A Domino Theory of Regionalism Revisited: The Role of Homogeneity/Heterogeneity in Institutions†

Renliang Liu Thanasis Stengos Yiguo Sun

March 16, 2021

Abstract

This paper proposes that country-pairs sharing homogeneity in domestic institutions tend to sign preferential trade agreements (PTAs) interdependently, and such an interdependence effect decreases with institutional differential. Using spatial econometrics, we use democracy and economic freedom as fundamental components of the formulation of the spatial weight ma-trix. The baseline results support our hypothesis from both the panel data over 1996 to 2017 through a probit model and the cross-sectional data through a spatial autoregres-sive probit model. We also employ a compound spatial weight matrix to account for the institutional differential and geographical vicinity simultaneously, the results of which confirm the role of domestic institutions in the evolution of PTAs. In addition, the findings are also robust to the correction of potential endogeneity in the spatial weight matrix and we provide not only novel evidence for the spatial effect of in-stitutional differential on the domino-like spread of PTAs, but also some insights on the relationship between regionalism and multilateralism.

Keywords: Domino theory, institutions, multilateralism, preferential trade agreements, probit model, regionalism, spatial econometrics

JEL Codes: C21, C25, F13, F15, O24

†Contact Information: renliang@uoguelph.ca (R. Liu); tstengos@uoguelph.ca (T. Stengos); yisun@uoguelph.ca (Y. Sun). All the authors are affiliated with Department of Economics and Finance, Gordon S. Lang School of Business and Economics, University of Guelph, Guelph, Ontario N1G 2W1, Canada.

1 Introduction

In the past decades, the world has witnessed a dramatic growth of economic regionalism through the formation of preferential trade agreements (PTAs). Despite the popularity of PTAs at a policy level, there is a long-standing debate in academia over its role in multilateralism. On the one side, some economists see PTAs as stumbling blocks towards global free trade since regionalism may reduce incentives (or increase costs) for those involved countries to pursue further multilateral liberalization and meanwhile reduce global trade and welfare (see, e.g., Krugman (1989), Bhagwati (1991), Frankel, Stein and Wei (1995), Krishna (1998), Schiff and Winters (2003), Limão (2007), and Bhagwati (2008)).[1]

However, on the other side, other economists also suggest that such regionalism could be stepping stones towards global free trade rather than milestones (see, e.g., Summers (1991), Estevadeordal, Freund and Ornelas (2008), and Tabakis and Zanardi (2019)).[2] A seminal work by Baldwin (1993) introduces a domino theory of regionalism, which suggests that PTAs could evolve in a domino-like way, and by which he suggests that regionalism can be a stepping block towards globally freer trade (also see Baldwin (2006)).[3] In particular, the domino theory both describes and explains the phenomenon of further expansion of PTAs, suggesting that existing PTAs can change the political balance of the countries that are initially outsiders and thus make them join existing PTAs or build up

[1] Specifically, Krugman (1989) delivers his concerns that regional trading blocs may undermine the process of multilateralism such as the WTO and reduce world welfare. Bhagwati (1991, 2008) addresses the danger of regionalism as it may both diminish world welfare and slow down the path towards global free trade. Frankel et al. (1995) also claim that the global trading system is in danger of being excessive regionalization. Krishna (1998) suggests that PTA-driven regionalism could reduce domestic incentives for those countries to pursue multilateral liberalization. Schiff and Winters (2003) find that regionalism is more likely to undermine full free trade and that it may increase the chances of trade wars. Focusing on the PTAs with non-trade objectives, Limão (2007) finds that, both theoretically and empirically, deep PTAs cause a stumbling block to global free trade since they increase the costs of multilateral tariff reductions especially for large countries.

[2] See Bagwell, Bown and Staiger (2016) for a comprehensive literature review on the comparison between PTA-led (especially deep PTAs) and WTO-led liberalization, where the authors argue that the WTO is not passé even in the era of PTAs.

[3] In a recent study, Baldwin (2016) documents the historical background and future expectation of the WTO and multilateralism, where he suggests that a two-pillar system is very likely to appear with one as the WTO and the other one as a system of multilateralized megaregionals (with a few large emerging markets probably excluded).

new ones.[4] Due to the lack of appropriate econometric tools, the domino-like evolution of PTAs has not been formally tested in an empirical manner until Egger and Larch (2008) (EL, hereafter) who adopt spatial econometric methods and construct the spatial weight matrix from geographic distance.[5] Their results confirm the interdependence effect between PTAs through geographical vicinity for country-pairs.[6] From a different perspective, Baldwin and Jaimovich (2012) derive a contagion index from a political economy model that basically measures bilateral trade reliance, and find that the self-reinforcing extension of PTAs also depends on such contagion of trade.[7] However, the existing literature in many ways overlooks the effect of homogeneity/heterogeneity in institutions on the domino-like expansion of PTAs.

To fill the gap, this article adopts an institutional view of Baldwin's domino theory by proposing and testing the hypothesis that PTA memberships are interdependent among countries with homogeneous classification of institutions, and such interdependence decreases with their institutional differential. There are two main reasons to shift our focus towards institutions. First, institutions, both economic and political, are well documented as fundamental determinants for economies to be successful, especially in the long run (see, e.g., North (1990), Acemoglu, Johnson and Robinson (2005a), and Acemoglu and Johnson (2005)). Following this argument, a large body of literature suggests that domestic institutions are not only crucial for an autarky but also for an open economy, as a source of comparative advantage (see, e.g., Dollar and Kraay (2003), Nunn

[4]Relevant theories have also been derived. For instance, Yi (1996) considers endogeneous formation of coalitions and different rules of customs union formation (open or unanimous) and thus presents additional theoretical foundation for the domino-like evolution of PTAs (only under 'open regionalism' rule in his paper). Using a model of strategic network formation, Goyal and Joshi (2006) find that bilateral agreements are consistent with the spirit of the WTO since the complete network or global free trade is stable in their model. Saggi and Yildiz (2010) use an equilibrium theory of trade agreements with both the degree and type of trade agreements being endogenous, and find that bilateralism can always yield global free trade.

[5]Early empirical literature on the domino theory of regionalism includes Sapir (2001) who focuses on the experience of Europe without employing spatial econometric methods.

[6]EL also use 'natural' trade flows to construct an alternative spatial weight matrix following the model of Anderson and van Wincoop (2003) to address the impact of trade costs.

[7]Baldwin and Jaimovich (2012) also include the variable of political distance in their right-hand-side control variables as a robustness check.

(2007), Levchenko (2007), Costinot (2009), and Chor (2010)).[8] Consistent with the views of Acemoglu et al. (2005a) and Acemoglu and Johnson (2005), democracy and economic freedom that cover property rights and contracts' protection are further found to be the most predominant institutional aspects that can drive trade flows and policies toward the formation of a preferential trade agreement (see, e.g., Ferrantino (1993), Maskus and Penubarti (1995), Evenett and Hoekman (2005), Endoh (2006), Levchenko (2007), Ma, Qu and Zhang (2010), Feenstra, Hong, Ma and Spencer (2013), and Sheng and Yang (2016) for economic institutions,[9] and Morrow, Siverson and Tabares (1998), Mansfield, Milner and Rosendorff (2000), Mansfield, Milner and Rosendorff (2002), Yu (2010), Liu and Ornelas (2014), Wong and Yu (2015), Davis and Wilf (2017), and Pinar and Stengos (2020) for political institutions).[10]

Second, preferential trade agreements have been evolving substantially in the new millennium since they tend to contain more "WTO-extra" provisions beyond the current WTO framework, which are believed to have very different causes and consequences compared with traditional agreements (see, e.g., Limão (2007), Horn, Mavroidis and Sapir (2010), Vicard (2012), Baier, Bergstrand and Feng (2014a), Orefice and Rocha (2014), Kohl, Brakman and Garretsen (2016), Lechner (2016), Hofmann, Osnago and Ruta (2017), and Mattoo, Mulabdic and Ruta (2017)).[11] The set of new policy areas include competition

[8]See Belloc (2006); Nunn and Trefler (2014) for a literature review on how institutions can work as a source of comparative advantage and take a role in international trade dynamics.

[9]Specifically, Levchenko (2007) finds that contract enforcement, property rights, and shareholder protection are among the most important determinants of trade flows. Both Ferrantino (1993) and Maskus and Penubarti (1995) highlight the role of intellectual property rights protection on international trade. Evenett and Hoekman (2005) focus on the welfare gains from fostering domestic competition and transparency in multilateral trade agreements. Endoh (2006) finds that countries with good quality of governance have more incentive to form PTAs. More generally, Dollar and Kraay (2003) examine the interrelationships among institutions, trade openness, and income, by documenting the joint role of trade and institutions in long-term economic growth. Ma et al. (2010) document the positive relationship between a good legal system and exports among firms that use more customized goods as intermediate inputs, while Feenstra et al. (2013) reach a similar conclusion based on Chinese province-firm data. Sheng and Yang (2016) highlight the role of institutional reform in the expansion of export variety.

[10]In brief, countries with high democratic level are found to cooperate more in the field of international business such as trade flows, policies (both bilateral, regional, and multilateral), and investment.

[11] Specifically, Horn et al. (2010) shed light on the development of the WTO-X provisions and their legally enforcement by particularly focusing on the EU and the US. Mattoo et al. (2017) find that deep trade agreements tend to have more trade creation and less trade diversion than shallow ones and might also

policy, environmental laws, labor market regulation, immigration, intellectual property rights, energy safety, among other non-trade issues. By contrast, those traditional provisions incorporated in the content of the WTO are referred to as "WTO-plus" provisions, such as tariff, customs, export taxes, anti-dumping, among other trade-related issues.[12] The increasing depth of trade agreements requires greater cooperation between trade partners on domestic institutions and policies, including harmonization of regulation and standards, free movement of goods and factors, and as a consequence, imposes greater restrictions on national sovereignty (see, e.g., Vicard (2012)).[13] Overall, the rise of the "WTO-extra" provisions in trade talks makes it more important to align domestic institutions and the formation of PTAs than ever before.

Against the background of institutions-preferential trade agreements nexus, we propose the institutional interdependence hypothesis of PTA memberships. We first document that countries sharing similar domestic institutions tend to have comparable domestic standards and rules that matter to international business, such as intellectual property rights, environmental protection, labour market regulation, and regulation on movement of goods and other factors. As a result, preferential trade agreements signed by those countries having homogeneous institutions naturally cover a great number of overlapping provisions that comply with their domestic standards and rules. Therefore, the pre-

increase trade with non-members. Vicard (2012) uncovers the political causes and consequences of the formation of deep trade agreements, in particular international security. Orefice and Rocha (2014) provide evidence on the two-way link between deep integration through forming PTAs and production networks trade, and find that deep agreements increase trade in production networks between member countries by around 35 percentage points, which can be even higher for the industries that require deeper integration. Furthermore they find that trade in production networks can motivate trade partners to sign deep PTAs. Kohl et al. (2016) shed some light on the impact of heterogeneity of PTAs on trade promotion, by showing that trade-related provisions are trade promoting but those deep provisions beyond the WTO mandate are not. In addition, they also find legal enforcement is key to the effectiveness of PTAs. From a political point of view, Lechner (2016) owes the inclusion of non-trade provisions in PTAs to the domestic battle between different interest groups. Baier et al. (2014a) find that deeper types of trade agreements could have larger partial effects on the intensive and extensive margins of trade than shallower ones.

[12]Back in the 1990s, there was a debate over what trade negotiators should negotiate about. For instance, Krugman (1997) suggested "the demand for harmonization is by and large ill-founded both in economics and in law; realistic political economy requires that we give it some credence, but not too much."

[13]In the context of shallow trade agreements, due to the limited enforcement of trade agreements in practice, PTA members still need to adopt cooperative domestic policies to make it in their own self-interest to pursue the pre-determined PTA goals (Ederington, 2001).

existing PTAs signed by those countries can work as templates for the potential new ones among the same group of countries, which can lower both ex-ante negotiation and ex-post enforcement costs for the potential PTAs among the group. However, such an externality effect of the PTAs among one group (sorted by domestic institutions) does not apply to other groups that differ in their classifications of domestic institutions as their policy fields of focus are also very different. Therefore, the documented institutional interdependence effect of PTA memberships decreases with institutional differential.

In testing the institutional interdependence hypothesis of PTA memberships, this paper also makes an important contribution to the empirical literature on trade agreement determinants, in particular the third-country effect of PTAs, as ignoring such effect may result in biased estimates of other factors including the domino effect driven by geographic vicinity (e.g., Baier and Bergstrand (2004), EL, Chen and Joshi (2010), Baldwin and Jaimovich (2012), Jaimovich (2012), and Baier, Bergstrand and Mariutto (2014b)). Methodologically, we consider both panel and cross-sectional data to address the short-to-medium and long term effects, respectively. For the panel data over 1996 to 2017 that covers 111 countries (or 6,105 country-pairs), we employ a probit model with an additional spatial lag term, and for the cross-sectional data that covers 142 countries (or 10,011 country-pairs), we employ a spatial autoregressive (SAR) probit model by following LeSage and Pace (2009). The spatial weight matrix is calculated from the differences of institutional scores between country-pairs that is distinguished from the literature that either use geographical or economic components. Precisely, we consider two institutional factors in the calculation of the spatial weights that cover levels of democratization and economic freedom, which address the most popular concerns over trade-related domestic institutions in the literature as remarked earlier. The approximate maximum likelihood estimation method we use to estimate the SAR probit model is based on a recent paper by Martinetti and Geniaux (2017), which also differs from the previous studies on PTA formation that mainly rely on Bayesian methods (e.g., EL and Jaimovich (2012)).

Both the results from the cross-sectional data through a SAR probit model and panel data through a probit model provide strong evidence for the institutional interdependence effect of PTAs as the estimate of the institution-based domino effect is both economically and statistically significant. In addition, the estimates of other control variables are highly consistent with the theoretical predictions derived in the literature. To further ensure that the interactive effect driven from institutional differential is independent of that driven from geographical distance, we employ the method of Case, Rosen and Hines Jr (1993) to account for the mixed domino effects in the cross-sectional setup. Using a compound spatial weight matrix generated from both institutional and geographical information, we confirm the importance of the homogeneous/heterogeneous classification of institutions in the domino-like evolution of PTAs, and find that the institutional differential accounts for around sixty percentage points of the total spatial effect while geographi-cal distance plays a relatively less important role. Such finding is also supported in the pooled-panel probit regressions when we include both the spatial lag terms with the spatial weight matrix allowing for the institutional differential and geographical distance as regressors.

Next, it is necessary for us to consider the issue of possible endogeneity both in the institutional spatial weights and explanatory variables, especially in the cross-sectional regressions, since the causality between the formation of PTAs, or international trade in general, and domestic institutions can run two-ways.[14] On one hand, institutional quality can partially determine the formation of preferential trade agreements through either direct or indirect channels such as spatial effects. On the other hand, the institutions of potential members or candidates of PTAs can be changed (usually improved in terms of democracy and market liberalization) during the negotiation of trade agreements (either bilateral, regional, or multilateral) for those countries to meet the accession criteria (e.g.,

[14]In the panel regressions, we take five-year lagged values of institutional scores into consideration, which makes the issue of endogeneity less serious.

the EU and the WTO) (See, for instance, Chow (2003)).[15] Therefore, to address these concerns in the SAR probit model, we construct a new estimation procedure allowing for endogeneity of both the spatial weights and the explanatory variables following Qu and Lee (2015). A small set of Monte Carlo simulations is conducted to assess the finite sample performance of the proposed estimator. We find that our main results findings are not sensitive to the correction of potential endogeneity.

The reminder of the paper is organised as follows. In Section 2 we proceed to formally define the institutional interdependence hypothesis of PTA memberships that we would test in this paper. Section 3 presents the baseline analysis that tests the hypothesis based on both the cross-sectional and panel data. Section 4 extends our analysis by accounting for both mixed domino effects and endogeneity issues in the SAR probit model. In Section 5, we conclude.

2 The Institutional Interdependence Hypothesis of PTA Memberships

As mentioned in the previous section, the traditional wisdom has it that the domino-like evolution of regionalism through PTAs mainly relies on geographical vicinity and, more generally, trade costs induced and proxied by geographical distance. As a result, when it comes to empirical analysis, only physical distance has been used to construct a spatial weight matrix.[16] This approach is reasonable in many scenarios especially for those countries that are identical, or at least similar, to each other with regard to institutional background. However, in the context of trade agreement negotiations among

[15] There is a rich set of literature investigating the institutional impact of international trade. For instance, Acemoglu, Johnson and Robinson (2005b) investigate the role of Atlantic trade on the change of institutions in Europe. Levchenko (2013) documents the positive impact of international trade on the quality of institutions including contract enforcement and property rights. Stefanadis (2010) finds that different types of ruling agents can cause different institutional impact of international trade.

[16] Baldwin and Jaimovich (2012) introduce a contagious index that measures trade reliance between countries to construct the spatial weight matrix in their empirical analysis.

countries having very heterogeneous institutions, in particular political regimes (e.g., democratic level) and economic systems (e.g., property rights, contracts protection, market access and competition), such an interdependence effect should not only depend on the pre-determined geographical characteristics but also, perhaps more importantly on institutional differences.

Under the principles and the relevant chapters of the WTO agreement, specifically GATT Article XXIV, WTO membership countries have relatively high degree of autonomy with respect to signing bilateral or regional PTAs to further lower trade barriers beyond the WTO agreement. In other words, Article XXIV allows the WTO member countries to form customs unions or free trade agreements where two or more countries will decide to completely eliminate all tariffs between each other, still, without eliminating tariffs on goods imported from the rest of the world. Given the freedom of forming PTAs, countries can sign PTAs focusing on different subjects beyond the most standard tariffs to accord with their different stages of development, in particular domestic institutions. Documented by recent studies, there is an increasing number of provisions that are beyond the traditional WTO mandate being added to trade agreements (e.g., provisions of environment, labor, and intellectual property rights). Those trade agreements with a large set of non-trade objectives are referred to as deep trade agreements, proven to have different determinants and consequences compared with the traditional ones (see, e.g., Limão (2007), Horn et al. (2010), Vicard (2012), Baier et al. (2014a), Orefice and Rocha (2014), Kohl et al. (2016), Lechner (2016), Hofmann et al. (2017), Mattoo et al. (2017), and Laget, Osnago, Rocha and Ruta (2020)). Referring to the dataset constructed by Hofmann et al. (2017), Figure 1 plots the depth of PTAs measured by the number of WTO-extra provisions against the year of entry for PTAs, which shows a clear growth trend of PTA depthness since 1990s.[17]

[17]The dataset records the detailed provisions of each PTA signed between 1958 and 2015. The non-trade or deep provisions include anti-corruption, competition policy, environment, labor market, intellectual property rights, energy safety, among many others. In their paper, Hofmann et al. (2017) also argue that preferential trade agreements became deeper over time.

However, both the distribution of the PTA depth and its trend are uneven across different groups of countries sorted by institutional context in terms of democracy and economic freedom. As shown in Table 1, the two indicators measuring the aggregate scores of institutions for country-pairs (i.e., PIVsum and EFWsum) have significantly positive correlations with the depth of PTAs, which is consistent with the findings of Vicard (2012) that institutional determinants including democracy are more relevant to deep trade agreements than to shallow ones.[18] In addition, such correlations are robust to two different measures of PTA depth such as the number and proportion of WTO-extra provisions. In other words, country-pairs with higher degree of democracy and more freedom of economic activities tend to concentrate more on non-trade (or WTO-extra) issues during trade negotiations in order to ensure that the ongoing trade deals are consistent with their domestic laws and values. On the other hand, the non-trade articles raised by countries with high degree of democracy and economic freedom are not equally acceptable by countries with low scores in these institutional aspects. It is partially because those countries lack both domestic laws and enforcement powers to execute the horizontal commitments in trade agreements, and furthermore it might even not be in their best interest to adopt such commitments in the first place as, for instance, low standards of environmental protection may distort competition and thus strengthen export competitiveness (Esty and Geradin, 1997). Overall, the data suggests that country-pairs with higher scores of democracy and economic freedom tend to sign deeper trade agreements on average, and vice versa.[19]

The above stylized facts can lead us one step further to the institution-based domino theory of regionalism. Given that country-pairs with similar institutions tend to share closer views on what to negotiate and include in trade agreements as we have docu-

[18] In Table 1, the two indicators measuring institutional differential between two sides of country-pairs (i.e., PIVdist and EFWdist) have significantly negative correlations with the depth of PTAs, which implies that not only the aggregate level of institutional scores is important to the depth of PTAs, but also the differential of the institutional scores between trading partners.

[19] Ravenhill (2010) makes similar conclusion by studying the evolution of regionalism in East Asia.

mented, pre-existing PTAs can thus be more useful templates for potential PTAs among country-pairs sharing similar institutions by reducing both the ex-ante negotiation and ex-post enforcement costs for potential new ones. For instance, due to the pre-existing trade agreement reached by Canada and the U.S. (i.e., CUSMA), a relatively deep one that covers a large set of topics other than trade-related ones such as intellectual property rights, state-owned enterprises, and anticorruption, the negotiators of the trade agreement between Canada and the European Union (i.e., CETA) could hold more consensus on what to discuss and include in their new trade agreement by taking CUSMA as a template or starting point. Moreover, from the Canadian point of view, the overlaps between CUSMA and CETA make it less costly to set up a new regulatory agency to monitor the enforcement of CETA since the Canadian authorities have already had previous experience in that respect. In contrast, it is very difficult for country-pairs with significant heterogeneity in their institutional environment to reach a consensus on what to negotiate and probably include in their trade agreements and it is common for them to hold opposite opinions on what to talk about. Thus, in the second case, heterogeneity of institutions may limit the impact of pre-existing PTAs on potential new ones.

Based on the above analysis, we formally propose the institutional interdependence hypothesis on the domino-like evolution of regionalism that PTAs can be interdependent based on institutional differences, as PTAs for example, achieved by countries with more liberal regimes can increase the chances of reaching other PTAs among the same group. Similar domino effects also exist among countries with lower degree of democracy and economic freedom, or less liberal regimes in general. We finally summarize the institutional interdependence hypothesis of PTA memberships as follows:

PTA memberships are interdependent based on the homogeneity/heterogeneity in domestic institutions between country-pairs, and such an interdependence effect decreases with institutional differential.

3 Benchmark Specifications and Results

In this section, we start to test the institutional interdependence hypothesis of PTA memberships defined in Section 2. Before anything else, we present some information about the variables and data we will use in the empirical analysis. We next concentrate on the hypothesis testing with cross-sectional data by a SAR probit model, and finally we test the hypothesis with panel data by a standard probit model.

3.1 Variables and data

We include the following explanatory variables in our specifications mainly by following the earlier empirical work on PTA determinants such as Baier and Bergstrand (2004), in addition to a spatial lag term that we would discuss in the next subsection with details:

- NATURAL: the log of the inverse of the great circle distance between the capitals of trade partners.
- REMOTE: the remoteness of a pair of continental trade partners from the rest of the world.
- GDPsum: the total bilateral market size measured by total GDP of trade partners.
- GDPsim: the similarity of trade partners in terms of GDP scale.
- DKL: the absolute difference in GDP per capita between trade partners.
- SQDKL: the square of DKL.
- DROWKL: the relative factor endowment difference between the rest of the world and a given pair of trade partners.
- PIVdist: the absolute differential of democracy between trade partners.
- EFWdist: the absolute differential of economic freedom between trade partners.

With respect to the data sources used in this study, we first take the regional trade agree-ments dataset used by EL which includes all the bilateral and regional trade agreements notified to the WTO from 1950 to 2017. Next, institutional indicators as the key variables of interest are obtained from the PolityIV dataset for the political one, the polity score of democracy (Marshall, Jaggers and Gurr, 2002), and the Fraser Institute dataset for the eco-nomic one, the index of economic freedom (Gwartney, Lawson, Hall and Murphy, 2020). In both indexes, a larger value represents a higher score. Specifically, the polity score cap-tures the regime authority spectrum on a scale ranging from -10 (hereditary monarchy) to $+10$ (consolidated democracy), and the economic freedom is measured in five broad areas which include size of government (e.g., tax, subsidies, and state ownership of as-sets), legal system and property rights (e.g., judicial independence, protection of property rights, and legal enforcement of contracts), sound money (e.g., inflation), freedom to trade internationally (e.g., trade barriers and capital controls) and regulation (e.g., credit, labor, and business regulations). The geographical data including latitude and longitude coor-dinates at country-level is taken from the GeoDist dataset of CEPII (Mayer and Zignago, 2011), based on which we can calculate the geographical distance between country-pairs. The data of other economic variables that might be PTA determinants are retrieved from the WDI dataset of the World Bank. Based on data availability, the balanced panel data we use covers 1996 to 2017 and is in four five-year intervals following the standard way in the literature, with totally 111 countries included. The cross-sectional data is taken from year 2017 and covers a larger sample of 142 countries. It is worth noting that there is a smaller coverage of countries in the panel data analysis than in the cross-sectional analysis because of the data availability of the institutional variables (i.e., democracy and economic freedom). The summary statistics of the key variables is presented in Tables 2 and 3 for the cross-sectional and panel data, respectively, and finally the countries in our sample are listed in the Table A.1 and Table A.2 in the appendix.

3.2 Testing the hypothesis with cross-sectional data: A spatial autoregressive probit model

In this subsection, we proceed to the cross-sectional analysis of the determinants of PTAs, especially testing the domino effect driven by institutional similarity between country-pairs. Given the empirical question, we find that SAR probit models (or SAR binary choice models) are well-fitted as such models deal with observations with spatial or network dependence. For instance, Case (1992), as an early attempt, uses SAR probit models to study the interdependence effect in the adoption of new technology in Indonesia. Calabrese and Elkink (2014) conduct a comprehensive literature review and a Monte Carlo study on the estimation methods in the literature. Billé and Arbia (2019) review the theoretical studies on spatial limited dependent variable models and their applications in health economics. Following the above literature and particularly LeSage and Pace (2009), we set up the spatial autoregressive probit model in vector form as follows:

$$\begin{aligned} &PTA^* = \rho W^I \cdot PTA^* + X\beta + \varepsilon, \\ &PTA = I[PTA^* > 0], \end{aligned} \tag{1}$$

where PTA^* and PTA are $n \times 1$ vectors, 0 is an $n \times 1$ vector of zeros, W^I is an $n \times n$ spatial weight matrix constructed from country-pair institutional differential, X is an $n \times k$ matrix of other explanatory variables. $I[A]$ is an indicator function, taking a value of one if event A occurs and zero otherwise. ρ is the so-called spatial lag parameter and $\rho = 0$ implies the non-existence of the interdependence of PTA memberships, while $\rho > 0$ implies the existence of the interdependence of PTA memberships with its effect decreasing with the institutional differential between country-pairs. Thus, a significantly positive estimate of ρ would verify the hypothesis we raise in Section 2. β is a vector of unknown parameters of other explanatory variables. Finally, ε is an $n \times 1$ vector of errors and n equals the number of country-pairs (or observations). Specifically, each element of the vector PTA^* can be read as the minimum value of utility differential for a pair of trading partners between having or not having a bilateral PTA. Specifically, $PTA_i^* = min(\Delta U_{i_1}, \Delta U_{i_2})$, where ΔU_{i_1} and ΔU_{i_2} denote the utility differences for country i_1 and country i_2 between signing or not signing a bilateral PTA (or joining or not joining a regional PTA), respectively. In the data, PTA_i^* is not directly observable. Instead, we can only observe PTA_i as the output of the indicator function $I[PTA_i^* > 0]$, which equals to 1 if and only if both countries i_1 and i_2 are better off in terms of welfare by having a preferential trade relation between each other. Thus, PTA is a binary vector which takes value 1 if there is a preferential trade relation within a country-pair and takes value 0 otherwise. As discussed earlier, the matrix X includes all the control variables including NATURAL, REMOTE, GDPsum GDPsim, DKL, SQDKL, DROWKL, PIVdist, EFWdist, and also a constant term.

Next, we proceed to set up the institution-based spatial weight matrix W^I, whose elements measure the interdependence effect of PTAs through institutional homogeneity between country-pairs. One notes that a typical spatial weight matrix is row-normalized and has all its diagonal elements being zeros. To be more precise, suppose that one country-pair i consists of two countries i_1 and i_2, and another country-pair j consists of two countries j_1 and j_2. Then the institutional differential calculated by the Euclidean distance formula between i and j is given by:

$$Distance_{ij}^I = \sqrt{(d_i - d_j)^2 + (e_i - e_j)^2}, \tag{2}$$

where d_i and d_j represent the aggregate scores of democracy for country-pairs i and j, respectively; e_i and e_j represent the aggregate scores of economic freedom for country-pairs i and j, respectively.[20] Specifically, $d_i = d_{i_1} + d_{i_2}$ and $e_i = e_{i_1} + e_{i_2}$, where d_{i_1} and

[20] In a robustness check, we take the institutional differential between country-pairs as the average of institutional differences between countries in different country-pairs. Following the above notation, we have $Distance_{i_1,j_1}^I = \sqrt{(d_{i_1} - d_{j_1})^2 + (e_{i_1} - e_{j_1})^2}$, which denotes the institutional difference between countries $i1$ and $j1$. Next, we have $Distance_{ij}^I = (\sum_{c_i} \sum_{c_j} Distance_{c_i,c_j}^I)/4$, where $c_i \in (i_1, i_2)$ denotes a country from the country-pair i, and $c_j \in (j_1, j_2)$ denotes a country from the country-pair j. We confirm that our findings remain qualitatively the same with the alternative form of the spatial weights.

d_{i_2} stand for the individual levels of democracy for countries i_1 and i_2, respectively, and e_{i_1} and e_{i_2} stand for the individual levels of economic freedom for countries i_1 and i_2, respectively. One should note that the two aggregate scores of institutional variables are standardized as z-scores from the original data in order to be more comparable with each other since their original scale ranges are very different. Next, we construct the spatial weights as a negative exponential function of the institutional differential:

$$w_{ij}^{I*} = e^{-Distance_{ij}^{I}}, i \neq j, \tag{3}$$

which denotes the $(i, j)^{th}$ element in the negative exponential distance weight matrix, with all the diagonal entries set to zero.[21] Due to the limitation of memory that a standard personal computer can hold, we replace the cells of the spatial weight matrix that are less than the value of 95th percentile with zero. As a result, there are 5% of the cells in the spatial weight matrix are non-zero.[22] As the final step to obtain the spatial weight matrix ready for regressions, we row-standardize the above negative exponential distance weight matrix as follows:

$$w_{ij}^{I} = w_{ij}^{I\,*} / \sum_{j} w_{ij}^{I\,*}, \tag{4}$$

which denotes the $(i, j)^{th}$ element in the final spatial weight matrix W^{I}, with the sum of each row equal to one.

After formalizing the SAR probit model and, more importantly, the corresponding spatial weight matrix, we move to estimation methods. Due to the shortcomings of Bayesian methods in spatial probit regressions including the issues of accuracy and extremely time consuming computation burden for large samples, we adopt the approximated maximum likelihood estimation method introduced in Martinetti and Geniaux

[21] In a robustness check, we use inverse distance weighting instead of negative exponential distance weighting and find that our findings are insensitive to such change. Due to space limitation, we did not report the corresponding results in the paper, which is available from the authors upon request.

[22] For the same reason, the cells of the spatial weight matrix that are less than the value of 98th percentile are replaced with zero in EL. The choice of the threshold value does not affect our main results.

(2017), where the vector of the regression errors follow a multivariate normal distribution whose parameters depend on the spatial structure of the observations. Thus, our estimation procedure is unlike the previous studies on the interdependence effect of PTAs, in particular with cross-sectional data, that rely on Bayesian methods such as EL and Jaimovich (2012). According to different simulation studies, the adopted approximated maximum likelihood method achieves impressive reduction in computation time and meanwhile performs relatively better than other methods in terms of accuracy.[23] Note that the efficiency estimation is of an utmost important concern for us as we have a fairly large spatial weight matrix of dimension $10,011 \times 10,011$.

In what follows, we begin to discuss the empirical results, as shown in Table 4. Column (1) indicates the theoretical predictions of the sign of each explanatory variable. Next to that, column (2) represents the results from the probit model without spatial effects. Column (4), represents the baseline results from the above SAR probit model, with the institutional differential spatial weight matrix. For the benefits of comparison with EL results, we also include column (3) to show the results examining geography-based domino effect.[24] As presented in columns (2), (3), and (4), all the estimates from our regressions have the signs in agreement with the theoretical predictions. In terms of the spatial effect of PTAs through institutional linkage, we find that $W^I \cdot PTA^*$ in column (4) has a positive estimate of coefficient ρ at the 1% significance level, indicating that the memberships of PTAs depend on each other between county-pairs, and such a dependence effect decreases with the institutional differential between country-pairs, which presents clear evidence verifying the hypothesis proposed in Section 2. In other words, if two country-pairs share very similar domestic institutions in terms of political regime and economic

[23]For instance, see Novkaniza and Djuraidah (2018) for a set of simulation studies using different estimation procedures of spatial probit models. In addition, Wilhelm and de Matos (2013) construct an alternative estimation procedure of spatial probit models based on the Bayesian method of LeSage (2000). Bivand and Piras (2015) present a comprehensive comparison between different implementations of estimation for spatial econometrics.

[24]To avoid confusion, the way of constructing the spatial weight matrix of geographical distance is shown in the next section when we move to the analysis of mixed domino effects which includes both institutional and geographical spatial weighting.

system, their decisions on whether or not to join (or establish) bilateral or regional trade agreements are highly correlated. By contrast, if one country-pair has very liberal institutions with both full democracy and full freedom of economic decisions such as market entry and competition, while the other country-pair has the institutions with relatively more centralized political regime and more government intervention in the economy, their decisions on PTA memberships are not that inter-related.

Regarding the estimates for the two control variables of PIVdist and EFWdist, we find that both estimates have the expected signs, which implies that countries experiencing a smaller institutional differential with regard to democracy and economic freedom do have a better chance to form PTAs with each other. In a robustness check, we also find that the observed estimate of the institution-determined spatial lag term could be significantly larger if we exclude the control variables of PIVdist and EFWdist in the right-hand side of equation. This set of results provides additional evidence for the institutional interdependence hypothesis since the omitted part that countries sharing more identical institutions are more likely to form PTAs can make the estimated coefficient of $W^I \cdot PTA^*$ upward biased. Hence, the inclusion of institutional differential between trading partners does correct the potential upward biasness of the estimate for the spatial lag term.

In the right panel of Table 4, we present the results of EL whose spatial weight matrix W^G is calculated from geographical distance between country-pairs. The comparison between columns (4) and (6) indicates that the estimation results in this paper are in general comparable to those in EL, although the spatial lag terms focused by the two papers are generated from different sources, i.e., institutions and geography. The above argument is also supported by the comparison between columns (3) and (6) when we also take domino effect through geographical vicinity into account. In addition, it is worth mentioning that all the control variables in our regressions have the estimates consistent with the predictions from economic theory, while there are a few exceptions in the results of EL. Overall, we provide clear evidence for the long-term domino-like evolution of regionalism that is associated with institutional homogeneity/heterogeneity between country-pairs based on the cross-sectional data of 2017, and our findings can be supplemented with the ones of EL as both papers provide empirical evidence for the domino theory of regionalism but from different perspectives.

Finally, we want to address that the estimate of the institutional interdependence effect of PTA memberships is also economically meaningful based on the marginal effects shown in Table 6. To be specific, the direct effect measures the average effect of the change in one of those explanatory variables on the left-hand binary variable (i.e., PTA), without taking spatial linkages into account. The indirect effect measures the average effect of the change in one of those explanatory variables on the variable of PTA through spatial linkages, while the total effect stands for the sum of the direct and indirect effects. In general, we find that the direct effects are on a par with the indirect effects for all the variables, which indicates the economic significance of our findings.

3.3 Testing the hypothesis with panel data: A probit model

In addition to the cross-sectional analysis, we also use the standard probit model with a dynamic spatial lag term with a time varying spatial weight matrix as one of the regressors to test the institutional interdependence hypothesis in the panel data. In this case, we use data on membership events for four years between 1996 and 2017 in a manner that $t = 2001, 2006, 2011, 2017$, and $t - 5 = 1996, 2001, 2006, 2012$. The intervals are five years with one exception to make full use of the sample. Adopting the common assumption in the literature that new memberships of PTAs in the present do not affect the probability of new memberships in the past such as five years ago, we have the panel probit model in vector form as follows:[25]

$$
\begin{aligned}
PTA^*{}_t &= \rho W^I_{t-5} \cdot PTA_{t-5} + X_{t-5}\beta + \varepsilon_t, \\
PTA_{t-5} &= I[PTA^*{}_{t-5} > 0],
\end{aligned} \tag{5}
$$

where $PTA^*{}_t$ and PTA_{t-5} are $n \times 1$ vectors, 0 is an $n \times 1$ vector of zeros, W^I_{t-5} is an $n \times n$ spatial weight matrix constructed from institutional differential observed five years ago, X_{t-5} is an $n \times k$ matrix of explanatory variables, ρ is an unknown parameter that tests our hypothesis by measuring the spatial effect, and β is a vector of unknown parameters of other explanatory variables. Finally ε_t is an error term and n stands for the number of country-pairs (or observations in each year). As before, PTA_t is a binary vector which takes value 1 if there is a preferential trade agreement within a country-pair in the year of t and takes value 0 otherwise, and $PTA^*{}_t$ is still an unobservable vector that measures the minimum welfare gains of trading partners. Again, X_{t-5} includes all the control variables that include NATURAL, REMOTE, GDPsum GDPsim, DKL, SQDKL, DROWKL, PIVdist, EFWdist, and a constant term. As defined in the cross-sectional case, the entries of the spatial weight matrix are inversely related to the institutional differential calculated from the indicators of democracy and economic freedom, with all the diagonal elements set to zeros. Particularly in the case of panel data analysis, we calculate the spatial lag term $W^I_{t-5} \cdot PTA_{t-5}$ beforehand and treat it as one of the regressors to account for the spatial interdependence effect in the probit model. The advantages of having a panel data model include simpler estimation and larger size of data. In addition, the regression results drawn from the panel data model in Equation (5) have more implications about the dynamic feature of the evolution of PTAs than the cross-sectional ones which mainly estimate the long-term effect.

We next briefly interpret the estimates drawn from the probit regressions with panel

[25] Similarly, new memberships in the present should not affect institutional indicators in the past, or five years ago. Thus, potential endogeneity driven by the two-way casualty between institutions and PTA memberships is a less serious concern in the analysis of panel data.

data as they are largely consistent with those from the SAR probit regressions with cross-sectional data. To start with, column (1) of Table 5 again shows the theoretical predictions for the signs of each explanatory variable. Column (2) presents the estimation results without the spatial effects. Columns (3) and (4) present the results with the spatial weight matrices of geographical vicinity (included for comparison) and institutional differential, respectively. First of all, we find that all the estimates in columns (2), (3), and (4) have the signs in agreement with the theoretical predictions. As for the effect of institutional interdependence (or domino effect), we observe a highly significant positive estimate for the coefficient of spatial lag term in column (4), which once again provides clear evidence for our hypothesis of the institutional interdependence of PTA memberships. Results in column (3) re-confirms the domino effect through geographical vicinity and such finding is consistent with EL. For robustness, we include both the spatial lag terms generated from institutional differential and geographical distance to ensure that the institution-based domino effect is robust to the inclusion of the geography-based domino effect, and we show the results in column (5). Not surprisingly, the institutional domino effect measured by the coefficient of the institutional spatial lag term $W^I_{t-5} \cdot PTA_{t-5}$ becomes somewhat smaller after the inclusion of the geographical spatial lag term $W^G_{t-5} \cdot PTA_{t-5}$, while remaining still statistically significant. Note that the inclusion of geography-based domino effect in the cross-sectional case is presented in the next section by implementing a compound spatial weight matrix in the SAR probit model. Finally, the estimation results of EL are presented in columns (6) and (7) for comparison only that we do not address much.

4 Spatial Autoregressive Probit Model Extensions

Section 4 contains two subsections. In Section 4.1, we first take a compound spatial weight matrix to account for the mixed domino effects from both the institutional similarity and geographical vicinity in the cross-sectional analysis as the earlier results from the panel data suggest that the two interdependence effects might work together in determining PTA memberships. In section 4.2, we propose a two-step estimation method to allow for the endogeneity of both the institutional spatial weight matrix and other control variables, so that we take the potential two-way causality between institutions and the formation of PTAs into consideration.

4.1 Mixed domino effects

As suggested by the probit regression results drawn from the panel data, it is very likely that the two domino effects, one from the institutional channel and the other one from the geographical channel, can work at the same time. Thus, we revise the SAR probit model, in particular the spatial weight matrix, to allow for the mixed domino effects in the analysis of cross-sectional data.

4.1.1 A compound spatial weight matrix

For the above purpose, we construct a compound spatial weight matrix as follows by following Case et al. (1993):

$$W^C = \alpha W^G + (1-\alpha)W^I, \alpha \in [0,1], \tag{6}$$

where W^I is the spatial weight matrix constructed from institutional components and W^G is the spatial weight matrix constructed from geographical components. We define α as the proportion of spatial weight constructed from geographic distance, ranging from 0 to 1. Thus, the contribution of institutional differential to the compound spatial weighting decreases with α, while that of geographical distance increases with α. In order to generate the spatial weights, we need to calculate the geographical distance between, for instance, country-pairs i and j as follows:

$$Distance_{ij}^{G} = (\sum_{c_i} \sum_{c_j} Distance_{c_i,c_j}^{G})/4 \tag{7}$$

where $c_i \in (i_1, i_2)$ denotes a country from the country-pair i and $c_j \in (j_1, j_2)$ denotes a country from the country-pair j, respectively. The term $Distance_{c_i,c_j}^{G}$ measure the geographical distance between countries c_i and c_j, which is calculated from the latitude and longitude coordinates of their capitals. Thus, the term $\sum_{c_i} \sum_{c_j} Distance_{c_i,c_j}^{G}$ measures the total geographical distance between the members of the two country-pairs i and j, and it is further divided by four to obtain the average distance between the two country-pairs.

Based upon the average geographical distance between country-pairs, a typical component of W^G without row-standardization is calculated from a negative exponential function :

$$w_{i,j}^{G*} = e^{-Distance_{ij}^{G}/500}, i \neq j, \tag{8}$$

which denotes the $(i, j)^{th}$ element in the negative exponential distance weight matrix, with all the diagonal entries set to zeros. Again, we replace the cells of the spatial weight matrix that are less than the value of 95th percentile with zeros to save memory space in estimation. Next, we make row-standarization for the above matrix to obtain the actual elements of the geographical spatial weight matrix W^G:

$$w_{i,j}^{G} = w_{i,j}^{G*} / \sum_{j} w_{i,j}^{G*}, \tag{9}$$

which denotes the $(i, j)^{th}$ element in the final spatial weight matrix W^G, with the sum of each row equal to one. Recall that the institution-based spatial weight matrix W^I is set up in Equations (2)-(4).

4.1.2 Empirical results

To find the optimal choice of α, we compare the goodness of fit measured by log-likelihood with different values of α. In particular, we calculate the log-likelihood associated with α ranging from 0.1 to 0.9 with an equal 0.1 increment, and choose the one that minimizes the negative log-likelihood function. We observe a U-curve relationship between the compound weight α and the negative log-likelihood, and find that the minimum value of the negative log-likelihood is achieved when α equals 0.4. To make our results more accurate, we have the second round of grid search by specifying α from 0.35 to 0.45 with an equal 0.1 increment. The second round of search confirms that the most appropriate α which minimizes the negative log-likelihood equals 0.40. The optimal choice of α indicates that around sixty percents of the mixed domino effects are driven by the differences of domestic institutions with respect to democracy and economic freedom, while the rest is driven by the geographical distance. Table 7 presents the estimation results with α ranging from 0.1 to 0.9 with 0.1 intervals, along with the associated values of negative log-likelihood. Overall, we find that the estimates of the mixed domino effects range from 0.391 to 0.438 as α changes, with most of the control variables having the expected effects with high degree of statistical significance. We then focus on the optimal set of results when α equals 0.4 in column (5), suggested by the grid search. In this case, with all the estimates of control variables consistent with the baseline ones, we find that the estimated coefficient of mixed domino effects is between those of pure domino effect from institutional differential and geographical distance as shown in Table 4, since the present estimated coefficient of mixed domino effects equals 0.410. Indeed, such finding is both expectable and reasonable as the mixed domino effects combine the separate ones.

4.2 Endogeneity issues

This section of the analysis is to deal with the potential endogeneity between the formation of PTAs and domestic institutions, particularly in the cross-sectional case, since there can be inverse causality between the two variables. To be specific, on the one hand, countries are required to have institutional reform in a variety of domestic policies includ-ing enforcement of intellectual property, labor, and environmental standards, in order to be eligible to join certain trade agreements (e.g., the EU and the WTO), and thus PTA memberships might affect domestic institutions. For instance, Chow (2003) discusses the required conditions for China to meet in order to enter the WTO such as adopting more liberal policy on market access for foreign companies. On the other hand, as remarked earlier in Section 1, institutions can play an important role in the formation of PTAs. By addressing the potential endogeneity issues, we can further ensure that our findings are robust.

4.2.1 An estimation procedure allowing for endogeneity

Following Qu and Lee (2015), we set up a new estimation procedure to deal with the endogeneity issues in the following model in vector form:

$$\begin{aligned} PTA^* &= \rho W^I \cdot PTA^* + X_1\beta + \varepsilon, \\ PTA &= I[PTA^* > 0], \end{aligned} \tag{10}$$

where the notation is generally consistent with that in Section 3. Specifically, $X_1 = [X_{11}, X_{12}]$, where X_{12} is endogenous that contains the institutional control variables (i.e., PIVdist, EFWdist) and X_{11} is exogenous that contains other control variables. The institutional spatial weight matrix W^I has its typical element $w_{ij}^I = h(z_{ij})$, where z_{ij} is the institutional differential between country-pairs calculated from the country-pair aggregate scores of democracy and economic freedom, following Equation (2), and h represents the negative exponential function. We let Z denote the two aggregate indicators of democracy and economic freedom that generate z_{ij} and we further assume that Z is endogeneous. Thus, the spatial weight matrix W^I is allowed to be endogeneous in our model because of the endogeneity of Z. Next, we generate an auxiliary model as follows:

$$X_{12} = X_2\Gamma_1 + u_1, \tag{11}$$

and

$$Z = X_3\Gamma_2 + u_2, \tag{12}$$

where X_2 and X_3 are exogenous variables, and X_{11}, X_2 and X_3 are allowed to share common variables. In other words, X_{11}, X_2, and X_3 are exogenous to u_1, u_2, and ε. Next, we make some regularity assumption about the distributions of the error terms u_1, u_2, and ε as follows:

$$\begin{pmatrix} u_{1i} \\ u_{2i} \\ \varepsilon_i \end{pmatrix} \overset{i.i.d.}{\sim} N(0, \Sigma), \tag{13}$$

where

$$\Sigma = \begin{pmatrix} \Sigma_{u_1}^2 & \Sigma'_{u_1u_2} & \Sigma'_{\varepsilon u_1} \\ \Sigma_{u_1u_2} & \Sigma_{u_2}^2 & \Sigma'_{\varepsilon u_2} \\ \Sigma_{\varepsilon u_1} & \Sigma_{\varepsilon u_2} & \sigma_\varepsilon^2 \end{pmatrix}. \tag{14}$$

Next, we let $\xi = \varepsilon - u_1\delta_1 - u_2\delta_2$, $\delta_1 = \Sigma^{-1}\Sigma_{\varepsilon u_1}$, and $\delta_2 = \Sigma^{-1}\Sigma_{\varepsilon u_2}$, and have the following specification:

$$PTA^* = \rho W^I \cdot PTA^* + X_1\beta + (X_{12} - X_2\Gamma_1)\delta_1 + (Z - X_3\Gamma_2)\delta_2 + \xi, \tag{15}$$

where $\sigma_{\xi}^{2} = \sigma_{\varepsilon}^{2} - \Sigma'_{\varepsilon u_1}\Sigma_{u_1}^{2^{-1}}\Sigma_{\varepsilon u_1} - \Sigma'_{\varepsilon u_2}\Sigma_{u_2}^{2^{-1}}\Sigma_{\varepsilon u_2}$.

After running the regressions based on Equations (11) and (12) (or first-step regressions), we obtain the estimates of Γ_1 and Γ_2 that are $\widehat{\Gamma}_1$ and $\widehat{\Gamma}_2$, respectively. Next, we replace Γ_1 and Γ_2 by $\widehat{\Gamma}_1$ and $\widehat{\Gamma}_2$, respectively, in Equation (15), and estimate the SAR probit model by approximate maximum likelihood estimation method in the second-step regression (Martinetti and Geniaux, 2017). A set of simulations related to the estimation method is presented.

4.2.2 Monte Carlo simulations

We present the main results of the Monte Carlo simulations that evaluate the finite sample performance of the estimation procedure introduced above that allows for the endogeneity in both the spatial weights and some explanatory variables. In the experiment, the sample size n can vary among 100, 250, 500, and 1,000 to measure the performance of the method in terms of estimation consistency. For each n, we repeat 1,000 Monte Carlo trials. The true values in the data generation process are set as $\beta_0 = 0$, $\beta_1 = 0.3$, $\beta_2 = 0.8$, $\Gamma_1 = 1$, $\Gamma_2 = 1$, and $\rho = 0.3$, without losing generality. Endogeneity is induced in the model by assuming that ε is correlated with u_1 and u_2. Specifically, w e assume $\varepsilon = u_1 + u_2 + u_3$, where u_1, u_2, and u_3 follow the standard normal distribution independently. The spatial weight matrix is calculated as the inverse Euclidean distance of the endogeneous variable Z, with row-standardization and all the diagonals and the entries that are less than the value of the 90th percentile set as zeros.[26] The estimated results given in Table 8 support the consistency of the proposed estimator. The Root Mean Square Er-ror (RMSE) drops quickly toward zero for all the four parameters of interest (β_0, β_1, β_2, ρ) as the sample size increases. In particular, the RMSE for the estimate of the interdepen-dence effect ρ decreases from 0.113 to 0.048 when sample sizes increases from 100 to 1,000. Overall, the results from the above simulations imply that the estimator given by the com-

[26] The simulation results are qualitatively unchanged if we choose other threshold percentiles such as 95th percentile.

structed two-step estimation method is consistent under the presence of endogeneity in both spatial weight matrix and explanatory variables.[27]

4.2.3 Empirical results

Applying the estimation method to our empirical question, we use the ten-year lagged values of institutional variables as instruments, since ten years should be a long enough period for most trade agreement negotiation to complete. For instance, Wong and Yu (2015) find that the average duration for countries to join the WTO is around seven years with a standard deviation equal to five years.[28]

The estimation results are presented in Table 9 with column (1) addressing the theoretical predictions. Column (2) shows the results given by the new estimation method that can deal with the potential endogeneity in both spatial weights and explanatory variables. It is worth noting that since we use the ten-year lagged values of democracy and economic freedom as instrumental variables, we have a smaller sample size than the baseline one. Thus, the results presented by the new estimator might be driven by the sample size. To deal with this issue we also run the baseline SAR probit regression with the same sample to address such concern, as shown in column (3). In general, our results confirm that endogeneity is not a big issue in shaping our findings as the estimates in column (2) are very close to those in column (3). In particular, the estimates of the spatial lag parameter, ρ, are almost the same in the two cases, with 0.294 given by the estimation allowing for endogeneity and 0.292 given by the baseline estimation.

[27] In another experiment, we show that the estimator we use in the baseline regressions is not consistent if there is endogeneity in either spatial weight matrix or explanatory variables.

[28] We also use other instruments such as five-year-lagged values of institutional variables, in which our results remain qualitatively unchanged.

5 Conclusion

Our paper presents an institutional view of the domino theory of regionalism by proposing and testing the institutional interdependence hypothesis of PTA memberships. We find that the homogeneity/heterogeneity of institutions such as democracy and economic freedom constitutes a new channel through which PTA memberships can be interdependent among different countries (or country-pairs).

Our findings have important policy implications for the progress of global trade liberalization, since we highlight the significance of domestic institutional background in the formation of bilateral or regional trade agreements. More importantly, our findings are related to the debate on the role of regionalism in multilateralism since some economists see regionalism through PTAs as stumbling blocks while others see it as building blocks towards multilateralism. According to our findings, the long-term effect of PTAs on global free trade, especially those deep PTAs, to some extent, depends on the homo-geneity/heterogeneity of domestic institutions including, in particular, democracy and economic freedom. More specifically, the domino effect of regionalism documented by Baldwin (1993) can either lead to multilateralism (or a two-pillar system as proposed in Baldwin (2016)) or several large trading blocs that share similar domestic institutions within each bloc, with the actual equilibrium determined by the global distribution of institutions. We leave the empirical investigation of the welfare ramifications of our study for future research.

References

Acemoglu, Daron and Simon Johnson, "Unbundling institutions," *Journal of Political Economy*, 2005, *113* (5), 949–995.

_ , _ , and James A Robinson, "Institutions as a fundamental cause of long-run growth," *Handbook of Economic Growth*, 2005, *1*, 385–472.

_ , _ , and _ , "The rise of Europe: Atlantic trade, institutional change, and economic growth," *American Economic Review*, 2005, *95* (3), 546–579.

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

Bagwell, Kyle, Chad P Bown, and Robert W Staiger, "Is the WTO passé?," *Journal of Economic Literature*, 2016, *54* (4), 1125–1231.

Baier, Scott L and Jeffrey H Bergstrand, "Economic determinants of free trade agreements," *Journal of International Economics*, 2004, *64* (1), 29–63.

_ , _ , and Michael Feng, "Economic integration agreements and the margins of international trade," *Journal of International Economics*, 2014, *93* (2), 339–350.

_ , _ , and Ronald Mariutto, "Economic determinants of free trade agreements revisited: Distinguishing sources of interdependence," *Review of International Economics*, 2014, *22* (1), 31–58.

Baldwin, Richard, "A domino theory of regionalism," *NBER Working Paper No.4465*, 1993.

_ , "Multilateralising regionalism: spaghetti bowls as building blocs on the path to global free trade," *The World Economy*, 2006, *29* (11), 1451–1518.

_ , "The World Trade Organization and the future of multilateralism," *Journal of Economic Perspectives*, 2016, *30* (1), 95–116.

_ and Dany Jaimovich, "Are free trade agreements contagious?," *Journal of International Economics*, 2012, *88* (1), 1–16.

Belloc, Marianna, "Institutions and international trade: A reconsideration of comparative advantage," *Journal of Economic Surveys*, 2006, *20* (1), 3–26.

Bhagwati, Jagdish, *The world trading system at risk*, Princeton University Press, 1991.

_ , *Termites in the trading system: How preferential agreements undermine free trade*, Oxford University Press, 2008.

Billé, Anna Gloria and Giuseppe Arbia, "Spatial limited dependent variable models: A review focused on specification, estimation, and health economics applications," *Journal of Economic Surveys*, 2019, *33* (5), 1531–1554.

Bivand, Roger and Gianfranco Piras, "Comparing implementations of estimation methods for spatial econometrics," *Journal of Statistical Software*, 2015, *63* (18).

Calabrese, Raffaella and Johan A Elkink, "Estimators of binary spatial autoregressive models: A Monte Carlo study," *Journal of Regional Science*, 2014, *54* (4), 664–687.

Case, Anne, "Neighborhood influence and technological change," *Regional science and urban economics*, 1992, *22* (3), 491–508.

Case, Anne C, Harvey S Rosen, and James R Hines Jr, "Budget spillovers and fiscal policy interdependence: Evidence from the states," *Journal of Public Economics*, 1993, *52* (3), 285–307.

Chen, Maggie Xiaoyang and Sumit Joshi, "Third-country effects on the formation of free trade agreements," *Journal of International Economics*, 2010, *82* (2), 238–248.

Chor, Davin, "Unpacking sources of comparative advantage: A quantitative approach," *Journal of International Economics*, 2010, *82* (2), 152–167.

Chow, Gregory C, "Impact of joining the WTO on China's economic, legal and political institutions," *Pacific Economic Review*, 2003, *8* (2), 105–115.

Costinot, Arnaud, "On the origins of comparative advantage," *Journal of International Economics*, 2009, *77* (2), 255–264.

Davis, Christina L and Meredith Wilf, "Joining the Club: Accession to the GATT/WTO," *The Journal of Politics*, 2017, *79* (3), 964–978.

Dollar, David and Aart Kraay, "Institutions, trade, and growth," *Journal of Monetary Economics*, 2003, *50* (1), 133–162.

Ederington, Josh, "International coordination of trade and domestic policies," *American Economic Review*, 2001, *91* (5), 1580–1593.

Egger, Peter and Mario Larch, "Interdependent preferential trade agreement memberships: An empirical analysis," *Journal of International Economics*, 2008, *76* (2), 384–399.

Endoh, Masahiro, "Quality of governance and the formation of preferential trade agreements," *Review of International Economics*, 2006, *14* (5), 758–772.

Estevadeordal, Antoni, Caroline Freund, and Emanuel Ornelas, "Does regionalism affect trade liberalization toward nonmembers?," *The Quarterly Journal of Economics*, 2008, *123* (4), 1531–1575.

Esty, Daniel C and Damien Geradin, "Market access, competitiveness, and harmonization: Environmental protection in regional trade agreements," *Harvard Environmental Law Review*, 1997, *21*, 265.

Evenett, Simon J and Bernard Hoekman, "Government procurement: Market access, transparency, and multilateral trade rules," *European Journal of Political Economy*, 2005, *21* (1), 163–183.

Feenstra, Robert C, Chang Hong, Hong Ma, and Barbara J Spencer, "Contractual versus non-contractual trade: The role of institutions in China," *Journal of Economic Behavior & Organization*, 2013, *94*, 281–294.

Ferrantino, Michael J, "The effect of intellectual property rights on international trade and investment," *Review of World Economics*, 1993, *129* (2), 300–331.

Frankel, Jeffrey, Ernesto Stein, and Shang-jin Wei, "Trading blocs and the Americas: The natural, the unnatural, and the super-natural," *Journal of Development Economics*, 1995, *47* (1), 61–95.

Goyal, Sanjeev and Sumit Joshi, "Bilateralism and free trade," *International Economic Review*, 2006, *47* (3), 749–778.

Gwartney, James, Robert Lawson, Joshua Hall, and Ryan Murphy, "Economic freedom dataset," *Economic Freedom of the World: Annual Report*, 2020.

Hofmann, Claudia, Alberto Osnago, and Michele Ruta, "Horizontal depth: A new database on the content of preferential trade agreements," *World Bank Policy Research Working Paper No.7981*, 2017.

Horn, Henrik, Petros C Mavroidis, and André Sapir, "Beyond the WTO? An anatomy of EU and US preferential trade agreements," *The World Economy*, 2010, *33* (11), 1565–1588.

Jaimovich, Dany, "A Bayesian spatial probit estimation of Free Trade Agreement contagion," *Applied Economics Letters*, 2012, *19* (6), 579–583.

Kohl, Tristan, Steven Brakman, and Harry Garretsen, "Do trade agreements stimulate international trade differently? Evidence from 296 trade agreements," *The World Economy*, 2016, *39* (1), 97–131.

Krishna, Pravin, "Regionalism and multilateralism: A political economy approach," *The Quarterly Journal of Economics*, 1998, *113* (1), 227–251.

Krugman, Paul, "Is bilateralism bad?," *NBER Working Paper No.2972*, 1989.

_ , "What should trade negotiators negotiate about?," *Journal of Economic Literature*, 1997, *35* (1), 113–120.

Laget, Edith, Alberto Osnago, Nadia Rocha, and Michele Ruta, "Deep trade agreements and global value chains," *Review of Industrial Organization*, 2020, *57*, 379–410.

Lechner, Lisa, "The domestic battle over the design of non-trade issues in preferential trade agreements," *Review of International Political Economy*, 2016, *23* (5), 840–871.

LeSage, James P, "Bayesian estimation of limited dependent variable spatial autoregressive models," *Geographical Analysis*, 2000, *32* (1), 19–35.

_ **and Robert Kelley Pace**, *Introduction to Spatial Econometrics*, Chapman and Hall/CRC, 2009.

Levchenko, Andrei A, "Institutional quality and international trade," *The Review of Economic Studies*, 2007, *74* (3), 791–819.

_ , "International trade and institutional change," *The Journal of Law, Economics, & Organization*, 2013, *29* (5), 1145–1181.

Limão, Nuno, "Are preferential trade agreements with non-trade objectives a stumbling block for multilateral liberalization?," *The Review of Economic Studies*, 2007, *74* (3), 821–855.

Liu, Xuepeng and Emanuel Ornelas, "Free trade agreements and the consolidation of democracy," *American Economic Journal: Macroeconomics*, 2014, *6* (2), 29–70.

Ma, Yue, Baozhi Qu, and Yifan Zhang, "Judicial quality, contract intensity and trade: Firm-level evidence from developing and transition countries," *Journal of Comparative Economics*, 2010, *38* (2), 146–159.

Mansfield, Edward D, Helen V Milner, and B Peter Rosendorff, "Free to trade: Democracies, autocracies, and international trade," *American Political Science Review*, 2000, *94* (2), 305–321.

_ , _ , and _ , "Why democracies cooperate more: Electoral control and international trade agreements," *International Organization*, 2002, *56* (3), 477–513.

Marshall, Monty G, Keith Jaggers, and Ted Robert Gurr, "Polity IV project: Political regime characteristics and transitions, 1800-2002," *University of Maryland*, 2002.

Martinetti, Davide and Ghislain Geniaux, "Approximate likelihood estimation of spatial probit models," *Regional Science and Urban Economics*, 2017, *64*, 30–45.

Maskus, Keith E and Mohan Penubarti, "How trade-related are intellectual property rights?," *Journal of International Economics*, 1995, *39* (3-4), 227–248.

Mattoo, Aaditya, Alen Mulabdic, and Michele Ruta, "Trade creation and trade diversion in deep agreements," *World Bank Policy Research Working Paper No.8206*, 2017.

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

Morrow, James D, Randolph M Siverson, and Tressa E Tabares, "The political determinants of international trade: The major powers, 1907–1990," *American Political Science Review*, 1998, *92* (3), 649–661.

North, Douglass C, *Institutions, institutional change and economic performance*, Cambridge

University Press, 1990.

Novkaniza, Fevi and Anik Djuraidah, "Comparison method of spatial autoregressive probit estimation," in "8th International Conference on Industrial Engineering and Operations Management, IEOM 2018" IEOM Society 2018, pp. 2016–2017.

Nunn, Nathan, "Relationship-specificity, incomplete contracts, and the pattern of trade," *The Quarterly Journal of Economics*, 2007, *122* (2), 569–600.

— **and Daniel Trefler**, "Domestic institutions as a source of comparative advantage," *Handbook of International Economics*, 2014, *4*, 263–315.

Orefice, Gianluca and Nadia Rocha, "Deep integration and production networks: An empirical analysis," *The World Economy*, 2014, *37* (1), 106–136.

Pinar, Mehmet and Thanasis Stengos, "Democracy in the neighbourhood and foreign direct investment," *Review of Development Economics*, 2020.

Qu, Xi and Lung-Fei Lee, "Estimating a spatial autoregressive model with an endogenous spatial weight matrix," *Journal of Econometrics*, 2015, *184* (2), 209–232.

Ravenhill, John, "The 'new East Asian regionalism': A political domino effect," *Review of International Political Economy*, 2010, *17* (2), 178–208.

Saggi, Kamal and Halis Murat Yildiz, "Bilateralism, multilateralism, and the quest for global free trade," *Journal of International Economics*, 2010, *81* (1), 26–37.

Sapir, Andre, "Domino effects in Western European regional trade, 1960–1992," *European Journal of Political Economy*, 2001, *17* (2), 377–388.

Schiff, Maurice and L Alan Winters, *Regional integration and development*, The World Bank, 2003.

Sheng, Liugang and Dennis Tao Yang, "Expanding export variety: The role of institutional reforms in developing countries," *Journal of Development Economics*, 2016, *118*, 45–58.

Stefanadis, Christodoulos, "Appropriation, property rights institutions, and international trade," *American Economic Journal: Economic Policy*, 2010, *2* (4), 148–72.

Summers, Lawrence, “Regionalism and the world trading system,” in “Policy implications of trade and currency zones,” Federal Reserve Bank of Kansas City Kansas City, 1991, pp. 295–301.

Tabakis, Chrysostomos and Maurizio Zanardi, “Preferential trade agreements and antidumping protection,” *Journal of International Economics*, 2019, *121*, 103246.

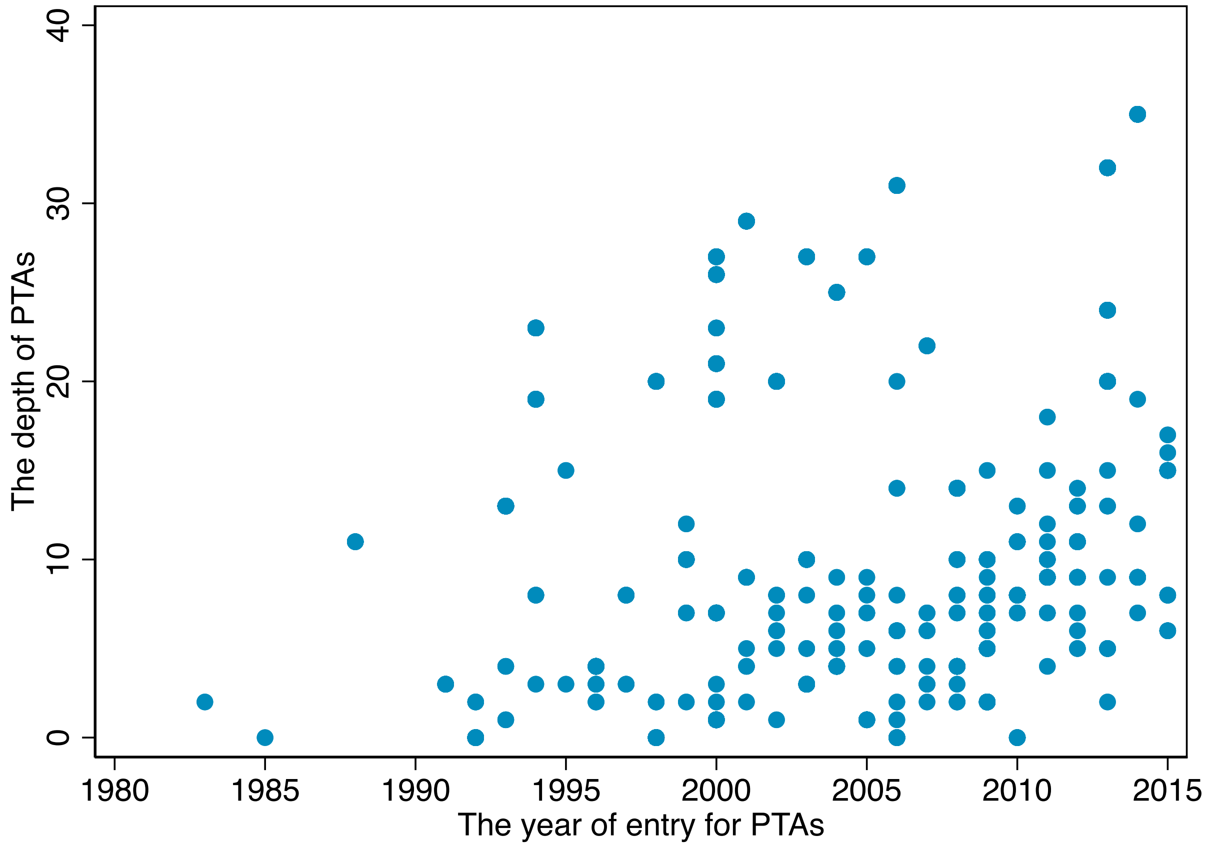
Vicard, Vincent, “Trade, conflict, and political integration: Explaining the heterogeneity of regional trade agreements,” *European Economic Review*, 2012, *56* (1), 54–71.

Wilhelm, Stefan and Miguel Godinho de Matos, “Estimating spatial probit models in R,” *The R Journal*, 2013, *5* (1), 130–143.

Wong, Ka-Fu and Miaojie Yu, “Democracy and accession to GATT/WTO,” *Review of Development Economics*, 2015, *19* (4), 843–859.

Yi, Sang-Seung, “Endogenous formation of customs unions under imperfect competition: Open regionalism is good,” *Journal of International Economics*, 1996, *41* (1-2), 153–177.

Yu, Miaojie, “Trade, democracy, and the gravity equation,” *Journal of Development Economics*, 2010, *91* (2), 289–300.



Note: This figure plots the evolution of the depth of PTAs measured by the number of the WTO-extra provisions and the year of entry over 1980 to 2015. The dataset is the same as in Hofmann et al. (2017).

Figure 1: The trend of PTA depthness

Table 1: Sample correlations between PTA depth and institutional indicators

	PIVsum	EFWsum	PIVdist	EFWdist
WTO-extra %	0.431***	0.260***	-0.305***	-0.121***
WTO-extra #	0.461***	0.344***	-0.317***	-0.173***

Note: This table presents the Pearson's correlations between the number (#) or proportion (%) of the WTO-extra provisions and institutional indicators. Specifically, PIVsum measures the summation of the trading partners' democracy levels, EFWsum measures the summation of the trading partners' economic freedom indices, PIVdist measures the absolute distance between the trading partners' democracy levels, and EFWdist measures the absolute distance between the trading partners' economic freedom indices. Finally, superscripts *, ** and *** represent statistical significance at the ten, five and one percent level, respectively.

Table 2: Summary statistics for the cross-sectional data

Variables	N	Mean	Std. dev.	Min	Max
NATURAL	10,011	-8.674	0.783	-9.894	-4.088
GDPsum	10,011	26.324	1.679	21.583	31.085
GDPsim	10,011	-2.031	1.390	-9.165	-0.693
DKL	10,011	1.749	1.237	0.000	6.229
SQDKL	10,011	4.589	5.635	0.000	38.806
REMOTE	10,011	8.978	0.120	8.786	9.487
DROWKL	10,011	1.353	0.631	0.017	3.462
PIVdist	10,011	6.136	5.490	0.000	20.000
EFWdist	10,011	0.955	0.710	0.000	4.200

Note: This table presents the summary statistics of the cross-sectionnal data that includes the variables of NATURAL (the log of the inverse of the great circle distance between two trade partners' capitals), REMOTE (the remoteness of a pair of continental trading partners from the rest of the world), GDPsum (Total bilateral market size denotes the total GDP of trade partners), RGDPsim (the similarity of two countries in terms of their GDP), DKL (the absolute difference in GDP per capita), SQDKL (the square of DKL), DROWKL (the relative factor endowment difference between the rest of the world and a given country-pair), PIVdist (the absolute distance of democracy between trading partners), and EFWdist (the absolute distance of economic freedom between trading partners).

Table 3: Summary statistics for the panel data

Variables	N	Mean	Std. dev.	Min	Max
NATURAL	24,420	-8.698	0.814	-9.892	-4.088
GDPsum	24,420	25.991	1.771	20.699	30.838
GDPsim	24,420	-2.085	1.427	-9.614	-0.693
DKL	24,420	1.815	1.280	0.000	6.152
SQDKL	24,420	4.935	5.952	0.000	37.850
REMOTE	24,420	8.988	0.128	8.786	9.487
DROWKL	24,420	1.440	0.664	0.013	3.640
PIVdist	24,420	5.785	5.312	0.000	20.000
EFWdist	24,420	1.143	0.886	0.000	5.770

Note: This table presents the summary statistics of the panel data. We have 6,105 observations for each year and totally 24,420 observations for four years. See Table 2 for the detailed information about the variables.

Table 4: Empirical results from the cross-sectional data

	Theory	Baseline			Estimates in EL	
		Non-spatial	W^G	W^I	Non-spatial	W^G
	(1)	(2)	(3)	(4)	(5)	(6)
$W^I \cdot PTA^*$	+			0.316***		
$W^G \cdot PTA^*$	+		0.430***			0.749***
NATURAL	+	0.595***	0.465***	0.583***	0.530***	0.736***
		(0.023)	(0.018)	(0.023)	(0.024)	(0.031)
GDPsum	+	0.200***	0.167***	0.179***	0.311***	0.233***
		(0.011)	(0.009)	(0.011)	(0.010)	(0.013)
GDPsim	+	0.193***	0.172***	0.179***	0.310***	0.246***
		(0.014)	(0.012)	(0.013)	(0.014)	(0.016)
DKL	+	0.133***	0.134***	0.134***	0.226***	0.070
		(0.040)	(0.038)	(0.039)	(0.045)	(0.050)
SQDKL	-	-0.025***	-0.033***	-0.029***	0.107***	0.055***
		(0.009)	(0.009)	(0.009)	(0.012)	(0.012)
REMOTE	+	0.653***	0.421***	0.719***	0.845***	0.576***
		(0.147)	(0.084)	(0.144)	(0.099)	(0.110)
DROWKL	-	-0.239***	-0.070***	-0.179***	0.111***	0.148***
		(0.029)	(0.021)	(0.027)	(0.024)	(0.031)
PIVdist	-	-0.060***	-0.055***	-0.044***		
		(0.003)	(0.002)	(0.002)		
EFWdist	-	-0.094***	-0.078***	-0.053***		
		(0.022)	(0.020)	(0.02)		
Negative Log-likelihood			4,687	4,949		
Year		2017	2017	2017	2005	2005
Countries		142	142	142	178	178
Observation		10,011	10,011	10,011	15,753	15,753

Note: The left panel presents the predicted signs from trade theory; the middle panel presents the baseline results with no spatial lag term, the geographical spatial lag term ($W^G \cdot PTA$), and the institutional spatial weight matrix ($W^I \cdot PTA$); the right panel presents the results of EL for comparison, with no spatial term and the geographical spatial lag term ($W^G \cdot PTA$). The spatial weight matrix W^I is calculated from two institutional indicators that of democracy and economic freedom, and W^G is calculated from geographical coodinates. The control variables include NATURAL (the log of the inverse of the great circle distance between two trade partners' capitals), REMOTE (the remoteness of a pair of continental trading partners from the rest of the world), GDPsum (Total bilateral market size denotes the total GDP of trade partners), RGDPsim (the similarity of two countries in terms of their GDP), DKL (the absolute difference in GDP per capita), SQDKL (the square of DKL), DROWKL (the relative factor endowment difference between the rest of the world and a given country-pair), PIVdist (the absolute distance of democracy between trading partners), and EFWdist (the absolute distance of economic freedom between trading partners). Standard errors are reported in parentheses. Finally, superscripts *, ** and *** represent statistical significance at the ten, five and one percent level, respectively.

Table 5: Empirical results from the panel data

	Theory	Baseline				Estimates in EL	
	Theory (1)	Non-spatial (2)	W^G (3)	W^I (4)	W^G & W^I (5)	Non-spatial (6)	W^G (7)
$W^I_{t-5} \cdot PTA_{t-5}$	+			2.451***	2.188***		
				(0.082)	(0.085)		
$W^G_{t-5} \cdot PTA_{t-5}$	+		1.082***		0.988***		2.509***
			(0.032)		(0.033)		(0.127)
NATURAL	+	0.690***	0.588***	0.688***	0.595***	0.581***	0.452***
		(0.015)	(0.015)	(0.015)	(0.015)	(0.020)	(0.023)
GDPsum	+	0.205***	0.182***	0.140***	0.125***	0.893***	0.773***
		(0.007)	(0.007)	(0.007)	(0.007)	(0.076)	(0.076)
GDPsim	+	0.197***	0.179***	0.158***	0.146***	0.026	0.040
		(0.009)	(0.009)	(0.009)	(0.010)	(0.059)	(0.061)
DKL	+	0.159***	0.191***	0.183***	0.209***	0.571***	0.583***
		(0.028)	(0.029)	(0.029)	(0.029)	(0.066)	(0.067)
SQDKL	-	-0.072***	-0.083***	-0.070***	-0.081***	0.106***	0.130***
		(0.007)	(0.007)	(0.007)	(0.007)	(0.025)	(0.025)
REMOTE	+	0.830***	0.805***	1.047***	0.987***	16.536***	15.668***
		(0.091)	(0.092)	(0.092)	(0.093)	(0.739)	(0.758)
DROWKL	-	-0.039**	-0.002	-0.095***	-0.056***	0.559***	0.460***
		(0.017)	(0.018)	(0.018)	(0.018)	(0.093)	(0.095)
PIVdist	-	-0.024***	-0.022***	-0.007***	-0.007***		
		(0.002)	(0.002)	(0.002)	(0.002)		
EFWdist	-	-0.128***	-0.131***	-0.097***	-0.104***		
		(0.013)	(0.013)	(0.013)	(0.013)		
Pseudo R^2		0.222	0.263	0.254	0.287		
Log-likelihood		-11104.243	-10524.768	-10648.345	-10182.191		
Year		1996-2017	1996-2017	1996-2017	1996-2017	1955-2005	1955-2005
Countries		111	111	111	111	146	146
Observations		24,420	24,420	24,420	24,420	93,323	93,323

Note: The left panel presents the predicted signs from trade theory; the middle panel presents the baseline results with no spatial lag term, the geographical spatial lag term ($W^G_{t-5} \cdot PTA_{t-5}$), the institutional spatial lag term ($W^I_{t-5} \cdot PTA_{t-5}$), and both the two spatial lag terms; the right panel presents the results of EL for comparison, with no spatial term and the geographical spatial lag term ($W^G_{t-5} \cdot PTA_{t-5}$). See Table 4 for the detailed information about the variables. One note that all the explanatory variables use the values observed in year $t-5$ in the panel data analysis. Standard errors are reported in parentheses. Superscripts *, ** and *** represent statistical significance at the ten, five and one percent level, respectively.

Table 6: Marginal effects

	Direct	Indirect	Total
NATURAL	0.163	0.075	0.238
GDPsum	0.050	0.023	0.073
GDPsim	0.050	0.023	0.073
DKL	0.037	0.017	0.055
SQDKL	-0.008	-0.004	-0.012
REMOTE	0.201	0.092	0.294
DROWKL	-0.050	-0.023	-0.073
PIVdist	-0.012	-0.006	-0.018
EFWdist	-0.015	-0.007	-0.022

Note: The table presents the marginal effects obtained from the SAR probit model with the institutional differential based spatial weight matrix using the cross-sectional data. See Table 4 for the detailed information about the variables. Direct effects measures the average over all the observations of the effects of the change of an explanatory variable of a single observation on the choice probability of that same observation. Indirect effects represent the average over all the observations of the effect of a change on an explanatory variable on the choice probability of the neighbouring observations. Total effects stand for the sum of direct and indirect impacts.

Table 7: Empirical results with compound spatial weight matrix

	Theory	$W^C = \alpha W^G + (1-\alpha)W^I$								
		$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.4$	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$W^C \cdot PTA^*$	+	0.431***	0.438***	0.416***	0.410***	0.406***	0.403***	0.400***	0.396***	0.391***
NATURAL	+	0.463***	0.441***	0.451***	0.461***	0.464***	0.466***	0.464***	0.466***	0.462***
		(0.019)	(0.018)	(0.018)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
GDPsum	+	0.179***	0.182***	0.179***	0.176***	0.177***	0.177***	0.177***	0.178***	0.178***
		(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
GDPsim	+	0.178***	0.180***	0.178***	0.177***	0.178***	0.177***	0.177***	0.177***	0.177***
		(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
DKL	+	0.126***	0.124***	0.137***	0.143***	0.141***	0.142***	0.141***	0.138***	0.137***
		(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)
SQDKL	-	-0.034***	-0.037***	-0.038***	-0.038***	-0.037***	-0.037***	-0.037***	-0.036***	-0.035***
		(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
REMOTE	+	0.270**	0.409***	0.531***	0.576***	0.593***	0.597***	0.601***	0.595***	0.585***
		(0.110)	(0.097)	(0.094)	(0.093)	(0.093)	(0.093)	(0.093)	(0.094)	(0.094)
DROWKL	-	-0.087***	-0.041*	-0.046**	-0.055**	-0.059**	-0.060**	-0.060**	-0.064***	-0.067***
		(0.024)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.024)
PIVdist	-	-0.039***	-0.044***	-0.049***	-0.052***	-0.054***	-0.055***	-0.056***	-0.057***	-0.057***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
EFWdist	-	-0.043**	-0.061***	-0.068***	-0.070***	-0.075***	-0.075***	-0.075***	-0.074***	-0.076***
		(0.020)	(0.020)	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Negative Log-likelihood		4,854	4,802	4,804	4,800	4,802	4,801	4,805	4,808	4,815
Year		2017	2017	2017	2017	2017	2017	2017	2017	2017
Countries		142	142	142	142	142	142	142	142	142
Observation		10,011	10,011	10,011	10,011	10,011	10,011	10,011	10,011	10,011

Note: The table presents the results obtained from the SAR probit model with compound spatial weight matrix W^C. The compound spatial weight matrix W^C is calculated from both the geographical spatial weight and institutional spatial weight as shown in Equation (6). See Table 4 for the detailed information about the variables. Standard errors are reported in parentheses. Superscripts *, ** and *** represent statistical significance at the ten, five and one percent level, respectively.

Table 8: Simulation results for the estimation method with endogeneity

	β_0	β_1	β_2	ρ
		$n = 100$		
Avg. est.	0.059964	0.396611	1.124607	0.281456
Avg. bias	0.059964	0.096611	0.324607	-0.01854
Avg. abs. bias	0.301337	0.266906	0.39073	0.088301
Std. dev.	2.01813	0.421497	3.039674	0.111779
RMSE	2.018012	0.432222	3.055446	0.113252
		$n = 250$		
Avg. est.	0.003735	0.31652	0.852911	0.295078
Avg. bias	0.003735	0.01652	0.052911	-0.00492
Avg. abs. bias	0.117819	0.118716	0.112546	0.065615
Std. dev.	0.149978	0.150731	0.141232	0.080698
RMSE	0.149949	0.151559	0.150752	0.080808
		$n = 500$		
Avg. est.	-0.00512	0.310667	0.827631	0.297883
Avg. bias	-0.00512	0.010667	0.027631	-0.00212
Avg. abs. bias	0.076735	0.07769	0.073923	0.048849
Std. dev.	0.098646	0.099238	0.091461	0.061608
RMSE	0.098729	0.099761	0.0955	0.061613
		$n = 1,000$		
Avg. est.	0.000227	0.303652	0.813672	0.302332
Avg. bias	0.000227	0.003652	0.013672	0.002332
Avg. abs. bias	0.055414	0.052707	0.048474	0.038313
Std. dev.	0.069681	0.065855	0.060237	0.047887
RMSE	0.069647	0.065923	0.061739	0.047919

Note: The table presents the simulation results for the estimation method that allows for the endogeneity of both the spatial weight matrix and explanatory variables. In the simulation, parameters n denotes the sample size which increases from 100 to 1,000, and we repeat 1,000 Monte Carlo trials for each n, $\beta_0 = 0$, $\beta_1 = 0.3$, $\beta_2 = 0.8$, and $\rho = 0.3$.

Table 9: Empirical results with endogeneity

	Theory (1)	Endogeneity (2)	Baseline (subsample) (3)
$W^I \cdot PTA^*$	+	0.294***	0.292***
NATURAL	+	0.554***	0.539***
		(0.023)	(0.022)
GDPsum	+	0.170***	0.162***
		(0.012)	(0.012)
GDPsim	+	0.176***	0.173***
		(0.016)	(0.016)
DKL	+	0.115***	0.108**
		(0.044)	(0.043)
SQDKL	-	-0.016	-0.018*
		(0.010)	(0.010)
REMOTE	+	0.072	0.066
		(0.053)	(0.052)
DROWKL	-	-0.239***	-0.227***
		(0.030)	(0.029)
PIVdist	-	-0.050***	-0.048***
		(0.004)	(0.002)
EFWdist	-	-0.179***	-0.091***
		(0.036)	(0.024)
Negative log-likelihood		4,027	4,071
Year		2017	2017
Countries		126	126
Observations		7,875	7,875

Note: The table presents the results obtained from the SAR probit model that allows for endogeneity in both the spatial weight matrix and explanatory variables. See Table 4 for the detailed information about the variables. Standard errors are reported in parentheses. Superscripts *, ** and *** represent statistical significance at the ten, five and one percent level, respectively.

Table A.1: Countries in the sample

Country	Cross-sectional	Panel	Country	Cross-sectional	Panel
Albania	$\surd$	$\surd$	Ecuador	$\surd$	$\surd$
Algeria	$\surd$	$\surd$	Egypt	$\surd$	$\surd$
Angola	$\surd$		El Salvador	$\surd$	$\surd$
Argentina	$\surd$	$\surd$	Estonia	$\surd$	$\surd$
Armenia	$\surd$		Ethiopia	$\surd$	
Australia	$\surd$	$\surd$	Fiji	$\surd$	$\surd$
Austria	$\surd$	$\surd$	Finland	$\surd$	$\surd$
Azerbaijan	$\surd$		France	$\surd$	$\surd$
Bahrain	$\surd$	$\surd$	Gabon	$\surd$	$\surd$
Bangladesh	$\surd$	$\surd$	Gambia	$\surd$	
Belarus	$\surd$		Georgia	$\surd$	
Belgium	$\surd$	$\surd$	Germany	$\surd$	$\surd$
Benin	$\surd$	$\surd$	Ghana	$\surd$	$\surd$
Bhutan	$\surd$		Greece	$\surd$	$\surd$
Bolivia	$\surd$	$\surd$	Guatemala	$\surd$	$\surd$
Botswana	$\surd$	$\surd$	Guinea	$\surd$	
Brazil	$\surd$	$\surd$	Guinea-Bissau	$\surd$	$\surd$
Bulgaria	$\surd$	$\surd$	Guyana	$\surd$	$\surd$
Burkina Faso	$\surd$		Haiti	$\surd$	$\surd$
Burundi	$\surd$	$\surd$	Honduras	$\surd$	$\surd$
Cambodia	$\surd$		Hungary	$\surd$	$\surd$
Cameroon	$\surd$	$\surd$	India	$\surd$	$\surd$
Canada	$\surd$	$\surd$	Indonesia	$\surd$	$\surd$
Cape Verde	$\surd$		Iran	$\surd$	$\surd$
Central African Republic	$\surd$	$\surd$	Iraq	$\surd$	
Chad	$\surd$	$\surd$	Ireland	$\surd$	$\surd$
Chile	$\surd$	$\surd$	Israel	$\surd$	$\surd$
China	$\surd$	$\surd$	Italy	$\surd$	$\surd$
Colombia	$\surd$	$\surd$	Jamaica	$\surd$	$\surd$
Congo	$\surd$	$\surd$	Japan	$\surd$	$\surd$
Costa Rica	$\surd$	$\surd$	Jordan	$\surd$	$\surd$
Croatia	$\surd$	$\surd$	Kazakhstan	$\surd$	
Cyprus	$\surd$	$\surd$	Kenya	$\surd$	$\surd$
Czech Republic	$\surd$	$\surd$	Kuwait	$\surd$	$\surd$
Denmark	$\surd$	$\surd$	Kyrgyz Republic	$\surd$	
Dominican Republic	$\surd$	$\surd$	Laos	$\surd$	

Note: The table presents the countries in our sample.

Table A.2: Countries in the sample (continued)

Country	Cross-sectional	Panel	Country	Cross-sectional	Pane
Latvia	√	√	Portugal	√	√
Lebanon	√		Qatar	√	
Lesotho	√		Russia	√	√
Liberia	√		Rwanda	√	√
Libya	√		Saudi Arabia	√	
Lithuania	√	√	Senegal	√	√
Luxembourg	√	√	Sierra Leone	√	√
Macedonia	√		Singapore	√	√
Madagascar	√	√	Slovak Republic	√	√
Malawi	√	√	Slovenia	√	√
Malaysia	√	√	South Africa	√	√
Mali	√	√	South Korea	√	√
Mauritania	√		Spain	√	√
Mauritius	√	√	Sri Lanka	√	√
Mexico	√	√	Suriname	√	
Moldova	√		Swaziland	√	
Mongolia	√		Sweden	√	√
Morocco	√	√	Switzerland	√	√
Mozambique	√		Tajikistan	√	
Myanmar	√		Tanzania	√	√
Namibia	√	√	Thailand	√	√
Nepal	√	√	Togo	√	√
Netherlands	√	√	Trinidad and Tobago	√	√
New Zealand	√	√	Tunisia	√	√
Nicaragua	√	√	Turkey	√	√
Niger	√	√	Uganda	√	√
Nigeria	√	√	Ukraine	√	√
Norway	√	√	United Arab Emirates	√	√
Oman	√	√	United Kingdom	√	√
Pakistan	√	√	United States	√	√
Panama	√	√	Uruguay	√	√
Papua New Guinea	√	√	Venezuela		√
Paraguay	√	√	Vietnam	√	
Peru	√	√	Zambia	√	√
Philippines	√	√	Zimbabwe	√	√
Poland	√	√			

Note: The table presents the countries in our sample.