The Indirect Effects of FDI on Trade: A Network Perspective

Paolo Sgrignoli[1], Rodolfo Metulini[2], Zhen Zhu[3] and Massimo Riccaboni*[3,4]

[1]Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy
[2]Department of Economics and Management, University of Brescia, Brescia, Italy
[3]LIME, IMT School for Advanced Studies Lucca, Lucca, Italy
[4]DMSI, KU Leuven, Leuven, Belgium

September 20, 2016

Abstract

The relationship between international trade and foreign direct investment (FDI) is one of the main features of globalization. Since FDI takes both direct and indirect channels from origin to destination countries due to factors such as global value chains, tax and investment treaties, and corporate strategies, in this paper we investigate the effects of FDI on trade from a network perspective. We use a unique data set of international corporate control as a measure of stock FDI to construct a corporate control network (CCN) where the nodes are the countries and the edges are the corporate control relationships. Based on the CCN, the network measures, i.e., the shortest path length and the communicability, are computed to capture the indirect channel of FDI. Empirically we find that corporate control has a positive effect on trade both directly and indirectly. The result is robust with different specifications and estimation strategies. Hence, our paper provides strong empirical evidence of the indirect effects of FDI on trade. Moreover, we identify a number of factors affecting whether corporate control and trade are complements or substitutes, including regional trade agreements and the region of Asia. We also find that the indirect effects are more pronounced for manufacturing sectors than for primary sectors such as oil extraction and agriculture.

Keywords: *Networks; Foreign direct investment; Corporate control*
JEL classification: C21; F10; F14; F23; L22

1 Introduction

The relationship between international trade and foreign direct investment (FDI), which is one of the main features of globalization, is complex and it is not limited to the issue of whether they are complementary or not (Fontagné, 1999). Previous studies on the effects of FDI on trade are by and large confined to a

*Email: massimo.riccaboni@imtlucca.it, corresponding author.

two-country setting, where bilateral trade is solely determined by the characteristics of the two countries considered. The reasoning behind is that the bilateral trade will be decreased if the bilateral FDI is horizontal[1] or will be increased if the bilateral FDI is vertical[2] (Markusen, 1997, 2004; Markusen and Maskus, 2003).

However, the empirical tests based on the two-country setting have concluded with mixed results. For example, relying on a cross-sectional firm survey data set from the 1970s in the United States, Lipsey and Weiss (1984) find that a US firm's outward FDI to a foreign area is positively associated with its exports to that foreign area. Based on a panel of China's bilateral data with 19 foreign areas during 1984-1998, Liu et al. (2001) also show that the inward FDI from a foreign area to China induces China's exports to that foreign area. Conversely, Belderbos and Sleuwaegen (1998) find a negative relationship between Japanese electronics firms' exports to Europe and their investment in Europe in the late 1980s when Europe adopted a strict antidumping policy. They also find that more trade is created if the investment is related to global value chains (GVCs). Moreover, Blonigen (2001) finds a mixed relationship between affiliate production and exports of Japanese automobile products in the United States from the late 1970s to the early 1990s. Finally, Amiti and Wakelin (2003) run a gravity model for every year in a panel of bilateral data between 36 countries during 1986-1994 and find that FDI has a positive (or negative) effect on trade when countries are different (or similar) in terms of factor endowments and when trade costs are low (or high).

Our paper investigates the effects of FDI on trade beyond the two-country setting. Before examining its effects on bilateral trade, we quantify FDI from a network perspective and consider both the direct and indirect channels of FDI. There are a number of reasons why a network perspective is needed for understanding FDI.

First, besides the simple dichotomy of horizontal and vertical FDI, the mixed nature of FDI has been noted in the literature and new terms such as "complex FDI" and "networked FDI" have been created to account for more structured forms of FDI such as export platforms and production networks[3] (Baldwin and Okubo, 2014; Helpman, 2006; Yeaple, 2003).

Second, due to the varying availability of tax and investment treaties across countries, a parent company may locate an affiliate in an intermediate country to control another affiliate in the final destination country so as to receive the most favorable tax and investment treatment in the host countries (Bénassy-Quéré et al., 2005; Gumpert et al.; Hines and Rice, 1994; Van't Riet et al., 2014). For example, as a well-connected country in terms of tax and investment treaties, the

[1] Horizontal FDI refers to a replicate of the same-stage (usually the final stage) production serving the local economy and hence it substitutes trade because there is no need to import the same goods from the source country anymore.

[2] Vertical FDI refers to different stages of production being located in different countries and hence it promotes trade because of the intermediate and final products flow between the source and the host countries.

[3] A closely related concept is global value chains (GVCs). Open policies and technological advances in the last few decades have led to the fragmentation of the production process across country borders and the building of a global production network (Antràs and Chor, 2013; Gereffi et al., 2005). It is of great economic importance, e.g., in 2010, the global production network accounted for approximately two thirds of both exports and imports of goods for the United States and France (UNCTAD, 2013).

Netherlands is the world's largest pass-through country for approximately 1600 billion euros of FDI in 2009 (Weyzig, 2013). Other famous examples include the FDI "round-tripping" through Hong Kong to China and through Mauritius to India (Wei, 2005).

Third, another driver of indirect FDI is corporate strategies. That is, a parent company may locate an affiliate in a intermediate foreign country if it is geographically, politically, or culturally closer to the final destination country so as to penetrate its market (Kalotay, 2012). For example, Russia indirectly invests in Central and Eastern European countries through Cyprus, taking advantage of the latter's accession in the European Union (Pelto et al., 2004). Note that this driver may, to some extent, overlap with the above driver of tax and investment treaties. The fact that Hong Kong and Mauritius serve as the gateways of inward FDI to China and India respectively is also due to the geographical and cultural closeness. In practice, multinational companies often divide a target region into several clusters of countries and pursue a regional management structure, where regional management centers are established with strategic and operational roles in each cluster (Amann et al., 2014). Moreover, FDI can be a consequence of experience accumulated and can be conceived as a sequential process when crossing multiple country borders (Kogut, 1983).

Our paper highlights the importance of indirect effects and it is therefore related to the recent literature of FDI and trade investigating the third-country effects (Baltagi et al., 2007; Blonigen et al., 2007; Garretsen and Peeters, 2009). However, our paper differs from these studies in at least two aspects. First, while they study the determinants of FDI, we are interested in the consequences of FDI on trade.[4] Second, to capture the third-country effects, they use spatial econometrics whereas we use the tools of network analysis.[5]

To take into account both the direct and indirect channels of FDI, we use a unique data set of international corporate control (as a measure of stock FDI) covering almost all countries over the world, while previous results are often based on a small sample of countries or case studies due to the paucity of FDI data.

Furthermore, we construct a corporate control network where the nodes are the countries and the edges (both directed and weighted) are the corporate control relationships. Based on the corporate control network, the shortest path length (which can be either direct or indirect) and the communicability (which is an overall measure of "communication" between nodes) between each pair of countries are computed. The shortest path length (or the communicability) complements (or substitutes) the direct corporate control intensity and together they provide a more complete accounting of the effects of FDI on trade.

Then we find, using the Heckman two-stage (H2S) gravity model, a positive effect of FDI on trade both for the direct corporate control intensity and the indirect measures, i.e., the shortest path length and the communicability. Therefore our paper provides strong empirical evidence of the importance of indirect effects of FDI on trade. We also identify a number of factors affecting

[4]Even more recently, (Park and Park, 2015) study the effects of FDI on trade by considering the third-country effects. They use spatial econometrics and focus on the inward FDI to China, whereas we use network analysis and a cross-country data set.

[5]In fact, we perform a spatial econometrics exercise with our data set and find a significant third-country effect, which motivates us to study the indirect effects of FDI on trade from a network perspective. The spatial econometrics result is available upon request.

whether FDI and trade are complements or substitutes, including regional trade agreements (RTAs) and the region of Asia. We also find that, compared with primary sectors such as oil extraction and agriculture, manufacturing sectors have more pronounced indirect effects of FDI on trade.

The remainder of this paper is structured as follows: Section 2 introduces FDI and trade networks and describes the network measures of indirect effects; Section 3 describes our data set and provides some exploratory analysis; Section 4 specifies our econometric methodology while Section 5 presents our main results; finally, Section 7 concludes the paper.

2 Networks of FDI and Trade

2.1 Global Systems as Networks

There is a significant body of literature developed recently in studying economic phenomena from a network perspective. For example, economic systems such as international trade (De Benedictis and Tajoli, 2011; Fagiolo et al., 2009; Frankel and Rose, 2002; Garlaschelli and Loffredo, 2005; Glick and Rose, 2002; Reyes et al., 2014; Serrano and Boguná, 2003) and corporate control (Altomonte and Rungi, 2013; Head and Ries, 2008; Vitali et al., 2011) can be considered as networks and their network properties can be used to understand other economic variables (Fagiolo and Mastrorillo, 2014; Ferrier et al., 2016; Riccaboni et al., 2012; Schiavo et al., 2010; Sgrignoli et al., 2015).

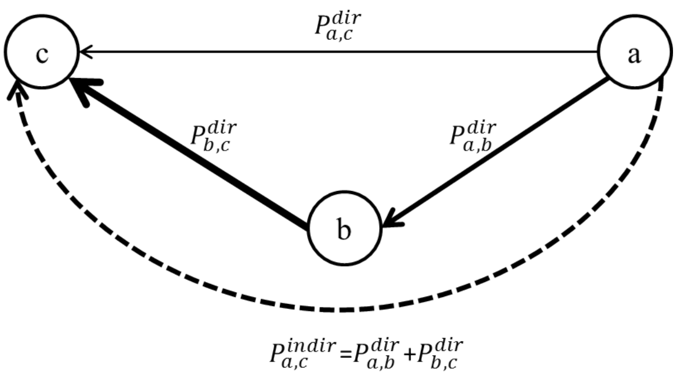
For the world system of either trade or corporate control, we can construct a network, where we identify countries as nodes and interaction channels between them as edges. For simplicity, we call them the world trade web (WTW) and the corporate control network (CCN) respectively. As a result, we have both networks composed of n nodes (countries). The two networks are based on two kinds of weighted directed edges, i.e., two $n \times n$ adjacency matrices, one corresponding to trade flows (T) and the other to corporate control links (C). The generic element of T (or C) represents the value of exports T_{ij} (or the number of corporate control ties C_{ij}) from country i to country j.

2.2 Network Measures of Indirect Effects

As discussed above, factors such as GVCs, tax and investment treaties, and corporate strategies, allow the indirect effects of FDI on trade between countries. To capture the indirect effects of FDI on trade, we use two types of network measures based on the CCN. One is the shortest path length (Brandes, 2001; Newman, 2001; Opsahl et al., 2010). Recall that C_{ij} is the edge weight from country i to country j on the CCN. Define the direct path length from i to j as $P_{ij}^{dir} = \frac{1}{C_{ij}^{\alpha}}$, where $\alpha \geq 0$. The indirect path length from i to j is then computed by adding up the direct path lengths from i to j:

$$P_{ij}^{indir} = P_{ih}^{dir} + \cdots + P_{gj}^{dir} = \frac{1}{C_{ih}^{\alpha}} + \cdots + \frac{1}{C_{gj}^{\alpha}} \tag{1}$$

Figure 1: Direct and indirect paths.



where $\alpha \geq 0$ and $h, \ldots, g$ are the intermediate countries between i and j. We follow (Ferrier et al., 2016) and choose $\alpha = 1$ as our benchmark[6].

Figure 1 shows the difference between the direct and indirect path lengths in a three-country example. Note that the indirect path $P_{a,c}^{indir}$ may be shorter than the direct path $P_{a,c}^{dir}$ according to Equation 1.

Finally, let $\mathbf{P}_{ij}^{indir}$ be the set of all possible indirect paths from i to j, the shortest path length from i to j is defined as:

$$spl_{ij} = \min \left\{ P_{ij}^{dir}, \mathbf{P}_{ij}^{indir} \right\} \tag{2}$$

Note that P_{ij}^{dir} may not exist from i to j and that there may be no paths from i to j at all (in this case $spl_{ij} = \infty$ and we treat it as a missing value).

The other network measure to capture the indirect effects is the communicability, which takes into account not only the shortest path but also all the other walks that connect one node to another (Estrada and Hatano, 2008).

Let A be the adjacency matrix of an undirected and unweighted network and $A_{ij} = A_{ji}$ equals 1 if there is an edge between i and j and otherwise equals 0. A well-known property of A is that the (i, j) entry of the sth power of A, $(A^s)_{ij}$, returns the number of walks of length s starting at i and ending at j. A walk of length s is defined as a sequence of nodes $v_0, v_1, \ldots, v_{s-1}, v_s$ such that, for each $i = 1, 2, \ldots, s$, a link exists from v_{i-1} to v_i. Note that the nodes involved in a walk are not necessarily different from each other (i.e., some nodes may be revisited).

Then the communicability is defined as:

$$cmb_{ij} = \sum_{s=0}^{\infty} \frac{(A^s)_{ij}}{s!} \tag{3}$$

where s is used to discount the number of walks of length s because less importance should be given to longer walks.

The eigenvalues of A in the non-increasing order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are also called the spectrum of the graph. And Equation 3 can be rewritten in terms of

[6]We also discount the importance of indirect links with respect to direct ones by choosing $\alpha = 0.5$ and the main regression results stay the same. See Tables A3 and A4 in the appendix.

the graph spectrum:

$$cmb_{ij} = \sum_{k=1}^{n} \phi_k(i)\phi_k(j)e^{\lambda_k} \tag{4}$$

where $\phi_k(i)$ is the ith element of the kth orthonormal eigenvector of the adjacency matrix associated with the eigenvalue λ_k.

Note that, unlike the shortest path length, the communicability is based on an undirected and unweighted version of the original network, i.e., $cmb_{ij} = cmb_{ji}$. While the shortest path may be either direct or indirect, the communicability takes into account all possible walks (including paths). As a result, it provides a simplified but comprehensive measure of the level of "communication" between nodes. In our econometric analysis below, we use the shortest path length as the main network measure and use the communicability as a robustness check.

3 Data and Descriptive Analysis

3.1 Data

In our empirical analysis we use corporate control data as a measure of stock FDI. Our corporate control data is taken from the ORBIS database compiled by Bureau Van Dijk for the year 2010[7]. We only consider the cross-country ownership relationships and a control is assumed if the parent company holds the voting rights majority (50.01%) of the affiliate in another country. The data thus indicates for each pair of countries the number of control links present between them for both directions. The data we end up with contains 36,461 ultimate parent multinational firms, controlling a total of 354,569 affiliates in 209 countries, for the year 2010. In our data, parent firms located in OECD economies hold around 84% of all the control links, the 63% of which are still in an OECD country. The headquarters located in European Union countries, in particular, control 46% of all affiliates, of which roughly three quarters are located outside the Union.

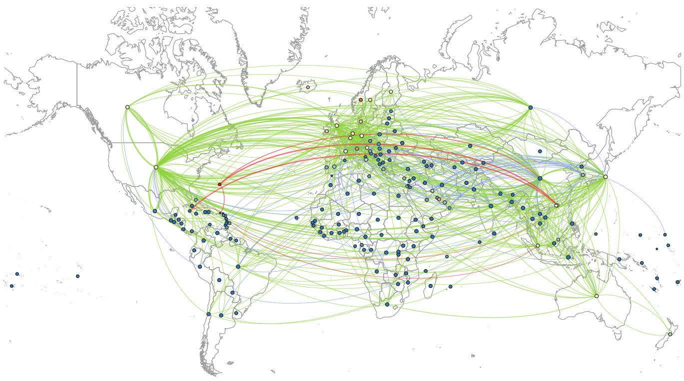
Our trade data is taken from the BACI dataset (Gaulier and Zignago, 2010), which originates from the data reported by over 150 countries to the United Nations Statistics Division (COMTRADE database) but also integrates new approaches to reconcile those reports, in order to have a single consistent figure for each bilateral flow. The version (Harmonized System 1996 or simply HS96) we use covers more than 200 countries and 5000 products, between 1998 and 2012.

Using the two data sets together, we obtain a data set of 191 countries (nodes) along with trade and corporate control relationships among them. Since the only year present in both datasets is 2010, we perform a cross-sectional analysis on that year.

We employ additional country-specific data such as real gross domestic product (GDP) per capita (gdp) and population (pop) from the World Bank. We also use bilateral country geographic, political, and socioeconomic data from the

[7]A related work using the ORBIS database is (Altomonte and Rungi, 2013). Unlike theirs, our data usage is restricted to the cross-country control links.

Figure 2: WTW and CCN in 2010. The figure plots the directed top 2.5% edges by weight. Blue edges represent trade-only relations and red ones represent corporate-control-only, while green ones indicate the presence of both. The edge thickness is proportional to the log of edge weight. Node size is proportional to the log of population (*pop*), while color (from blue to red) is proportional to the log of real GDP per capita (*gdp*).



CEPII GeoDist dataset (Mayer and Zignago, 2011). The latter includes information about between-country geographical distance (*dist*[8]), contiguity (*contig*, i.e., whether two countries share a border), colony relationship (*colony*, i.e., whether one of the two countries has ever been a colony of the other), whether two countries have ever been unified (*smctry*), and ethnical language commonality (*comlang*, i.e., if spoken by at least 9% of population). We use the above variables to implement the gravity model in Section 5.

3.2 Exploratory Analysis

To give an idea of how the trade and FDI networks look like, in Figure 2 we plot the top 2.5%[9] directed edges in terms of edge weight, according to the following criteria: the edge color identifies whether the relationship is solely in trade (blue), solely in corporate control (red), or in both (green); the edge thickness is proportional to the log of edge weight[10]; and the node size and color are proportional, respectively, to the log of population and the log of real GDP per capita.

It is straigtforward to see that most of the significant connections are characterized by both trade and corporate control (green) and the most intensive interaction happens between Europe and the United States. Another interesting finding is that the corporate-control-only connections (red) are primarily associated with "tax havens" such as Bermuda and Cayman Islands.[11] Therefore,

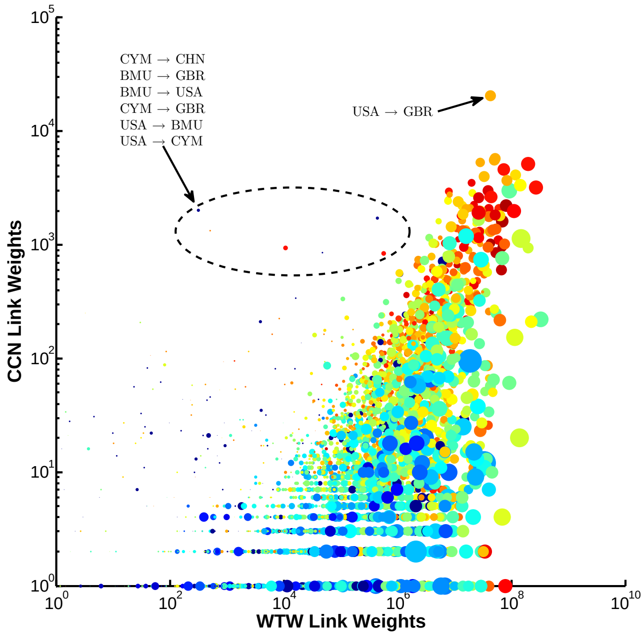
[8]We employ the great-circle definition of country distances. Results do not change if we use alternative distance definitions.

[9]We have also checked the thresholds of 1% and 5% and they convey the same information.

[10]In the cases where both trade and corporate control are present (green edges), the weight is calculated as the mean of the two after normalization.

[11]Other "tax havens" identified by the red links include British Virgin Islands, Hong Kong, and Singapore. Note that there is a link between Bermuda and China and every other red link is between "tax havens."

Figure 3: WTW vs CCN edge weights for the aggregate level. The dot color (blue to red) is proportional to the log of $\frac{gdp_i \times gdp_j}{dist_{ij}}$. The dot size is proportional to the log of $\frac{pop_i \times pop_j}{dist_{ij}}$.



what Figure 2 suggests is that FDI and trade are strongly correlated and that FDI has a preference over "tax havens".

As another exploratory analysis, Figure 3 shows the log-log scatter plots of the WTW and the CCN edge weights. Each dot is an element in the space (T_{ij}, C_{ij}), i.e., the space of the two networks edge weights. The dot color is proportional to the log of $\frac{gdp_i \times gdp_j}{dist_{ij}}$ and the dot size is proportional to the log of $\frac{pop_i \times pop_j}{dist_{ij}}$. The rationale behind this analysis resides in the well-known empirical success of the gravity model for FDI, but especially for trade, i.e., both goods and investments flows are well explained by a gravity-like equation involving country sizes (e.g., *gdp* and *pop*) and geographical distance. In its simplest form, the gravity model of trade prescribes direct proportionality to countries' sizes and inverse proportionality to their geographical distance, i.e.,

$$F_{ij} \propto \frac{s_i s_j}{d_{ij}} , \tag{5}$$

where F_{ij} represents the flow between country i and country j, s_i and s_j represent their respective sizes, and d_{ij} is the geographical distance between the two.

If the gravity rules, one should expect that most of the variation in the cloud of points can be explained by larger country sizes and smaller distances. In our case *gdp* (dot color) plays a more evident role as richer pairs of countries tend to be located in the north-east part of the plot, whereas *pop* (dot size) has less explanatory power. Furthermore, Figure 3 suggests a positive relationship between the edge weights in the two networks, as a high level of exports is in general associated with a high number of corporate control links.

It is also interesting to notice which the outlier edges are. We find again that most of them are "tax havens," where there are intense incoming and outgoing corporate control links and relatively low flows of goods (in Figure 3 we highlight only a few).

4 Econometric Specifications

As stated in Section 1, besides the direct effects of FDI on trade between countries, the indirect effects of FDI on trade are possible due to factors such as GVCs, tax and investment treaties, and corporate strategies. Therefore, when explaining trade, we introduce the network measures of the indirect effects. In particular, we either complement the direct corporate control intensity with the shortest path length or substitute it with the communicability. We do not include both the direct corporate control intensity and the communicability at the same time because the two variables are highly correlated[12] in our data and would therefore produce biased estimations. Therefore, we only include the communicability in the regressions to account for an overall measure of FDI relationships between countries.

Moreover, what we expect to see is a strong correlation between corporate control links and trade flows in the two possible ways, i.e., both when the two relationships are in the same direction (corporate control export and trade export, e.g., a parent company exports inputs to its foreign affiliates) and when they happen in opposite directions (corporate control export and trade import, e.g., a parent company imports processed inputs from its foreign affiliates). We also intend to test the effects of a group of factors that may affect whether corporate control and trade are substitutes or complements, including regional trade agreements (RTAs) and the region of Asia.

An important feature of today's globalized and integrated economy is the proliferation of regional trade agreements (RTAs). According to the WTO data, the number of RTAs has risen from less than 100 in the early 1990s to over 300 today and more than half of the world trade is governed by at least one RTA (Damuri, 2012; WTO, 2011). While one might be tempted to think that trade is naturally of a global span, it is in fact very regional[13]. Without a widespread harmonization of trade and investment agreements (WTO, 2013), it is reasonable to suspect different interactions between trade and corporate control depending on whether an RTA is in place.

Finally, another important aspect to test is the special case of Asia. As one can observe in Figure 2, the trade relationships are particularly concentrated in Asia. Therefore, we want to test if being in Asia has any significant effects on trade and on the relation between corporate control and trade.

Now we turn to the econometric specifications. In the following analyses

[12] The reason why they are so correlated is that the communicability takes into account both the direct link and the indirect links and assigns the largest weight to the direct link. See Table 2 for the correlation coefficients among the variables we used in the regressions.

[13] The word "regional" here refers to big international economic blocks such as EU, NAFTA, and ASEAN.

we use the Heckman two-step (H2S) (Heckman, 1979; Helpman et al., 2008) gravity equation to model the relation between trade and corporate control (as a measure of stock FDI).

Since the seminal works by (Linnemann, 1966) and (Tinbergen, 1962), the gravity model has been widely used in empirical studies on trade because of its excellent fit with the data (Egger, 2002; Frankel and Rose, 2002). Moreover, a lot of effort has been put into refining it and giving it a consistent economic foundation (Bergstrand, 1985; Van Wincoop and Anderson, 2003). Generalizing Equation 5 above, and applying a log transformation of both sides, we can write the formula in a linear multivariate form:

$$logF_{ij} = \beta_0 + \beta_1 logs_i + \beta_2 logs_j + \beta_3 logd_{ij} + \ldots + \epsilon_{ij} \qquad (6)$$

where ϵ_{ij} is the stochastic residual term, usually assumed to be i.i.d. and $\sim N(0, \sigma^2)$, and "..." indicates the possibility to add more regressors (e.g., country- or edge-specific characteristics) to the model specification.

Empirically, the size variables (s_i and s_j) typically include GDP per capita[14] (*gdp*) and population (*pop*) while the impedance factors (d_{ij}) typically include geographical distance (*dist*). Since we focus on the effects of FDI on trade, besides the traditional variables of a gravity model, we control for both the direct corporate control intensity (CC) and the shortest path length (*spl*).

The two-step sample selection estimators are used to model bilateral trade in the presence of zero flows, as they allow us to remove the effects of the extensive margin of trade in order to correctly estimate the intensive margin effects, in contrast with other biased approaches which calculate coefficients that combine both the extensive and intensive margins (Head and Mayer, 2013). Helpman et al. (2008) also show that traditional estimates are biased and that most of the bias is not due to selection but rather to the omission of the extensive margin. Moreover, (Linders and De Groot, 2006) conclude that censored or truncated regression and replacement of zero flows with arbitrary numbers are not preferable as these approaches may yield misleading results and they rely on ad-hoc assumptions and artificial censoring. Sample selection models, on the other hand, allow zero flows and the size of potential trade to be explained jointly and are proved to be the most reasonable choice.

In particular (Helpman et al., 2008) provide a theoretical framework jointly determining both the set of trading partners and their trade volumes, using the H2S selection model (Heckman, 1979). They develop a trade model in which firms face fixed and variable costs of exporting and productivity varies depending on both firms and destinations. Furthermore, trade channels depend on the profitability. Therefore for any pair of countries there may be no firm productive enough to profitably export. As a result, the model is consistent with zero trade flows in both directions between some countries, as well as positive, though asymmetric, trade flows in both directions between others.

Following this literature we carry out all the econometric analyses in this paper using the H2S, which involves first a probit model to estimate the probability of a positive trade flow between any pair of countries and a second step that estimates the log-linear specification of the gravity equation on the positive-flow observations, with a selection correction.[15]

[14] Unlike in the previous exploratory analysis, where real GDP per capita is used, we use nominal GDP per capita in the regressions as the dependent variable is in current US dollars.

[15] Many alternative models have been introduced in the literature that employ the two-step

5 Results

In our baseline model we consider the logarithm[16] of directed bilateral trade flows ($\ln trade$) as the dependent variable and the logarithm of the number of direct corporate control links ($\ln CC$) and the logarithm of the shortest path length ($\ln spl$), as a proxy of indirect FDI, as the key explanatory variables. In addition, the model includes the logarithm of geographical distance (the great-circle definition, $\ln dist$), and the traditional gravity dummy variables including contiguity (*contig*), colony relations (*colony*), whether they have ever been unified (*smcrtry*), and common language (if spoken by at least 9% of the population, *comlang*). Table 1 shows the summary statistics for the variables used in the regressions. Note that we rescale the communicability (*cmb*) values because of their large magnitude and that 90% of the shortest paths available are indirect[17].

Table 2 reports the correlation coefficients among the variables used in the regressions. Note that $\ln CC$ is correlated with both $\ln spl$ (negatively) and $\ln cmb$ (positively). Therefore, we either control $\ln CC$ and $\ln spl$ at the same time to account for the direct and indirect effects of FDI respectively or, as a robustness check, control $\ln cmb$ only to have a comprehensive measure of the FDI effects.

The results of the baseline models are presented in Table 3. We have three specifications to account for FDI, $\ln CC$ only, $\ln spl$ only, or both $\ln CC$ and $\ln spl$.

To estimate the first step of the H2S selection model, i.e., the probability that a dyad trade relationship exists (extensive margin), we use $\ln dist$[18] as the explanatory variable. As expected, we find that the probability is negatively correlated with the geographical distance between the two countries. In the second step, all the explanatory variables' coefficients have the expected signs and significance, i.e., $\ln gdp$ and $\ln pop$ both have positive and significant coefficients, $\ln dist$ has a negative and significant coefficient, and all the dummies have positive and significant coefficients.[19].

Most importantly, we find that the number of corporate control links, $\ln CC$, has a positive and significant effect in explaining trade and the shortest path length, $\ln spl$, has a negative effect in explaining trade. That is, two countries trade more both if they have more direct corporate control links and if they are closer to each other by an indirect path on the CCN[20].

techniques to address the issue of zero flows. Among them the most common ones are the zero inflated Poisson pseudo maximum likelihood (ZIPPML) and the zero inflated negative binomial maximum likelihood (ZINBML) as well as exponential conditional expectations (ECE), although there has been a long debate in the literature about the appropriateness of different models (Martin and Pham, 2015; Silva et al., 2015).

[16]We use the natural logarithm, although the regressions can be run with other bases such as $\log_{10}$. The only difference is that dummy variables coefficients are reduced approximately to half of the value. But all significance levels are unchanged.

[17]This justifies our use of the shortest path length as a measure of the indirect effects. If most of the shortest paths are rather direct, the variable $\ln spl$ will be redundant since we already have the direct corporate control intensity, $\ln CC$.

[18]As far as we know, a unanimous way to model the extensive margin still does not exist.

[19]Our results are robust to additional controls such as common religion, common colonial ties, and landlocking effects

[20]The shortest path can also be a direct one. But as stated above, 90% of time the shortest paths are indirect.

Note that our choice of $\ln spl$ rather than spl provides an interesting interpretation of the result. If all the shortest paths of the CCN are the direct ones and when $\alpha = 1$, $\ln spl$ would be simply $\ln \frac{1}{CC} = -\ln CC$, according to Equations 1 and 2. As a result, the coefficients estimated for either $\ln CC$ or $\ln spl$ alone would be of the same magnitude but with the opposite sign. However, as stated above, about 90% of the shortest paths are indirect. Therefore, by comparing the magnitudes of the coefficients estimated between $\ln CC$ and $\ln spl$, we learn that the indirect effects are slightly larger than the direct ones.

We consider also the case when corporate control and trade are in different directions. The right panel of Table 3 shows the regression result by replacing $\ln CC$ with $\ln CC\_inv$ (i.e., the trade importer country controls affiliates in the trade exporter country) and $\ln spl$ with $\ln spl\_inv$. With this model specification the absolute values of the estimated coefficients of corporate control ($\ln CC\_inv$) and of the shortest path length ($\ln spl\_inv$) both increase, meaning that the number of inverse corporate control links and the inverse shortest path length have larger effects on trade.

Furthermore, in the first column of both panels of Table 4 we introduce the interactions of $\ln CC$ and $\ln spl$ with $\ln dist$ into our baseline specification. We find that $\ln spl$ flips its sign after the interaction terms are introduced. Both $\ln CC$ and $\ln spl$ have opposite signs with respect to their interaction terms. Note that the indirect effects of FDI on trade increase with distance while the direct effects of FDI on trade decrease with distance. Hence, it is possible to identify the critical value of $dist$ for which $\ln CC$ (or $\ln spl$) starts to affect trade negatively. To do so, we rewrite the model equation as

$$\begin{aligned} \ln trade = \beta_0 + \beta_1 \ln CC + \beta_2 \ln spl + \beta_3 \ln dist \\ + \beta_4 \ln CC \ln dist + \beta_5 \ln spl \ln dist + \ldots \end{aligned} \tag{7}$$

where, for simplicity, "…" indicates other possible factors. The value of $dist$, $\overline{dist}$, we look for is the one that solves $\frac{\partial \ln trade}{\partial \ln CC} = \beta_1 + \beta_4 \ln \overline{dist} = 0$ or $\frac{\partial \ln trade}{\partial \ln spl} = \beta_2 + \beta_5 \ln \overline{dist} = 0$.

For the first column of the left panel of Table 3, $\overline{dist}$ is about 90334.995 kilometers for $\ln CC$ and is 2473.526 kilometers for $\ln spl$. In the regression sample, no country pairs are with distances longer than 90334.995 kilometers and 11.83% of country pairs are with distances shorter than 2473.526 kilometers. Therefore, for the majority of country pairs, trade benifits from corporate control relationships, both directly and indirectly.

We also explore other factors of explaining trade. In the second column of both panels of Table 4 we analyze the influence of regional trade agreements (RTAs). We are interested in knowing if belonging to a common RTA fosters trade and whether it affects the relation between trade and corporate control. As expected, the presence of RTA (rta) has a positive and significant coefficient. However, both $\ln CC$ and $\ln spl$ have opposite signs with respect to their interaction terms with rta, meaning that the presence of RTAs reduces the positive (both direct and indirect) effects of corporate control on trade. This result makes intuitive sense because trade depends less on corporate control links once an RTA, which is composed of explicit arrangements to encourage trade, enters into force. The same result also holds for the inverse direction (the right panel).

Next we shift our attention to Asia. Previous studies have shown that Asia is an active participant of GVCs (Baldwin, 2008; Zhu et al., 2014). Therefore, we consider the 10 ASEAN countries[21] plus China and add a dummy, $ASEAN + China$, identifying them as trade exporters as well as the interactions of $\ln CC$ and $\ln spl$ with $ASEAN + China$. The third column of both panels of Table 4 shows the regression result. $ASEAN + China$ has a postive and significant coefficient in both directions while the only significant interaction term is with $\ln CC$ and when both trade and corporate control are in the same direction, meaning that the 11 Asian countries considered carry out more exports than the average and export more to the countries where they have controlled affiliates.

We also explicitly consider the heterogeneity of the indirect effects across sectors. For instance, compared with primary sectors such as oil extraction and agriculture, manufacturing sectors are more involved with GVCs (Los et al., 2015). Hence we expect more pronounced indirect effects for manufacturing sectors. To test this, we exploit the sectoral information of our data set[22] and decompose it into 6 2-digit NAICS (North American Industry Classification System) sectors.[23] Table 5 shows the result, where we control for both the intercept effects by introducing sector dummies and the slope effects by interacting sector dummies with the number of corporate control links (i.e., $\ln CC$ in the second column) and the shortest path length variable (i.e., $\ln spl$ in the third column), where the benchmark NAICS sector is 21, Mining, Quarrying, and Oil and Gas Extraction.[24] We find that, the manufacturing sectors (31-33) not only have larger intercept effects but also, as expected, have more pronounced indirect effects of FDI on trade, when compared with the primary sectors (11, 21, and 22). For example, the sector-specific coefficient of $\ln spl$ is about zero for 11 (i.e., $-0.058 + 0.058 = 0$) and turns much more negative only for the manufacturing sectors (e.g., $-0.058 - 0.268 = -0.326$ for 31).

6 Robustness Checks

In this section, we perform a number of robustness checks. First of all, there is an endogeneity issue that a mutual causal linkage may exist between FDI and trade.[25] Also, Both trade and FDI can be studied in a gravity model with common determinants such as distance and colonial ties (Kleinert and Toubal, 2010). To address this issue, we specify the following simultaneous equations model[26] and run a 3SLS (three-stage least squares) regression (Mitze et al.,

[21] These are Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand, and Vietnam.

[22] We thank the anonymous reviewer for the suggestion to study the sectoral heterogeneity.

[23] They are: 11, Agriculture, Forestry, Fishing and Hunting; 21, Mining, Quarrying, and Oil and Gas Extraction; 22, Utilities; and 31-33, Manufacturing. The detailed information of how we convert HS96 to NAICS is available upon request.

[24] We also run the regressions separately for every restricted sample for each sector. The result is very similar and is available upon request.

[25] We thank the anonymous reviewer for raising this point.

[26] We remove *colony* from the first equation and remove *smctry* from the second equation to satisfy the identification condition.

2010; Wooldridge, 2015):

$$\begin{cases} \ln trade = \beta_0 + \beta_1 \ln CC + \beta_2 \ln spl + \beta_3 \ln gdp\_o + \beta_4 \ln gdp\_d \\ \qquad + \beta_5 \ln pop\_o + \beta_6 \ln pop\_d + \beta_7 \ln dist \\ \qquad + \beta_8 contig + \beta_9 smctry + \beta_{10} comlang + \epsilon \\ \ln CC = \gamma_0 + \gamma_1 \ln trade + \gamma_2 \ln gdp\_o + \gamma_3 \ln gdp\_d \\ \qquad + \gamma_4 \ln pop\_o + \gamma_5 \ln pop\_d + \gamma_6 \ln dist \\ \qquad + \gamma_7 contig + \gamma_8 colony + \gamma_9 comlang + \upsilon \end{cases} \tag{8}$$

where the endogenous variables are $\ln trade$ and $\ln CC$ and all the other variables are assumed to be exogenous.

The reduced form result[27] is reported in Table A1. Again, the coefficient of $\ln spl$ is significant and negative, which is in line with our previous results.

Another drawback of our previous analysis is that we measure the strength of FDI between countries by counting the number of corporate control links, which may differ from each other in terms of firm and investment sizes. We mitigate this problem by using an alternative data source of stock FDI from the UNCTAD (United Nations Conference on Trade and Development).[28] Unlike our measure, the alternative data source has the exact magnitude[29] of stock FDI in millions of US dollars. We recompute the shortest path length variable based on the UNCTAD data set and report the regression results in Table A2. As before, a negative and significant coefficient of $\ln spl$ is estimated. Note that the UNCTAD data set renders much fewer observations than our previous data set does, which explains why we prefer the number of corporate controls computed on the firm-level ORBIS database as a proxy of stock FDI.

Yet another potential issue with our measure is that aggregating the number of corporate control links across countries may overestimate the strength of indirect control. For example, if some firms in country a control some affiliates in country b and some firms in country b have some affiliates in country c, we assume that there is an indirect control from country a to country c. However, country a's affiliates in country b are not necessarily the same ones controlling the affiliates in country c. Therefore, we need to discount the importance of indirect links with respect to direct ones. To do so, we recompute the shortest path length with $\alpha = 0.5$. As a result, 84.2% of the shortest paths are indirect, as opposed to 90% if $\alpha = 1$. However, the main regression results are still robust with $\alpha = 0.5$ (see Tables A3 and A4 in the appendix).[30]

As another robustness check, we run the regressions by replacing $\ln CC$ and $\ln spl$ with the overall measure of FDI "communication" between countries, the communicability, $\ln cmb$. The main result stays the same and is reported in the appendix (Table A5).

Last but not least, we also estimate the baseline models using the Poisson pseudo maximum likelihood (PPML) and the zero inflated Poisson pseudo maximum likelihood (ZIPPML). Results are reported in the appendix (Table A6)

[27]The original form result is available upon request.

[28]The stock FDI data set is downloaded from UNCTAD's Bilateral FDI Statistics for the year 2010.

[29]Sometimes the numbers reported by the FDI origin and host countries may be different. In these cases, we take the average.

[30]The result of the interaction term between $ASEAN + China$ and $\ln CC$ is not robust with $\alpha = 0.5$ (see Table A4 in the appendix).

and the signs of the estimated coefficients are indeed robust with respect to different estimation methods.

7 Conclusions

In this paper we investigate the effects of FDI on trade from a network perspective. We use a unique data set of international corporate control as a measure of stock FDI. We first construct the networks of trade (WTW) and corporate control (CCN) and find a significant correlation between them. Most importantly, factors such as GVCs, tax and investment treaties, and corporate strategies, allow the indirect effects of FDI on trade between countries.

Within the H2S gravity model, we either complement the direct corporate control intensity with the shortest path length or substitute it with the communicability to have a comprehensive accounting of the effects of FDI on trade. We find that in general corporate control (as a measure of stock FDI) has a positive effect on trade both directly and indirectly. This result is robust with respect to different specifications and estimation strategies, therefore providing strong empirical evidence of the indirect effects of FDI on trade. We also identify a number of factors affecting whether corporate control and trade are complements or substitutes, including regional trade agreements (RTAs) and the region of Asia. Moreover, we find that the indirect effects are more pronounced for manufacturing sectors than for primary sectors such as oil extraction and agriculture.

To extend our work, we may consider the heterogeneity of the corporate control links in the future if more firm-level data become available. Currently we weight the edges of the CCN simply by counting the number of corporate control links between countries. This may be problematic if the links are of very different importance in terms of, for example, economic size. Another potential improvement with more firm-level information is that the real indirect corporate control paths can be traced out if we focus on the same firms in the intermediate countries.

Acknowledgments

The authors acknowledge the funding from the Italian Ministry of Education, University and Research (MIUR) through the National Research Program of Italy (PNR), the CRISIS Lab project, and the project “The global virtual-water network: social, economic, and environmental implications” (FIRB - RBFR12BA3Y). They would like to acknowledge helpful comments from Rene Belderbos, Giorgio Fagiolo, and Enrico De Angelis and from the participants in the ECCS 2014 conference in Lucca (Italy). Our special thanks are due to Armando Rungi (IMT, Lucca) for his insightful suggestions and his generous support of our data preparation. M.R. acknowledges the funding from the Multiplex FP7 project (Foundational Research on MULTIlevel comPLEX networks and systems).

References

Carlo Altomonte and Armando Rungi. Business groups as hierarchies of firms: Determinants of vertical integration and performance. Technical Report 155, European Central Bank, 2013.

Bruno Amann, Jacques Jaussaud, and Johannes Schaaper. Clusters and regional management structures by western mncs in asia: overcoming the distance challenge. *Management International Review*, 54(6):879–906, 2014.

Mary Amiti and Katharine Wakelin. Investment liberalization and international trade. *Journal of International Economics*, 61(1):101–126, 2003.

Pol Antràs and Davin Chor. Organizing the global value chain. *Econometrica*, 81(6):2127–2204, 2013. ISSN 0012-9682. doi: 10.3982/ECTA10813.

Richard Baldwin and Toshihiro Okubo. Networked fdi: Sales and sourcing patterns of japanese foreign affiliates. *The World Economy*, 37(8):1051–1080, 2014.

Richard E Baldwin. The spoke trap: hub and spoke bilateralism in east asia. In B Eichengreen, YC Park, and C Wyplosz, editors, *China, Asia, and the New World Economy*, pages 51–86. Oxford University Press, 2008.

Badi H Baltagi, Peter Egger, and Michael Pfaffermayr. Estimating models of complex fdi: Are there third-country effects? *Journal of Econometrics*, 140 (1):260–281, 2007.

René Belderbos and Leo Sleuwaegen. Tariff jumping dfi and export substitution: Japanese electronics firms in europe. *International Journal of Industrial Organization*, 16(5):601–638, 1998.

Agnès Bénassy-Quéré, Lionel Fontagné, and Amina Lahrèche-Révil. How does fdi react to corporate taxation? *International Tax and Public Finance*, 12 (5):583–603, 2005.

Jeffrey H Bergstrand. The gravity equation in international trade: some microeconomic foundations and empirical evidence. *The review of economics and statistics*, 67(3):474–481, August 1985.

Bruce A Blonigen. In search of substitution between foreign production and exports. *Journal of international economics*, 53(1):81–104, 2001.

Bruce A Blonigen, Ronald B Davies, Glen R Waddell, and Helen T Naughton. Fdi in space: Spatial autoregressive relationships in foreign direct investment. *European Economic Review*, 51(5):1303–1325, 2007.

Ulrik Brandes. A faster algorithm for betweenness centrality*. *Journal of Mathematical Sociology*, 25(2):163–177, 2001.

Yose Rizal Damuri. 21st century regionalism and production sharing practice. Technical report, Graduate Institute Geneva, Geneva, 2012.

Luca De Benedictis and Lucia Tajoli. The world trade network. *The World Economy*, 34(8):1417–1454, 2011.

Peter Egger. An econometric view on the estimation of gravity models and the calculation of trade potentials. *The World Economy*, 25(2):297–312, 2002.

Ernesto Estrada and Naomichi Hatano. Communicability in complex networks. *Physical Review E*, 77(3):036111, 2008.

Giorgio Fagiolo and Marina Mastrorillo. Does human migration affect international trade? a complex-network perspective. *PLoS ONE*, 9(5):e97331, May 2014. ISSN 1932-6203. doi: 10.1371/journal.pone.0097331.

Giorgio Fagiolo, Javier Reyes, and Stefano Schiavo. World-trade web: Topological properties, dynamics, and evolution. *Physical Review E*, 79(3):036115, 2009.

Gary Ferrier, Javier Reyes, and Zhen Zhu. Technology diffusion on the international trade network. *Journal of Public Economic Theory*, 18(2):291–312, 2016.

Lionel Fontagné. Foreign direct investment and international trade: Complements or substitutes? OECD Science, Technology and Industry Working Papers 1999/03, 1999.

Jeffrey Frankel and Andrew Rose. An estimate of the effect of common currencies on trade and income. *The Quarterly Journal of Economics*, 117(2): 437–466, 2002.

Diego Garlaschelli and Maria I Loffredo. Structure and evolution of the world trade network. *Physica A: Statistical Mechanics and its Applications*, 355(1): 138–144, 2005.

Harry Garretsen and Jolanda Peeters. Fdi and the relevance of spatial linkages: do third-country effects matter for dutch fdi? *Review of World Economics*, 145(2):319–338, 2009.

Guillaume Gaulier and Soledad Zignago. BACI: International trade database at the product-level. the 1994-2007 version. Working Papers 2010-23, CEPII, October 2010.

Gary Gereffi, John Humphrey, and Timothy Sturgeon. The governance of global value chains. *Review of International Political Economy*, 12(1):78–104, February 2005. ISSN 0969-2290. doi: 10.1080/09692290500049805.

Reuven Glick and Andrew K Rose. Does a currency union affect trade? the time-series evidence. *European Economic Review*, 46(6):1125–1151, 2002.

Anna Gumpert, James R Hines Jr, and Monika Schnitzer. Multinational firms and tax havens. *Review of Economics and Statistics*, page to appear.

Keith Head and Thierry Mayer. Gravity equations: Workhorse, toolkit, and cookbook. Sciences Po Economics Discussion Papers 2013-02, Sciences Po Departement of Economics, January 2013.

Keith Head and John Ries. Fdi as an outcome of the market for corporate control: Theory and evidence. *Journal of International Economics*, 74(1): 2–20, 2008.

James J Heckman. Sample selection bias as a specification error. *Econometrica*, 47(1):153, January 1979. ISSN 00129682. doi: 10.2307/1912352.

Elhanan Helpman. Trade, fdi, and the organization of firms. *Journal of Economic Literature*, 44(3):589–630, 2006.

Elhanan Helpman, Marc Melitz, and Yona Rubinstein. Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics*, 123(2):441–487, 2008.

James R Hines and Eric M Rice. Fiscal paradise: Foreign tax havens and american business. *The Quarterly Journal of Economics*, 109(1):149–182, 1994.

Kalman Kalotay. Indirect fdi. *The Journal of World Investment & Trade*, 13 (4):542–555, 2012.

Jörn Kleinert and Farid Toubal. Gravity for fdi. *Review of International Economics*, 18(1):1–13, 2010.

Bruce Kogut. Foreign direct investment as a sequential process. In *The Multinational in the 1980s*. MIT Press, 1983.

Gert-Jan M Linders and Henri LF De Groot. Estimation of the gravity equation in the presence of zero flows. Technical report, Tinbergen Institute Discussion Paper, 2006.

Hans Linnemann. *An econometric study of international trade flows*. North-Holland Publishing Company Amsterdam, 1966.

Robert E Lipsey and Merle Yahr Weiss. Foreign production and exports of individual firms. *The Review of Economics and Statistics*, pages 304–308, 1984.

Xiaming Liu, Chengang Wang, and Yingqi Wei. Causal links between foreign direct investment and trade in china. *China Economic Review*, 12(2):190–202, 2001.

Bart Los, Marcel P Timmer, and Gaaitzen J Vries. How global are global value chains? a new approach to measure international fragmentation. *Journal of Regional Science*, 55(1):66–92, 2015.

James R Markusen. Trade versus investment liberalization. Technical report, National Bureau of Economic Research, 1997.

James R Markusen. *Multinational firms and the theory of international trade*. MIT press, 2004.

James R Markusen and Keith E Maskus. General-equilibrium approaches to the multinational enterprise: A review of theory and evidence. *Handbook of International Trade*, page 320, 2003.

Will Martin and Cong S Pham. Estimating the gravity model when zero trade flows are frequent and economically determined. 2015.

Thierry Mayer and Soledad Zignago. Notes on CEPII's distances measures: The GeoDist database. Working paper 25, CEPII, 2011.

Timo Mitze, Björn Alecke, and Gerhard Untiedt. Trade-fdi linkages in a simultaneous equations system of gravity models for german regional data. *International Economics*, 122:121–162, 2010.

Mark EJ Newman. Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Physical review E*, 64(1):016132, 2001.

Tore Opsahl, Filip Agneessens, and John Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3): 245–251, 2010.

Innwon Park and Soonchan Park. Modes of foreign direct investment and patterns of trade: An alternative empirical approach. *The World Economy*, 38 (8):1225–1245, 2015.

Elina Pelto, Peeter Vahtra, and Kari Liuhto. Cyp-rus investment flows to central and eastern europe-russia's direct and indirect investments via cyprus to cee. *Journal of Business Economics and Management*, 5(1):3–13, 2004.

Javier Reyes, Rossitza Wooster, and Stuart Shirrell. Regional trade agreements and the pattern of trade: A networks approach. *The World Economy*, 37(8): 1128–1151, 2014.

Massimo Riccaboni, Alessandro Rossi, and Stefano Schiavo. Global networks of trade and bits. *Journal of Economic Interaction and Coordination*, 8(1): 33–56, August 2012. ISSN 1860-711X. doi: 10.1007/s11403-012-0101-x.

Stefano Schiavo, Javier Reyes, and Giorgio Fagiolo. International trade and financial integration: a weighted network analysis. *Quantitative Finance*, 10 (4):389–399, 2010. doi: 10.1080/14697680902882420.

Angeles M Serrano and Marián Boguná. Topology of the world trade web. *Physical Review E*, 68(1):015101, 2003.

Paolo Sgrignoli, Rodolfo Metulini, Stefano Schiavo, and Massimo Riccaboni. The relation between global migration and trade networks. *Physica A: Statistical Mechanics and its Applications*, 417:245–260, January 2015. doi: 10.1016/j.physa.2014.09.037.

João Silva, MC Santos, Silvana Tenreyro, and Frank Windmeijer. Testing competing models for non-negative data with many zeros. *Journal of Econometric Methods*, 4(1):29–46, 2015.

Jan Tinbergen. *Shaping the World Economy: Suggestions for an International Economic Policy.* The Twentieth Century Fund, New York, 1962.

UNCTAD. World investment report: Global value chains: Investment and trade for development. Technical report, United Nations, 2013.

Eric Van Wincoop and James E Anderson. Gravity with gravitas: a solution to the border puzzle. *American Economic Review*, 93(1):170–192, 2003.

Maarten Van't Riet, Arjan Lejour, et al. Ranking the stars: Network analysis of bilateral tax treaties. Technical report, CPB Netherlands Bureau for Economic Policy Analysis, 2014.

Stefania Vitali, James B Glattfelder, and Stefano Battiston. The network of global corporate control. *PloS one*, 6(10):e25995, 2011.

Wenhui Wei. China and india: Any difference in their fdi performances? *Journal of Asian Economics*, 16(4):719–736, 2005.

Francis Weyzig. Tax treaty shopping: structural determinants of foreign direct investment routed through the netherlands. *International Tax and Public Finance*, 20(6):910–937, 2013.

Jeffrey M Wooldridge. *Introductory econometrics: A modern approach*. Nelson Education, 2015.

WTO. World trade report 2011 - The WTO and preferential trade agreements: From co-existence to coherence. Technical report, World Trade Organization, 2011.

WTO. World trade report 2013 - Factors shaping the future of world trade. Technical report, World Trade Organization, 2013.

Stephen Ross Yeaple. The complex integration strategies of multinationals and cross country dependencies in the structure of foreign direct investment. *Journal of International Economics*, 60(2):293–314, 2003.

Zhen Zhu, Federica Cerina, Alessandro Chessa, Guido Caldarelli, and Massimo Riccaboni. The rise of china in the international trade network: A community core detection approach. *PLoS ONE*, 9(8):e105496, 08 2014.

Table 1: Summary statistics for the variables used in the regressions.

Variable	Description	# of Observations	Mean	Std. Dev.	Min	Max
trade	Value of trade in thousands of current USD	37,442	372,915.9	4,169,913	0	328,000,000
CC	Number of corporate control links	37,442	10.17	160.97	0	20,711
gdp	GDP per capita in thousands of current USD	36,477	11.79	16.92	0	88.41
pop	Population	36,477	36,200,000	135,000,000	20,470	1,340,000,000
contig	1 if the two countries are contiguous	36,672	.0153	.123	0	1
comlang	1 if the two countries have a common ethnical language	36,672	.153	.360	0	1
colony	1 if the two countries have a colonial relation	36,672	.0107	.103	0	1
smctry	1 if the two countries were/are the same country	36,672	.008	.089	0	1
dist	Distance, great circle formula, most important cities/agglomerations	36,672	8,075.14	4,591.68	1.047.89	19,904.45
rta	1 if the two countries have a regional trade agreement in force	35,532	.069	.253	0	1
$ASEAN + China$	1 if the exporter country is an ASEAN country or China	37442	.056701	.2312735	0	1
spl	Shortest path length (weighted, directed) based on the CCN	23,936	.344	.452	.00005	3.834
cmb	Communicability (unweighted, undirected) based on the CCN	37,442	.265	.302	.0004	2.042

Table 2: Matrix of the pairwise correlations among the variables used in the regressions.

	$\ln trade$	$\ln CC$	$\ln CC\_inv$	$\ln gdp$	$\ln pop$	*comlang*	*contig*	*smctry*	*colony*	$\ln dist$	*rta*	$\ln cmd$	$\ln spl$	$\ln spl\_inv$
$\ln CC$	0.547													
$\ln CC\_inv$	0.516	0.570												
$\ln gdp$	0.295	0.445	0.176											
$\ln pop$	0.413	0.174	0.138	-0.228										
comlang	0.017	0.088	0.085	-0.027	-0.065									
contig	0.189	0.174	0.172	0.007	0.046	0.110								
smctry	0.088	0.065	0.066	-0.007	-0.019	0.104	0.311							
colony	0.137	0.195	0.192	0.045	0.042	0.174	0.130	0.059						
$\ln dist$	-0.266	-0.227	-0.224	-0.089	0.089	-0.087	-0.350	-0.275	-0.071					
rta	0.283	0.303	0.302	0.169	-0.061	0.022	0.189	0.141	0.063	-0.555				
$\ln cmd$	0.652	0.594	0.597	0.436	0.218	-0.044	0.044	-0.020	0.086	-0.020	0.216			
$\ln spl$	-0.506	-0.716	-0.434	-0.649	-0.131	-0.004	-0.051	0.000	-0.061	0.019	-0.210	-0.722		
$\ln spl\_inv$	-0.437	-0.433	-0.717	-0.138	-0.079	-0.002	-0.049	-0.004	-0.059	0.018	-0.208	-0.728	0.446	
$ASEAN + China$	0.156	0.000	0.041	-0.122	0.283	-0.037	-0.004	-0.016	-0.023	0.143	-0.078	0.070	-0.100	-0.038

Table 3: Regressions for the baseline models: (1) CC only; (2) spl only; (3) CC and spl. The left panel reports the cases where $trade$ and CC (or spl) are in the same direction. The right panel reports the cases where $trade$ and CC (or spl) are in the opposite directions. Note that the network measure of the indirect effects is spl.

	Baseline (1)	Baseline (2)	Baseline (3)	Baseline (inv) (1)	Baseline (inv) (2)	Baseline (inv) (3)
$\ln trade$ (**Dep. Var.**)						
$\ln CC$	0.230 ***		0.103 ***			
	(0.024)		(0.018)			
$\ln CC\_inv$				0.276 ***		0.127 ***
				(0.023)		(0.026)
$\ln spl$		-0.251 ***	-0.203 ***			
		(0.013)	(0.016)			
$\ln spl\_inv$					-0.306 ***	-0.246 ***
					(0.020)	(0.022)
$\ln gdp\_o$	1.158 ***	0.933 ***	0.924 ***	1.190 ***	1.171 ***	1.160 ***
	(0.018)	(0.017)	(0.017)	(0.016)	(0.018)	(0.017)
$\ln gdp\_d$	0.832 ***	0.817 ***	0.809 ***	0.773 ***	0.604 ***	0.594 ***
	(0.016)	(0.012)	(0.012)	(0.017)	(0.025)	(0.024)
$\ln pop\_o$	1.097 ***	1.059 ***	1.047 ***	1.100 ***	1.110 ***	1.099 ***
	(0.013)	(0.010)	(0.010)	(0.012)	(0.013)	(0.013)
$\ln pop\_d$	0.862 ***	0.870 ***	0.861 ***	0.845 ***	0.831 ***	0.818 ***
	(0.012)	(0.009)	(0.009)	(0.012)	(0.014)	(0.014)
$contig$	0.566 ***	0.429 ***	0.399 ***	0.554 ***	0.417 *	0.378 *
	(0.214)	(0.140)	(0.139)	(0.210)	(0.232)	(0.223)
$colony$	0.536 ***	0.679 ***	0.568 ***	0.501 ***	0.809 ***	0.673 ***
	(0.184)	(0.125)	(0.126)	(0.180)	(0.189)	(0.184)
$smctry$	1.083 ***	0.801 ***	0.816 ***	1.080 **	0.767 **	0.787 **
	(0.295)	(0.207)	(0.205)	(0.289)	(0.349)	(0.335)
$comlang$	0.837 ***	0.810 ***	0.782 ***	0.825 ***	0.835 ***	0.800 ***
	(0.066)	(0.047)	(0.047)	(0.065)	(0.070)	(0.067)
$\ln dist$	-2.115 ***	-1.608 ***	-1.543 ***	-2.084 ***	-2.152 ***	-2.070 ***
	(0.276)	(0.208)	(0.206)	(0.270)	(0.352)	(0.339)
Cons.	-11.694 ***	-13.626 ***	-13.667 ***	-11.659 ***	-11.547 ***	-11.615 ***
	(1.602)	(1.133)	(1.121)	(1.568)	(1.918)	(1.843)
$trade\_dummy$ (**Dep. Var.**)						
$\ln dist$	-0.400 ***	-0.384 ***	-0.384 ***	-0.400 ***	-0.386 ***	-0.386 ***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Cons.	3.792 ***	3.479 ***	3.479 ***	3.792 ***	3.477 ***	3.477 ***
	(0.087)	(0.091)	(0.091)	(0.087)	(0.091)	(0.091)
lambda	4.809 ***	1.720 *	1.578 *	4.708 ***	4.625 ***	4.443 ***
	(1.313)	(0.961)	(0.952)	(1.285)	(1.598)	(1.536)
N	34433	29516	29516	34433	29020	29020

Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Regressions with $dist$, rta, and $ASEAN+China$, with the interactions of CC and spl with $dist$, rta, and $ASEAN+China$. The left panel reports the cases where $trade$ and CC (or spl) are in the same direction. The right panel reports the cases where $trade$ and CC (or spl) are in the opposite directions. Note that the network measure of the indirect effects is spl.

	$dist$	rta	$ASEAN+China$	$dist$ (inv)	rta (inv)	$ASEAN+China$ (inv)
$\ln trade$ (**Dep. Var.**)						
$\ln dist$	-1.899 ***	-1.201 ***	-1.400 ***	-2.460 ***	-1.682 ***	-1.886 ***
	(0.209)	(0.207)	(0.195)	(0.367)	(0.279)	(0.257)
$\ln CC$	0.702 ***	0.167 ***	0.128 ***			
	(0.170)	(0.021)	(0.018)			
$\ln CC\_inv$				0.340	0.198 ***	0.116 ***
				(0.275)	(0.022)	(0.020)
$(\ln dist) \times (\ln CC)$	-0.061 ***					
	(0.020)					
$(\ln dist) \times (\ln CC\_inv)$				-0.018		
				(0.031)		
$\ln spl$	1.580 ***	-0.215 ***	-0.144 ***			
	(0.160)	(0.016)	(0.016)			
$\ln spl\_inv$				1.054***	-.255 ***	-0.243 ***
				(0.255)	(0.018)	(0.017)
$(\ln dist) \times (\ln spl)$	-0.202 ***			1.054 ***		
	(0.018)			(0.256)		
$(\ln dist) \times (\ln spl\_inv)$				-0.147***		
				(0.029)		
rta		1.676***			1.880 ***	
		(0.112)			(0.136)	
$rta \times (\ln CC)$		-0.022				
		(0.043)				
$rta \times (\ln CC\_inv)$					-0.045	
					(0.051)	
$rta \times (\ln spl)$		0.339***				
		(0.042)				
$rta \times (\ln spl\_inv)$					0.096 **	
					(0.049)	
$ASEAN+China$			1.599 ***			1.265 ***
			(0.137)			(0.135)
$(ASEAN+China) \times (\ln spl)$			0.051			
			(0.048)			
$(ASEAN+China) \times (\ln CC)$			0.149 **			
			(0.063)			
$(ASEAN+China) \times (\ln spl\_inv)$						-0.018
						(0.054)
$(ASEAN+China) \times (\ln CC\_inv)$						-0.016
						(0.066)
$\ln gdp\_o$	0.929 ***	0.927 ***	0.983 ***	1.166 ***	1.145 ***	1.169 ***
	(0.017)	(0.017)	(0.017)	(0.018)	(0.014)	(0.013)
$\ln gdp\_d$	0.817 ***	0.796 ***	0.825 ***	0.595 ***	0.593 ***	0.600 ***
	(0.012)	(0.012)	(0.012)	(0.025)	(0.019)	(0.018)
$\ln pop\_o$	1.041 ***	1.046 ***	1.002 ***	1.095 ***	1.093 ***	1.067 ***
	(0.010)	(0.010)	(0.010)	(0.013)	(0.010)	(0.010)
$\ln pop\_d$	0.858 ***	0.856 ***	0.873 ***	0.811 ***	0.815 ***	0.823 ***
	(0.009)	(0.009)	(0.009)	(0.014)	(0.011)	(0.011)
$contig$	0.398 ***	0.424 ***	0.343 **	0.408 *	0.430 **	0.349 **
	(0.136)	(0.135)	(0.131)	(0.231)	(0.175)	(0.169)
$colony$	0.544 ***	0.479 ***	0.604 ***	0.644 ***	0.572 ***	0.714 ***
	(0.125)	(0.125)	(0.123)	(0.189)	(0.142)	(0.140)
$smctry$	0.643 ***	0.554 ***	0.801 **	0.662 *	0.514 **	0.794 ***
	(0.197)	(0.194)	(0.191)	(0.345)	(0.260)	(0.254)
$comlang$	0.704 ***	0.715 ***	0.801 ***	0.730 ***	0.735 ***	0.825 ***
	(0.047)	(0.047)	(0.046)	(0.069)	(0.052)	(0.051)
Cons.	-10.084 ***	-15.926 ***	-13.867 ***	-8.155 ***	-14.236 ***	-12.089 ***
	(1.144)	(1.122)	(1.055)	(2.006)	(1.511)	(1.398)
$trade\_dummy$ (**Dep. Var.**)						
$\ln dist$	-0.384 ***	-0.378 ***	-0.384 ***	-0.386 ***	-0.380 ***	-0.386 ***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Cons.	3.479 ***	3.417 ***	3.479 ***	3.477 ***	3.415 ***	3.477 ***
	(0.091)	(0.091)	(0.091)	(0.091)	(0.092)	(0.091)
lambda	1.099	0.607	0.712	4.568 ***	3.435 ***	3.339 ***
	(0.981)	(0.959)	(0.900)	(1.679)	(1.261)	(1.167)
N	29516	29193	29516	29020	28711	29020

Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Heckman two step model with sector fixed effects and the interactions of sector dummies (denoted by 2-digit numbers) with CC and spl. The benchmark NAICS sector is 21, Mining, Quarrying, and Oil and Gas Extraction.

	CC	spl
$\ln trade$ (**Dep. Var.**)		
$[NAICS]$		
11	1.412 ***	1.588 ***
	(0.037)	(0.061)
22	-0.417 ***	-0.167 **
	(0.040)	(0.065)
31	2.749 ***	2.881 ***
	(0.036)	(0.057)
32	2.724 ***	2.792 ***
	(0.036)	(0.057)
33	2.834 ***	2.948 ***
	(0.036)	(0.057)
$\ln CC$	-0.399 ***	0.040 **
	(0.050)	(0.017)
$\ln spl$		-0.058 **
		(0.023)
$[NAICS \times (\ln CC)]$		
11	-0.214 **	
	(0.099)	
22	0.216 ***	
	(0.072)	
31	0.468 ***	
	(0.059)	
32	0.732 ***	
	(0.054)	
33	0.865 ***	
	(0.052)	
$[NAICS \times (\ln spl)]$		
11		0.058 ***
		(0.021)
22		-0.067 ***
		(0.022)
31		-0.268 ***
		(0.020)
32		-0.413 ***
		(0.020)
33		-0.487 ***
		(0.020)
$\ln gdp\_o$	0.940 ***	0.632 ***
	(0.006)	(0.011)
$\ln gdp\_d$	0.715 ***	0.746 ***
	(0.006)	(0.007)
$\ln pop\_o$	0.946 ***	0.902 ***
	(0.005)	(0.006)
$\ln pop\_d$	0.717 ***	0.753 ***
	(0.005)	(0.005)
contig	1.108 ***	0.774 ***
	(0.057)	(0.065)
colony	0.581 ***	0.716 ***
	(0.058)	(0.061)
smctry	0.856 ***	0.399 ***
	(0.079)	(0.096)
comlang	0.755 ***	0.659 ***
	(0.024)	(0.026)
$\ln dist$	-0.813 ***	-1.038 ***
	(0.018)	(0.016)
Cons.	-17.370 ***	-15.175 ***
	(0.161)	(0.197)
$trade\_dummy$ (**Dep. Var.**)		
$\ln dist$	-0.361 ***	-0.309***
	(0.004)	(0.004)
Cons.	2.846 ***	2.168***
	(0.033)	(0.036)
lambda	-1.218 ***	-0.701***
	(0.063)	(0.048)
N	213031	188935

Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A1: Simultaneous equations model (SEM) with three-stage least-squares regression. The reduced model is reported.

	$\ln trade$ (**Dep. Var.**)	$\ln CC$ (**Dep. Var.**)
$\ln spl$	-0.218 ***	
	(0.013)	
$\ln gdp\_o$	0.962 ***	0.499 ***
	(0.017)	(0.007)
$\ln \_gdp\_d$	0.832 ***	0.270 ***
	(0.012)	(0.006)
$\ln \_pop\_o$	1.067 ***	0.236 ***
	(0.010)	(0.005)
$\ln \_pop\_d$	0.877 ***	0.190 ***
	(0.009)	(0.004)
contig	0.515 ***	0.861 ***
	(0.126)	(0.065)
colony	0.662 ***	0.516 ***
	(0.124)	(0.066)
smctry	0.906 ***	0.029
	(0.185)	(0.097)
comlang	0.821 ***	0.410 ***
	(0.046)	(0.024)
$\ln dist$	-1.237 ***	-0.310 ***
	(0.021)	(0.011)
Cons.	-15.862 ***	-5.228 ***
	(0.303)	(0.146)
N	15870	15870

Endogenous variables: $\ln trade$ and $\ln CC$; Exogenous variables: $\ln spl$, $\ln gdp\_o$, $\ln gdp\_d$, $\ln pop\_o$, $\ln pop\_d$, *contig*, *colony*, *smctry*, *comlang*, and $\ln dist$. Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Poisson pseudo maximum likelihood (PPML) estimates with bilateral stock FDI data from the UNCTAD. Note that we exclude all the negative FDI values before calculating the centrality.

	FDI_stock (1)	FDI_stock (2)	FDI_stock (3)
$\ln trade$ (**Dep. Var.**)			
$\ln FDI\_stock$	0.156 ***		0.140 ***
	(0.000)		(0.000)
$\ln spl$		-0.174 ***	-0.060 ***
		(0.000)	(0.000)
$\ln gdp\_o$	0.307 ***	0.519 ***	0.267 ***
	(0.000)	(0.000)	(0.000)
$\ln gdp\_d$	0.612 ***	0.696 ***	0.589 ***
	(0.000)	(0.000)	(0.000)
$\ln pop\_o$	0.647 ***	0.740 ***	0.626 ***
	(0.000)	(0.000)	(0.000)
$\ln pop\_d$	0.644 ***	0.733 ***	0.632 ***
	(0.000)	(0.000)	(0.000)
contig	0.438 ***	0.372 ***	0.425 ***
	(0.000)	(0.000)	(0.000)
colony	-0.260 ***	-0.142 ***	-0.254 ***
	(0.000)	(0.000)	(0.000)
smctry	0.532 ***	0.633 ***	0.496 ***
	(0.000)	(0.000)	(0.000)
comlang	0.213 ***	0.310 ***	0.218 ***
	(0.000)	(0.000)	(0.000)
$\ln dist$	-0.463 ***	-0.568 ***	-0.463 ***
	(0.000)	(0.000)	(0.000)
Cons.	-7.544 ***	-11.400 ***	-7.205 ***
	(0.000)	(0.000)	(0.000)
N	4810	26220	4810

Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Regressions for the baseline models: (1) CC only; (2) spl only; (3) CC and spl. The left panel reports the cases where $trade$ and CC (or spl) are in the same direction. The right panel reports the cases where $trade$ and CC (or spl) are in the opposite directions. Note that the network measure of the indirect effects is spl, which is computed with $\alpha = 0.5$ in Equation 1.

	Baseline (1)	Baseline (2)	Baseline (3)	Baseline (inv) (1)	Baseline (inv) (2)	Baseline (inv) (3)
$\ln trade$ (**Dep. Var.**)						
$\ln CC$	0.230 ***		0.079 ***			
	(0.024)		(0.018)			
$\ln CC\_inv$				0.276 ***		0.074 ***
				(0.023)		(0.025)
$\ln spl$		-0.563 ***	-0.480 ***			
		(0.028)	(0.034)			
$\ln spl\_inv$					-0.734 ***	-0.655 ***
					(0.039)	(0.045)
$\ln gdp\_o$	1.158 ***	0.900 ***	0.897 ***	1.190 ***	1.138 ***	1.135 ***
	(0.018)	(0.018)	(0.018)	(0.016)	(0.017)	(0.017)
$\ln gdp\_d$	0.832 ***	0.798 ***	0.794 ***	0.773 ***	0.544 ***	0.543 ***
	(0.158)	(0.012)	(0.012)	(0.017)	(0.024)	(0.024)
$\ln pop\_o$	1.097 ***	1.044 ***	1.037 ***	1.100 ***	1.091 ***	1.086 ***
	(0.013)	(0.010)	(0.010)	(0.012)	(0.013)	(0.012)
$\ln pop\_d$	0.862 ***	0.859 ***	0.853 ***	0.845 ***	0.808 ***	0.802 ***
	(0.012)	(0.009)	(0.010)	(0.012)	(0.013)	(0.013)
$contig$	0.566 ***	0.397 ***	0.378 ***	0.554 ***	0.371 *	0.353 *
	(0.214)	(0.138)	(0.138)	(0.210)	(0.214)	(0.210)
$colony$	0.536 ***	0.648 ***	0.568 ***	0.501 ***	0.777 ***	0.701 ***
	(0.184)	(0.124)	(0.126)	(0.180)	(0.174)	(0.173)
$smctry$	1.084 ***	0.807 ***	0.818 ***	1.080 ***	0.771 **	0.782 **
	(0.295)	(0.204)	(0.202)	(0.289)	(0.323)	(0.316)
$comlang$	0.837 ***	0.769 ***	0.753 ***	0.825 ***	0.779 ***	0.764 ***
	(0.066)	(0.047)	(0.047)	(0.065)	(0.065)	(0.064)
$\ln dist$	-2.115 ***	-1.568 ***	-1.526 ***	-2.084 ***	-2.082 ***	-2.038 ***
	(0.276)	(0.205)	(0.204)	(0.270)	(0.326)	(0.320)
Cons.	-11.694 ***	-13.211 ***	-13.304 ***	-11.659 ***	-10.963 ***	-11.064 ***
	(1.602)	(1.116)	(1.110)	(1.568)	(1.774)	(1.740)
$trade\_dummy$ (**Dep. Var.**)						
$\ln dist$	-0.400 ***	-0.384 ***	-0.384 ***	-0.400 ***	-0.386 ***	-0.386 ***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Cons.	3.792 ***	3.479 ***	4.479 ***	3.792 ***	3.478 ***	3.477 ***
	(0.087)	(0.091)	(0.090)	(0.087)	(0.091)	(0.091)
lambda	4.808 ***	1.523	1.444	4.708 ***	4.278 ***	4.195 ***
	(1.313)	(0.948)	(0.944)	(1.285)	(1.478)	(1.450)
N	34433	29516	29516	34433	29020	29020

Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Regressions with $dist$, rta, and $ASEAN + China$, with the interactions of CC and spl with $dist$, rta, and $ASEAN + China$. The left panel reports the cases where $trade$ and CC (or spl) are in the same direction. The right panel reports the cases where $trade$ and CC (or spl) are in the opposite directions. Note that the network measure of the indirect effects is spl, which is computed with $\alpha = 0.5$ in Equation 1.

	$dist$	rta	$ASEAN + China$	$dist$ (inv)	rta (inv)	$ASEAN + China$ (inv)
$\ln trade$ (**Dep. Var.**)						
$\ln dist$	-1.726 ***	-1.168 ***	-1.381 ***	-2.353 ***	-1.651 ***	-1.855 ***
	(0.206)	(0.206)	(0.194)	(0.351)	(0.269)	(0.251)
$\ln CC$	1.027 ***	0.140 ***	0.111 ***			
	(0.183)	(0.021)	(0.019)			
$\ln CC\_inv$				0.376	0.145 ***	0.063 ***
				(0.284)	(0.022)	(0.020)
$(\ln dist) \times (\ln CC)$	-0.101 ***					
	(0.021)					
$(\ln dist) \times (\ln CC\_inv)$				-0.028		
				(0.032)		
$\ln spl$	3.709 ***	-0.508 ***	-0.342 ***			
	(0.340)	(0.035)	(0.035)			
$\ln spl\_inv$				2.070 ***	-0.674 ***	-0.651 ***
				(0.521)	(0.037)	(0.037)
$(\ln dist) \times (\ln spl)$	-0.476 ***					
	(0.038)					
$(\ln dist) \times (\ln spl\_inv)$				-0.310 ***		
				(0.058)		
rta		1.557 ***			1.711 ***	
		(0.096)			(0.115)	
$rta \times (\ln CC)$		0.047				
		(0.047)				
$rta \times (\ln CC\_inv)$					0.008	
					(0.055)	
$rta \times (\ln spl)$		0.814 ***				
		(0.092)				
$rta \times (\ln spl\_inv)$					0.120	
					(0.105)	
$ASEAN + China$			1.437 ***			1.253 ***
			(0.124)			(0.114)
$(ASEAN + China) \times (\ln spl)$			-0.024			
			(0.109)			
$(ASEAN + China) \times (\ln CC)$			0.105			
			(0.067)			
$(ASEAN + China) \times (\ln spl\_inv)$						-0.025
						(0.067)
$(ASEAN + China) \times (\ln CC\_inv)$						-0.066
						(0.114)
$\ln gdp\_o$	0.902 ***	0.902 ***	0.962 ***	1.140 ***	1.122 ***	1.143 ***
	(0.018)	(0.018)	(0.018)	(0.017)	(0.014)	(0.013)
$\ln gdp\_d$	0.804 ***	0.784 ***	0.813 ***	0.542 ***	0.543 ***	0.548 ***
	(0.012)	(0.012)	(0.012)	(0.024)	(0.019)	(0.019)
$\ln pop\_o$	1.030 ***	1.036 ***	(0.995 ***	1.082 ***	1.081 ***	1.054 ***
	(0.010)	(0.010)	(0.010)	(0.012)	(0.010)	(0.010)
$\ln pop\_d$	0.851 ***	0.849 ***	0.866 ***	0.795 ***	0.800 ***	0.807 ***
	(0.009)	(0.009)	(0.009)	(0.014)	(0.011)	(0.011)
$contig$	0.411 ***	0.417 ***	0.331 **	0.419 *	0.423 **	0.324 **
	(0.133)	(0.134)	(0.131)	(0.219)	(0.169)	(0.165)
$colony$	0.555 ***	0.486 ***	0.606 ***	0.678 ***	0.610 ***	0.744 ***
	(0.124)	(0.124)	(0.123)	(0.179)	(0.139)	(0.138)
$smctry$	0.642 ***	0.531 ***	0.802 ***	0.664 **	0.498 **	0.789 ***
	(0.193)	(0.193)	(0.190)	(0.328)	(0.251)	(0.247)
$comlang$	0.670 ***	0.688 ***	0.779 ***	0.692 ***	0.702 ***	0.789 ***
	(0.047)	(0.047)	(0.046)	(0.066)	(0.051)	(0.051)
Cons.	-10.848 ***	-15.675 ***	-13.607 ***	-8.276 ***	-13.711 ***	-11.530 ***
	(1.126)	(1.116)	(1.050)	(1.916)	(1.460)	(1.362)
$trade\_dummy$ (**Dep. Var.**)						
$\ln dist$	-0.384 ***	-0.378 ***	-0.384 ***	-0.386 ***	-0.380 ***	-0.386 ***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Cons.	3.479 ***	3.418 ***	3.479 ***	4.478 ***	3.415 ***	3.478 ***
	(0.091)	(0.091)	(0.091)	(0.091)	(0.092)	(0.091)
lambda	0.637	0.417	0.586	4.341 ***	3.189 ***	3.092 ***
	(0.964)	(0.953)	(0.897)	(1.600)	(1.218)	(1.138)
N	29516	29193	29516	29020	28711	29020

Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Regressions for the baseline models and for the specifications with the interactions of *cmb* with *dist*, *rta*, and $ASEAN + China$. Note that the network measure of the indirect effects is *cmb*.

	Baseline	*dist*	*rta*	$ASEAN + China$
$\ln trade$ (**Dep. Var.**)				
$\ln cmd$	0.852 ***	-1.432 ***	0.839 ***	0.811 ***
	(0.042)	(0.225)	(0.029)	(0.033)
$\ln dist$	-2.234 ***	-1.739 ***	-1.768 ***	-2.064 ***
	(0.284)	(0.246)	(0.205)	(0.223)
$(\ln dist) \times (\ln cmd)$		0.269 ***		
		(0.026)		
rta			0.110	
			(0.082)	
$ASEAN + China$				1.771 ***
				(0.117)
$rta \times (\ln cmd)$			-0.690 ***	
			(0.059)	
$(ASEAN + China) \times (\ln cmd)$				0.372 ***
				(0.080)
$\ln gdp\_o$	0.950 ***	0.960 ***	0.976 ***	0.969 ***
	(0.022)	(0.018)	(0.015)	(0.017)
$\ln gdp\_d$	0.581 ***	0.589 ***	0.605 ***	0.587 ***
	(0.021)	(0.018)	(0.015)	(0.017)
$\ln pop\_o$	0.973 ***	0.971 ***	0.993 ***	0.941 ***
	(0.015)	(0.013)	(0.010)	(0.012)
$\ln pop\_d$	0.726 ***	0.724 ***	0.745 ***	0.732 ***
	(0.015)	(0.013)	(0.010)	(0.012)
contig	0.665 ***	0.754 ***	0.610 ***	0.628 ***
	(0.222)	(0.189)	(0.154)	(0.173)
colony	0.898 ***	0.966 ***	0.889 ***	0.945 ***
	(0.189)	(0.162)	(0.129)	(0.148)
smctry	0.976 ***	0.763 ***	0.557 ***	0.971 ***
	(0.305)	(0.262)	(0.210)	(0.238)
comlang	0.799 ***	0.736 ***	0.749 ***	0.821 ***
	(0.068)	(0.059)	(0.047)	(0.053)
Cons.	-4.511 ***	-8.249 ***	-8.376 ***	-5.077 ***
	(1.694)	(1.483)	(1.220)	(1.324)
$trade\_dummy$ (**Dep. Var.**)				
$\ln dist$	-0.400 ***	-0.400 ***	-0.395 ***	-0.400 ***
	(0.010)	(0.010)	(0.010)	(0.010)
Cons.	3.792 ***	3.792 ***	3.732 ***	3.792 ***
	(0.087)	(0.087)	(0.087)	(0.087)
lambda	4.976 ***	4.252 ***	3.392 ***	3.888 ***
	(1.355)	(1.158)	(0.969)	(1.062)
N	34433	34433	33991	34433

Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Diagnostic checks using the baseline models: (1) CC only; (2) spl only; (3) CC and spl. OLS, Poisson pseudo maximum likelihood (PPML) and zero inflated Poisson pseudo maximum likelihood (ZIPPML) are used. The network measure of the indirect effects is spl.

	OLS (1)	OLS (2)	OLS (3)	PPML (1)	PPML (2)	PPML (3)	ZIPPML (1)	ZIPPML (2)	ZIPPML (3)
$\ln trade$ (**Dep. Var.**)									
$\ln CC$	0.237 ***		0.104 ***	0.086 ***		0.114 ***	0.089 ***		0.119 ***
	(0.016)		(0.018)	(0.000)		(0.000)	(0.000)		(0.000)
$\ln spl$		-0.253 ***	-0.205 ***		-0.079 ***	0.024 ***		-0.078 ***	0.030 ***
		(0.013)	(0.016)		(0.000)	(0.000)		(0.000)	(0.000)
$\ln gdp\_o$	1.160 ***	0.932 ***	0.922 ***	0.616 ***	0.614 ***	0.564 ***	0.601 ***	0.610 ***	0.557 ***
	(0.012)	(0.017)	(0.017)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln gdp\_d$	0.836 ***	0.818 ***	0.809 ***	0.742 ***	0.756 ***	0.732 ***	0.731 ***	0.751 ***	0.724 ***
	(0.011)	(0.012)	(0.012)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln pop\_o$	1.093 ***	1.058 ***	1.046 ***	0.793 ***	0.790 ***	0.768 ***	0.783 ***	0.786 ***	0.762 ***
	(0.009)	(0.010)	(0.010)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln pop\_d$	0.858 ***	0.869 ***	0.860 ***	0.754 ***	0.767 ***	0.741 ***	0.745 ***	0.762 ***	0.735 ***
	(0.008)	(0.009)	(0.009)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
contig	0.789 ***	0.503 ***	0.466 ***	0.428 ***	0.434 ***	0.441 ***	0.435 ***	0.439 ***	0.447 ***
	(0.117)	(0.126)	(0.126)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
colony	0.519 ***	0.677 ***	0.566 ***	-0.164 ***	-0.119 ***	-0.177 ***	-0.170 ***	-0.122 ***	-0.183 ***
	(0.121)	(0.124)	(0.125)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
smctry	1.322 ***	0.899 ***	0.906 ***	0.751 ***	0.683 ***	0.714 ***	0.723 ***	0.667 ***	0.697 ***
	(0.160)	(0.185)	(0.185)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
comlang	0.844 ***	0.810 ***	0.782 ***	0.249 ***	0.303 ***	0.240 ***	0.256 ***	0.309 ***	0.244 ***
	(0.044)	(0.046)	(0.047)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln dist$	-1.123 ***	-1.240 ***	-1.204 ***	-0.567 ***	-0.596 ***	-0.554 ***	-0.565 ***	-0.595 ***	-0.551 ***
	(0.021)	(0.021)	(0.022)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cons.	-17.209 ***	-15.558 ***	-15.440 ***	-11.850 ***	-11.900 ***	-11.060 ***	-11.450 ***	-11.710 ***	-10.800 ***
	(0.272)	(0.303)	(0.304)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$trade\_dummy$ (**Dep. Var.**)									
Cons.							-7.375 ***	-8.654 ***	-8.656 ***
							(0.177)	(0.267)	(0.267)
$\ln dist$							0.752 ***	0.827 ***	0.828 ***
							(0.020)	(0.030)	(0.030)
N	20776	15858	15858	30779	20126	20126	30779	20126	20126

Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.