The Political Economy of Preferential Trade Agreements: An Empirical Investigation*

Giovanni Facchini†
Peri Silva‡
Gerald Willmann§

March 31, 2020

Abstract

In this paper, we develop a political economy model to study the decision of representative democracies to join a preferential trading agreement (PTA), distinguishing between free trade areas (FTA) and customs unions (CU). Our theoretical analysis suggests that income inequality and bilateral trade imbalances are important factors in determining the formation of PTAs, while the patterns of geographic specialization determine whether a CU or an FTA will emerge in equilibrium. Using a sample of 187 countries over the period 1960-2015, our empirical analysis lends strong support to these predictions: Income inequality and trade imbalances both reduce the likelihood of PTA formation, while geographical specialization increases the conditional probability of an FTA (over a CU).
JEL classification numbers: F13
Keywords: Free Trade Areas, Customs Unions, Trade Imbalances, Income Inequality

*We are grateful to seminar participants at the 10th Australasian Trade Workshop at the University of Sydney, the DEGIT XX meetings in Geneva, the 2015 meetings of the Southern Economic Association, the SAET 2016 Conference, the Fall of 2017 edition of the Midwest trade conference at SMU, the Federal University of Paraiba (UFPB, Brazil), the Marxe School of Public and International Affairs, the University of Nottingham, Southern Illinois University, and at Xavier University for helpful comments and discussions.

†University of Nottingham, CEPR, CESifo, GEP, and IZA; email: Giovanni.Facchini@nottingham.ac.uk.
‡Kansas State University, UFPB, GEP and Ld'A, email: pdasilva@k-state.edu.
§University of Bielefeld, CESifo, GEP and Ifw Kiel, email: gerald@willmann.com.

1 Introduction

The last decades have seen a rapid increase in the number of preferential trading agreements (PTAs) and as of September 2019, the World Trade Organization (WTO) has been notified of 695 PTAs, 481 of which are currently in force.[1] Most countries are members of more than one PTA, and only two countries – South Sudan and Somalia – are currently not engaged in any form of preferential trade liberalization. At the same time, while these agreements are pervasive, they do take different forms. In particular, the formation of free trade areas (FTAs) is more common than that of customs unions (CUs), with eight FTAs in force for each CU.[2] What drives a country's decision to form a PTA? Which factors shape the choice of the type of PTA to be established? The goal of this paper is to provide answers to these questions by developing a political economy model of the formation of PTAs, which allows us to characterize the factors that affect the decision to form a PTA, and those that matter in the choice of its type (FTA or CU). We then assess the predictions of our theoretical framework on a large sample of countries spanning the period 1960-2015.

Our theoretical analysis is based on a three-country, multiple good setting in which two prospective members strategically interact to choose the tariff levels applied *vis-à-vis* each other and the rest of the world, whereas the rest of the world implements most favorite nation (MFN) tariffs. The underlying economic structure is the oligopolistic trade model used in many studies of regional trade agreements (e.g. Krishna 1998, Freund 2000, Ornelas 2005b), in which even 'small' countries are able to influence their import prices because markets are segmented and firms are price setters. In each country, individuals derive income from labor supply, and from the profits generated by the oligopolistic firms in which they own a stake. Importantly, firm ownership is unevenly distributed in the population. Building upon this structure, we model the working of a representative democracy, where the citizenry in each prospective member chooses the trade policy regime (PTA or multilateral) and elected representatives determine the actual tariffs to be implemented. This framework extends the model developed by Facchini, Silva and Willmann (2013) by allowing for multiple goods and, more importantly, for trade imbalances between prospective member countries.

[1] Notice that WTO information on PTAs are based on notifications rather than on the "physical number" of PTAs. Thus, a PTA that includes both goods and services counts as two notifications. As of September 2019, the current number of physical PTAs is 302. This information is available at http://www.wto.org/english/tratop_e/region_e/region_e.htm.

[2] Our calculations use information from the WTO's Regional Preferential Agreement database available at http://rtais.wto.org/UI/publicsummarytable.aspx.

The choice of trade policy regime is modeled by means of a four-stage game. In the first stage, each prospective member holds a sequence of votes to choose between a non-discriminatory MFN trade policy, a free trade area or a customs union. In the second stage, voters choose the representative, who will then select the tariff policy in stage three. Under a MFN trade regime, the policy will be non-discriminatory. If instead a preferential agreement is in place, trade will be free between member countries. Moreover, external tariffs will be coordinated if a CU is formed, whereas the members will set external trade policies unilaterally in the case of an FTA. In the last stage of the game, firms compete on quantities, taking as given the trade policies set in the third stage.

Our analysis indicates that trade imbalances and income inequality play a relevant role in shaping the decision to form a PTA (of either type). To understand the role of trade imbalances, note that in our model, preferential access received by a prospective member tends to increase that country's aggregate welfare by raising the profits of the firms owned by local residents and based there. At the same time, granting preferential access tends to reduce it, as the decline in profits and tariff revenues outweighs the increase in consumer surplus. If bilateral trade is unbalanced, the degree of market access exchanged between prospective members is unequal. In particular, the greater the trade imbalances, the less politically viable is the formation of a PTA in the country facing a trade deficit, and, as a result, the less likely will be a PTA to emerge in equilibrium. As for the role played by income inequality, note that as wealth becomes more concentrated, the oligopolistic profits becomes less relevant for the median voter and, as a result, the PTA formation may not be politically viable.

Our results also indicate that — if a PTA is established — the patterns of geographic specialization are an important determinant of the choice between an FTA and a CU. In order to understand this point, note that in our model strategic delegation arises only in the CU but not in the FTA of MFN regime. Its extent is greater, the more misaligned are the interests of the median voters in the two prospective member countries, i.e. the more specialized is their production structure. Greater strategic delegation leads to higher external tariffs being chosen under a CU, making this type of agreement less desirable than an FTA from the point of view of the median voter.

Our empirical analysis takes these predictions to the data. In particular, we have assembled a large panel dataset covering 187 countries over the period 1960-2015. Following the spirit of our theoretical model, the decision to form a CU or FTA is described as a two-stage process,

in which a country pair first chooses whether to establish a PTA, and subsequently determines the agreement type. This idea is implemented using a Probit model with sample selection (Van de Ven and Van Pragg 1981). The econometric results lend support to our theoretical predictions. In particular we find that the greater are bilateral trade imbalances between two countries and the less equal is the income distribution in each country, the less likely it is for them to form a PTA. Furthermore, regarding the choice between an FTA and a CU, we find that the greater is geographic specialization, the more likely it is for an FTA to be formed instead of a CU. Our findings are robust to the inclusion of additional controls in both the selection and the latent equations, to alternative definitions of the key dependent variables and to focusing on specific subsamples of the data.

Our paper is related to two main strands of the literature. First, we build on the empirical studies that have investigated the economic determinants of the formation of PTAs. In their pioneering contribution, Baier and Bergstrand (2004) show that economic size, size asymmetry, distance, and degree of remoteness play an important role in explaining the emergence of a PTA between a pair of countries. Egger and Larch (2008) extend this analysis by accounting for the *domino effect* suggested by Baldwin (1995), i.e. they investigate how the formation of a PTA between two countries can induce other trading partners to either join this existing agreement or to create their own PTA. More recently, Baldwin and Jaimovichi (2012) build on this idea and develop a theoretically-grounded measure of interdependence among PTAs.[3] Our empirical analysis extends this literature by additionally accounting for the role of income inequality and trade imbalances in the decision to form a PTA, and by explicitly considering the factors affecting the choice between an FTA and a CU.

Second, our paper is also related to the theoretical body of work that has emphasized the role of politics in the formation of PTAs (see Freund and Ornelas (2010) for a recent review). In an early contribution, Grossman and Helpman (1995) develop a lobbying model, in which the governments of prospective member countries trade off aggregate welfare against campaign contributions in their decision to join an FTA. Importantly, throughout their analysis they assume the external tariffs to be constant, and show that the formation of an FTA is politically feasible if trade is balanced, and trade diversion is pervasive. Ornelas (2005)

[3]Other important papers in this literature are Chen and Joshi (2010) and Bergstrand and Egger (2013). In particular, Chen and Joshi allow for the possibility of hub-and-spoke patterns to emerge, whereas Bergstrand and Egger consider instead the determinants of the joint formation of PTAs and bilateral investment agreements (BITs). More recently, Baier, Bergstrand and Moriutto (2014) have investigated in greater detail the role played by the domino effect.

extends this framework by endogenizing the determination of external tariffs. By eliminating intra-bloc barriers, the creation of an FTA lowers the incentives of import competing firms to lobby for higher external tariffs, inducing a reduction in the rents from lobbying (tariff complementarity). This reduces the political viability of welfare decreasing FTAs, contrary to the earlier findings by Grossman and Helpman (1995). Facchini, Silva and Willmann (2013) complement their analysis by modeling the working of a representative democracy and explicitly considering the choice between the formation of an FTA and a CU.[4] This paper's contribution to the literature is thus two-fold: on the one hand we extend our previous theoretical work by modeling the important role played by trade imbalances; on the other, we assess the predictions of this rich framework empirically on a panel dataset spanning all preferential trade agreements in place over the period 1960–2015.

The rest of the paper is organized as follows. Section 2 presents the basic setup of the model, while Section 3 characterizes the conditions for the political viability of the establishment of a PTA, and for the choice between an FTA and a CU. In Section 4, we present our main predictions and describe our dataset. Section 5 presents the results of our empirical analysis, whereas in Section 6 we provide additional evidence. Section 7 concludes.

2 The Model

To study the formation of preferential trade agreements, we extend the standard oligopolistic model of trade that has been used in several analyses of regionalism (Krishna 1998, Freund 2000, Ornelas 2005b, 2007). Our setting allows us to consider how the decision to form a PTA and the choice of its type depend on: (i) bilateral trade imbalances; (ii) the degree of geographic specialization and (iii) the income distribution within each prospective member country.

Consider a three–country, $n+1$–good economy, where A and B are prospective members, while F is an aggregate entity that stands for the rest of the world. Good 0, the numéraire, is freely traded and produced in all countries, using only labor according to the identity production technology $X_0 = L_0$. As a result wages equal to 1. Goods 1 through n are instead produced by oligopolistic firms competing on quantities, using only labor under a

[4]Richardson (1994) builds a stylized model to study the choice between joining an FTA and a CU, emphasizing that a lobby might prefer an FTA than a CU, since "...in an FTA a domestic industry needs to lobby only the domestic government..., whereas, in a CU,...a larger legislative group be courted".

constant returns to scale technology. For simplicity the corresponding constant marginal cost is normalized to zero.

Each non numéraire good is produced in F by a firm of size one. A measure $\alpha \in (0.5, 1]$ of each firm in a fraction $\phi \in \{1/n, 2/n, ..., 1\}$ of the oligopolistic industries is located in A, whereas a measure $(1-\alpha)$ of these is based in B. Correspondingly, for the remaining $(1-\phi)n$ oligopolistic industries, a measure α is located in B and $(1-\alpha)$ in A. For tractability, we order sectors such that industries $i = 1, ..., \phi n$ are primarily located in A, whereas industries $j = \phi n+1, ..., n$ are located in B.[5] As industries are mirror images of each other, the parameter α thus captures the extent of geographic specialization in the patterns of production, whereas ϕ describes the share of numéraire exporting industries relative to importing industries for A, and hence also the pervasiveness of bilateral trade imbalances in the non-numeraire sectors between A and B. Since the numéraire sector does not affect the political balance of power, in the remainder of the paper *trade imbalances* will refer to imbalances in the exchange of non-numeraire goods (see Grossman and Helpman (1995) for a similar approach).[6] Introducing notation that will be useful later on, let $x^i_{A,B}$ be the quantity of good i produced by a firm located in A and consumed in B. Since a measure α of firms in industries 1 through ϕn are located in A, the amount of good i, produced in A, and consumed in B is given by $\alpha x^i_{A,B}$ for $i = 1, ..., n\phi$.

The population in each country consists of a continuum of individuals l of mass one. Each supplies one unit of labor, but individuals differ in the stake $\gamma_{s,l}$ they own of the profitable oligopolistic firms in country s. We assume that the oligopolistic sector's distribution of profits is the same in A and B, and normalize the fraction of profits received by the average voter to one ($\overline{\gamma} = 1$). Typical income distributions then imply that the share γ^m received by the median voter is such that $\gamma^m \leqslant 1$ (Alesina and Rodrik 1994, Dutt and Mitra 2002), and higher values of γ^m indicate greater equality in the distribution of income.

Preferences are identical across countries and individuals, and can be described by the following quasi-linear, quadratic, and additively separable, utility function:

$$u(x) = x^0 + \sum_{i=1}^{n} u_i(x^i) \tag{1}$$

[5]For example, consider a situation where n equals 10 and ϕ equals 6/10. In this case, country A has a greater measure of firms in goods 1 through 6, while country B has a greater measure of firms in goods 7 through 10.

[6]Alternatively, good 0 could be thought of as capturing financial flows among countries.

where $u_i(x^i) = Hx^i - \frac{x^{i^2}}{2}$, $H > 0$ for all i and thus the demand for good i is given by $x^i = H - p^i$. The assumptions on the supply and demand sides of the model[7] ensure that markets are segmented.

We model trade policy by assuming that each prospective member country can apply tariffs on imports from the other two countries. Denote by $t_{s,d}$ the tariff vector – whose components are $t^i_{s,d}$ for $i = 1...n$ – applied by country $d \in \{F, A, B\}$ on imports from country $s \in \{F, A, B\}$, where clearly $t_{d,d} = 0$. Country d's entire tariff matrix is then denoted by $\mathbf{t}_d = (t_{F,d}, t_{A,d}, t_{B,d})$, and the tariffs applied by the various countries are given by the stacked matrices, i.e. $\mathbf{t} = (\mathbf{t}_F, \mathbf{t}_A, \mathbf{t}_B)$. Note that different trade policy regimes impose different restrictions on these tariff matrices. If a preferential agreement between A and B is in place, then $t_{A,B} = t_{B,A} = 0$ – i.e. all goods traded between the two countries are given duty free access.[8] Furthermore, if the PTA takes the form of a customs union, then A and B must set the same tariffs on imports from F. Otherwise, A and B will apply MFN tariffs on imports. Notice also that, by assumption, F always applies MFN tariffs on imports from A and B.

Having laid out the primitives of the model, focusing on country A the indirect utility of agent l is given by:

$$\begin{aligned} v_A\left(\mathbf{t},\gamma_{A,l}\right) &= 1 + \gamma_{A,l}\sum_{i=1}^{n\phi}\alpha\pi^i_A(\mathbf{t}) + \gamma_{A,l}\sum_{j=n\phi+1}^{n}(1-\alpha)\,\pi^j_A(\mathbf{t}) \\ &+ \sum_{i=1}^{n\phi} t^i_{F,A}x^i_{F,A}(\mathbf{t}_A) + \sum_{j=n\phi+1}^{n} t^j_{F,A}x^j_{F,A}(\mathbf{t}_A) \\ &+ \sum_{i=1}^{n\phi}(1-\alpha)\,t^i_{B,A}x^i_{B,A}(\mathbf{t}_A) + \sum_{j=n\phi+1}^{n}\alpha t^j_{B,A}x^j_{B,A}(\mathbf{t}_A) \\ &+ \sum_{i=1}^{n\phi}\left[u\left(x^i_A(\mathbf{t}_A)\right) - p^i_A(\mathbf{t}_A)\,x^i_A(\mathbf{t}_A)\right] + \sum_{j=n\phi+1}^{n}\left[u(x^j_A(\mathbf{t}_A)) - p^j_A(\mathbf{t}_A)\,x^j_A(\mathbf{t}_A)\right] \end{aligned} \tag{2}$$

where $\pi^i_A(\mathbf{t}) = \sum_d \left[p^i_d - t^i_{A,d}\right] x^i_{A,d} = \sum_d \pi^i_{A,d}(\mathbf{t})$ represents the profits generated by a firm producing good i located in country A. Notice that in the case of industries i where production is geographically concentrated in country A, total sales in A are described by $x^i_A = x^i_{F,A} +$

[7]More specifically, marginal costs are constant and the preferences' separability and quasi-linearity rule out income and cross price effects.

[8]This assumption is standard in the literature and in line with GATT's Article XXIV. Saggi, Wong, and Yildiz (2019) relax it by considering the incentives faced by potential members of an FTA if they are allowed to jointly choose the degree of internal liberalization.

$\alpha x^i_{A,A} + (1-\alpha)\, x^i_{B,A}$, whereas total sales in A of the output of industries j where production is geographically concentrated in country B, are given by $x^j_A = x^j_{F,A} + (1-\alpha)x^j_{A,A} + \alpha x^j_{B,A}$. The first line in expression (2) represents labor income and profits accruing to individual l from industries concentrated in A and B respectively. The second and third lines represent tariff revenues collected by country A on imports from different sources, while the last line describes consumer surplus. As mentioned above, tariff revenue is rebated lump-sum to the citizenry, and is kept by the importing country. The indirect utility of an individual based in country B is defined analogously.

Trade policy setting is modeled as a four stage game among the three countries. In the first stage, each prospective member holds a sequence of votes to choose between a non–discriminatory MFN trade policy, a free trade area or a customs union. In the second stage, the population of each country elects a representative who will, in the third stage, decide the countries' tariff policy. In stage four, firms compete in quantities, taking as given the trade policy set in the third stage. We solve the model backwards, starting from the last stage.

Focusing on country A (a similar analysis applies to B) the equilibrium quantities and prices for industries where production is geographically concentrated in country A ($i = 1, ..., n\phi$) and B ($j = n\phi + 1, ..., n$) are respectively given by:

$$x^i_{A,A} = \frac{\left[H + (1-\alpha)\, t^i_{B,A} + t^i_{F,A}\right]}{3} \qquad x^j_{A,A} = \frac{\left[H + \alpha t^j_{B,A} + t^j_{F,A}\right]}{3} \tag{3}$$

$$x^i_{F,A} = \frac{\left[H + (1-\alpha)\, t^i_{B,A} - 2t^i_{F,A}\right]}{3} \qquad x^j_{F,A} = \frac{\left[H + \alpha t^j_{B,A} - 2t^j_{F,A}\right]}{3}$$

$$x^i_{B,A} = \frac{\left[H - (2+\alpha)\, t^i_{B,A} + t^i_{F,A}\right]}{3} \qquad x^j_{B,A} = \frac{\left[H - (3-\alpha)\, t^j_{B,A} + t^j_{F,A}\right]}{3}$$

$$p^i_A = \frac{\left[H + (1-\alpha)\, t^i_{B,A} + t^i_{F,A}\right]}{3} \qquad p^j_A = \frac{\left[H + \alpha t^j_{B,A} + t^j_{F,A}\right]}{3} \tag{4}$$

Market segmentation implies that the price of goods in A depends only on the trade policies adopted by that country. Moreover, it also implies that equilibrium prices and quantities do not depend upon the bilateral trade imbalance parameter ϕ.

3 The PTA formation process

In this section we analyze the determination of tariffs under the three trade regimes (stages 2 and 3), and subsequently the choice of the trade regime itself (stage 1 of the game). In doing so, we will highlight the effects of bilateral net trade positions, of the patterns of geographic specialization, and the role played by the shape of the income distribution.

3.1 Trade Imbalances

As pointed out already by Grossman and Helpman (1995), bilateral trade imbalances between prospective member countries are likely to be important for the decision to join a preferential trading agreement. To highlight their role while keeping the analysis tractable, we focus on a situation where perfect geographic specialization prevails ($\alpha = 1$). In this case, goods for which production is geographically concentrated in A (B) are exported by A (B) and only imported — not produced — by the other prospective member. Recall that if $\phi = 0.5$ A and B have the same number of exporting industries, and hence bilateral trade in these industries is balanced. If $\phi > 0.5$, A has more export industries than B and runs a trade surplus *vis-à-vis* B, which in turn runs a trade deficit.[9] These trade imbalances increase in ϕ.

When analyzing tariff setting, we take the choice of trade regime (resulting from stage 1) as given. Our framework calls for the population of each country to elect a citizen, who will choose the tariff level to be applied on imports. The objective of each representative is to maximize her own welfare, given the tariffs chosen by other countries. We denote the share of the representative's profit (that is, her identity) by using '$\hat{}$', and continue to focus our analysis on country A — the problem for B being analogous. We start by considering first the two regimes in which trade policy is set non-cooperatively, namely MFN and FTA.

Assuming that an interior solution exists, the tariff vector chosen by representative $\hat{\gamma}_A$ in the third stage of the game is

$$\mathbf{t}_A^{MFN/FTA}(\hat{\gamma}_A, \hat{\gamma}_B) = \arg\max_{\mathbf{t}_A} v_A(\mathbf{t}, \hat{\gamma}_A) \text{ s.t.} \begin{cases} \mathbf{t}_{F,A} = \mathbf{t}_{B,A} & \text{for MFN} \\ \mathbf{t}_{B,A} = 0 & \text{for FTA} \end{cases} \tag{5}$$

i.e. it depends on her identity, but not on that of the other countries' representative.

[9] Allowing $\phi < 0.5$ would simply reverse the roles of surplus and deficit country, but we prefer to let A always denote the surplus and B the deficit country.

We are now ready to determine who will be the country's representative. Note that the voters' problem is one–dimensional and satisfies the single-crossing property (see Facchini, Silva, Willmann 2013). We can thus apply the median voter theorem and the choice of $\hat{\gamma}_A$ is given by the solution to:

$$\max_{\widehat{\gamma}_A} v_A \left(\langle \mathbf{t}_A^{MFN/FTA} \left(\widehat{\gamma}_A \right), \mathbf{t}_B^{MFN/FTA} \left(\widehat{\gamma}_B \right), \mathbf{t}_F \rangle, \gamma^m \right) \tag{6}$$

For the MFN and FTA cases we then have the following result:

Lemma 1 *If geographic specialization is perfect and tariffs are set non-cooperatively then strategic delegation does not arise in equilibrium, i.e.* $\hat{\gamma}_c = \gamma^m$, $\forall c \in \{A, B\}$. *Furthermore, if an FTA is formed, tariffs applied to the non-member are (weakly) lower than under the MFN arrangement.*

Proof. See Appendix A.[10] ■

Intuitively, since the markets for goods i and j are segmented, the equilibrium prices in A and B bear no relationship with each other. Moreover, in this non–cooperative setting, the tariffs applied by A and B can differ. The median voter is better off representing her own interests rather than delegating to someone else, as she does not have any influence on the partner's decisions. As for the second part of the lemma, tariff complementarity follows the same logic as in Saggi (2006) and Ornelas (2007), i.e. it is the result of the successful effort of the median voter to attenuate the degree of trade diversion generated by the preferential access granted to the partner country.

Turning to the CU regime, the difference with the FTA is that member countries now cooperate in setting a common trade policy. In particular, the external tariff maximizes the joint surplus[11] of the two countries' representatives and is given by:

$$\mathbf{t}_A^{CU}(\hat{\gamma}_A, \hat{\gamma}_B) = \arg\max_{\mathbf{t}_A} [v_A \left(\mathbf{t}, \widehat{\gamma}_A \right) + v_B \left(\mathbf{t}, \widehat{\gamma}_B \right)] \text{ s.t. } \mathbf{t}_{F,A} = \mathbf{t}_{F,B} \tag{7}$$

[10] This result holds more generally, i.e. also for $\alpha < 1$. See Appendix D for more details.

[11] This assumption is intuitively appealing – given that A and B are of equal size and in line with much of the existing literature (e.g. see Ornelas (2007), Saggi, Woodland and Yildiz (2013), Lake and Yildiz (2016) etc.). Notice though that Syropoulos (2003) has shown that the nature of the sharing rule of a CU with respect to tariff revenue can affect both the tariff preferences and the trade patterns of CU members in ways that can prevent the implementation of jointly optimal tariffs. An important insight of his analysis is that CU members have an incentive to influence their common tariffs, not just for external terms-of-trade reasons, but also for internal distributional purposes. Given the focus of our analysis, we abstract from such considerations. The issues of the delegation of tariff-setting authority within a CU and the choice of weights in the social welfare function have been analyzed instead by Gatsios and Karp (1991) and Melatos and Woodland (2007).

As before, in the second stage country A's representative will be chosen by the median voter as the solution to the following problem

$$\max_{\widehat{\gamma}_A} v_A \left(\langle \mathbf{t}_A^{CU} \left(\widehat{\gamma}_A, \widehat{\gamma}_B \right), \mathbf{t}_B^{CU} \left(\widehat{\gamma}_A, \widehat{\gamma}_B \right), \mathbf{t}_F \rangle, \gamma^m \right) \tag{8}$$

and similarly for country B. For the CU we then have the following result:

Lemma 2 *If geographic specialization is perfect, strategic delegation arises in a CU, and the elected representative has an ownership share twice that of the median voter. Moreover, the common external tariff is higher than the tariff applied by each member of an FTA.*

Proof. See Appendix A. ■

Intuitively, a tariff implemented on imports of good $i = 1, ...n\phi$ benefits country A more than country B. Therefore, when setting the common external tariff, the representatives choose an intermediate level of protection that internalizes the negative spill-over on B. Anticipating this effect, the median voter in A chooses a representative who is more protectionist than herself, and the same happens in B. Given that the representatives are more protectionist than in the case of the FTA, the resulting common external tariff is higher.

We now proceed to study the first stage, where the trade policy regime is chosen. In both countries, we envisage a sequence of two referenda. In the first one, voters decide between the MFN regime – i.e. the status quo – versus an FTA. The FTA arises only if it is preferred by the electorate in both A and B. The second referendum then puts the outcome of the first vote up against a CU. Again, the CU arises only if it is preferred by the electorate in both countries. The case that they both prefer either type of PTA over MFN, but then prefer different types of PTA's does not arise in the model and therefore the order of votes on FTA and CU does not matter. To determine the outcome of the first stage, we need to understand which regime is preferred by the decisive median voters. To this end, it is helpful to first compare the welfare implications of each regime. When measuring welfare, we weigh equally the utility of all individuals, focusing on the average voter's indirect utility function, $v_c(\mathbf{t}, \overline{\gamma}), \forall c \in \{A, B\}$.

While trade imbalances do not affect equilibrium tariffs and strategic delegation, they do impact welfare. In our oligopolistic trade framework, countries tend to benefit from preferential access to the partner's market, whereas they tend to lose from granting such preferential access to their own domestic market. When bilateral trade between A and B is balanced,

the overall welfare effect of a PTA tends to be positive due to the increased profit generated by receiving preferential access. With trade imbalances, however, the exchange of market access becomes asymmetric. In particular, if $\phi > 0.5$, A has more export sectors than B, and therefore receives greater preferential access from B than it grants in return. This clearly affects the welfare implications of a PTA for the two member countries. We can thus obtain the following result:

Lemma 3 *If geographic specialization is perfect, the larger the trade imbalances, the larger (smaller) the welfare benefit of a PTA for the partner country running a bilateral trade surplus (deficit).*

Proof. See Appendix B. ■

Lemma 3 analytically establishes a piecewise monotonic relationship between the extent of trade imbalances and the welfare effects of a PTA. To understand the intuition, Figure 1 illustrates the welfare ranking for the surplus country A (top panel) and the deficit country B (bottom panel), as we vary both trade imbalances (ϕ) and income equality (γ^m). When the exchange of market access is almost balanced (ϕ close to 0.5), the FTA welfare-dominates the MFN regime in both A and B, for all values of γ^m. As ϕ increases and A experiences an ever larger trade surplus *vis-à-vis* B, this result continues to hold for the surplus country A, but is reversed for the deficit country. A similar argument applies to the case of a CU. Under a balanced exchange of market access, a CU is welfare-enhancing relative to the MFN regime, unless γ^m is very high. As the exchange of market access becomes less balanced, the range of γ^m for which a CU increases welfare relative to the MFN regime becomes greater (smaller) in the surplus (deficit) country.

[INSERT FIGURE 1 HERE]

We can now turn to the solution of the first stage of the game, in which the trade policy regime is chosen by the median voters. For a PTA to be politically viable, the median voters' welfare must increase as the economy moves from a MFN regime to the PTA. To understand the role of the various forces at play, it is useful to decompose the change in the median voter's indirect utility as follows:

$$\begin{aligned} \Delta v_A\left(\mathbf{t}^{MFN}, \mathbf{t}^{PTA}, \gamma^m\right) &= v_A(\mathbf{t}^{PTA}, \gamma^m) - v_A(\mathbf{t}^{MFN}, \gamma^m) \\ &= \underbrace{\Delta v_A\left(\mathbf{t}^{MFN}, \mathbf{t}^{PTA}, \overline{\gamma}\right)}_{\text{Social welfare}} - \underbrace{(1-\gamma^m)}_{\text{Inequality}}\underbrace{\left(\Delta \pi_A\left(\mathbf{t}^{MFN}, \mathbf{t}^{PTA}\right)\right)}_{\text{Profits}} \end{aligned} \tag{9}$$

where 'Δ' represents the change from the MFN regime to a PTA and $\pi_A(\mathbf{t}) = \alpha \sum_{i=1}^{n\phi} \pi_A^i(\mathbf{t}) + (1-\alpha) \sum_{j=n\phi+1}^{n} \pi_A^j(\mathbf{t})$.[12] Since the profits of member countries' firms increase if they are granted preferential access under a PTA, equation (9) highlights that politically viable PTAs must be welfare increasing. We can now establish the following result:

Proposition 1 *If geographic specialization is perfect, the median voter's utility gain from a PTA increases (decreases) with the extent of the trade imbalances for the surplus (deficit) country.*

Proof. See Appendix C. ■

The piecewise nature of this analytical result parallels the findings for welfare (see Lemma 3). Figure 2 illustrates the median voter's ranking of the three regimes for the surplus country A (top panel) and the deficit country B (bottom panel). To understand the role played by trade imbalances, note that for balanced trade, i.e. if $\phi = 0.5$, both countries (symmetric in this case) support an FTA if equality is high, as both median voters benefit from the additional profits from market access generated by the PTA. This is illustrated at the lower edges of the two panels in Figure 2, for which the rankings in both countries coincide. If instead trade is not balanced (ϕ increases above 0.5), Proposition 1 implies that the median voter in A will tend to increasingly prefer a PTA over the MFN regime as ϕ (the surplus) increases, for any given level of equality, whereas the opposite holds in country B. Importantly, for any PTA to be politically viable, both median voters must prefer it over the MFN regime. Our proposition then implies that for any given level of equality the median voter in the deficit country is less inclined towards a PTA than her counterpart in the surplus country. Hence the support of the former is crucial for the final outcome.

[INSERT FIGURE 2 HERE]

3.2 Geographic Specialization

We can now study the effect of varying the degree of geographic specialization. Recall that a measure $\alpha \in [0.5, 1]$ of firms in industries concentrated in A is located in that country (and similarly for firms in industries concentrated in B). α then determines the degree of

[12] Similarly, $\Delta v_A\left(\mathbf{t}^{FTA}, \mathbf{t}^{CU}, \gamma^m\right)$ represents the change in country A's median voter's indirect utility function as the economy moves from an FTA to the CU regime.

geographic specialization, e.g. if $\alpha = 0.5$ A's and B's production structures are identical, whereas if $\alpha = 1$ the two countries are completely specialized in different industries. To keep the analysis tractable, we restrict attention to the case of balanced trade, i.e. $\phi = 0.5$.

As in the previous section, we first analyze the tariff determination and then the choice of trade policy regime. Focusing on the non–cooperative setting, it is easy to show that Lemma 1 continues to hold (see Appendix D). Turning to the case of a CU, we can establish the following result:

Lemma 4 *If trade is balanced, strategic delegation arises in a CU, and the representative's ownership share is higher than that of the median voter. Furthermore, strategic delegation increases with the degree of geographic specialization.*

Proof. See Appendix D. ■

A useful intermediate step to study the choice of trade policy regime is to compare social welfare under the three alternatives. Two features of our model are important in this exercise. First, income inequality drives a wedge between the trade policy preferences of the median and the average voters. Second, the median delegates power only under the CU, and strategic delegation increases in geographic specialization. This implies a positive relationship between geographic specialization and the common external tariffs for a CU (see equation 31 in Appendix D).

Figure 3 illustrates the welfare ranking of the different regimes for each prospective member country.[13] If γ^m is close to 1 (the relevant case for a PTA to arise as we will establish below), as the degree of geographic specialization increases, the FTA (and finally even the MFN regime in second place) start to dominate the CU. Intuitively, the higher is geographical specialization, the more pronounced becomes strategic delegation in the CU regime, and hence the less attractive is the CU because of its high common external tariff.

[INSERT FIGURE 3 HERE]

Moving on to the median voters' decision, we have the following result:

Proposition 2 *If trade is balanced, the preference of the median voter for a CU relative to an FTA decreases with the degree of geographic specialization.*

[13]Note that since trade is balanced, Figure 3 applies to both A and B. See Appendix B for more details.

Proof. See Appendix E. ■

Intuitively, strategic delegation under the CU increases in geographic specialization as established in Lemma 4, leading to a more protectionist representative and hence to a higher common external tariff. The median voter in the first stage anticipates this, and thus tends to prefer an FTA over a CU the higher is the extent of geographic specialization. This result is illustrated in Figure 4 (see also Appendix B and E for more details).

[INSERT FIGURE 4 HERE]

3.3 Income distribution

In the previous subsections we investigated the role of trade imbalances and geographic specialization, and found that their effects depended on the shape of the income distribution. Here we study the role of income equality itself, that is how varying degrees of income equality affect the equilibrium policy predictions of our model. In doing so, for reasons of tractability, we again focus on the two cases discussed so far: varying degrees of trade imbalance for geographically specialized member countries, and varying degrees of geographical specialization under balanced trade.

Before presenting our results, recall that a greater value of γ^m indicates a more equal income distribution and that a lower γ^m implies that the median assigns less weight to profits in her personal welfare consideration — relative to the average welfare. Moreover, recall that in a customs union (and only there), if geographical specialization is imperfect (i.e. $\alpha > 0.5$) the median strategically delegates to a representative with a multiple of her ownership share. The higher is γ^m, the more protectionist is then the representative she selects.

3.3.1 Income distribution and trade imbalances

Focusing on the welfare of the average voter, Figure 1 shows that if trade is balanced she prefers a PTA over the MFN regime. Moreover, for high levels of γ^m (i.e. high equality), she prefers an FTA, as strategic delegation would lead to very protectionist representatives under a CU, who would reduce welfare by choosing high tariffs. If equality is instead lower, and strategic delegation hence not as detrimental, the average voter prefers a CU as joint tariff setting internalizes cross country spillovers. As trade becomes unbalanced – i.e. as ϕ increases – the choice between FTA and CU continues to depend on income inequality, but for country

B the negative effect of the trade deficit — receiving little access while giving much — starts to make a PTA unattractive.

Remember though that it is the median voter who is decisive in the political equilibrium. As discussed, the difference in personal welfare between the median and the average is due to the weight on profit, which is lower for the median. The less equal is the income distribution (the lower γ^m), the more this difference matters. Figure 2 depicts the policy preferences of the median voter. Note that for $\gamma^m = 1$, i.e. at the right edge, her preferences coincide with those of the average voter depicted in Figure 1. As the distribution of income becomes less equal, the median cares less and less about profits. Since a PTA increases market access and profits, the median finds a PTA less attractive the lower is income equality.

This effect is at work under both PTA regimes. For the FTA this is all, and hence $\Delta v_c(\mathbf{t}^{MFN}, \mathbf{t}^{FTA}, \gamma^m), \forall c \in \{A, B\}$ increases in γ^m. Turning to the CU case, there is an additional effect, namely that strategic delegation leads to the choice of more protectionist representatives, the more so the higher is γ^m. This would work against the first effect, but note that the CU does not actually emerge in equilibrium given country B's median voter preferences (see the bottom of Figure 2). In essence, this additional negative effect not only makes the CU less attractive relative to the MFN, but also to the FTA from the point of view of the country running a trade deficit.

3.3.2 Income distribution and geographic specialization

We turn now to the case of imperfect geographic specialization (i.e. $\alpha < 1$), assuming trade to be balanced to keep the analysis tractable. The effect of income inequality on the welfare of the average voter is as follows. If geographic specialization is pronounced, the effect is similar to the previous case. Income equality increases the detrimental effect of strategic delegation, and thus decreases the relative attractiveness of a CU compared to an FTA. As geographical specialization decreases (i.e. α approaches 0.5), the reason for strategic delegation disappears, because the distributions of profit and consumer surplus are aligned across the two countries. Consequently, the detrimental effect of strategic delegation in terms of higher tariffs also disappears, and the CU becomes the preferred choice for the entire range of γ^m.

Consider now the median voter, who differs from the average because of the lower weights on profits in her objective function. In this case, we obtain the following result. For pronounced geographical specialization, a less equal income distribution makes the FTA less

attractive for the median because of the lower emphasis on profit. Given that the CU is less attractive than an FTA because of strategic delegation and of greater benefits derived from profits, the median then prefers MFN for low γ^m. For less pronounced geographical specialization, which renders CU's less harmful, the decreasing attractiveness of PTAs for lower levels of income equality leads to the choice of the MFN regime over a CU. Income inequality thus determines the choice between MFN and PTA, whereas the choice in terms of the type of PTA depends on the extent of geographical specialization.

4 Main Predictions and Dataset

The main results of the theoretical analysis can be summarized in a series of hypotheses that can be empirically assessed. Importantly, we can distinguish between factors that directly affect the decision to form a PTA, and those that instead impact the type of PTA that will be chosen. In this section we start by formulating these hypotheses, and will then present the data employed in the analysis.

4.1 Main Predictions

Building on the analysis carried out in Section 3, and focusing on Figure 2 to understand the effects of trade imbalances, and on Figure 4 for the role of geographic specialization, our discussion so far has highlighted how the viability of a PTA crucially depends on the support it gains in the prospective member country with a trade deficit. In particular, Proposition 1 indicates that the greater the trade imbalances, the less likely will be a PTA to emerge in equilibrium, as the exchange of preferential access between the member countries becomes more unequal. Turning to the role of the income distribution, the discussion following Figure 4 indicates that no PTA will emerge in equilibrium if the level of income equality is too low. We can summarize these results in the following:

Hypothesis 1 *The greater the trade imbalances, the lower is the likelihood that a PTA will emerge in equilibrium.*

Hypothesis 2 *If income equality is sufficiently low then a PTA will not emerge in equilibrium.*

While the pervasiveness of trade imbalances and income inequality are behind the decision to establish a PTA, our model suggests that these factors do not affect the popularity of FTAs relative to CUs. Proposition 2 indicates that the choice of one PTA regime over the other depends instead on the extent of geographic specialization. This factor plays an important role because it determines the extent of strategic delegation in a CU, which may lead to the common external tariffs being inefficiently high. In fact, if the degree of geographic specialization is very high (α close to 1), the elected representative will be significantly more protectionist than the median voter in the CU regime, whereas no strategic delegation occurs in an FTA (see equation (31) in the Appendix). This might make the FTA the equilibrium choice as shown in the upper-right region in Figure 4. If geographic specialization is instead low (α close to 0.5), a CU will emerge. These results are summarized in the following:

Hypothesis 3 *If a PTA is formed, the higher (lower) is the degree of geographic specialization the more likely is an FTA (CU) to emerge in equilibrium.*

Moreover, Figure 4 also suggests that, for intermediate levels of geographic specialization, the formation of an FTA becomes politically viable if the degree of income equality is sufficiently high. Otherwise, a CU may be formed. We will assess this ancillary prediction, alongside with other robustness checks, in the empirical section.

4.2 Dataset

To assess the implications of our model, we have collected a large dyadic panel dataset with country-pair information spanning 187 countries over the period 1960-2015, at five-year intervals. We follow Egger and Larch (2008) and Baier, Bergstrand and Feng (2014) in focusing on data at this frequency.[14] The reason behind our choice is that preferential trading agreements are typically accompanied by long implementation periods, and data at five-year intervals are more likely to account for this than higher frequency data. Descriptive statistics for the variables used in this study can be found in Table 1. The four columns reflect the different dimensions of the dataset that we want to explore. In particular, column 1 provides the average and standard deviation for each variable in the entire sample, whereas column 2 reports the same information focusing on country-pairs belonging to the same PTA. Column

[14]See also by Baier and Bergstrand (2007) and Baier, Yotov and Zylkin (2019).

3 restricts the attention to country-pairs belonging to the same FTA, and column 4 focuses on country-pairs in the same CU.

[Table 1 here]

To capture the presence of a preferential trade agreement between a country pair, we have used information from Baier, Bergstrand, and Feng (2014) and updated their dataset to 2015 using the WTO's website on Regional Trade Agreements.[15] Baier, Bergstrand and Feng (2014) have classified agreements based on *de jure* characteristics, distinguishing between one-way agreements, two-way partial agreements, free trade areas, customs unions, common markets and economic unions. One-way and two-way partial agreements typically involve selective sectoral trade liberalization, and are not captured in our theoretical model. For this reason, we don't explicitly consider them in the empirical analysis. Furthermore, as pointed out by Baier, Bergstrand, and Feng (2014), the number of common markets and economic unions is very limited, and for this reason we follow their strategy and consider them together with classic customs unions. As a result, our empirical work will focus exclusively on FTAs and CUs, i.e. agreements in which trade among members is substantially duty free. In the case of CUs, member countries must have additionally agreed upon and implemented a common external tariff for the vast majority of products.[16]

In particular, we construct two variables. The first, PTA_{abt}, takes a value of one if at time t a preferential trade agreement is in place between countries a and b. The second, FTA_{abt}, characterizes instead different types of agreements, and takes a value of one if at time t a Free Trade Area is in place between countries a and b, and zero if instead a CU is in force. Columns 1 and 2 of Table 1 indicate that 5108 observations, or about 7.5 percent of the total, represent full-fledged preferential trade agreements taking the form of CUs or FTAs. Note also that according to Table 1, about 60.7 percent of these observations are represented by country pairs belonging to an FTA, while the rest belongs to a CU. As several recent efforts have been carried out to collect information on existing preferential trading agreements, we

[15] The relevant information can be obtained at "https://rtais.wto.org/UI/PublicMaintainRTAHome.aspx". For a recent database including measures of preferential trade liberalization beyond goods, see Hoffman, Osnago and Ruta (2017).

[16] This requirement is important as not all negotiated agreements have been implemented. For example MERCOSUR members have agreed and implemented a common external tariff for more than 80 percent of the products they trade, and as a result MERCOSUR is described as a CU in our dataset. On the other hand, members of the Andean Community have agreed to implement a common external tariff but have failed to follow through with that decision. As such, the Andean community is not described as a CU in our dataset.

have assessed the robustness of our results also using alternative datasets used in the literature (see Section 6).

Turning to the determinants of the formation of a PTA emphasized in the theoretical model, our measure of trade imbalances IMB_{abt} is built using information on bilateral trade flows from the IMF's direction of trade database[17] and is defined as:

$$IMB_{abt} = \frac{|Exp_{abt} - Exp_{bat}|}{|Exp_{abt} + Exp_{abt}|} \tag{10}$$

where Exp_{abt} is the value of exports from country a to country b at time t etc. This measure can range between zero, when trade is balanced, and 1, when trade is unidirectional. Our dataset highlights that trade between country pairs is typically highly unbalanced, with a gap between bilateral exports averaging 69% of total bilateral trade. However, the same figure is substantially lower for countries belonging to the same FTA or CU, reaching only 46% of total bilateral trade, or, equivalently, about 67% of the average trade imbalance recorded for the entire sample. This is in line with Hypothesis 1, suggesting that the likelihood of PTA formation increases, the lower are the trade imbalances between prospective member countries.

Turning to the shape of the income distribution, we capture it using the net Gini coefficient ($INEQ_{abt}$) taken from Solt's (2019) Standardized World Income Inequality Database.[18] In particular, we use the highest net Gini coefficient within a country-pair as our model suggests that - ceteris paribus - the least equalitarian country in each pair will find the creation of a PTA less politically sustainable. A comparison between columns 1 and 2 of Table 1 suggests that the average Gini coefficient in the least equalitarian country in each pair for the entire sample (41.15) is higher than the corresponding figure for countries belonging to the same CU or FTA (38.18). Again, this is broadly consistent with Hypothesis 2 from our theoretical model, suggesting that for a PTA to be established, member countries should be sufficiently equalitarian.

As for the main determinant of the PTA type choice highlighted in our model, we measure the degree of geographic specialization using information on the *share* of total value

[17]This is the same source used by Subramanian and Wei (2007), among others.

[18]Solt (2019) standardized previous data on inequality from different multilateral institutions, making information available for 196 countries starting from 1960. The net Gini coefficient takes into account possible income redistribution promoted by national governments through the tax system. Solt (2009) finds that the degree of inequality on a net-basis is significantly lower than on a gross-basis in particular in developed countries.

added generated from agricultural, manufacturing and service activities in the gross domestic products for each country. More specifically, consider a pair formed by country a and b and denote the service, industry and agriculture share of GDP in country i by SER_i, IND_i, and AGR_i respectively, where $i \in \{a, b\}$. Then, the degree of geographical specialization between countries a and b is defined as:

$$GEO_{abt} = |SER_{at} - SER_{bt}| + |IND_{at} - IND_{bt}| + |AGR_{at} - AGR_{bt}| .$$

This index can take values between $[0, 2]$, with a greater value indicating greater specialization. Our choice of indicator is inspired by the index of regional industry specialization described by Krugman (1991), and has the advantage of requiring information that is available from the World Bank's World Development Indicators dataset over a long time period and for the large number of countries included in our analysis. Column 1 of Table 1 suggests that on average the country–pairs involved in our sample differ in their reliance on a particular economic activity by 43.87 percentage points. Country pairs involved in a PTA are more similar (the corresponding figure is 28 percentage points). More importantly, a comparison between columns 3 and 4 reveals that the extent of geographic specialization for members of an FTA is 32.41 percentage points, which is far greater than the degree of geographic specialization of CU members which is equal to 21.28 percentage points. This is in line with Hypothesis 3, which suggests that the extent of geographic specialization should be greater among members of an FTA than among members of a CU.

In our analysis we will also control for a series of additional drivers that have been shown in the literature (see Baier and Bergstrand 2004, Egger and Larch 2008) to play a significant role in the formation of a PTA. More specifically, we include information on the total economic size of each country-pair ($GDPSUM_{abt}$), the inverse of the distance between two trade partners ($NATURAL_{ab}$), an indicator for whether countries in a pair are located on the same continent ($DCONT_{ab}$), the weighted average of the distance between the two countries and third-country trade partners ($REMOTE_{abt}$), the similarity in the economic size between two trade partners ($GDPSIM_{abt}$), the relative factor endowment asymmetry between two trade partners (DKL_{abt}), its squared-value ($SDKL_{abt}$), and the average relative asymmetry in factor endowments between each country in a country-pair and other trade partners ($DROWKL_{abt}$). The recent literature has also pointed out that the formation of a PTA between countries in a pair may either encourage the formation of other PTAs or may lead to the enlargement of

existing agreements. To account for this possibility, we additionally control for the index of interdependence ($INTERD_{abt}$) among PTAs proposed by Egger and Larch (2008), and, as a robustness test, we also control for the theory-based measure of interdependence developed by Baldwin and Jaimovichi (2012).[19] We represent this group of additional drivers of the formation of PTAs by the matrix $\mathbf{X}$ and these variables are constructed using data from the CEPII website. More details on the exact definitions of each of these variables can be found in Table A1 of the appendix.

5 Empirical Analysis

This section has two main objectives. First, we will lay out the econometric strategy implemented to assess the predictions of our theoretical analysis. Second, we will present our main results.

5.1 Specification

Following the spirit of our theoretical framework and the existing empirical literature, we can model the formation of a preferential trade agreement as a two-step process, where countries first decide whether to form a PTA (Hypotheses 1 and 2) and then agree on its type (Hypothesis 3), i.e. on whether the PTA will be an FTA or a CU. Thus, we have a combination of self-selection into a PTA in the first stage, and a binary decision about its type (CU or FTA) in the second stage, a setting which can be empirically examined using the probit model in the presence of selection developed by Van de Ven and Van Pragg (1981).

Our strategy represents a natural extension of the econometric approaches followed in the literature. For instance, Baier and Bergstrand (2004) specify a probit model on a cross-sectional dataset to investigate the determinants of the formation of preferential trade agreements. Egger and Larch (2008) specify a similar model, but on a panel dataset, to investigate the role played by interdependence in the formation of PTAs. A similar methodology has also been implemented by Bergstrand and Egger (2013) to analyze the determinants of bilateral investment treaties. As it is well known, in the context of a binary response model, using (country-pair) fixed effects to account for unobservables may give rise to the incidental parameters problem. To address this concern, Chamberlain (1980) suggests to use instead the

[19] See Table A1 for the exact definition.

average of time-variant explanatory variables to obtain consistent estimates of the parameters of interest. Following Egger and Larch (2008) and Baldwin and Jaimovich (2012) we implement this strategy in all our specifications.[20]

The first stage decision is described by the following specification:

$$PTA_{abt} = \alpha_0 + \alpha_1 INEQ_{ab,t-5} + \alpha_2 IMB_{ab,t-5} + \boldsymbol{\beta}\mathbf{X}_{ab,t-5} + \gamma\overline{\mathbf{Z}}_{ab} + \epsilon_{abt} \tag{11}$$

where PTA_{abt} is a binary variable that takes a value of 1 if a country-pair ab is part of the same CU or FTA in year t, and zero otherwise, and IMB_{abt} and $INEQ_{abt}$ are respectively our measures of trade imbalances and income inequality. Matrix $\mathbf{X}$ is the set of additional drivers of the formation of a PTA considered in the literature (see section 4.2), and the $\overline{\mathbf{Z}}$ matrix includes time-invariant averages of all the controls used in the specification.[21]

As the establishment of a preferential agreement between a pair of countries is likely to affect their overall economic structure, using contemporaneous characteristics of the country pair might lead to parameter estimates that are biased due to reverse causality. To mitigate this concern, we follow Egger and Larch (2008) and Bergstrand and Egger (2013) among others,[22] and lag all right hand side variables. In most specifications we also include year fixed effects to control for common time specific shocks. Our theoretical model provides predictions on the expected sign of the coefficients α_1 and α_2. In particular, Hypotheses (1) and (2) suggest that the greater is the trade imbalance (IMB_{abt}) within a country-pair, and the greater is the degree of income inequality ($INEQ_{abt}$), the less likely it is for a PTA to emerge in equilibrium. As a result, we expect $\alpha_1 < 0$ and $\alpha_2 < 0$.

The second stage decision is then captured by the following binary model:

$$FTA_{abt} = \theta_0 + \theta_1 GEO_{ab,t-5} + \delta\overline{\mathbf{Z}'}_{ab} + v_{abt} \tag{12}$$

where FTA_{abt} is a binary variable that equals 1 if an FTA is in place for country-pair ab in year t, and zero if instead a CU is in force. GEO_{abt} is a measure of the degree of geographic special-

[20] In their study of third countries' impacts on the formation of PTAs, Chen and Joshi (2010) use instead a linear probability model to allow for a rich set of country fixed effects. In a robustness check we implement a variation of their approach, obtaining broadly similar results. See section 6.

[21] I.e. the time-invarying averages of IMB_{ab}, $INEQ_{ab}$, and the control variables included in Matrix X.

[22] In a robustness check, we also report results for a specification in which we lag our right hand side variables by 10 years in order to control for the fact that some PTAs may have a longer phase-in process, obtaining similar results.

ization for a country-pair and matrix $\overline{\mathbf{Z}'}_{ab}$ includes time-invariant average of the controls used in the specification.[23] Our theoretical model provides a clear prediction on the expected sign of θ_1. Hypothesis (3) indicates that, if a PTA is formed, the higher the degree of geographic specialization (GEO_{abt}), the more likely is an FTA to emerge as a political equilibrium. As a result, we expect $\theta_1 > 0$. Also in this case, the explanatory variables are lagged to mitigate reverse causality concerns. The error terms ϵ_{abt} and v_{abt} are assumed to be bivariate, zero mean normally distributed with correlation coefficient ρ. In all our estimates, standard errors are clustered at the country–pair level to allow for the possibility that observations in our dyadic regressions might not be independent.

5.2 Main Results

Table 2 reports our main results. The top panel focuses on the PTA formation decision, whereas the bottom one considers the choice of PTA type. We start by presenting in columns (1) through (3) three parsimonious specifications. The first replicates – using our data – the probit model estimated before in the literature (see e.g. Baier and Bergstrand 2004 and Egger and Larch 2008) to explain the decision to form a PTA – where the drivers are those included in matrix $\mathbf{X}$ (see Section 4.2). In column (2) we explain instead the decision to form a PTA accounting only for country-pair trade imbalances and income inequality as suggested by Hypotheses (1) and (2). Finally, in column (3) we estimate the same model of column (1), but we additionally account for the two drivers highlighted in our theoretical model. These specifications allow us to assess the additional role played by the two new determinants of the formation of a PTA highlighted in our theoretical analysis.[24] In columns (4) and (5) we estimate instead a different specification, namely the probit model with sample selection presented in Section 5.1, where we separately model the choice between establishing an FTA and a CU controlling for geographic specialization in the latent equation. In column (4) we do not include year fixed effects, whereas in our benchmark specification in column (5) we do so to account for common time-varying unobservable shocks.[25] To quantify the economic magnitudes involved, in column (6), we report the corresponding marginal effects. The latter

[23]Hence in our benchmark model we also control for the country-pair average of the variable GEO_{abt} in equation (12). In Table 3 we assess the robustness of our estimates by including additional controls in (12), and as a result $\overline{\mathbf{Z}'}_{ab}$ will include the time-invarying average of these additional controls as well.

[24]We would like to thank a referee for suggesting to include these three specifications.

[25]Note that in all specifications we also include time invarying averages of the controls. See the discussion in the previous section.

capture the change in the probability of forming a PTA (respectively an FTA) due to an infinitesimal change in each independent, continuous variable, and a discrete change in the probability for dichotomous variables.

[Table 2 here]

The results shown in column (1) broadly confirm patterns that have already been uncovered in the existing literature. In particular, we find that a PTA is more likely to emerge if two countries are geographically closer ($NATURAL$) to each other, if they belong to the same continent ($DCONT$), if other country-pairs are part of pre-existing PTAs ($INTERD$), if their total market size ($GDPSUM$) is larger, if they are more similar in terms of their economic size ($GDPSIM$) and if their factor endowments (DKL) are more dissimilar. As previous studies have also concluded, the effect of the latter is non-linear and increasing, but only up to a point (the sign of $SDKL$ is negative). Differently from earlier findings (e.g. Baier and Bergstrand 2004), our analysis indicates that the likelihood of establishing an agreement increases with the relative factor endowment difference between the rest of the world and a given country-pair ($DROWKL$).

The findings in column (2) provide support for the importance of inequality and trade imbalances as determinants of the formation of a PTA. Our estimates indicate that the higher a country-pair's trade imbalances and the less equalitarian is their income distribution, the lower is the probability that a PTA will be formed, and both these results are highly significant, with a pseudo–$R^2 = 0.11$. Finally, the results in column (3) indicate that adding these controls also improves the fit of the model compared to the baseline specification of column (1).[26] [27]

Our theoretical model indicates that the drivers of the decision to form a PTA differ from those involved in the choice of the type of agreement. For this reason, as discussed before, in columns (4) and (5) we estimate a probit model with sample selection. The Wald test reported at the bottom of the table indicates that the probit model with sample selection performs better than estimating equations (11) and (12) separately. Furthermore, the empirical findings shown in column (4) provide broad support for our theoretical predictions. Focusing on the

[26] The pseudo–R^2 improves from 0.49 to 0.51, or by approximately 3 percent.

[27] Baier, Yotov, and Zylkin (2019) find that the formation of PTAs has heterogeneous effects on bilateral trade flows depending on the geographic conditions and economic characteristics of countries in each pair. Likewise, different economic and geographic characteristics explain the formation of PTAs, as evidenced by the results shown in Table 2. Of course, these findings are related since country-pair characteristics affect trade flows and in turn will also affect the decision to form a PTA.

determinants of the formation of a PTA (upper panel), we again confirm that an increase in bilateral trade imbalances tends to significantly reduce the likelihood that a PTA will be put in place. Similarly, an increase in income inequality is negatively related to the probability that a PTA will be established between two countries.

Turning to the choice of the agreement type (bottom panel of Table 2), the results shown in column (4) indicate that, if a PTA has been formed, an FTA is more likely to emerge the more pronounced is the pattern of geographic specialization. These results provide strong support for the predictions of our theoretical model summarized in Hypothesis 3. Notice that the patterns uncovered in column (4) are confirmed and reinforced when we account additionally for time-varying common shocks in column (5). In particular, the direct effect of inequality in the PTA formation equation is now statistically significant at the 5% level. Moreover, the effects we have identified are economically important, as illustrated by the results reported in column (6). For instance, a one standard deviation increase in our measure of bilateral trade imbalances decreases the probability that a country-pair forms a PTA by about 0.52 percentage point – a large effect given that in our sample the probability of a country pair belonging to a PTA is only 7.5 percent. The same holds when we consider the determinants of the choice between an FTA and a CU. In particular, a one standard deviation increase in our measure of geographic specialization leads to an increase of 5.51 percentage points in the likelihood that an FTA – rather than a CU – will emerge in equilibrium.

The results we have reviewed so far indicate that the data support the basic predictions of our theoretical model. At the same time, it is interesting to investigate how well does our benchmark specification predict the actual formation of PTAs and their type. The former can be studied by using the fitted probabilities from the selection equation, and the latter by considering the fitted probabilities from the latent equation. As we pointed out in Section 4.2, the formation of a PTA is not a very common event – out of 67740 country-pair observations in our sample, only 5108 or 7.5 percent of the total have a PTA in place. Moreover, among country–pairs with a PTA, 60.7 percent of the observations are represented by FTAs and 39.3 percent by CUs. Following Bergstrand and Egger (2013), we use this a priori information about the proportion of events (PTA formation and FTA/CU formation) and non–events to form cutoff probabilities for the percent of correctly predicted, both for "true positives" and "true negatives." Focusing on the selection equation, our model successfully predicts 87 percent of the observations involving country pairs actually belonging to a PTA. Moreover,

our benchmark specification is also able to predict 86.6 percent of the observations involving country pairs that do not belong to a PTA. Turning to the choice between an FTA and a CU (described by the latent equation), our model is able to correctly predict 85.4 percent of the 2007 country-pairs that belong to the same CU, whereas it can correctly predict 67 percent of the 3101 country-pairs that belong to the same FTA. Overall, the empirical benchmark model correctly predicts 74.2 percent of the choice between an FTA and a CU for the country-pairs that have decided to form a PTA.

6 Additional Evidence

In this section, we extend our empirical analysis in three directions. First, we provide evidence on the role that additional factors might play in explaining the emergence of a PTA and the choice of its type. Next, we assess the robustness of our findings to the use of alternative measures/data sources for both our dependent and key explanatory variables. Finally, we investigate whether our findings continue to hold if we focus on specific sub–samples of the data.

In Table 3, we consider additional factors that might affect the formation of a PTA and/or the choice between an FTA and a CU. As pointed out in Hypothesis 2, the effect of income equality on the likelihood that a PTA will emerge in equilibrium might involve a threshold effect, rather than being monotonic as assumed so far in the empirical analysis. To account for this possibility, we pursue two alternative strategies. In column (1) we replace our continuous measure of inequality based on the Gini coefficient used in column (5) of Table 2 with an indicator variable equal to one if a country pair's highest Gini coefficient is in the top quartile;[28] in column (2) we allow instead the effect of inequality to be non-linear by introducing also a quadratic term.

The results from these specifications continue to support the theoretical predictions concerning the role of trade imbalances and geographic specialization. Moreover, they are also in line with Hypothesis 2. In particular, the findings in column (1) indicate that if inequality is sufficiently high, a PTA will be less likely to emerge in equilibrium; our estimates in column (2) indicate instead the presence of an inverted U–shaped relationship between the likelihood

[28]We have experimented also with alternative definitions of the threshold, e.g. the top tertile, and the results are very similar. The findings are available upon request.

of a PTA emerging in equilibrium and inequality, confirming that above a threshold level,[29] additional inequality will reduce the likelihood that a PTA will emerge in equilibrium.

An important result of our theoretical framework – summarized by Hypothesis 3 – is that the greater is the degree of geographic specialization, the more likely it is that an FTA will emerge in equilibrium if the two countries decide to form a PTA. The results described above provide strong support to this prediction. One ancillary prediction of our theoretical model is that the effect of geographic specialization will be stronger, the more equalitarian is the income distribution in the country pairs (see Section 4.1). We assess this prediction in column (3) (see in particular the bottom panel), finding support also for this additional implication of our model.

The literature has pointed out (e.g., Bond and Syropoulos 1996) that the number of existing members might affect the decision to enter a PTA. For this reason, in column (4) we additionally control for the number of other PTA members in the selection equation (upper panel).[30] While we find that larger PTAs are more attractive, accounting for the number of existing member countries does not affect our results. In column (5) of the same Table we also investigate whether the number of existing members might affect the choice of PTA type. While our results indicate that a larger number of existing members makes it less likely for an FTA to be formed compared to a CU, our main findings remain unaffected.

As we already discussed in the introduction, the literature on the choice between different types of preferential trade agreements is sparse. One interesting contribution is the paper by Lake and Yildiz (2016), who consider a three-country model in the presence of geographical asymmetries. In their setting some countries are located closer to each other than others, and geographically closer countries face lower trade costs than those further away. Their theoretical analysis indicates that there is a distance cut-off above which an FTA is the only viable choice of PTA. This suggests that the greater the geographical distance between the countries in a pair, the more likely it will be for an FTA rather than a CU to emerge in equilibrium. We assess this prediction — which is complementary to ours — in column (6),

[29]Note that the cutoff GINI coefficient for the change in the direction of the effect of income inequality equals 35.7, which represents a level far below the average in our sample (See Table 1). This result confirms that the non-monotonic effect of income distribution is related to economies with very equalitarian distributions of income.

[30]We would like to thank a referee for suggesting this specification. More specifically, we control for the natural log of 1 plus the number of other PTA partners, which is defined as the total number of member countries minus 2 for country-pairs that belong to the same PTA, and it equals zero otherwise. Using a measure without the natural log yields similar results, and these results are available upon request.

where in the bottom panel (latent equation), we control for the inverse of the distance between trade partners (captured by $NATURAL$). Interestingly, we find evidence corroborating this theoretical result: for the average country-pair and year, if they enter a PTA, more closely located countries are more likely to form a CU rather than an FTA. Importantly though, accounting for this additional factor does not affect our main results.

[Table 3 here]

The role of alternative sources of asymmetries across potential member countries in the formation of CUs or FTAs – like those which lead to different market sizes – has also been considered in the literature (Melatos and Woodland 2007).For this reason, in column (7), we additionally control for the degree of symmetry in market sizes for countries in each pair. Our findings indicate that – as expected – more similarly-sized countries are more likely to form a CU rather than an FTA since finding common ground in setting a common external tariff becomes easier. However, accounting for this additional driver does not affect our main results.[31]

Note that our theoretical model suggests that trade imbalances should be an important determinant of the decision to form a PTA. At the same time, it does not provide a clear prediction for their role in affecting the choice between a CU or an FTA. In column (8) of Table 3, we investigate whether they do play a role in the choice of PTA type by controlling for the extent of trade imbalances in the latent equation. Our results indicate that trade imbalances do not have a significant effect on the choice of PTA type, which is in line with Hypothesis 1. Importantly, the inclusion of this control does not affect our main results.

Finally, our benchmark model controls for country-pair characteristics by using the time-invariant averages of our control variables in line with other papers in this literature (see Section 5.1 for details). However, we can test the robustness of our results to the presence of country-fixed effects by following a control function approach based on Wooldridge (2015). In particular, we fit a linear probability model with country fixed effects for the selection equation. We implement a similar specification, in the latent equation, additionally controlling for the estimated error term of the selection equation. The findings indicate that our main results are robust.

[31]Note that our results are robust also to the inclusion of each of the controls used in the selection equation in the latent equation separately, as well as to the simultaneous inclusion of all PTA formation controls in the selection equation. These results are available upon request.

In Table 4, we investigate the robustness of our results to the use of alternative measures for both our dependent and explanatory variables. As already discussed earlier in the paper, the number of preferential trading agreements has rapidly increased over time. Importantly, various efforts have been carried out to collect systematic information on the nature of the preferential agreements in force. Some of the existing databases focus more on *de jure* criteria, whereas others emphasize more *de facto* considerations. It is, therefore, important to assess the robustness of our analysis to the use of alternative datasets proposed in the literature. In column (1) of Table 4, we present the results when our left-hand side variables (PTA formation decision and choice between a CU and an FTA) are constructed using the database collected by Mattevi (2005), whereas in column (2), we use instead the recent dataset collected by Egger and Larch (2008).[32] While some differences exist, the information contained in the data collected by Baier, Bergstrand, and Feng (2014) and in these alternative sources are broadly similar.[33] Importantly, using these alternative sources does not affect our results: the qualitative patterns we have uncovered in column 5 of Table 2 continue to hold, and even the magnitudes of the effects of our main explanatory variables are broadly comparable.

So far, we have employed five–year lagged values for our explanatory variables to address reverse causality concerns. In column (3), we assess the robustness of our findings to the introduction of ten–year lags to capture longer-term determinants of the preferential trading agreement formation process. Once again, our results are broadly unaffected.

Our analysis has shown that trade imbalances between prospective member countries may prevent the formation of a PTA. Our theoretical model indicates that the fate of a preferential trade agreement rests with the country running a bilateral trade deficit. To capture more precisely this idea, we have experimented with an alternative definition of trade imbalances emphasizing the net trade position of each country:

$$IMB_{abt} = \frac{|Exp_{abt} - Exp_{bat}|}{|I_a \times (IMP_{at} + Exp_{at}) + (1 - I_a) \times (IMP_{bt} + Exp_{bt})|} \tag{13}$$

where Exp_{abt} is the value of exports from country a to country b at time t, Exp_{at} (IMP_{at}) is

[32] In particular, we have used the updated version of the dataset used in Egger and Larch (2008), providing information on PTA formation up to 2015, which can be downloaded from 'https://www.ewf.uni-bayreuth.de/en/research/RTA-data/index.html". We have updated the information in Mattevi (2005) using the WTO's Regional Trade Agreements website.

[33] In particular, while in our baseline sample, 7.5 percent of the total observations are represented by FTAs and CUs, the same is true for respectively 10.3 and 7.1 percent of the total observations in Egger and Larch (2008) and Mattevi (2005).

the value of country a's total exports (imports) at time t, I_a is a dummy variable that equals one if country a runs a bilateral trade deficit with country b ($Exp_{abt} < Exp_{bat}$) and etc. The results are reported in column (4) of Table 4 and are broadly comparable with our benchmark findings.

Our theoretical model focuses on the role of bilateral trade imbalance in non-numéraire industries, where firms compete in a Cournot fashion. It is clear that imperfect competition is more likely to occur in industries characterized by differentiated rather than homogenous goods. To better capture this idea, we exploit Rauch's (1999) classification of goods in three broad categories, i.e., those traded in organized exchanges, those that are "reference priced" and those that are truly differentiated, and construct a measure of trade imbalances focusing only on the latter.[34] The results using this alternative measure are reported in column (5) and, once again, are qualitatively comparable to our baseline findings. To tackle the same issue, we have also experimented with an alternative definition of trade imbalances, focusing on their pervasiveness in the manufacturing sector[35] – where markets are more likely to be imperfectly competitive. Our findings are reported in column (6) and are once again in line with our benchmark.

Our proxy for geographic specialization builds on a comparison of the distribution of value added across the main sectors of the economy. In the context of our theoretical model, the focus is on the measure of firms in oligopolistic industries located in each country, which might be better captured by a comparison of the importance of the manufacturing sector across countries. For this reason, in column (7) we experiment with an alternative measure based on a comparison of the share of the manufacturing sector in value added. Our results are broadly unaffected.

[Table 4 here]

One important message emerging from the literature is that the formation of PTAs may affect other non-member countries' decisions to join an existing agreement or to form a new one. As explained in Section 4.2, in our benchmark specification, we control for this PTA driver using the measure of interdependence proposed by Egger and Larch (2008). More recently,

[34]The classification used in column (5) to identify differentiated goods is labeled as "liberal" in Rauch (1999). Using the alternative, "conservative" classification yields qualitatively similar results.

[35]Historical data on bilateral trade flows at the industry level is available using the first revision of the Standard International Trade Classification (SITC). We define manufactured products as bilateral trade flows in industries 5 through 8 using the 1-digit of the SITC.

Baldwin and Jaimovich (2012) develop a theory-based alternative. According to their measure, the effect of a given country a on the probability of forming a PTA with country b depends on the share of country a's imports from b in the latter country's total trade, as well as on the share of country a's preferential trade with other trading partners. In column (8), we use this alternative measure of interdependence and our results are qualitatively unaffected.

In Table 5, we investigate the robustness of our findings to different sample structures. It is well know that a significant fraction of the PTA considered in our analysis came into force since the end of the 20th century. In column (1), we thus study whether our model can help to explain the formation of only these most recent agreements, focusing on the years 2000 and beyond. While the magnitude of the coefficients is affected, the basic patterns we had uncovered in our benchmark specification continue to hold.

As we have argued before, the formation of FTAs is more popular than the formation of CUs – and out of 2007 country pairs that are members of a CU, 1167 or 58 percent of the total is related to the European Union (EU). Many observers have argued that the EU is much more than a simple trade agreement and that political considerations played a key role in its establishment and, a fortiori, in the six subsequent rounds of negotiations that have led its membership to involve 28 countries by 2015. For this reason, in column (2), we investigate whether our main predictions, in particular concerning the choice of PTA type, continue to hold if we drop the substantial number of CU country-pairs related to the formation of the EU from our sample. Our results indicate that this is indeed the case.

A similar concern might more generally apply to very large country groupings which have entered in preferential trade agreements. To assess the robustness of our results, beside country-pairs belonging to the EU, in column (3) we exclude from our analysis also country-pairs belonging to the Association of Southeast Asian Nations (ASEAN) and the Pan-Arab FTA, two agreements with more than ten members.[36] In column (4), we alternatively exclude from the analysis country-pairs that belong to the largest agreements in terms of the size of member countries' economies (EU, ASEAN, and NAFTA). The results confirm that our main findings continue to hold.

[Table 5 here]

[36] In 2015, the ASEAN and the Pan-Arab FTA had ten and sixteen members, respectively.

7 Conclusion

In this paper we have developed a representative democracy political economy model to shed new light into the process of forming a PTA, which allows us to distinguish between those factors that affect the decision to form a PTA, and those that matter for the choice of its type (FTA or CU). Our analysis highlights the important role played by bilateral trade imbalances and the shape of the income distribution in the decision to form a PTA. In particular, only if bilateral trade between prospective members is sufficiently balanced, and income inequality is sufficiently low, trade partners find a PTA to be politically viable. As for the choice of the agreement type, i.e. whether an FTA or a CU will emerge, our framework suggests that this depends on the extent geographic specialization between prospective member countries.

We then proceed to assess empirically the implications of our theory, using a sample of 187 countries covering the period 1960–2015. Our empirical analysis finds strong support for the predictions of the model. In particular, the greater are bilateral trade imbalances and income inequality, the less likely is a country-pair to have a PTA in place. Furthermore, we also find that the more pronounced is the pattern of geographic specialization between the two member countries, the more likely is an FTA to emerge between them instead of a CU.

We can think of at least two directions along which our research could be extended. First, our representative democracy approach emphasizes the role played by the median voter in determining the political viability of a trade regime. Real world decision making in democratic settings is typically more complex though, and involves taking into account both the concerns expressed by the majority of voters, and those put forward by organized pressure groups. The role played by lobbies has received significant attention in the literature on the determinants of the formation of FTAs (e.g. Grossman and Helpman 1995, Ornelas 2005), but to the best of our knowledge it has not yet been systematically explored in the context of the formation of CUs (Freund and Ornelas 2010). Extending the theoretical analysis to consider the general role of lobbying in the formation of a PTA is a very promising avenue for future research, but one that goes beyond the scope of this paper.[37] Second, several examples have emerged in which preferential liberalization has moved beyond trade, to involve "deep integration". Several data collection efforts are ongoing to describe and characterize this phenomenon (Mattoo, Rocha and Ruta 2019) but little is known when it comes to the factors shaping the decision to

[37]Note also that as pointed out by Dutt and Mitra (2002) lobbying activities are the likely reason why trade policies are typically biased against trade liberalization – and thus complement the variation in trade policy that is explained by models based on the median voter like ours.

undertake deep integration, and in particular to the political economy forces affecting it. While this question goes beyond the scope of our analysis, answering it will be key to understand the future working of the global economy, especially if the GATT-WTO multilateral approach to trade liberalization slows to a standstill.

References

[1] Alesina, A. and D. Rodrik (1994). "Distributive Politics and Economic Growth ", *The Quarterly Journal of Economics*, vol. 109, pp. 465–490.

[2] Atkinson, A. B. and A. Brandolini (2001). "Promises and Pitfalls in the Use of Secondary Data-Sets: Income Inequality in OECD Countries as a Case Study", *Journal of Economic Literature,* vol. 39, pp. 771-800.

[3] Baier, S. and J. Bergstrand (2004) "Economic Determinants of Free Trade Agreements", *Journal of International Economics*, vol. 64, pp. 29–63.

[4] Baier, S. and J. Bergstrand (2007) "Do free trade agreements actually increase members' international trade?", *Journal of International Economics*, vol. 71, pp. 72-95.

[5] Baier, S., J. Bergstrand and M. Feng (2014) "Economic integration agreements and the margins of international trade ", *Journal of International Economics*, vol. 93, pp. 339–350.

[6] Baier, S., J. Bergstrand and R. Mariutto (2014) "Economic Determinants of Free Trade Agreements Revisited: Distinguishing Sources of Interdependence", *Review of International Economics*, vol. 22, pp. 31-58.

[7] Baier, S., Y. Yotov, T. Zylkin (2019) "On the widely differing effects of free trade agreements: Lessons from twenty years of trade integration", *Journal of International Economics*, vol. 116, pp. 206-226.

[8] Baldwin, R. (1995) "A Domino Theory of Regionalism", In Baldwin, R., Haaparanta, P., Kiander, J. (Editors), Expanding Membership in the European Union. Cambridge University Press, Cambridge, UK.

[9] Baldwin, R. and D. Jaimovich (2012) "Are Free Trade Agreements contagious?," *Journal of International Economics*, vol. 88, pp. 1-16.

[10] Bergstrand, J. and P. Egger (2013) "What Determines BITs?", *Journal of International Economics*, vol. 99, pp. 107-122.

[11] Bond, E. and C. Syropoulos (1996) "The size of trading blocs market power and world welfare effects", *Journal of International Economics*, vol. 40, pp. 411-437.

[12] Chamberlain, G. (1980) "Analysis of Covariance with Qualitative Data", *The Review of Economic Studies*, vol. 47, pp. 225-38.

[13] Chase, K. (2006) "Multilateralism compromised: the mysterious origins of GATT Article XXIV ", *World Trade Review*, vol. 5, pp. 1–30.

[14] Chen, M. X. and S. Joshi, (2010) "Third-country effects on the formation of free trade agreements", *Journal of International Economics*, vol. 82, pp. 238-248.

[15] Dutt, P. and D. Mitra, (2002) "Endogenous trade policy through majority voting: an empirical investigation ", *Journal of International Economics*, vol. 58, pp. 107-133.

[16] Egger, P. and M. Larch (2008) "Interdependent Preferential Trade Agreement Memberships: An Empirical Analysis", *Journal of International Economics*, vol. 76, pp. 384–399.

[17] Estevadeordal, A., C. Freund and E. Ornelas (2008) "Does regionalism affect trade liberalization towards non–members?" *Quarterly Journal of Economics*, vol. 123, pp. 1531-1575.

[18] Facchini, G., P. A. Silva and G. Willmann (2013) "The customs union issue: Why do we observe so few of them?", *Journal of International Economics*, vol. 90, pp. 136-147.

[19] Foletti, L., M. Fugazza, Nicita, A. and M. Olarreaga (2011) "Smoke in the (Tariff) Water", *World Development*, vol. 34, pp. 248-264.

[20] Freund, C., (2000). "Multilateralism and the endogenous formation of preferential trade agreements "*Journal of International Economics* vol. 52, pp. 359-376.

[21] Freund, C. and E. Ornelas (2010) "Regional Trade Agreements", *Annual Review of Economics*, vol. 2, pp. 139-166.

[22] Gatsios, K. and L. (1991) "Delegation Games in Customs Unions ", *Review of Economic Studies*, vol. 58, pp. 391–97.

[23] Grossman, G. and E. Helpman (1995) "The Politics of Free-Trade Agreements", *American Economic Review*, vol. 85, pp. 667-690.

[24] Hoffman, C., A. Osnago, and M. Ruta (2017) "Horizontal depth : a new database on the content of preferential trade agreements", World Bank's Policy Research Working Paper 7981.

[25] Krishna, P., (1998) "Regionalism and multilateralism: a political economy approach", *Quarterly Journal of Economics*, vol. 113, pp. 227-250.

[26] Krugman, P. R. (1991) "Geography and Trade", MIT Press, Cambridge MA.

[27] Lake, J. and H. M. Yildiz (2016) "On the different geographic characteristics of Free Trade Agreements and Customs Unions", *Journal of International Economics*, vol. 113, pp. 213-233.

[28] Mattevi, E. (2005) "Does membership in preferential trading arrangements affect volumes of trade?", Master's Thesis, Free University of Bolzano.

[29] Mattoo, A., N. Rocha and M. Ruta (2019) "The evolution of deep trade agreements", mimeo, The World Bank.

[30] Melatos, M. and A. Woodland (2007) "Endogenous Trade Bloc Formation in an Asymmetric World", *European Economic Review*, vol. 51, pp. 901-924.

[31] Ornelas, E. (2005) "Rent Destruction and the Political Viability of Free Trade Agreements", *Quarterly Journal of Economics*, vol. 120, pp. 1475-1506.

[32] Ornelas, E., (2005b) "Trade creating free trade areas and the undermining of multilateralism "*European Economic Review* vol. 49, pp. 1717-1735.

[33] Ornelas, E. (2007) "Exchanging Market Access at the Outsiders' Expense - the Case of Customs Unions", *Canadian Journal of Economics,* vol. 40, pp. 207-224.

[34] Rauch, J. E. (1999) "Networks Versus Markets in International Trade", *Journal of International Economics,* vol. 48, pp. 7-35.

[35] Richardson, M. (1994) "Why a free trade area? The tariff also raises" *Economics and Politics* 6, pp. 79–96.

[36] Saggi, K. (2006) "Preferential Trade Agreements and Multilateral Tariff Cooperation," *International Economic Review*, vol. 47, pp. 29-57.

[37] Saggi, K., A. Woodland and H. M. Yildiz (2013) "On the Relationship between Preferential and Multilateral Trade Liberalization: The Case of Customs Unions ", *American Economic Journal: Microeconomics*, vol. 5, pp. 63-99.

[38] Saggi, K., W. Wong and H. M. Yildiz (2019) "Should the WTO require free trade agreements to eliminate internal tariffs?", *Journal of International Economics*, vol. 118, pp. 316-330.

[39] Subramanian, A. and S.J. Wei (2007) "The WTO promotes trade, strongly but unevenly," *Journal of International Economics*, vol. 72, pp. 151-175.

[40] Solt, F. (2009) "Standardizing the World Income Inequality Database", *Social Science Quarterly*, vol. 90, pp. 231-242.

[41] Solt, F. (2019) "Measuring income inequality across countries and over time: The Standardized World Income Inequality database", mimeo, November 2019.

[42] Syropoulos, C. (2003) "Rules for the Disposition of Tariff Revenues and the Determination of Common External Tariffs in Customs Unions" *Journal of International Economics*, vol. 60, pp. 387–416.

[43] Van de Ven, W. and B. Van Praag (1981) "The demand for deductibles in private health insurance : A probit model with sample selection," *Journal of Econometrics*, vol. 17(2), pp. 229-252.

[44] Wooldridge, J. (2015) "Control Function Methods in Applied Econometrics", *Journal of Human Resources*, vol. 50, pp. 420-445.

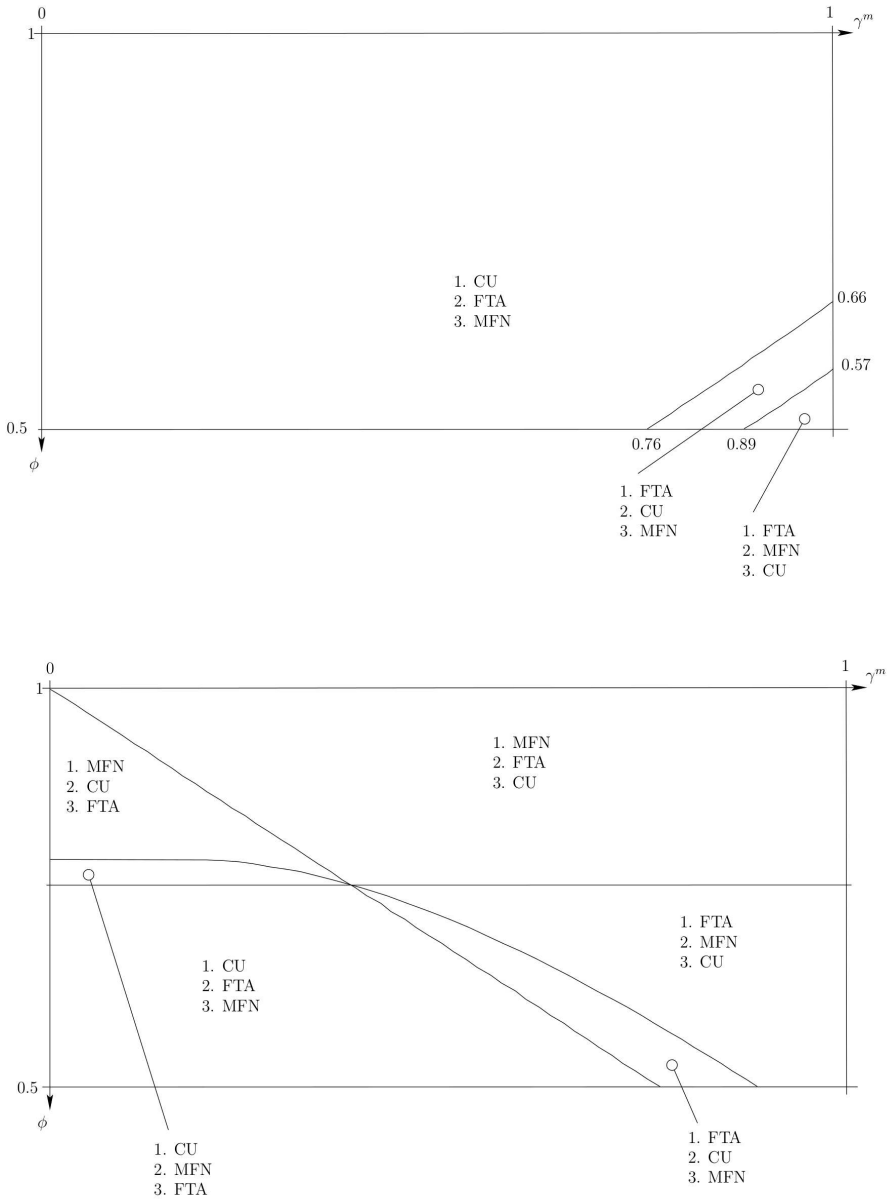

Figure 1: Welfare Ranking in country A (top) and B (bottom)

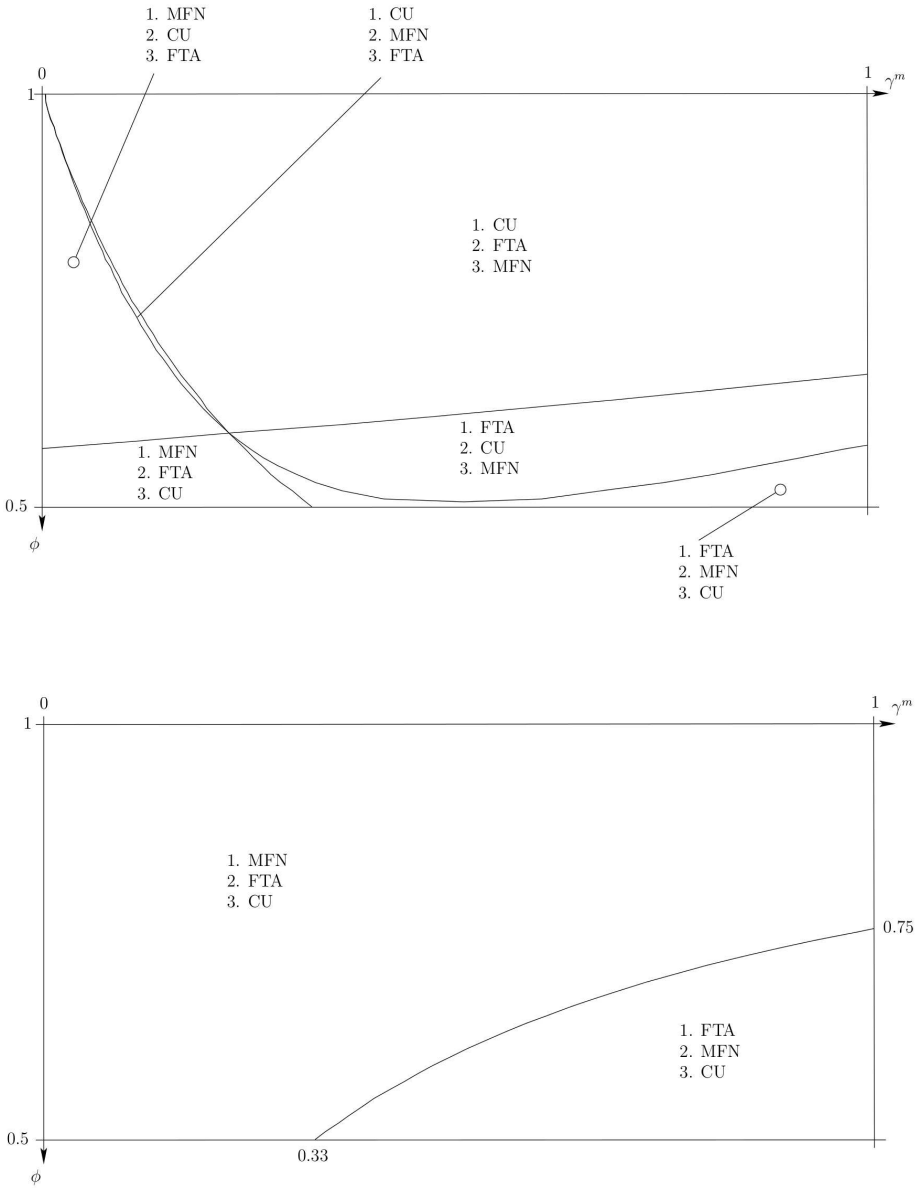

Figure 2: The median voter's rankings in country A (top) and B (bottom)

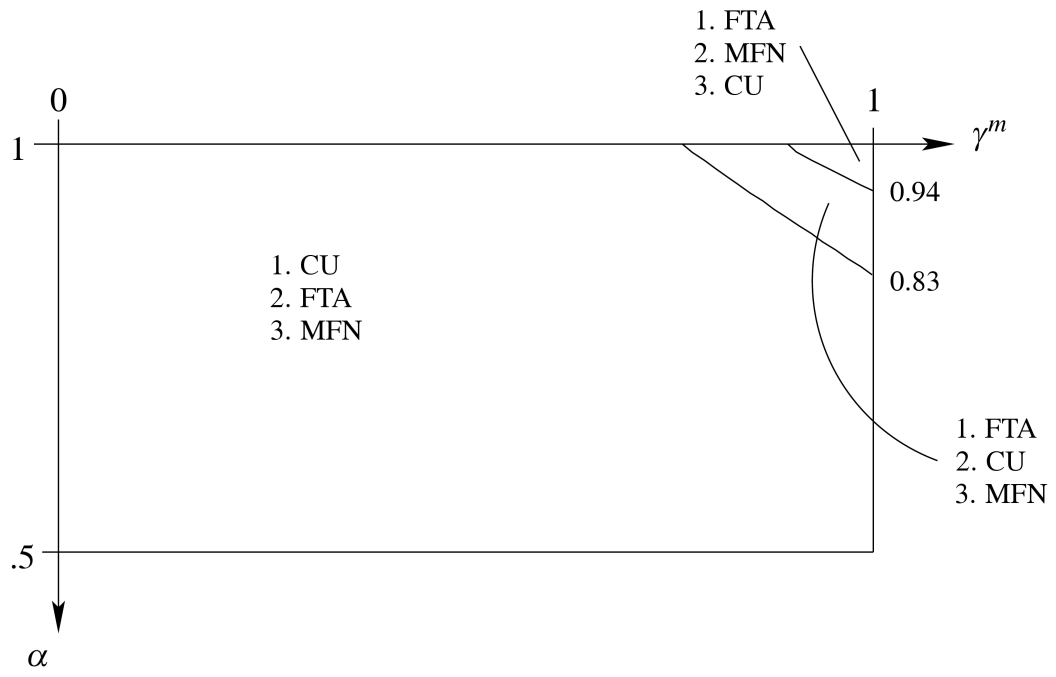

Figure 3: Welfare rankings

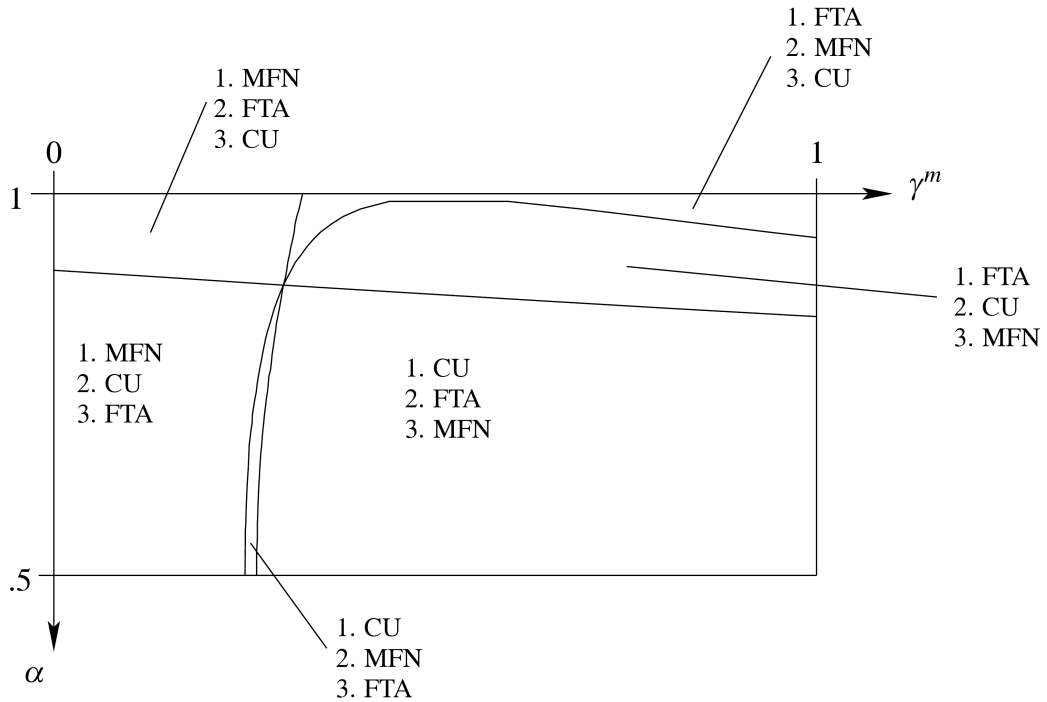

Figure 4: The median voter's rankings

Table 1: Descriptive Statistics

Main Variables	**(1)**	**(2)**	**(3)**	**(4)**
	Entire Sample	**PTA**	**FTA**	**CU**
Geographic Specialization (GEO)	43.87	28.04	32.41	21.28
	(28.60)	(29.02)	(33.95)	(16.96)
Inequality (INEQ)	41.15	38.18	40.14	35.15
	(8.52)	(8.09)	(7.89)	(7.44)
Trade-Imbalance (IMB)	0.69	0.46	0.49	0.39
	(0.33)	(0.32)	(0.32)	(0.31)
Matrix X Elements				
INTERD	0.13	0.37	0.30	0.47
	(0.14)	(0.20)	(0.17)	(0.19)
NATURAL	-8.64	-7.51	-7.80	-7.04
	(0.81)	(0.94)	(0.94)	(0.73)
DCONT	0.25	0.75	0.59	0.99
	(0.43)	(0.43)	(0.49)	(0.04)
REMOTE	8.92	8.73	8.81	8.59
	(0.22)	(0.33)	(0.31)	(0.32)
GDPSUM	11.02	11.25	11.28	11.23
	(0.79)	(0.85)	(0.73)	(0.99)
GDPSIM	-2.47	-1.73	-1.81	-1.62
	(1.75)	(1.09)	(1.16)	(0.99)
DKL	1.83	1.21	1.45	0.82
	(1.29)	(0.92)	(0.99)	(0.65)
SDKL	5.01	2.31	3.09	1.09
	(6.01)	(3.06)	(3.52)	(1.51)
DROWKL	1.16	1.21	1.12	1.35
	(0.58)	(0.59)	(0.58)	(0.59)
Number of observations	67740	5108	3101	2007

The table reports average values and standard deviations (in brackets). NATURAL is the natural logarithm of the inverse of the distance between countries in a country-pair; DCONT is a dummy variable equal to one if both countries in a country-pair are located in the same continent and zero otherwise; REMOTE is the country-pair simple average of the natural logarithm of the average of the distance between each country in a country-pair and its trade partners; GDPSUM is the natural logarithm of the sum of the total GDP of countries in a country-pair; GDPSIM is the natural logarithm of 1 minus the squared value of the share of each country's GDP in the total GDP of a country-pair; DKL is the absolute value of the difference of the log of the per-capita income for countries in a country-pair; SDKL is the squared value of DKL; DROWKL is the simple average of the absolute value of the difference between the log of the per-capita income of a country in a country-pair and the log of the average per-capita income of its trade partners. See Section 4.2 for the exact definitions of GEO, INEQ and IMB.

Table 2: Main Results

	Predicted Sign	(1)	(2)	(3)	(4)	(5)	Marginal Effects
PTA decision (selection)							
INEQ	-		-0.009**	-0.009**	-0.005*	-0.007**	-0.0005**
			(0.002)	(0.003)	(0.003)	(0.003)	(0.0002)
IMB	-		-0.393**	-0.279**	-0.234**	-0.202**	-0.0157**
			(0.029)	(0.042)	(0.037)	(0.038)	(0.0029)
Matrix X Elements							
INTERD	+	2.736**		3.161**	2.738**	2.957**	0.2301**
		(0.138)		(0.139)	(0.127)	(0.161)	(0.0125)
NATURAL	+	0.353**		0.280**	0.257**	0.257**	0.0200**
		(0.028)		(0.029)	(0.029)	(0.030)	(0.0023)
DCONT	+	0.704**		0.631**	0.692**	0.686**	0.0534**
		(0.042)		(0.044)	(0.039)	(0.039)	(0.0031)
REMOTE	+	-0.056		-0.058	-0.163*	-1.688*	-0.0131**
		(0.090)		(0.092)	(0.086)	(0.088)	(0.0068)
GDPSUM	+	1.034**		0.909**	0.905**	0.992**	0.0772**
		(0.053)		(0.051)	(0.046)	(0.081)	(0.0062)
GDPSIM	+	0.251**		0.217**	0.238**	0.276**	0.0214**
		(0.047)		(0.051)	(0.045)	(0.045)	(0.0035)
DKL	+	0.309**		0.342**	0.22**	0.207**	0.0161**
		(0.059)		(0.062)	(0.055)	(0.057)	(0.0044)
SDKL	-	-0.044**		-0.059**	-0.051**	-0.061**	-0.0047**
		(0.016)		(0.017)	(0.016)	(0.016)	(0.0012)
DROWKL	-	0.176**		0.128**	0.261**	0.164**	0.0128**
		(0.047)		(0.048)	(0.044)	(0.052)	(0.0041)
CU-FTA decision (latent)							
GEO	+				0.005**	0.006**	0.0019**
					(0.001)	(0.001)	(0.0004)
# of Obs.		67740	67740	67740	67740	67740	67740
# of Obs. with PTAs		5108	5108	5108	5108	5108	5108
Pseudo-R^2		0.495	0.111	0.514			
Bayesian Criterion		18472	32258	17811	23000	22773	22773
Wald indep. Equations					632.71**	588.56**	588.56**
Year Fixed Effects		No	No	No	No	Yes	Yes

Specification (1) - (3) are estimated using a standard probit model while specifications (4) - (5) are estimated using a probit model with sample selection. Standard errors for estimated coefficients are shown in parentheses and are clustered at the country-pair level. Each specification includes time-invariant averages of the control variables. "**" and "*" denote significance at 5 and 10 percent levels respectively.

Table 3: Additional drivers of PTA Formation

	Predicted Sign	(1) INEQ Dummy	(2) INEQ Squared	(3) INEQ in latent	(4) # other members	(5) # other members	(6) Lake and Yildiz (2016)	(7) Melatos and Woodland (2007)	(8) IMB in latent	(9) Country FE
PTA decision (selection)										
IMB	-	-0.194**	-0.202**	-0.202**	-0.210**	-0.237**	-0.215**	-0.218**	-0.261**	-0.015**
		(0.037)	(0.038)	(0.039)	(0.043)	(0.044)	(0.039)	(0.039)	(0.042)	(0.003)
INEQ	-	-0.260**	0.035**	-0.013**	-0.008**	-0.009**	-0.008**	-0.008**	-0.007**	-0.002**
		(0.039)	(0.017)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.0001)
$INEQ^2$			-0.0005**							
			(0.0002)							
Matrix X Elements										
INTERD	+	2.823**	2.904**	3.056**	2.915**	3.299**	3.079**	3.103**	2.941**	1.173**
		(0.160)	(0.161)	(0.164)	(0.164)	(0.180)	(0.164)	(0.163)	(0.158)	(0.154)
NATURAL	+	0.255**	0.259**	0.263**	0.201**	0.134**	0.223**	0.223**	0.249**	0.039**
		(0.030)	(0.030)	(0.030)	(0.021)	(0.035)	(0.029)	(0.029)	(0.029)	(0.002)
DCONT	+	0.696**	0.685**	0.673**	0.133**	0.579**	0.708**	0.712**	0.691**	0.076**
		(0.038)	(0.039)	(0.039)	(0.034)	(0.044)	(0.039)	(0.039)	(0.037)	(0.003)
REMOTE	+	-0.106	-0.160*	-0.142**	-0.096	-0.069	-0.163*	-0.153*	-0.196**	-0.111**
		(0.085)	(0.089)	(0.088)	(0.096)	(0.099)	(0.090)	(0.090)	(0.087)	(0.009)
GDPSUM	+	0.975**	0.979**	0.999**	1.222**	1.276**	1.025**	1.013**	0.939**	0.031**
		(0.082)	(0.080)	(0.083)	(0.094)	(0.099)	(0.083)	(0.083)	(0.082)	(0.008)
GDPSIM	+	0.265**	0.271**	0.265**	0.347**	0.358**	0.284**	0.185**	0.269**	-0.002
		(0.044)	(0.045)	(0.047)	(0.047)	(0.051)	(0.047)	(0.050)	(0.044)	(0.003)
DKL	+	0.203**	0.204**	0.224**	0.243**	0.281**	0.228**	0.247**	0.197**	0.032**
		(0.056)	(0.057)	(0.059)	(0.063)	(0.066)	(0.059)	(0.058)	(0.057)	(0.004)
SDKL	-	-0.060**	-0.061**	-0.062**	-0.052**	-0.052**	-0.062**	-0.067**	-0.059**	-0.011**
		(0.016)	(0.016)	(0.016)	(0.017)	(0.017)	(0.016)	(0.016)	(0.016)	(0.001)
DROWKL	-	0.141**	0.157**	0.132**	0.113**	-0.064	0.140**	0.143**	0.146**	0.010**
		(0.052)	(0.052)	(0.053)	(0.055)	(0.057)	(0.053)	(0.053)	(0.050)	(0.004)
# other members	+-				0.201**	0.159**				
					(0.021)	(0.026)				
CU-FTA decision (latent)										
GEO	+	0.006**	0.006**	0.021**	0.006**	0.010**	0.006**	0.006**	0.011**	0.002**
		(0.001)	(0.001)	(0.007)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.0002)
GEO*INEQ	-			-0.0004**						
				(0.0002)						
INEQ	+-			-0.018**						
				(0.009)						
NATURAL	-						-0.224**	-0.234**		
							(0.043)	(0.043)		
GDPSIM	-							-0.730**		
								(0.086)		
GEO*IMB	+-								-0.008**	
									(0.002)	
IMB	+-								-0.117	
									(0.112)	
Est. Errors (sel.)										0.837**
										(0.027)
# other member	+-					-0.816**				
						(0.027)				
# of obs.		67740	67740	67740	67740	76740	67740	67740	67740	67740
# of obs. PTAs		5108	5108	5108	5108	5108	5108	5108	5108	5108
Wald indep.		591.09**	600.26**	346.36**	446.97**	225.56**	330.82**	353.65**	346.36**	
Year FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Probit models with sample selection. Standard errors are shown in parentheses and are clustered at the country-pair level. Each specification includes time-invariant averages of the control variables. "**" and "*" denote significance at 5 and 10 percent levels, respectively. The binary measure of income inequality used in column (1) defines a country-pair with a low degree of income equality if the pair's highest Gini coefficient is greater than the 75 percentile of the distribution of income inequality across country-pairs for a given year.

Table 4: Alternative measures of key variables

	Predicted Sign	(1) Mattevi (2005)	(2) EL (2008)	(3) 10-Year Lags	(4) Alt. IMB Deficit	(5) Alt. IMB Diff.	(6) Alt. IMB Manuf.	(7) GEO Manuf.	(8) BJ Interd. (2012)
PTA decision (selection)									
INEQ	-	-0.012**	-0.005**	-0.008**	-0.005**	-0.006*	-0.007**	-0.007**	-0.011**
		(0.003)	(0.002)	(0.004)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
IMB	-	-0.206**	-0.069*	-0.125**	-0.038**	-0.081**	-0.149**	-0.203**	-0.133**
		(0.038)	(0.029)	(0.040)	(0.018)	(0.034)	(0.039)	(0.037)	(0.037)
Matrix X Elements									
INTERD	+	3.083**	0.145**	0.201**	2.762**	2.816**	2.849**	2.940**	9.457**
		(0.173)	(0.027)	(0.179)	(0.154)	(0.167)	(0.172)	(0.160)	(2.419)
NATURAL	+	0.282**	0.355**	0.337**	0.284**	0.275**	0.294**	0.259**	0.401**
		(0.028)	(0.025)	(0.031)	(0.029)	(0.030)	(0.029)	(0.030)	(0.031)
DCONT	+	0.575**	0.476**	0.663**	0.751**	0.656**	0.658**	0.685**	0.749**
		(0.039)	(0.035)	(0.039)	(0.037)	(0.038)	(0.039)	(0.038)	(0.038)
REMOTE	+	-0.129	-0.733**	-0.152**	-0.271**	-0.336**	-0.277**	-0.173*	-0.633**
		(0.087)	(0.071)	(0.091)	(0.089)	(0.092)	(0.092)	(0.088)	(0.089)
GDPSUM	+	0.844**	0.471**	0.671**	1.231**	0.814**	1.389**	0.994**	0.512**
		(0.080)	(0.052)	(0.066)	(0.077)	(0.087)	(0.094)	(0.081)	(0.062)
GDPSIM	+	0.248**	0.109**	0.125**	0.356**	0.276**	0.382**	0.278**	0.178**
		(0.048)	(0.030)	(0.047)	(0.042)	(0.046)	(0.051)	(0.045)	(0.042)
DKL	+	0.093*	0.282**	0.298**	0.181**	0.199**	0.200**	0.198**	0.162**
		(0.056)	(0.046)	(0.063)	(0.052)	(0.062)	(0.063)	(0.057)	(0.055)
SDKL	-	-0.049	-0.091**	-0.075**	-0.043**	-0.055**	-0.053**	-0.060**	-0.062**
		(0.016)	(0.011)	(0.018)	(0.015)	(0.016)	(0.017)	(0.016)	(0.015)
DROWKL	-	0.289**	0.315**	-0.172**	0.211**	0.183**	0.169**	0.169**	0.185**
		(0.054)	(0.038)	(0.057)	(0.045)	(0.052)	(0.055)	(0.052)	(0.048)
CU-FTA decision (latent)									
GEO	+	0.004**	0.003**	0.006**	0.005**	0.005**	0.005**	0.010**	0.005**
		(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
# of obs.		67780	67415	58270	73579	67314	58932	67740	65218
# of obs. PTAs		4820	6945	4749	5164	4862	5011	5108	4707
Wald indep.		506.53**	295.67**	476.47**	600.97**	511.52**	484.80**	592.21**	363.90**
Year FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Probit models with sample selection. Standard errors for estimated coefficients are shown in parentheses and are clustered at the country-pair level. Each specification includes time-invariant averages of the control variables. "**" and "*" denote significance at 5 and 10 percent level respectively.

Table 5: Different Sample Structures

	Predicted Sign	(1) Years 2000s	(2) NO EU	(3) NO EU, ASEAN and PAN ARAB	(4) NO EU, NAFTA, and ASEAN
PTA decision (selection)					
INEQ	-	-0.007*	-0.009**	-0.005*	-0.009**
		(0.004)	(0.003)	(0.003)	(0.003)
IMB	-	-0.161**	-0.169**	-0.140**	-0.166**
		(0.043)	(0.039)	(0.040)	(0.039)
Matrix X Elements					
INTERD	+	3.895**	2.507**	2.409**	2.519**
		(0.196)	(0.175)	(0.187)	(0.185)
NATURAL	+	0.309**	0.289**	0.237**	0.270**
		(0.031)	(0.032)	(0.033)	(0.032)
DCONT	+	0.601**	0.536**	0.557**	0.519**
		(0.039)	(0.042)	(0.042)	(0.041)
REMOTE	+	0.171**	0.086	0.096	-0.007**
		(0.083)	(0.091)	(0.092)	(0.093)
GDPSUM	+	0.648**	1.017**	1.068**	0.974**
		(0.089)	(0.086)	(0.089)	(0.086)
GDPSIM	+	0.295**	0.253**	0.316**	0.265**
		(0.053)	(0.047)	(0.049)	(0.048)
DKL	+	0.067	0.251**	0.264**	0.258**
		(0.069)	(0.060)	(0.064)	(0.061)
SDKL	-	-0.049**	-0.065**	-0.050**	-0.067**
		(0.018)	(0.016)	(0.018)	(0.017)
DROWKL	-	0.334**	0.083**	0.158**	0.098**
		(0.067)	(0.055)	(0.056)	(0.056)
CU-FTA decision (latent)					
GEO	+	0.004**	0.002**	0.003**	0.003**
		(0.001)	(0.001)	(0.001)	(0.001)
# of obs.		38443	66573	66188	66406
# of obs. PTAs		4470	3941	3556	3774
Wald indep.		497.51**	212.3**	222.61**	225.47**
Year FE		Yes	Yes	Yes	Yes

Probit models with sample selection. Standard errors for estimated coefficients are shown in parentheses and are clustered at the country-pair level. Each specification includes time-invariant averages of the control variables. "**" and "*" denotes significance at 5 and 10 percent level respectively.

Online Appendix

In this appendix we provide analytical details used in deriving Lemmata 1-4 and Propositions 1-2. The expressions related to the value of the indirect utility functions of the average and median voters were derived using Mathematica and the original file using this software is available upon request.

Appendix A

Proof of Lemma 1. We start by solving, for a given $\hat{\gamma}_A$, the MFN tariff determination problem. The first order conditions for problem 5 are given by:

$$\begin{aligned} -\frac{\partial p_A^i}{\partial t_A^i} x_A^i + x_{F,A}^i + t_A^i \frac{\partial x_{F,A}^i}{\partial t_A^i} + \widehat{\gamma}_A \frac{\partial \pi_{A,A}^i}{\partial t_A^i} &= 0 \qquad \text{for } i = 1, ..., n\phi \\ -\frac{\partial p_A^j}{\partial t_A^j} x_A^j + \left(x_{F,A}^j + x_{B,A}^j\right) + t_A^j \left(\frac{\partial x_{F,A}^j}{\partial t_A^j} + \frac{\partial x_{B,A}^j}{\partial t_A^j}\right) &= 0 \qquad \text{for } j = n\phi + 1, ..., n \end{aligned} \tag{14}$$

Using equilibrium prices and quantities from (3) and (4) we obtain

$$\begin{aligned} t_A^{MFN,i} &= \frac{H\left(1 + 2\widehat{\gamma}_A\right)}{11 - 2\widehat{\gamma}_A} \qquad \text{for } i = 1, ..., n\phi \\ t_A^{MFN,j} &= \frac{H}{4} \qquad \text{for } j = n\phi + 1, ..., n \end{aligned} \tag{15}$$

Importantly, equation (15) indicates that the equilibrium tariffs for country A depend only on the identity of that country's representative and on whether the country produces or not that particular good. Moreover, they do not depend on ϕ, i.e. the share of industries in which country A produces and exports goods.[38] As for the choice of the representative in stage 2 of the game, as shown by Facchini, Silva and Willmann (2013), the median voter cannot do better than representing the country himself, i.e. $\widehat{\gamma}_A = \gamma^m$, as this maximizes equation (6). The equilibrium MFN tariffs are then:

$$\begin{aligned} t_A^{MFN,i} &= \frac{H\left(1 + 2\gamma^m\right)}{11 - 2\gamma^m} \qquad \text{for } i = 1, ..., n\phi \\ t_A^{MFN,j} &= \frac{H}{4} \qquad \text{for } j = n\phi + 1, ..., n \end{aligned} \tag{16}$$

[38] Notice that expressions (15) rely on the assumption that the marginal cost of production equals zero. Otherwise, the numerator of expression $t_A^{MFN,i}$ would be the product of the difference between H and the marginal cost with expression $1+2\widehat{\gamma}_A$. A similar rational applies to the other tariffs discussed in this appendix.

We can now turn to the case of FTAs. In this case, free trade prevails between member countries ($t_{A,B}^{FTA,i} = t_{B,A}^{FTA,i} = 0$) and prospective members can set external tariffs independently. The solution to problem 5 is given by:

$$\begin{aligned} t_{F,A}^{FTA,i} &= \frac{H\left(1+2\widehat{\gamma}_A\right)}{11-2\widehat{\gamma}_A} \qquad \text{for } i = 1,...,n\phi \\ t_{F,A}^{FTA,j} &= \frac{H}{11} \qquad \text{for } j = n\phi+1,...,n \end{aligned}$$

Also in this case, the median voter in each country does not delegate power for the same reasons discussed for the MFN regime. Thus, the equilibrium external tariffs in the FTA case are given by:

$$\begin{aligned} t_{F,A}^{FTA,i} &= \frac{H\left(1+2\gamma^m\right)}{\left(11-2\gamma^m\right)} \qquad \text{for } i = 1,...,n\phi \\ t_{F,A}^{FTA,j} &= \frac{H}{11} \qquad \text{for } j = n\phi+1,...,n \end{aligned} \tag{17}$$

Comparing expressions (16) and (17) establishes the second part of Lemma 1. ■

Proof of Lemma 2. The first order conditions of problem 7 for goods $i = 1,...,n\phi$ are given by

$$-\frac{\partial p_A^i}{\partial t^i}x_A^i + x_{F,A}^i + t^i\frac{\partial x_{F,A}^i}{\partial t^i} + \widehat{\gamma}_A\left(\frac{\partial \pi_{A,A}^i}{\partial t^i} + \frac{\partial \pi_{A,B}^i}{\partial t^i}\right) - \frac{\partial p_B^i}{\partial t^i}x_B^i + \ x_{F,B}^i + t^i\frac{\partial x_{F,B}^i}{\partial t^i} = 0 \tag{18}$$

and for goods $j = n\phi+1,...,n$ by

$$-\frac{\partial p_A^j}{\partial t^j}x_A^j + x_{F,A}^j + t^j\frac{\partial x_{F,A}^j}{\partial t^j} - \frac{\partial p_B^j}{\partial t^j}x_B^j + x_{F,B}^j + t^j\frac{\partial x_{F,B}^j}{\partial t^j} + \widehat{\gamma}_B\left(\frac{\partial \pi_{B,A}^j}{\partial t^j} + \frac{\partial \pi_{B,B}^j}{\partial t^j}\right) = 0 \tag{19}$$

Using the symmetry of the demand structure between A and B, we have that $x_A^i = x_B^i$, $x_A^j = x_B^j$, $\pi_{A,A}^i = \pi_{A,B}^i$, $\pi_{B,A}^j = \pi_{B,B}^j$, and $\frac{\partial x_{F,A}^i}{\partial t^i} = \frac{\partial x_{F,B}^i}{\partial t^i}$. We therefore obtain the following common external tariffs:

$$\begin{aligned} t^{CU,i} &= \frac{H\left(1+2\widehat{\gamma}_A\right)}{\left(11-2\widehat{\gamma}_A\right)} \qquad \text{for } i = 1,...,n\phi \\ t^{CU,j} &= \frac{H\left(1+2\widehat{\gamma}_B\right)}{\left(11-2\widehat{\gamma}_B\right)} \qquad \text{for } j = n\phi+1,...,n \end{aligned} \tag{20}$$

It is clear from (20) that only the identity of country A's representative matters in deter-

mining the equilibrium common external tariff in goods 1 through $n\phi$, while only the identity of country B's representative matter in determining the common external tariff for the remaining goods. Importantly, the share of products produced and exported by a prospective member country does not affect the common trade policy. Turning now to the selection of the representatives, as shown by Facchini, Silva and Willmann (2013), strategic delegation occurs and in particular we have that:

$$\widehat{\gamma}_A = \widehat{\gamma}_B = 2\gamma^m. \tag{21}$$

Substituting equation (21) in equation (20) we obtain the common external tariff:

$$t^{CU,i} = t^{CU,j} = \frac{H\left(1+4\gamma^m\right)}{\left(11-4\gamma^m\right)} \quad \text{for any } i \text{ and } j \tag{22}$$

which implies that common external tariffs are higher than external tariffs under an FTA. ■

Appendix B

Proof of Lemma 3 We characterize welfare in each country by assessing the equilibrium value of the average voter's indirect utility function under the different trade policy regimes (i.e. using equilibrium tariffs – see expressions (16), (17) and (22)), along with quantities and prices (see expressions (3) and (4)). Let the difference in welfare between a PTA and the MFN regimes for country A be given by $\Delta v_A(\mathbf{t}^{MFN},\mathbf{t}^{PTA},\overline{\gamma})= v_A(\mathbf{t}^{PTA},\overline{\gamma}) - v_A(\mathbf{t}^{MFN},\overline{\gamma})$, and similarly for B. We can calculate these changes as follows:

$$\begin{aligned}
\Delta v_A(t^{MFN},t^{CU},\overline{\gamma}) &= \begin{array}{c} H^2 n[\gamma^{m^3}(8576-3424\phi)+1452\gamma^m(4+7\phi)+ \\ 64\gamma^{m^4}(-12+11\phi)-12\gamma^{m^2}(2068+35\phi)+ \\ 121(-44+179\phi)]/[16(121-66\gamma^m+8\gamma^{m^2})^2] \end{array} \\
\Delta v_A(t^{MFN},t^{FTA},\overline{\gamma}) &= \frac{H^2 n(-44+179\phi)}{1936} \\
\Delta v_B(t^{MFN},t^{CU},\overline{\gamma}) &= \begin{array}{c} H^2 n[16335-21659\phi-1452\gamma^m(-11+7\phi) \\ -64\gamma^{m^4}(1+11\phi)+12\gamma^{m^2}(-2103+35\phi)+ \\ 32\gamma^{m^3}(161+107\phi)]/[16(121-66\gamma^m+8\gamma^{m^2})^2] \end{array} \\
\Delta v_B(t^{MFN},t^{FTA},\overline{\gamma}) &= \frac{H^2 n(135-179\phi)}{1936}
\end{aligned} \tag{23}$$

where it is clear that the expressions for country B's change in welfare can be obtained by replacing ϕ with $1-\phi$ in the corresponding expressions for country A. To establish Lemma 3 notice that $\frac{\partial \Delta v_A(\mathbf{t}^{MFN},\mathbf{t}^{FTA},\overline{\gamma})}{\partial \phi} = -\frac{\partial \Delta v_B(\mathbf{t}^{MFN},\mathbf{t}^{FTA},\overline{\gamma})}{\partial \phi} = 179H^2n/1936$ which establishes the result for the case of the FTA. Turning to the CU, note that:

$$\frac{\partial \Delta v_A(\mathbf{t}^{MFN},\mathbf{t}^{CU},\overline{\gamma})}{\partial \phi} = \frac{H^2n[21659+10164\gamma^m-420\gamma^{m^2}-3424\gamma^{m^3}+704\gamma^{m^4}]}{16(121-66\gamma^m+8\gamma^{m^2})^2} > 0 \qquad (24)$$

since $\gamma^m \epsilon [0,1]$. Furthermore, $\frac{\partial \Delta v_B(\mathbf{t}^{MFN},\mathbf{t}^{CU},\overline{\gamma})}{\partial \phi} = -\frac{\partial \Delta v_A(\mathbf{t}^{MFN},\mathbf{t}^{CU},\overline{\gamma})}{\partial \phi} < 0$, thus establishing the result. ■

Constructing Figures 1-4. We focus on the case of Figure 1 to provide a map on how to construct the Figures 1-4 in the main text. In particular, we focus on the top panel of Figure 1, which describes the welfare ranking of the trade policy regimes for the prospective member country running a trade surplus (country A). A similar approach can be used to construct the bottom panel of Figure 1 as well as to construct Figures 2-4. As explained in the paper, the welfare ranking depends on how the different trade policy regimes affect the indirect utility function of the average voter.

First, we need to compare welfare under the FTA and CU regimes and then consider each PTA against the MFN regime. As for the former, we have the following expression:

$$\Delta v_A(t^{FTA},t^{CU},\overline{\gamma}) = \frac{H^2n32\gamma^m[7986\phi+16\gamma^{m^3}(-11+9\phi)-121\gamma^m(44+19\phi)-}{88\gamma^{m^2}(-22+5\phi)-121(44+179\phi)]/[121(121-66\gamma^m+8\gamma^{m^2})^2} \qquad (25)$$

As for the other two comparisons, they are described in expression (23). As can be seen from Figure 1, the difference in welfare levels between trade regimes depends on the parameters capturing the degree of bilateral trade imbalance (ϕ) and the degree of income inequality (γ^m).[39] We proceed by varying the parameter capturing the degree of trade imbalance (vertical axis) between 0.5 and 1.0 using 0.01 incremental values, and then calculate, for each value of this parameter, the threshold in terms of the degree of income inequality (horizontal axis) needed to set the difference in welfare level between the two trade regimes to zero. The table below shows the thresholds for the three pair-wise welfare comparisons:

[Table A2 – Welfare Ranking]

[39] In general, the sign of the difference between the value of the indirect utility functions does not depend on H since this parameter multiplies all other terms of these expressions.

In line with the top panel of Figure 1, the information provided by Table A2 indicates that, in the presence of a balanced distribution of market access ($\phi = 0.5$), the FTA welfare dominates the CU if inequality is sufficiently low ($\gamma^m > 0.76$) as indicated by column 1. Similarly, the presence of balanced bilateral trade also implies that a CU welfare dominates the MFN regime if the degree of income equality is sufficiently high ($\gamma^m > 0.89$) as indicated by column 2. As we move up along the vertical axis, towards a more unequal distribution of market access, we notice from Table A2 that the threshold indicating the degree of income equality needed so that the FTA (CU) welfare dominates the CU (MFN) also rises, implying increasing values for the parameter γ^m. We can then conclude that above a certain degree of unbalanced market access, the CU regime welfare dominates the FTA regime ($\phi > 0.66$) as well as the MFN regime ($\phi > 0.57$). Varying the degree of trade imbalances allows us then to derive the two lines found in the bottom right corner of Figure 1, separating areas where the average voter prefers the CU regime over the FTA regime (and vice-versa), as well as separating areas where the average voter either prefers the CU regime or the MFN regime. In line with Figure 1, the FTA regime welfare dominates the MFN regime regardless of the level of income equality according to column 3 of Table A2.

The bottom panel of Figure 1 follows the same approach used in the top panel of that figure while focusing on the case of the country running a bilateral trade deficit (country B). In the case of Figure 2, we use similar approach but the trade regime comparison relies on the indirect utility function of the median (rather than the average) voter since we consider the political viability of trade agreements in that case. Similarly, Figures 3 and 4 follow the approach used in Figures 1 and 2, respectively. The only difference between these two groups of figures is that Figures 3 and 4 rely on setting different values of the degree of geographic specialization (α) in order to obtain the threshold of the parameter measuring the degree of income inequality (γ^m) since we assume the presence of balanced trade in Figures 3 and 4.

Appendix C

Proof of Proposition 1 We characterize the political viability of a PTA by assessing the value of the median voter's indirect utility function under the different trade policy regimes (i.e. using equilibrium tariffs – see expressions (16), (17) and (22)), along with quantities and prices (see expressions (3) and (4)). Let the change in country A's median voter's indirect utility be given by $\Delta v_A(\mathbf{t}^{MFN},\mathbf{t}^{PTA},\gamma^m)= v_A(\mathbf{t}^{PTA},\gamma^m) - v_A(\mathbf{t}^{MFN},\gamma^m)$ as outlined in expression (9),

and similarly for country B. We can calculate these changes as follows:

$$\begin{aligned}\Delta v_A(t^{MFN}, t^{CU}, \gamma^m) &= H^2 n[\gamma^{m^3}(384-736\phi)+484(-1+\phi)+32\gamma^{m^4}\phi \\ &\quad +55\gamma^m(8+19\phi)+\gamma^{m^2}(-2176+2362\phi)] \\ &\quad /[16(11-4\gamma^m)^2(11-2\gamma^m)] \\ \Delta v_A(t^{MFN}, t^{FTA}, \gamma^m) &= \frac{H^2 n[-44+(44+135\gamma^m)\phi]}{1936}\end{aligned} \tag{26}$$

where the expressions for country B can be obtained by replacing ϕ with $1-\phi$ in expressions (26). It is easy to show that $\frac{\partial \Delta v_A(\mathbf{t}^{MFN}, \mathbf{t}^{FTA}, \gamma^m)}{\partial \phi} = -\frac{\partial \Delta v_B(\mathbf{t}^{MFN}, \mathbf{t}^{FTA}, \gamma^m)}{\partial \phi} = H^2 n(44+135\gamma^m)/1936$ which establishes the result for FTAs. Turning to the CU, note that

$$\frac{\partial \Delta v_A(\mathbf{t}^{MFN}, \mathbf{t}^{CU}, \gamma^m)}{\partial \phi} = \frac{H^2 n[484+1045\gamma^m+2362\gamma^{m^2}-736\gamma^{m^3}+32\gamma^{m^4}]}{16(11-4\gamma^m)^2(11-2\gamma^m)} > 0 \tag{27}$$

since $\gamma^m \epsilon [0,1]$. Furthermore, $\frac{\partial \Delta v_B(\mathbf{t}^{MFN}, \mathbf{t}^{CU}, \gamma^m)}{\partial \phi} = -\frac{\partial \Delta v_A(\mathbf{t}^{MFN}, \mathbf{t}^{CU}, \gamma^m)}{\partial \phi} < 0$, thus establishing the result. ■

Appendix D

In this appendix, we proceed in two steps. First, we prove the claim in footnote 10, i.e. that Lemma 1 generalizes to varying degrees of geographic specialization. Second, we show the proof of Lemma 4, which is related to common external tariffs under varying degrees of geographic specialization.

Focusing on country A (the results are analogous for country B), and following the same logic as in section 3.1, it follows immediately that no strategic delegation will occur in equilibrium in the MFN regime. As a result, the MFN tariff is given by

$$\begin{aligned}t_A^{MFN,i} &= \frac{H(1+2\alpha\gamma^m)}{4+7\alpha-2\alpha\gamma^m(2-\alpha)} \qquad \text{for } i=1,...,\frac{n}{2} \\ t_A^{MFN,j} &= \frac{H(1+2(1-\alpha)\gamma^m)}{11-7\alpha-(1+\alpha)2(1-\alpha)\gamma^m} \qquad \text{for } j=\frac{n}{2}+1,...,n\end{aligned} \tag{28}$$

and the symmetric production structure of our model implies that $t_A^{MFN,i} = t_B^{MFN,j}$ and $t_A^{MFN,j} = t_B^{MFN,i}$. Note that as long as all goods are produced in both A and B, income inequality matters in determining the level of the MFN tariffs applied to all goods. Further-

more, if sectors are equally spread across the member countries ($\alpha = 1/2$), the tariffs applied on each good are identical.

We can now turn to the FTA regime. Also in this case, no strategic delegation occurs and the equilibrium tariffs are given by:

$$\begin{aligned} t_{F,A}^{FTA,i} &= \frac{H\left(1+2\alpha\gamma^m\right)}{11-2\alpha\gamma^m} \qquad \text{for } i=1,...,\frac{n}{2} \\ t_{F,A}^{FTA,j} &= \frac{H\left[1+2\left(1-\alpha\right)\gamma^m\right]}{\left[11-2\left(1-\alpha\right)\gamma^m\right]} \quad \text{for } j=\frac{n}{2}+1,...,n \end{aligned} \tag{29}$$

and given the symmetry of the model, $t_{F,A}^{FTA,i} = t_{F,B}^{FTA,j}$ and $t_{F,A}^{FTA,j} = t_{F,B}^{FTA,i}$, and if $\alpha = 1/2$ all the tariffs are identical. Comparing equations (28) to (29) establishes the result. ■

Proof of Lemma 4. The solution to problem 7 is given by the following first order conditions:

$$\begin{aligned} t^{CU,i} &= \frac{H\left\{1+2\left[\alpha\widehat{\gamma}_A+\left(1-\alpha\right)\widehat{\gamma}_B\right]\right\}}{\left\{11-2\left[\alpha\widehat{\gamma}_A+\left(1-\alpha\right)\widehat{\gamma}_B\right]\right\}} \quad \text{for } i=1,...,\frac{n}{2} \\ t^{CU,j} &= \frac{H\left\{1+2\left[\left(1-\alpha\right)\widehat{\gamma}_A+\alpha\widehat{\gamma}_B\right]\right\}}{\left\{11-2\left[\left(1-\alpha\right)\widehat{\gamma}_A+\alpha\widehat{\gamma}_B\right]\right\}} \quad \text{for } j=\frac{n}{2}+1,...,n \end{aligned} \tag{30}$$

It is clear from (30) that the greater the share of profits received by the elected representatives, the higher is the tariff applied to imports from the non-member country. Turning now to the selection of the representative, the solution to problem 8 is given by:

$$\widehat{\gamma}_A = 2\gamma^m\left(1-2\alpha+2\alpha^2\right) \tag{31}$$

and $\frac{\partial\widehat{\gamma}}{\partial\alpha} > 0$ if $\alpha > \frac{1}{2}$. Finally, the equilibrium common external tariffs are given by:

$$\begin{aligned} t^{CU,i} &= \frac{H\left[1+4\gamma^m\left(1-2\alpha+2\alpha^2\right)\right]}{\left[11-4\gamma^m\left(1-2\alpha+2\alpha^2\right)\right]} \qquad \text{for } i=1,...,\frac{n}{2} \\ t^{CU,j} &= \frac{H\left[1+4\gamma^m\left(1-2\alpha+2\alpha^2\right)\right]}{\left[11-4\gamma^m\left(1-2\alpha+2\alpha^2\right)\right]} \qquad \text{for } j=\frac{n}{2}+1,...,n \end{aligned} \tag{32}$$

Note that common external tariffs continue to be higher than the external tariffs under the FTA regime. ■

Appendix E

Proof of Proposition 2 We need to evaluate the change in the median voter's indirect utility from the FTA regime to the CU, which is given by the following expression:

$$\begin{aligned}\Delta v(t^{FTA}, t^{CU}, \gamma^m) \;=\; & \gamma^{m^2}[-20.8 + 291.2\alpha - 956.7\alpha^2 + 1331\alpha^3 - 665.6\alpha^4 + \\ & \gamma^m(-11.3 + 15.1\alpha + 45.4\alpha^2 - 121\alpha^3 + 60.5\alpha^4) + \\ & \gamma^{m^2}(5.5 - 35.8\alpha + 114.1\alpha^2 - 222.8\alpha^3 + 276.4\alpha^4 \\ & -198\alpha^5 + 66\alpha^6) + \gamma^{m^3}(-0.5 + 4\alpha - 14.8\alpha^2 + 31.5\alpha^3 - \\ & 40.8\alpha^4 + 30\alpha^5 - 10\alpha^6)]/[(5.5 - \gamma^m\alpha)^2(5.5 + \gamma^m(\alpha - 1))^2 \\ & (-1.4 + \gamma^m(0.5 - \alpha + \alpha^2))^2]\end{aligned} \tag{33}$$

with an identical expression applying to country B. We can now compute:

$$\begin{aligned}\partial\Delta v(t^{FTA}, t^{CU}, \gamma_A^m)/\partial\alpha \;=\; & \gamma^{m^2}[12110.3 - 79581.8\alpha + 166084\alpha^2 - 110723\alpha^3 + \\ & \gamma^m(-4718.3 + 35859\alpha - 79267.3\alpha^2 + 52844.9\alpha^3) + \\ & \gamma^{m^2}(-514.7 + 1887.3\alpha - 10809.2\alpha^2 + 34568\alpha^3 - 41177\alpha^4 + 16471\alpha^5) + \\ & \gamma^{m^3}(571.9 - 3889\alpha - 13102\alpha^2 - 25621\alpha^3 + 27659.8\alpha^4 - 15772.4\alpha^5 + \\ & 4658.5\alpha^6 - 1331\alpha^7) + \gamma^{m^4}(-124.8 + 952.9\alpha - 3380.4\alpha^2 + 6942.4\alpha^3 - \\ & 8621.3\alpha^4 + 6624.8\alpha^5 - 3176.3\alpha^6 + 907.5\alpha^7) + \gamma^{m^5}(12.4 - 107.3\alpha + \\ & 437.3\alpha^2 - 1089\alpha^3 + 1839.8\alpha^4 - 2260.5\alpha^5 + 2079\alpha^6 - 1386\alpha^7 + 594\alpha^8 \\ & -132\alpha^9) + \gamma^{m^6}(-0.5 + 5\alpha - 24\alpha^2 + 72.5\alpha^3 - 152.5\alpha^4 + 232.5\alpha^5 - \\ & 255.5\alpha^6 + 193\alpha^7 - 90\alpha^8 + 20\alpha^9)]/ \\ & [(5.5 - \gamma^m\alpha)^3(5.5 + \gamma^m(\alpha - 1))^3(-1.4 + \gamma^m(0.5 - \alpha + \alpha^2))^3]\end{aligned} \tag{34}$$

This derivative is never positive and becomes increasingly negative for high levels of geographic specialization (higher α) combined with high levels of income equality (high γ^m) (see Figure A1). ■

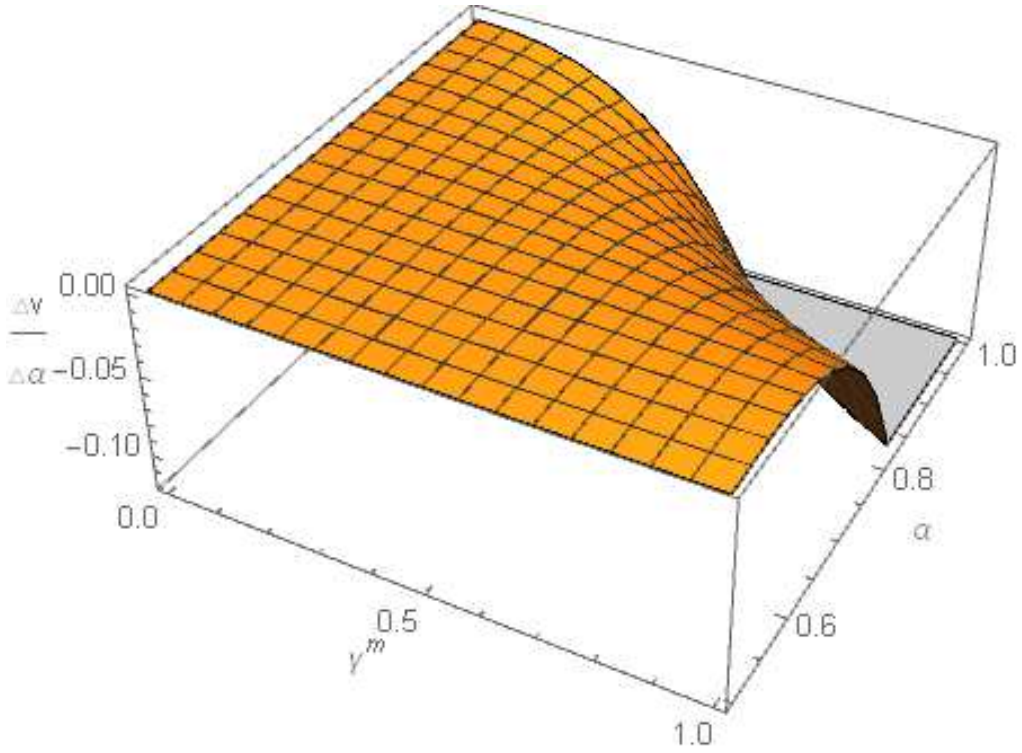

Figure A1: The behavior of $\partial \Delta v(t^{FTA}, t^{CU}, \gamma_A^m)/\partial \alpha$

Table A1: Definition of Control Variables

Matrix X	
Natural (+)	$\log(1/\text{distance}_{ab})$
DCONT (+)	equals one if countries in a country-pair are located in the same continent and zero otherwise
REMOTE (+)	$0.5\{\log[\sum_{k\neq b}\text{distance}_{ak}/(n_t-1)]+\log[\sum_{k\neq a}\text{distance}_{bk}/(n_t-1)]\}$
GDPSUM (+)	$\log(\text{GDP}_{at}+\text{GDP}_{bt})$
GDPSIM (+)	$\log(1-(\text{GDP}_{at}/(\text{GDP}_{at}+\text{GDP}_{bt}))^2-(\text{GDP}_{bt}/(\text{GDP}_{at}+\text{GDP}_{bt}))^2)$
DKL (+)	$\lvert\log(\text{GDPPC}_{at})-\log(\text{GDPPC}_{bt})\rvert$
SDKL (-)	$\lvert\log(\text{GDPPC}_{at})-\log(\text{GDPPC}_{bt})\rvert^2$
DROWKL (-)	$0.5\{\lvert\log[\sum_{k\neq a}\text{GDPPC}_{kt}/(n_t-1)]-\log(\text{GDPPC}_{at})\rvert+\lvert\log[\sum_{k\neq b}\text{GDPPC}_{kt}/(n_t-1)]-\log(\text{GDPPC}_{bt})\rvert\}$
INTERD (+)	average PTA membership of third countries as in Egger and Larch (2008)

*S*ign in parentheses indicates the predicted effect of a variable on the likelihood of PTA formation. The variable GDP_{at} represents country a's gross domestic product in year t, GDPPC_{at} represents country a's GDP per capita in year t, and FDI_{abt} represents the inward stock of FDI received by country a originating in country b at time t. REMOTE corresponds to the country-pair simple average of the natural logarithm of the average of the distance between each country in a country-pair and its trade partners; DROWKL corresponds to the simple average of the absolute value of the difference between the log of the per-capita income of a country in a country-pair and the log of the average per-capita income of its trade partners.

Table A2 --Welfare Ranking (Country A)

ϕ	(1) $\gamma^m_{FTA,CU}$	(2) $\gamma^m_{CU,MFN}$	(3) $\gamma^m_{FTA,MFN}$
0.50	0.7646	0.8874	FTA > MFN
0.51	0.7795	0.9027	FTA > MFN
0.52	0.7945	0.9178	FTA > MFN
0.53	0.8093	0.9328	FTA > MFN
0.54	0.8242	0.9476	FTA > MFN
0.55	0.839	0.9623	FTA > MFN
0.56	0.8537	0.9769	FTA > MFN
0.57	0.8684	0.9913	FTA > MFN
0.58	0.883	CU > MFN	FTA > MFN
0.59	0.8977	CU > MFN	FTA > MFN
0.6	0.9122	CU > MFN	FTA > MFN
0.61	0.9267	CU > MFN	FTA > MFN
0.62	0.9412	CU > MFN	FTA > MFN
0.63	0.9556	CU > MFN	FTA > MFN
0.64	0.9699	CU > MFN	FTA > MFN
0.65	0.9842	CU > MFN	FTA > MFN
0.66	0.9984	CU > MFN	FTA > MFN
0.67	CU > FTA	CU > MFN	FTA > MFN
≥0.68	CU > FTA	CU > MFN	FTA > MFN

Columns 1-3 report the share of income received by the median voter (γ^m) that makes the average voter indifferent between two trade regimes given the level of bilateral trade imbalance (ϕ). In this case, $\gamma^m_{FTA,CU}$ represents the share of the median voter income that makes the average voter indifferent between the FTA and the CU regimes. A similar notation applies in columns 2 and 3 to represent the share of the median voter income that makes the average voter indifferent between the CU and the MFN regimes as well as the FTA and the MFN regimes.