Head, 2008). In our framework, we can interpret this result by arguing that, as the size of the network increases over time, because the number of relevant trading countries increases, geographical distance continues to matter, as variable costs per kilometer might decline, but overall trade costs to stay connected with the entire network remain high.

Our analysis shows that while the role of distance did not decline in time, its impact on the relation between country's economic size or income and country connectivity has somewhat changed. In fact, we observe that the correlation levels between nodes' strength and countries' average income were higher in 1970 than in 2000. This confirms an increased participation to global trade, especially by low- and middle-income countries, probably due to the secular decline in trade costs and in other trade barriers.

The reduction in the correlation seems to be the result of a change of the ITN affecting especially countries in middle-distanced subnetworks: for these countries, the correlation between their trade flows and average incomes has become lower and less significant, possibly because of the historical reduction in trade costs. Instead, as distance increases beyond this middle range, trade costs are still relevant and therefore higher incomes—capable of overcoming such trade costs—are still correlated with higher trade flows.

Much more evident are the changes over time in other features of the network structure along different distances. As shown in Figure 10, the property of being assortative at short distances but disassortative (or at least much less assortative) at large distances has characterized the ITN since the 1970s. But over time, the extent of the disassortativity at longer distances has increased substantially. In the

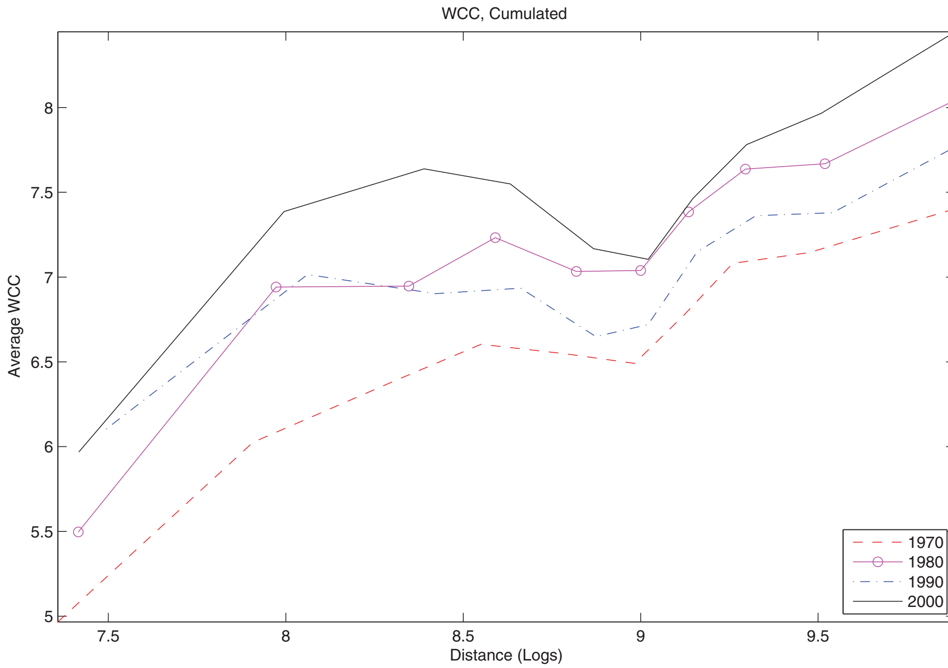

Fig. 11. Clustering. Years: 1970, 1980, 1990, 2000. (color online)

past, when the potential number of trading countries was smaller and trade costs were higher, the tendency to link to similar countries was stronger, and overall assortativity prevailed. Nowadays, as the heterogeneity of countries involved in international trade has increased, the disassortativity for large-distanced countries is much higher.

Also, clustering has increased remarkably over the years, even if the pattern over distance is somewhat similar in time (see Figure 11). This can again be related to increase in the number of preferential trade agreements signed between countries. Interestingly, the figure shows a clear increase in the clustering coefficient also for far away countries. This evidence is in line with the more recent tendency of countries to sign preferential trade agreements not so much with neighboring partners, but also with countries in other continents (see WTO, 2011). These results confirm that even if globalization and technological improvements have not eliminated the role of distance as a hindering factor to trade, its impact on the overall structure of the ITN has changed over the years.

6 Conclusions

In this paper, we explored whether the topological architecture of the ITN changes in geographical space and along time. We employed geographical distance between countries in the world to filter trade relationships in the ITN to build subnetworks of countries who trade with partners located at similar geographical distances.

Our main result is that geographical distance matters also from a complex network perspectives, as we find a relevant effect of distance on the ITN topological properties, and we show that such an effect is highly non-linear.

Our results highlight that the role of distance and trade costs in affecting trade patterns is not only based on linear bilateral distance in miles, and that the relevance of both a fixed and variable component of trade costs can jointly generate the type of non-linearity observed here.

These results are very important to better understand the phenomenon of globalization. In fact, globalization per se should imply that a country's position on the globe is irrelevant for its economic relations, as in a globalized system there is not a more central position by definition. Instead, the relevance of geographical distance in determining the connectivity of countries and the topological structure of the system highlight that this definition of globalization does not fully apply to what happened in the last part of the 20 century.

This is confirmed by the analysis showing that the geographical pattern of trade and the geographical extensive margin are very different for groups of countries at different distance ranges: while the aggregated ITN over all distance ranges is disassortative, shorter distance subnetworks are assortative. Furthermore, the trade intensity and number of triangular trade relationships decreases as distance increases. Both results confirm that the structure of the ITN changes dramatically in geographical space.

The results on the structure of the network also show the role of specific countries working as hubs of the system, thereby assigning a more central position to some specific nodes. These results indicate that the role of distance is different for countries with different economic size, which should have a different capacity to overcome the trade costs that distance imply. Large countries with very high trade volumes can exploit the economies of scale associated with their size created by the presence of fixed costs. Therefore, they can trade profitably also with far-away countries, and they can play the role of hubs of the system.

The fact that the ITN is assortative when distance is small and disassortative when it is large opens up interesting policy insights. First of all, this confirms that economic shock transmission through trade, depending on the network structure, is relatively independent from the shock size per se: the propagation can be very different for apparently similar initial trade fluctuations according to which node is hit. Therefore, a country exposure does not only depend on its trade over GDP ratio, or on its distance from trade partners. Indeed, it is well-known (Newman, 2002) that exogenous shocks hitting assortative networks percolate more easily among their nodes, as strongly-connected nodes tend to concentrate in a core group, which, in the case of ITN, tends to be geographically concentrated. However, these assortative clusters are also relatively resilient to such shocks because of the redundancy in connectivity patterns. Therefore, groups of countries in the ITN that are geographically close are at the same time more exposed to shocks and more resilient to them with respect to countries that are more distant, which on the contrary are less exposed to shocks but when they are hit take much longer to recover after a shock.

Second, the choice of trade partners should not focus only on bilateral links between countries. Not every trade link provides a similar access to the world market. Access to the global market is facilitated for a country that is distant from the main importers if it links to a hub of the system, possibly by joining the production chain of such a hub, or to a lesser extent through a trade agreement that ease the access to the market of a central player. The results are also relevant to assess the scope of trade agreements. In the past, trade agreements occurred, especially between countries belonging to the same region, perceived as "natural trading partners." But insofar as a trade agreement facilitates trade with a given partner, its marginal benefit can be higher when occurring between far away countries with higher trade obstacles to overcome.

Comparing the results of the analysis over time, we note that the average trade partner in the middle-distanced group has increased its relative strength or share of total trade between 1970 and 2000, hinting to an increase participation in international trade of such a group of countries. This is confirmed by the correlation structure between network statistics and per-capita GDP.

The effect of distance over time is not trivial. Over the time period examined, the size of the network increases, as the number of connected countries increase, and the diameter of the network grows. Therefore, distance remains relevant (as shown also in gravity models), in spite of a (relative) decline in trade costs, because the incidence of such costs has increased for most countries with the increase of their openness and of their number of trade links.

This study can be extended in many ways. First, one may want to explore the topological properties of the ITN by explicitly embedding the network in a spatial structure and use methodologies developed in the literature on spatial networks (Barthelemy, 2011). Second, a more theoretical explanation building on the interplay between fixed (sunk) and variable trade costs may be conceived, so as to develop a proper network formation model able to replicate, for example, the structural breaks detected in, e.g., assortativity patterns when we move from smaller to higher distances. Furthermore, the empirical evidence in the paper is consistent with Chaney (2013), the negative correlation between *outDegree* and *ANND* is saying that only countries that are exporting to a lot of destinations are able to reach remote markets. The model by Chaney also suggest an empirical test, if the mechanism driving countries distribution of destination markets is the one suggested by the model ones may expect a positive correlation between country productivity (eventually proxied by income per capita) and the average distance of its export.

Acknowledgments

This paper is a revised version of *The International Trade Network in Space and Time* that was published as Scuola Superiore Sant'Anna LEM Working Paper no. 2012/17. The authors thank the participants to the ARS'11 Conference in Naples, NAIT workshop in Macerata, ETSG Conference in Leuven, and ITSG in Ancona for their insightful comments, and Johan Koskinen, Pietro Panzarasa, Sara Gorgoni, Thomas Chaney, Pat Doreian and two anonymous referees for comments to previous versions of the paper.

Conflicts of interest

Angela Abbate, Luca De Benedictis, Giorgio Fagiolo and Lucia Tajoli have nothing to disclose.

References

Abeysinghe, T., & Forbes, K. (2005). Trade linkages and output-multiplier effects: A structural VAR approach with a focus on Asia. *Review of International Economics*, **13**, 356–375.

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Review of Modern Physics*, **74**, 47–97.

Anderson, J. E. (2011). The gravity model. *Annual Review of Economics*, **3**, 133–160.

Anderson, J. E., & van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, **93**, 170–192.

Anderson, J. E., & van Wincoop, E. (2004). Trade costs. *Journal of Economic Literature*, **42**, 691–751.

Anderson, J. E., & Yotov, Y. V. (2012). The gold standard gravity. *NBER Working Papers* 17835.

Artis, M., Galvao, A. B., & Marcellino, M. (2007). The transmission mechanism in a changing world. *Journal of Applied Econometrics*, **22**, 39–61.

Arvis, J.-F., Duval, Y., Shepherd, B., Utoktham, C., & Raj, A. (2016). Trade costs in the developing world: 1996–2010. *World Trade Review*, **15**, 451–474.

Baltagi, B. H., Egger, P., & Pfaffermayr, M. (2007). Estimating models of complex FDI: Are there third-country effects? *Journal of Econometrics*, **140**, 260–281.

Barigozzi, M., Fagiolo, G., & Garlaschelli, D. (2010). The multi-network of international trade: A commodity-specific analysis. *Physical Review E*, **81**, 046104.

Barthelemy, M. (2011). Spatial networks. *Physics Reports*, **499**, 1–101.

Bernard, A., Jensen, J., Redding, S., & Schott, P. (2007). Firms in international trade. *Journal of Economic Perspectives*, **21**, 105–130.

Bhattacharya, K., Mukherjee, G., & Manna, S. (2007). The international trade network. In A. Chatterjee, & B. Chakrabarti (Eds.), *Econophysics of markets and business networks.* Milan, Italy: Springer-Verlag.

Bhattacharya, K., Mukherjee, G., Sarämaki, J., Kaski, K., & Manna, S. (2008). The international trade network: Weighted network analysis and modeling. *Journal of Statistical Mechanics: Theory Experiment A*, **2**, P02002.

Brun, J.-F., Carrère, C., Guillaumont, P., & de Melo, J. (2005). Has distance died? Evidence from a panel gravity model. *World Bank Economic Review*, **19**, 99–120.

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *American Economic Review*, **98**, 1707–1721.

Chaney, T. (2013). The gravity equation in international trade: An explanation. *NBER Working Papers* 19285.

Chaney, T. (2016). Networks in international trade. In Y. Bramoullé, A. Galeotti, & B. Rogers (Eds.), *The Oxford handbook of the economics of networks* (pp. 759–798). Oxford, UK: Oxford University Press.

Chen, M. X., & Joshi, S. (2010). Third-country effects on the formation of free trade agreements. *Journal of International Economics*, **82**, 238–248.

De Benedictis, L., Nenci, S., Santoni, G., Tajoli, L., & Vicarelli, C. (2014). Network analysis of world trade using the BACI-CEPII dataset. *Global Economy Journal*, **14**, 287–343.

De Benedictis, L., & Taglioni, D. (2011). The gravity model in international trade. In L. De Benedictis, & L. Salvatici (Eds.), *The trade impact of European union preferential policies,* chapter 4 (pp. 55–89). Berlin Heidelberg: Springer.

De Benedictis, L., & Tajoli, L. (2011). The world trade network. *The World Economy*, **34**, 1417–1454.

De Groot, H. L., Linders, G.-J., Rietveld, P., & Subramanian, U. (2004). The institutional determinants of bilateral trade patterns. *Kyklos*, **57**, 103–123.

Disdier, A., & Head, K. (2008). The puzzling persistence of the distance effect on bilateral trade. *Review of Economics and Statistics*, **90**, 37–48.

Dorogovtsev, S., & Mendes, J. (2003). *Evolution of networks: From biological nets to the internet and WWW*. Oxford: Oxford University Press.

Duenas, M., & Fagiolo, G. (2013). Modeling the international-trade network: A gravity approach. *Journal of Economic Interaction and Coordination*, **8**, 155–178.

Eaton, J., & Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, **70**, 1741–1779.

Egger, P., & Larch, M. (2008). Interdependent preferential trade agreement memberships: An empirical analysis. *Journal of International Economics*, **76**, 384–399.

Eichengreen, B., & Irwin, D. A. (1995). Trade blocs, currency blocs and the reorientation of world trade in the 1930s. *Journal of International Economics*, **38**, 1–24.

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

Ekholm, K., Forslid, R., & Markusen, J. R. (2007). Export-platform foreign direct investment. *Journal of the European Economic Association*, **5**, 776–795.

Fagiolo, G. (2007). Clustering in complex directed networks. *Physical Review E*, **76**, 026107.

Fagiolo, G. (2010). The international-trade network: Gravity equations and topological properties. *Journal of Economic Interaction and Coordination*, **5**, 1–25.

Fagiolo, G. (2015). The international trade network: Empirics and modeling. In J. Nicoll Victor, M. Lubell, & A. H. Montgomery (Eds.), *Oxford handbook of political networks,* chapter 29. Oxford, UK: Oxford University Press.

Fagiolo, G., Schiavo, S., & Reyes, J. (2008). On the topological properties of the world trade web: A weighted network analysis. *Physica A*, **387**, 3868–3873.

Fagiolo, G., Schiavo, S., & Reyes, J. (2009). World-trade web: Topological properties, dynamics, and evolution. *Physical Review E*, **79**, 036115.

Fagiolo, G., Schiavo, S., & Reyes, J. (2010). The evolution of the world trade web: A weighted-network approach. *Journal of Evolutionary Economics*, **20**, 479–514.

Feenstra, R. (2002). Border effects and the gravity equation: Consistent methods for estimation. *Scottish Journal of Political Economy*, **49**, 491–506.

Forbes, K. (2002). Are trade linkages important determinants of country vulnerability to crises? In S. Edwards and J. A. Frankel (Eds.), *Preventing currency crises in emerging markets*. Chicago: University of Chicago Press.

Fujita, M., & Krugman, P. R. (2004). The new economic geography: Past, present and the future. *Papers in Regional Science*, **83**, 139–164.

Garlaschelli, D., Di Matteo, T., Aste, T., Caldarelli, G. & Loffredo, M. (2007). Interplay between topology and dynamics in the World Trade Web. *The European Physical Journal B*, **57**, 1434–6028.

Garlaschelli, D., & Loffredo, M. (2004). Fitness-dependent topological properties of the world trade web. *Physical Review Letters*, **93**, 188701.

Garlaschelli, D., & Loffredo, M. (2005). Structure and evolution of the world trade network. *Physica A*, **355**, 138–44.

Haggett, P., & Chorley, R. (1969). *Network analysis in geography*. London, UK: Edward Arnold.

Head, K., & Mayer, T. (2010). Gravity, market potential and economic development. *Journal of Economic Geography*, **11**(2), 281–294.

Head, K., & Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. *Handbook of International Economics*, **4**, 131–195.

Helliwell, J. F., & Padmore, T. (1985). Empirical studies of macroeconomic interdependence. In R. Jones, & P. Kenen (Eds.), *Handbook of International Economics*. North Holland, Amsterdam, The Netherlands: Elsevier Science Publishers B.V.

Helpman, E., Melitz, M., & Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *Quarterly Journal of Economics*, **123**, 441–487.

Hillberry, R., & Hummels, D. (2008). Trade responses to geographic frictions: A decomposition using micro-data. *European Economic Review*, **52**, 527–550.

Hummels, D. (2007). Transportation costs and international trade in the second era of globalization. *Journal of Economic Perspectives*, **21**, 131–154.

Iapadre, P. L., & Tajoli, L. (2014). Emerging countries and trade regionalization. A network analysis. *Journal of Policy Modeling*, **36**, S89–S110.

Jackson, M. O. (2010). *Social and economic networks*. Princeton, NJ: Princeton University Press.

Kali, R., Méndez, F., & Reyes, J. (2007). Trade structure and economic growth. *Journal of International Trade & Economic Development*, **16**, 245–269.
Kali, R., & Reyes, J. (2007). The architecture of globalization: A network approach to international economic integration. *Journal of International Business Studies*, **38**, 595–620.
Krugman, P. (1995). Growing world trade: Causes and consequences. *Brookings Papers on Economic Activity*, **26**, 327–377.
Krugman, P. R. (1998). What's new about the new economic geography? *Oxford Review of Economic Policy*, **14**, 7–17.
Krugman, P. R., & Venables, A. (1995). Globalization and the inequality of nations. *Quarterly Journal of Economics*, **110**, 857–880.
Lawless, M. (2010). Deconstructing gravity: Trade costs and extensive and intensive margins. *Canadian Journal of Economics*, **43**, 1149–1172.
Lendle, A., Olarreaga, M., Schropp, S., & Vézina, P.-L. (2016). There goes gravity: eBay and the death of distance. *The Economic Journal*, **126**, 406–441.
Li, X., Jin, Y. Y., & Chen, G. (2003). Complexity and synchronization of the World trade Web. *Physica A: Statistical Mechanics and its Applications*, **328**, 287–96.
McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, **27**, 415–444.
Melitz, J., & Toubal, F. (2014). Native language, spoken language, translation and trade. *Journal of International Economics*, **93**, 351–363.
Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, **71**, 1695–1725.
Newman, M. E. J. (2002). Assortative mixing in networks. *Physical Review Letters*, **89**, 87–101.
Newman, M. E. J. (2003). Mixing patterns in networks. *Physical Review E*, **67**, 026126.
Obstfeld, M., & Rogoff, M. (2001). The six major puzzles in international macroeconomics: Is there a common cause?. *NBER Macroeconomics Annual 2000*, **15**, 339–412.
Piccardi, C., & Tajoli, L. (2015). Are preferential agreements significant for the world trade structure? A network community analysis. *Kyklos*, **68**, 220–239.
Redding, S. J. (2011). Theories of heterogeneous firms and trade. *Annual Review of Economics*, **3**, 77–105.
Redding, S. J., & Venables, A. J. (2004). Economic geography and international inequality. *Journal of International Economics*, **62**, 53–82.
Reichardt, J., & White, D. (2007). Role models for complex networks. *The European Physical Journal B*, **60**, 217–224.
Reyes, J., Schiavo, S., & Fagiolo, G. (2008). Assessing the evolution of international economic integration using random-walk betweenness centrality: The cases of East Asia and Latin America. *Advances in Complex Systems*, **11**, 685–702.
Santos Silva, J. M. C., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, **88**, 641–658.
Serrano, A., & Boguñá, M. (2003). Topology of the world trade web. *Physical Review E*, **68**, 015101(R).
Serrano, A., Boguñá, M., & Vespignani, A. (2007). Patterns of dominant flows in the world trade web. *Journal of Economic and Coordination*, **2**, 111–124.
Squartini, T., Fagiolo, G., & Garlaschelli, D. (2011a). Randomizing world trade. I: A binary network analysis. *Physical Review E*, **84**, 046117.
Squartini, T., Fagiolo, G., & Garlaschelli, D. (2011b). Randomizing world trade. II: A weighted network analysis. *Physical Review E*, **84**, 046118.
Subramanian, A., & Wei, S.-J. (2007). The WTO promotes trade, strongly but unevenly. *Journal of international Economics*, **72**, 151–175.
Tzekina, I., Danthi, K., & Rockmore, D. (2008). Evolution of community structure in the world trade web. *The European Physical Journal B—Condensed Matter*, **63**, 541–545.
Wei, S. (1996). Intra-national versus International trade: How Stubborn are Nations in Global Integration?. *NBER Working Paper* 5531.
Wilson, A. (2000). *Complex spatial systems: The modelling foundations of urban and regional analysis*. London: Routledge.
WTO (2011). *The WTO and preferential trade agreements: From co-existence to coherence. World Trade Report 2011*. Geneva: WTO Publications.