The Economic Impact of Deepening Trade Agreements*

Lionel Fontagné (Bank of France, CEPII and PSE)† Nadia Rocha (Word Bank)‡

Michele Ruta (Word Bank)§ Gianluca Santoni (CEPII)¶

Draft: September 21, 2022

Abstract

This paper explores the economic impacts of preferential trade agreements, conditional on their level of ambition. We cluster 278 agreements, encompassing 910 provisions over 18 policy areas and estimate the trade elasticity for the different clusters. We then use these elasticities in a series of general equilibrium counterfactual situations for endowment economies, revealing that deepening existing agreements (the intensive margin of regional integration) could boost world trade by 3.9 percent and world GDP by 0.9 percent. The expected gains from deepening agreements within or across regions vary depending on the initial depth of agreements and the size of regional markets.

Key Words: Preferential Trade Agreements; Deep Integration; Structural Gravity; General Equilibrium.

JEL Codes: F14, F15.

---

*The views expressed in this paper are the authors' and do not necessarily reflect those of the institutions they belong to. Lionel Fontagné thanks the support of the EUR grant ANR-17-EURE-0001. A previous version circulated under the title "A General Equilibrium Assessment of the Economic Impact of Deep Trade Agreements". We thank seminar participants at the World Bank Economics of Deep Trade Agreements Seminar Series and two anonymous referees for insightful comments. The data underlying this article are available at Mendeley with Reserved DOI: 10.17632/d63grb8csw.1

†Banque de France, 31 rue Croix des Petits-Champs, 75049 PARIS cedex 01 (France). Email: lionel.fontagne@banque-france.fr.

‡Email: nadia.rocha@worldbank.org.

§Email: michele.ruta@worldbank.org.

¶CEPII, 20 avenue de Segur 75007 Paris. Email: gianluca.santoni@cepii.fr

Introduction

The content of Preferential Trade Agreements (PTAs) has largely changed over time. Trade agreements in the 1950s focused on few policy areas, mostly regulating border measures such as tariffs and quotas, and included a limited number of regulatory requirements and commitments in these areas. While many modern PTAs deal primarily with border measures, a growing number of the recent trade agreements, like the Comprehensive and Progressive Agreement for Trans-Pacific Partnership (CPTPP) or the African Continental Free Trade Area (AfCFTA), are "deep"—i.e. they are complex legal documents covering a large array of border and behind the border policies and including liberalizing commitments, transparency, enforcement and other regulatory requirements. Despite the heterogeneous content of modern trade agreements, the large body of economic literature on PTAs mostly relies on dummy variables to identify their trade effects (Limão 2016). This approach fails to capture the multidimensional nature of the depth of trade agreements and can lead to measurement error bias in assessing their economic effects.

This paper takes a novel look at the economic impact of PTAs taking into account the *depth* of these agreements. We build on the recent strand of the structural gravity literature (Anderson & Van Wincoop 2003) and explore the uneven economic impacts deriving from the diverse provisions contained in PTAs—the *intensive margin* of regional integration. Different provisions have different effects on trade depending on how they affect trade costs. They may also affect differently PTA members and non-members. For instance, while discriminatory provisions such as rules on antidumping duties or export taxes ease restrictions to trade between members only, non-discriminatory provisions such as rules on subsidies or competition policy reduce trade costs for both members and non-members by allowing all foreign firms to benefit from undistorted competition in the PTA members' markets. In other words, the varying content of PTAs may be an important determinant of the uneven trade effects observed in Baier, Yotov & Zylkin (2019). The purpose of this paper is to provide a new method to assess these differential effects and to offer a quantification using new data on the content of PTAs.

The quantification analysis relies on a comprehensive characterization of the provisions included in PTAs based on a new database compiled by the World Bank (Mattoo, Rocha & Ruta 2020). We use information on all policy domains (excluding tariffs) covering objectives, substantive commitments, regulatory requirements and enforcement procedures included in legal texts and annexes of the 278 PTAs in force and notified to the WTO up to 2018. Examples of policy areas covered in the database include competition policy, State-Owned Enterprises (SOEs), subsidies, public procurement, Technical Barriers to Trade (TBTs), Sanitary and Phytosanitary Standards (SPS), labor rights, and environmental rules.

Such very rich set of information has to be collapsed in broad categories of PTAs in order to be tractable in a general equilibrium framework accounting for the complex impact of PTAs with various levels of ambition on the world matrix of bilateral trade costs. We go beyond the dummy approach – which does not capture depth

– used in the large majority of the literature, and the use of synthetic indicators of depth such as the count of provisions in PTAs (Mattoo, Mulabdic & Ruta 2017). We start by defining statistically significant groupings. We rely on a clustering approach to identify groups of trade agreements based on the provisions' content. In doing so, we opt for the iterative "k-means++" algorithm developed by Arthur & Vassilvitskii (2007) which ensures greater accuracy by randomizing starting points at each replication. Given the underlying distribution of provisions in each of the 18 policy areas in the data, the silhouette width criterion (which evaluates cluster fit on within-group cohesion and between-group separation) recommends 3 clusters.

This method defines a natural grouping of agreements grounded on a transparent statistical approach based solely on the content of the treaties. Since PTAs within a group are expected to have a similar distribution of provisions by policy areas (as clustering maximizes within-group cohesion) their impact on trade costs within members and between members and third countries is expected to be different. Moreover, as PTAs in different groups are expected to have a statistically different distribution of provisions (as clustering maximizes between group separation) the marginal contribution of each policy area is expected to be different across clusters. In order to gauge the differential contribution of policy areas in the classification, we provide evidence on the policy markers of different clusters.

In a second step we estimate an explicit bilateral trade function taking stock of the classification of agreements. Once more, we let the data speak and estimate the mean impact on trade for the PTAs belonging to each cluster. Note that a large part of the literature estimating the impact of PTAs on trade may be flawed because it does not control for the appropriate benchmark in terms of trade cost, i.e. domestic sales (Yotov 2021). Importantly, we take the gravity equation seriously and integrate internal trade flows in the estimation of the trade impacts of the different types of PTAs. We compile the largest database for which internal and international trade is available in a panel, based on the last release of UNIDO data.

A challenge when assessing the trade impact of PTAs is endogeneity. Countries self-select in signing a PTA due to unobservable bilateral linkages (Baier & Bergstrand 2004). By the same token, the networks of firms and their joint involvement in Global Value Chains (GVCs) contribute to shape the geography of PTAs. "Lobbying for globalization" (Blanga-Gubbay, Conconi & Parenti 2020) is ascertained in the US case: large firms benefit more from regional integration because it reduces the trade frictions between affiliates and improves market access; consequently they spend more on lobbying for PTAs.[1] Hence, the endogeneity concerns that have been raised for PTAs in general are likely to bind also for their *depth*, which is a challenge to be addressed here. We address this issue in different ways. First, we rely on a conservative fixed effects strategy. Second, we control for the indirect intensity of bilateral GVC relationships, which captures only the bilateral income generated

[1] As evidenced in the case of US anti-dumping duties (Bown, Conconi, Erbahar & Trimarchi 2021), trade frictions propagate throughout global production networks. Similarly Blanchard, Bown & Johnson (2016) and Bown, Erbahar & Zanardi (2021) show respectively that offshoring shapes the optimal trade policy of a country and that the importing countries tend to remove anti-dumping duties from their main partner countries in GVCs.

throughout production linkages with third countries (i.e. excluding the potentially endogenous component related to direct trade). Third, we reallocate randomly country-pairs across groups of PTAs of different content and show after a large number of replications that the distribution of parameters for each group of PTAs is not statistically different.

The last step of the analysis illustrates how to perform counterfactual exercises in a theoretical consistent way, taking stock of general equilibrium effects. We firstly move all the *existing* PTAs away from their cluster towards the most ambitious one (i.e. a PTA cluster associated to higher trade) and and examine the economic effects of such policy reform of preferential trade agreements. Since we are not changing the network of agreements but only their content, this exercise captures a variation in the intensive margin of PTAs. We find that deepening all existing trade agreements could boost world trade by 3.9 percent and world GDP by 0.9 percent relative to the baseline. Few countries are negatively affected by the deepening of trade cooperation, but overall gains are significantly positive for most of the countries, illustrating the importance of the intensive PTA margin for international trade cooperation.

We then simulate the economic impacts of deepening all the existing agreements of each region from their current level of ambition to the highest level of ambition, and repeat the exercise sequentially for PTAs signed within or between regions. Countries in the East Asia and Pacific region would mostly benefit from deepening preferential trade agreements within the region, while countries in the Middle East and North Africa region would benefit more from deepening agreements with partners outside the region. All other regions fall somewhere these two extremes. Given the diversity of policy preferences may be larger across than within regions, the low hanging fruits when it comes to deepening trade agreements may be mostly regional. But the gains of these different integration strategies (within or between regions) would ultimately differ across countries.

The last step simulates the impact of the extensive margin of regional integration for PTAs of different levels of ambition.

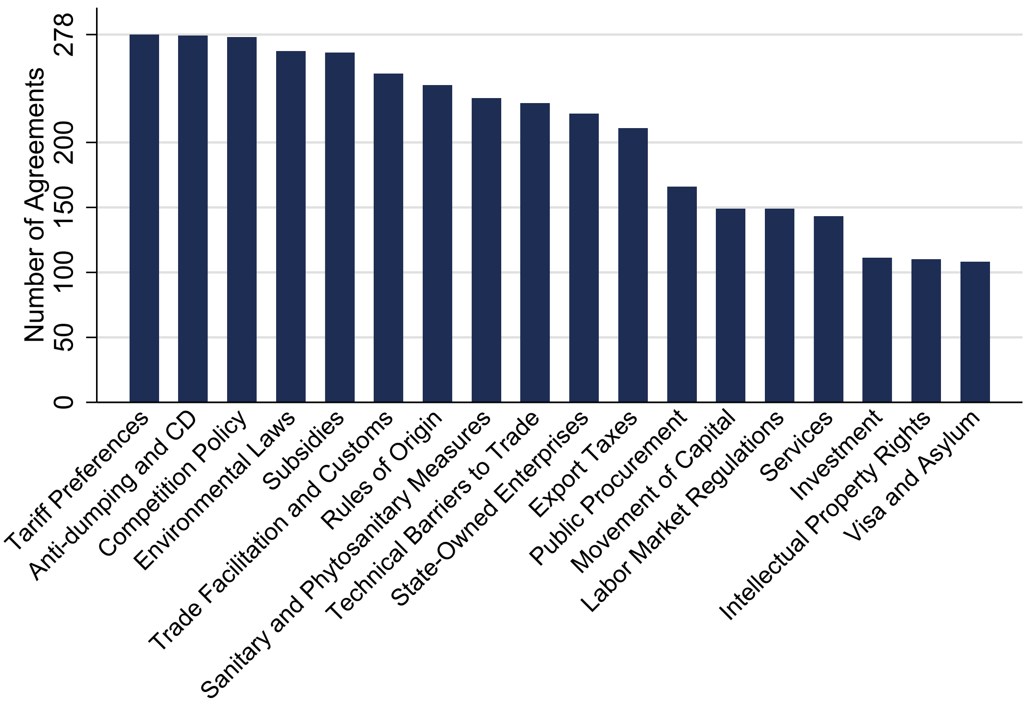
The rest of the paper is structured as follows. Section 1 describes the rich data set used in our exercise and characterizes the clusters of PTAs in terms of commonalities in their content. Section 2 presents the methodology used to measure the ex post impact of the different types of PTAs and characterizes these clusters in terms of their impact on trade. Section 3 addresses the issue of endogeneity in the relationship between trade and deep PTAs. Section 4 shows how to exploit the information on the trade impact of PTAs of different types in a general equilibrium framework for an endowment economy in order to simulate counterfactuals for the world economy. Section 5 concludes.

1 Clustering PTAs based on their content

1.1 Data on the provisions contained in PTAs

The analysis in this paper is based on a new database on the detailed content of deep trade integration -i.e. the depth of commitments that countries take in PTAs (Mattoo et al. 2020).[2] In particular, this database provides information on the content of 18 policy areas most frequently covered in PTAs (Figure 1). The list of policy areas mapped includes border measures such as anti-dumping duties or export taxes, and behind the border measures such as technical barriers to trade, competition policy and environmental law, among others.

Figure 1: Classifications of PTAs by objective of policy area



Source: Deep Integration Handbook, (Mattoo et al. 2020)

The analysis focuses on a sample of 278 trade agreements that have been signed between 1978 and 2018 and that are currently in force and notified to the WTO. For each agreement and policy area, the database provides a series of questions covering aspects such as stated objectives and substantive commitments, as well as aspects relating to transparency, procedures and enforcement. The number of provisions coded (910 in total) varies by policy area, reflecting differences in terms of coverage and complexity across policy areas that are negotiated in the agreements (Table 1). The share of provisions included in each policy area across agreements ranges between 7 percent on average for anti-dumping to more than 30 percent on average for policy areas such as competition

[2]The methodology and data are available in this link `https://datatopics.worldbank.org/dta/index.html`

policy or services. More than one-half of the agreements considered for this analysis cover 50 percent or more of the provisions included in policy areas such as subsidies, sanitary and phytosanitary measures, competition policy, rules of origin and services. The share of agreements covering 50 percent of provisions is much lower (below 30 percent) for policy areas such as anti-dumping, labor market regulation, intellectual property rights and visa and asylum.

The rich information on the content of PTAs poses the challenge of how to aggregate it to define and quantify the agreements' overall depth. Different approaches can be contemplated from the simple count, or the coverage ratio of the provisions included in an agreement, where it is assumed that the relative importance of each provision is the same across policy areas and agreements. Alternative methods to assign different weights to different provisions according to their commonality or explanatory power across agreements include principal components analysis or machine learning algorithms. In this analysis we use an agnostic statistical procedure to classify PTAs into an optimal set of groups where agreements present both the maximum similarity in terms of provisions included within groups, and the maximum difference between groups.

1.2 The classification algorithm

The 278 agreements covered in the dataset include a large number of clauses that have been grouped in 910 type of provisions across 18 policy areas. Not all provisions are present in all the agreements and when they are their content may be of various ambition. We now describe how this extremely detailed information can be used to partition PTAs into distinct groups (clusters) based on the similarity of the underlying treaties (presence or not of individual provisions as well as their ambition). The purpose of the classification is to group PTAs up to the point where the dissimilarity between clusters and the similarity within clusters is maximal. The typical approach to solve this problem in a multi-dimensional space is to rely on Euclidian distance.[3]

We cluster groups of PTAs based on the proximity of their scores obtained for the different categories of provisions. These scores are a simple metric of the ambition of the PTAs for each area. We then identify the optimal number of clusters, i.e. the number that simultaneously maximizes the distance between groups and the cohesion within them, by a statistical criterion: the silhouette width (Rousseeuw 1987). The silhouette width measures the separation between clusters by evaluating how similar agreements within a cluster are to each other with respect to those in the nearest group. Figure A2 in the Appendix reports the average silhouette width by number of clusters. The silhouette reaches its maximum when the data are partitioned across three clusters.

[3]An analogy may help: an electric car producer may want to partition its potential clients in different groups in order to better target advertising investments. Information is needed on age, gender, income, location in an urban or rural area, daily commuting time and time spent surfing on the social networks with a smartphone. The algorithm would, e.g., identify the group of high-income technophile consumers, which would differ from consumers who make the automobile an ostentatious object of consumption, and also distant from the group environmentalist consumers who prioritize the reduction of CO2 emissions. Other analogies could be found in medicine or biology.

To identify the composition of each cluster, we rely on a state-of-the-art statistical classification method, the "k-means++" clustering algorithm, which is a non-hierarchical iterative clustering method. The algorithm partitions the data into a number of pre-determined clusters (defined on the silhouette width) based on the dissimilarity matrix measuring the Euclidean distance between agreements across the 18 policy areas covered in each of the treaties under analysis. In this approach, groups' centers are defined randomly and, at each iteration, the group center is chosen based on a probability proportional to the minimal distance to the closest previously defined center, ensuring a greater accuracy of the resulting classification.[4] Eventually, the 278 PTAs are partitioned across three clusters as follow: Cluster #1 (29 agreements), Cluster #2 (96 agreements) and Cluster #3 (153 agreements). The complete list of PTAs within each cluster is available here. These clusters are going to be used in the next section to assess the impact of different PTAs on trade.

Formally, our data can be represented as a matrix composed by 278 agreements (columns) and 910 provisions (rows). Each cell informs us on the presence or ambition of a given provision in a given agreement. Notice than the rows will eventually be grouped in "areas", e.g. competition or services policy, in order to reduce the dimensionality of the matrix.[5] There are 18 such areas.

For sake of clarity, let us first give an example of how the data is structured in the services policy area. The provision covering the obligations needed for a juridical person "to be considered a service supplier of a party to the agreement" allows for 6 different options across the 278 agreements. These options range from the most restrictive as "being incorporated under the domestic law of the party and have substantive business operations in the territory of a member" (coded with value 1); to the most liberal as "being owned or controlled by natural persons of the other party" (coded with value 6). The PTA between Australia and New Zealand (year 1983) by including the most liberal formulation of the provision (i.e. "being owned or controlled by natural persons of the other party") is an illustration of deeper cooperation in the service domain than, for example, the Andean Community (1988); whose founding treaty mentioned the stricter requirement of being incorporated under the domestic law of the party and have substantive business operations in the territory of a member.

In order to obtain the score that will help cluster the agreements, we firstly need to normalize the data as within a policy area provisions are coded by increasing level of ambition (1, 2, ...n) – see our previous example, while other provisions are coded with a dichotomic variable (1, 0). We calculate, for each provision in the matrix composed by 278 agreements (columns) and 910 provisions (rows), the mean of each provision across agreements. This normalization has the double advantage of: (i) harmonizing the measuring scale across provisions (not all provisions are binary); ii) factoring in the frequency of the provision across agreements (if a provision is not present in a certain agreement it is coded as zero). The normalized scores assigned to each of the 910 provisions

[4]We implement the classification using the "kmeapp" function in R. Although we allow for a fairly large number of iterations, i.e. 5,000, the algorithm converges to a stable classification after few rounds. As a robustness we test alternative classification methods in Section 2.2

[5]The implementation of the k-mean clustering poses some difficulties for data sets with high levels of dimensionality.

are then aggregated using the simple average across all the provisions that are included in each of the 18 policy areas mapped. Lastly, the "k-means++" algorithm is applied to a reduced matrix composed by 18 rows (defined by the policy domains) and 278 columns (defined by the agreements).

Table 2 provides an example to illustrate how the data on the content of preferential trade agreements are normalized and aggregated in order to generate a set of clusters. We consider a hypothetical situation where we have a total of 5 PTAs that need to be grouped into 3 clusters which include only one policy area (area xx), which comprises 3 coded provisions (a, b, c). These provisions take values 1 to 3 according to their level of ambition and take value zero if they are not included in the agreement. The first step is to construct a matrix with normalized scores capturing the average occurrence of each provision across PTAs. Consider provision "a". This provision is present in PTA-1 and is coded with the highest level of ambition (score=3). Provision "a" is also present in PTA-2, but with the lowest level of ambition (score=1). Provision "a" is absent from PTAs-3 to 5 (code=0). The average occurrence of provision "a" across the PTAs is therefore 0.8 ((3+1+0+0+0)/5). This value also captures the frequency of provision "a" across agreements. A similar exercise can be done with provisions "b" and "c", where the average occurrence is equal to 0.8 and 1, respectively. As a second step, we normalize the score provided to each provision by dividing the current score by the average occurrence (frequency). The normalized scores for provisions "a", "b" and "c" included in agreements PTA-1 to PTA-5 are provided in Table 3. The normalized score of provision "a" in agreement PTA-1 is equal to 3.75 (3/0.8) and captures the relative occurrence and intensity (gradient 1, 2, . . . n) of this provision in this agreement. Finally, we aggregate all the provisions' scores within each policy area to reduce the dimensionality of the matrix.[6]

1.3 The marginal effect of provisions on the probability of an agreement

In order to provide an illustration of the policy content of PTAs in each cluster, we use a simple linear probability model to ask, what is the marginal effect of the 18 policy areas on the probability of an agreement being in clusters #1 or #3 with respect to cluster #2? Specifically, we run two separate regressions for the probability of an agreement being classified in cluster #1 or #3. The total number of observations in each regression is accordingly 278. In the right-hand side, we include the 18 scores by policy domain used in the clustering exercise. As additional controls we also include a dummy for the decade in which the agreement has been ratified and a series of dummies for the income level of participants: high-high, low-low, high-low income. In order to fix the reference group, we include in both regressions a dummy variable equal to 1 for the agreements classified in the cluster #2.

[6] Table A1 in the Appendix provides a subset of actual data in the raw and normalized matrix for the area of services. Note that provision # 850 is coded 5 in EFTA (high ambition) and only 2 in the Chile-Japan PTA, which illustrates the (1,2,. . . n) coding afore mentioned. In Cluster #1 the final score is 0.840; it is 0.846 in cluster 3 and 1.362 in Cluster 2: these are also the scores indicated in table 5, row "Services", last three columns. The same exercise pertains for the 17 other policy areas and for other PTAs in order to obtain the score for each area and each PTA (an 18 x 278 matrix) thereafter used for the clustering.

Table 1: Content of preferential trade agreements, 1978-2018

Policy Area	Number Provisions	Average Coverage by Policy Domain	Agreements with zero Provisions	Agreements with $< 25\%$ Coverage	Agreements between $25 - 50\%$ Coverage	Agreements between $50 - 75\%$ Coverage	Agreements with $> 25\%$ Coverage
Anti-dumping	36	0.10	2	267	9	0	0
Competition Policy	35	0.37	2	85	121	59	11
Countervailing duties	14	0.21	1	194	67	16	0
Environment	48	0.16	11	196	51	20	0
Export Taxes	46	0.20	68	94	106	10	0
IPR	120	0.07	170	78	28	1	1
Investment	56	0.18	168	6	49	55	0
Labor Market	18	0.16	128	98	9	18	25
Migration	30	0.13	169	31	59	19	0
Movement of Capitals	81	0.21	132	26	78	42	0
Public Procurement	95	0.15	111	98	14	54	1
Rules of Origin	38	0.36	35	35	126	82	0
SPS	53	0.14	47	168	61	2	0
STE	52	0.23	57	66	150	4	1
Services	62	0.33	136	3	21	78	40
Subsidies	36	0.33	10	51	176	41	0
TBT	34	0.18	51	136	83	8	0
Trade Facitlitation	52	0.25	30	120	90	38	0

Note: *Coverage* report the share of non-zero provisions within each cluster and policy Area.

Table 2: Practical Example of Provisions aggregation (raw matrix)

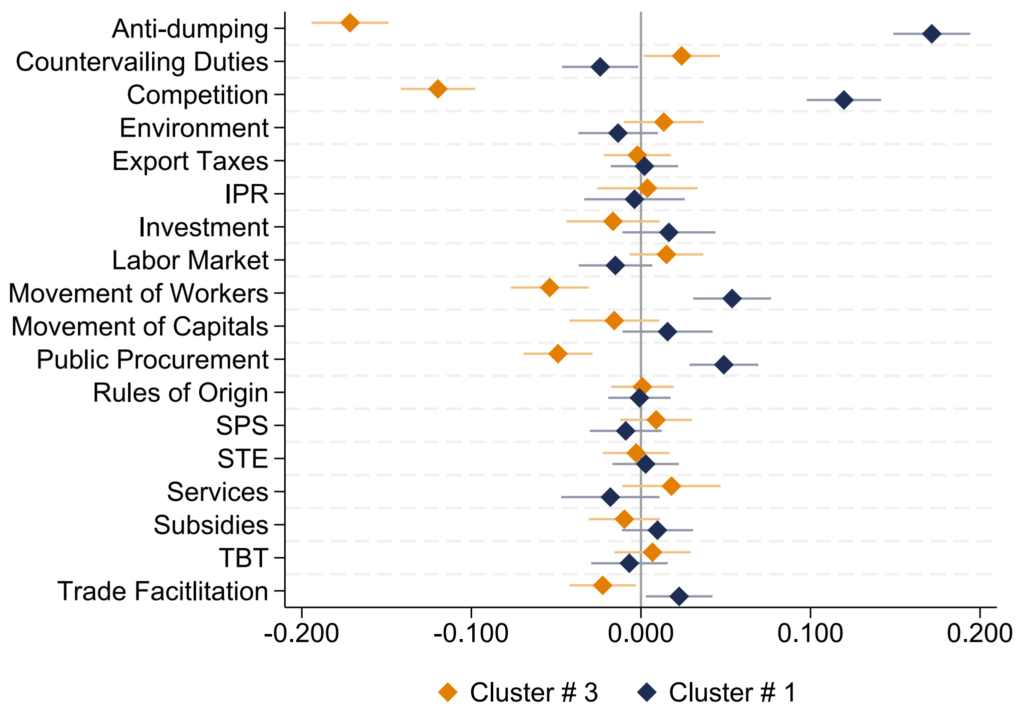
Policy Area	Provision	Cluster # 1 PTA-1 (1)	Cluster # 2 PTA-2 (2)	 PTA-3 (3)	Cluster # 3 PTA-4 (4)	 PTA-5 (5)	Average (6)
Area xx	a	3	1	0	0	0	0.8
Area xx	b	2	1	0	1	0	0.8
Area xx	c	1	1	1	1	1	1
Coverage by PTA		1	1	0.33	0.66	0.33	
Coverage by Cluster		**1.00**	**0.67**		**0.5**		

Table 3: Practical Example of Provisions aggregation (normalized matrix)

Policy Area	Provision	Cluster # 1 PTA-1 (1)	Cluster # 2 PTA-2 (2)	 PTA-3 (3)	Cluster # 3 PTA-4 (4)	 PTA-5 (5)	Average (6)
Area xx	a	3.75	1.25	0.00	0.00	0.00	
Area xx	b	2.50	1.25	0.00	1.25	0.00	
Area xx	c	1.00	1.00	1.00	1.00	1.00	
Score by PTA		2.42	1.17	0.33	0.75	0.33	
Score by cluster		**2.42**	**0.75**		**0.54**		

Taking all these elements on board, we plot in Figure 2 the 18 provision areas on the vertical axis and on the horizontal axis the marginal effect of each on the probability for a PTA of belonging to cluster #1 and #3, respectively. We observe that anti-dumping and competition provisions play an important role in cluster #1 relative to provisions in areas such as labor regulation. Following that line of reasoning, deepening a PTA classified in cluster #3 into a PTA classified in cluster #1 – the type of counterfactual we will consider below – would on average require to introduce ambitious provisions on anti-dumping, competition and technical regulation. Public procurement and movement of workers are also the type of provisions discriminating between those two type of PTAs.

Figure 2: Marginal Probability of being in Cluster # 1 and # 3



Notes: Marginal effect of the 18 different provision areas from a linear probability model for being in the corresponding cluster. The model controls for the decade of signature of the agreements. Cluster # 2 is the reference group.

2 Ex-post quantification of the impact of different types of PTAs

In this section, we estimate the trade impact of the 3 clusters of PTAs identified with the "k-means++" algorithm in a gravity framework in order to recover the partial effect on trade of PTAs associated with the different clusters. We also test for the robustness of our results to alternative classifications.

2.1 The structural gravity estimation

Specifically, we estimate, using PPML with panel data, the following structural gravity model:[7]

$$X_{ij,t} = exp\left(\sum_{z=1}^{3}\beta_z PTA^z_{ij,t} + \beta \text{Transitory} PTA_{ij,t} + \sum_{T=1978}^{2000}\beta_T INTL\ BRDR_{ij} * T + \pi_{i,t} + \chi_{j,t} + \mu_{ij}\right) + \epsilon_{ij,t} \tag{1}$$

Where $X_{ij,t}$ includes both intra-national and international yearly manufacturing trade flows from 1978 to 2018.[8] Including intra-national sales is critical as the domestic economy is the most appropriate benchmark for trade integration (Yotov 2012). Otherwise the estimated coefficients would suffer from a missing variable bias. The data on inter- and intra-national trade come from, respectively, the UN-ComTrade and UNIDO-Indstat database.[9] Following standard practice, we fill in intra-national flows using linear interpolation between non-missing data and extrapolating remaining missing values using gross output to value added ratios as in Head & Mayer (2021). In each year only countries with non-missing intra-national trade flows enter the estimation sample.

Our main variable of interest, $PTA_{ij,t}$, is split across cluster groups: $PTA^{z=1}_{ij,t}$, $PTA^{z=2}_{ij,t}$ and $PTA^{z=3}_{ij,t}$. The dummy $INTL\ BRDR_{ij}$ takes the value of one in case of an international trade flow. We interact this dummy with decades indexed by T (leaving the period after 2010 as reference). Exporter-time, $\pi_{i,t}$, and importer-time, $\chi_{j,t}$, fixed effects control for time-varying Multilateral Resistance Terms, while directional bilateral fixed effects, μ_{ij}, control for time-invariant unobserved characteristics of the country pair potentially leading to self-selection into PTAs (Baier & Bergstrand 2007).[10] As an additional control, we also include the variable $TransitoryPTAs_{ij,t}$, to control for agreements that are no longer in force.

The results presented in Table 4 provide the elasticity of bilateral trade to the different types of PTAS (as

[7] $X_{ij,t}$ is in levels in columns 1 to 4 of Table 4 and in shares of absorption at destination – namely $\frac{X_{ij,t}}{\sum_i X_{ij,t}}$ – in the remaining estimation tables. See discussion below.

[8] An issue is that the adjustment of exporters to the inception of a PTA is not instantaneous. While Anderson & Yotov (2016) use four-year intervals, with our data it would be difficult to use such strategy as each individual agreement would hardly be implemented at the first date of a given interval. Averaging across agreements with different commencement date would not adequately capture the impact of the depth of agreements on trade volumes. Accordingly we adopt a specification relying on sequential data, as suggested by Egger, Larch & Yotov (2022).

[9] ComTrade is used to retrieve information on international trade flows from 1978 up to 2018 whereas Indstat provides data on gross output used to compute domestic sales (as the difference between production ant total exports). Due to the limited information on international trade, transition economies enter the estimation sample in 1992, after the collapse of the Soviet Union.

[10] Our results are robust to symmetric country-pair fixed effects a shown in Table A2 in the Appendix.

grouped by the clustering). This elasticity is here estimated within sample alternatively for the raw data and for the sample including extrapolated data. The set of elasticities obtained for the extrapolated data will be introduced in a second step as a parameter in the estibration procedure. In doing so, we are able to exploit a squared dataset of 112 countries in the general equilibrium counterfactual evaluation for year 2018. Restricting the sample to non extrapolated domestic sales would constraint the analysis to only 61 countries, unevenly distributed across regions. We favor the broader dataset as it improves coverage and reliability of the regional general equilibrium results.

The first five and the seventh columns exploit the raw UNIDO data we rely on. There are 338,685 observations for 136 countries for which we observe internal flows at least once over the estimation period (of which 61 countries in year 2018). In order to expand internal flows coverage, in column 6 we rely on a broad sample obtained by extrapolating domestic sales using output to value added ratios, as in Head & Mayer (2021), which leads to 591,320 "observations" for 143 countries (of which 112 in year 2018).[11]

Column (1) is the standard estimation strategy with importer-time, exporter-time and dyadic fixed effects, and controlling for internal flows. Column (2) replicates column (1) by adding the control for transitory PTAs, which is not significantly affect the PTAs coefficient. Column (3) adds the control for internal flows (with a dummy taking the value of one in the case of international flows interacted with a decade specific indicator variable).[12] Comparing with columns (1) and (2) shows that this dramatically reduces the trade impact of PTAs. The second result is that the negative impact on commerce of crossing the border decreases progressively, which is the other side of the "globalization" coin.

The impact of PTAs shown in column (3) is then split by clusters of PTAs. The estimated elasticity ranges from 0.112 to 0.606, with cluster #1 having the largest impact on trade. We refer to this cluster as the one of (revealed) "deep" PTAs, the second as "medium" and the third as "shallow" PTAs. The comparison of Column (3) and Column (4) shows a key result of our paper: using a well-specified gravity equation, the usual approach relying on a single dummy for the presence of PTAs misses the differentiated impact of PTAs having different ambition. The average point estimate for the usual dummy approach is 0.273, while we get 0.112, 0.227 and 0.606 respectively for cluster #3, #2 and #1.[13] The impact of such a difference in quantifying the impact of regional integration is indeed of primary importance: the use of a single dummy variable would overestimate the impact on trade and GDP of the new low-ambition agreements, while underestimating this impact for the high-ambition agreements.[14]

[11] The list of 61 and 51 additional countries is given in the Appendix Table A5.

[12] The years 2001 to 2018 represent the excluded period.

[13] Assuming a value for the elasticity of substitution of $\sigma = 5$ (see e.g. estimations in Fontagné, Guimbard & Orefice (2022)) the baseline PTAs estimates can be expressed as tariff equivalent effects. Deep PTAs are equivalent to a tariff reduction of 11.4 percent (i.e. [exp(0.606 / (-5)) - 1]*100 = 11.4); Medium PTAs are equivalent to a reduction of 4.4 percent and Shallow ones are equivalent to a reduction of 2.2 percent.

[14] This is illustrated in TableA3 in the Appendix by comparing columns (3) and (4) with columns (7) and (8) respectively. In this table showing the results of the extensive margin of regional integration for East Asia and Pacific region countries, the impact of the PTAs of *medium* ambition is very close to the impacts that would be obtained with a simple dummy approach, meaning that

One issue to be addressed is that if the bilateral factors are incompletely captured or proxied, they will be loaded into the importer and exporter fixed effects used later in the counterfactuals. This may be the case here since the bilateral factors are time invariant (so as not to capture variation in signed agreements). In this case, the model would fit large trade flows quite well but not so well for small flows (Egger & Nigai 2015). Correcting for this potential bias caused by unobservable trade costs that operate through general equilibrium constraints allows the impact of EPAs to be correctly identified (Kharel 2019).[15] As our database also includes domestic sales, starting from column (5) we are able to control for the relative size of the trading partners and use an estimator on shares. The PPML estimator, in fact, assigns more weight to countries with large imports volumes in the identification of the parameters (Eaton, Kortum & Sotelo 2013, Head & Mayer 2014). A way to eliminate differences in the penalization of large and small trade flows is to normalize trade flows by destination country total absorption (Eaton et al. 2013), this is what is done for specifications in columns (5) and (7) of Table 4 with the raw data, and in columns (6) and (8) using the extrapolated data.[16]

While Columns (1-5) and Column (7) rely on raw data, Column (6) and Column (8) report the estimations results on extrapolated data. Both the points estimates of each cluster and the difference between them are very similar across the two estimation samples. In order to ensure a broader and more representative set of countries in the General Equilibrium simulations, we opt for the sample and estimated elasticities in Column 6 ot Table 4.

Finally, an important econometric issue is that our 3-way fixed effect panel PPML procedure with time-invariant country-pair, time-varying exporter and importer fixed effects may be subject to an incidental parameter problem when the number of periods is small (such as ours), which biases the estimated parameter for PTAs in gravity equations as well as their confidence intervals.[17] Accordingly, in Column (7) we replicate Column (5), using raw data, by relying on the fix developed by Weidner & Zylkin (2021) which confirms that point estimates and standard errors are both larger when the incidental parameter problem is properly addressed. We do the same in Column (8), by replicating the estimates in Column (6), using extrapolated data. Importantly, the statistical significance of our estimated parameters is confirmed.

Our estimate in Column (6) yield a robust set of elasticities for a broad set of countries to simulate the welfare effect of trade policy shocks. Indeed, the identified trade elasticities to PTAs in Column (6) will be the parameters introduced in the general equilibrium gravity model, jointly with the country-pair fixed effects and the international border effects. Two types of counterfactuals for regional integration can be contemplated.

the usual approach is not capturing the richness of the agreements signed (e.g. 2.29% additional total exports instead of 2.30%) for Australia. In contrast, Australia's signing of new superficial (column (5) and (6)) or ambitious (resp. (1)and (2)) PTAs would make a difference, as the impact on exports would be 1.31% versus 4.56%, respectively, which would not be captured by the usual single-dummy approach.

[15] We are indebted to an anonymous referee for stressing this issue

[16] As shown by Sotelo (2019) poisson estimation on market share variable with country fixed effects is equivalent to the Multinomial PML proposed in (Eaton et al. 2013).

[17] In fact, the usual clustering procedures provide biased (too narrow) standard errors in such setting.

One set of counterfactuals on the intensive margin and one on the extensive margin. On the intensive margin – which is the key issue addressed in this paper – we simulate the effect of deepening existing agreements. This can be done for all PTAs in the world, or for all PTAs involving countries in a given region (with other countries within the same region, or alternatively with countries in different regions). Practically, in the counterfactual policy scenario for these simulations, positive entries in either $PTA_{z=3}$ or $PTA_{z=2}$ are switched to zero while the corresponding entries in $PTA_{z=1}$ are set to 1.[18]

On the extensive margin, we simulate the effect of "signing" missing agreements with all trade partners for the East-Asia-Pacific (EAP) region. This is done assuming that new agreements are alternatively: shallow (all zero entries in are switched to $PTA_{z=3}$, other entries are unchanged), medium (resp. to $PTA_{z=2}$) or deep (to $PTA_{z=1}$). We also provide, for sake of comparison with the usual single-dummy approach, a counterfactual where countries in this region sign PTAs with all their trading partners, with no consideration of the depth of the agreements signed.

Evidence by new (Canada and the EU: CETA) or renegotiated (renegotiation of the NAFTA: USMCA) PTAs suggests that the complexity of PTAs has increased through time. As the clustering algorithm centers on the provision content and not on the timing of the agreement we test if the trade effect of deep PTAs (i.e. cluster #1) has changed over time. The estimation results by decade are reported in Figure A1 in the Appendix. Until the end of the 2000s, the estimated elasticity of deep PTAs appears to be fairly stable, but it increased during the 2010s. However, confidence intervals are larger in the last decade and tend to overlap with the overall effect. For this reason, we favor the average elasticity estimated over the whole period in our quantitative exercise.

[18] Multilateral resistances terms and factory-gate prices are indexes subject to a normalisation. This imposes to choose a reference country not directly impacted by the counterfactual and to restrict the exercise to cross-sectional comparisons across countries. See Chapter 2 of Yotov, Piermartini, Monteiro & Larch (2017) for a detailed discussion. As we have chosen South-Africa as numeraire, the change in exports of the Sub-Saharan Africa region is underestimated because we do not change the vector of PTAs for South Africa.

Table 4: PPML: Gravity Estimations of the elasticity of trade to PTAs by Cluster

Dep Var:	X_{ijt}				X_{ijt}/X_{jt}			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$PTA_{ij,t}$	0.569	0.588	0.273					
	(0.047)	(0.054)	(0.040)					
$PTA^{K\#1}_{ij,t}$				0.606	0.525	0.484	0.542	0.492
				(0.046)	(0.041)	(0.073)	(0.052)	(0.091)
$PTA^{K\#2}_{ij,t}$				0.227	0.205	0.260	0.241	0.308
				(0.045)	(0.030)	(0.035)	(0.037)	(0.046)
$PTA^{K\#3}_{ij,t}$				0.112	0.096	0.135	0.109	0.154
				(0.054)	(0.041)	(0.040)	(0.048)	(0.049)
Transitory $PTAs_{ij,t}$		0.082	0.104	0.203	0.206	0.130	0.233	0.145
		(0.050)	(0.046)	(0.054)	(0.037)	(0.043)	(0.046)	(0.053)
INTL BRDR∗1980			-0.855	-0.843	-1.077	-0.989	-1.063	-0.975
			(0.044)	(0.043)	(0.048)	(0.062)	(0.076)	(0.094)
INTL BRDR∗1990			-0.528	-0.529	-0.792	-0.774	-0.787	-0.762
			(0.034)	(0.034)	(0.038)	(0.045)	(0.056)	(0.064)
INTL BRDR∗2000			-0.168	-0.169	-0.261	-0.315	-0.256	-0.302
			(0.027)	(0.027)	(0.028)	(0.033)	(0.043)	(0.047)
Intra-National flows	Raw	Raw	Raw	Raw	Raw	Extrapolated	Raw	Extrapolated
Period	1978-2018	1978-2018	1978-2018	1978-2018	1978-2018	1978-2018	1978-2018	1978-2018
Frequency	yearly	yearly	yearly	yearly	yearly	yearly	yearly	yearly
N. Country ID	136	136	136	136	136	143	136	143
Observations	338,685	338,685	338,685	338,685	338,685	591,320	338,685	591,320
FEs	it, jt, ij	it, jt, ij	it, jt, ij	it, jt, ij	it, jt, ij	it, jt, ij	it, jt, ij three-way correction	it, jt, ij three-way correction

Notes: Exporter-time (it), Importer-time (jt) and Exporter-Importer (ij) fixed effects are always included. From column 1 to column 5, standard errors in parentheses are clustered by country-pair. In column 7 and 8 both standard errors and point estimates are corrected using the Weidner & Zylkin (2021) procedure, implemented in Stata with *ppml_fe_bias*. In column 5 and column 7 missing values in domestic sales are linearly interpolated and the remaining missing values are extrapolated using the evolution of a country total exports.

2.2 Alternative classification algorithm

As our exercise relies on the "k-means++" algorithm, it is important to check whether our results are robust to alternative classification algorithms. Changing algorithm may alter the results in two dimensions: the composition of PTAs in each cluster; the estimated elasticity for each cluster in the structural gravity. Both dimensions have implications for the counterfactuals performed afterwards. In order to consistently test the different approaches, we first compute the clusters using the alternative methods and then re-estimate a new set of trade elasticity for each cluster. In this application, we consider three different clustering routines: an alternative iterative algorithm (the Partitioning Around Medoids, PAM); a non-iterative hierarchical clustering procedure; and last, we manually reclassify "borderline PTAs". [19]

The list of PTAs in each of the three clusters is provided here. The main differences are as follows. With the genetic algorithm Group 1 is restricted to 24 PTAs, compared to 29 with our preferred algorithm, 33 with the hierarchical one and 34 when we manually reclassify "borderline PTAs" in columns (5) and (6) (see note of Table 5). All in all the k-means++ algorithm sounds as a good compromise that properly select the 29 most ambitions PTAs and populates the least ambitious group (153 PTAs). [20]

Our estimations are robust to these alternative classification algorithms, as shown in Table 5 where we give the results of the PPML in share with raw data. These results, to be compared with Column (5) and Column (7) in Table 4, are not statistically different for deep PTAs in cluster#1. This is of first order importance since our main counterfactuals exploit the intensive margin of regional integration by shifting existing to the "Deep" cluster: as agreements with the higher trade effects (i.e. cluster #1) largely overlap with the ones identified by the preferred routine hence the magnitude of the trade premium associated with the most ambitious group is hardly affected.

We lastly recalculate the estimated effects after reallocating the few "borderline" agreements to the closest alternative clusters in Column (5) and Column and (6) of Table 5. We manually re-classify "borderline" agreements as visualized in Figure 3 reporting the position of the 278 agreements over the cluster space. The coordinates represent the first two principal components extracted from the 18 features used in the clustering algorithm. The color and shape of each point represents the different clusters. While the separation between clusters is clear-cut with very few agreements at the "border" of their partition, and agreements in Cluster #1 (the "Deep" ones) standing apart from the rest of the sample, there are a couple of instance where the separation is less clear-cut. These are the PTAs that we reclassify. Overall, our results prove again to be robust to this reclassification.

[19] The reclassification is based on a visual inspection of the cluster space obtained with the "k-means++" algorithm, reported in Figure 3.

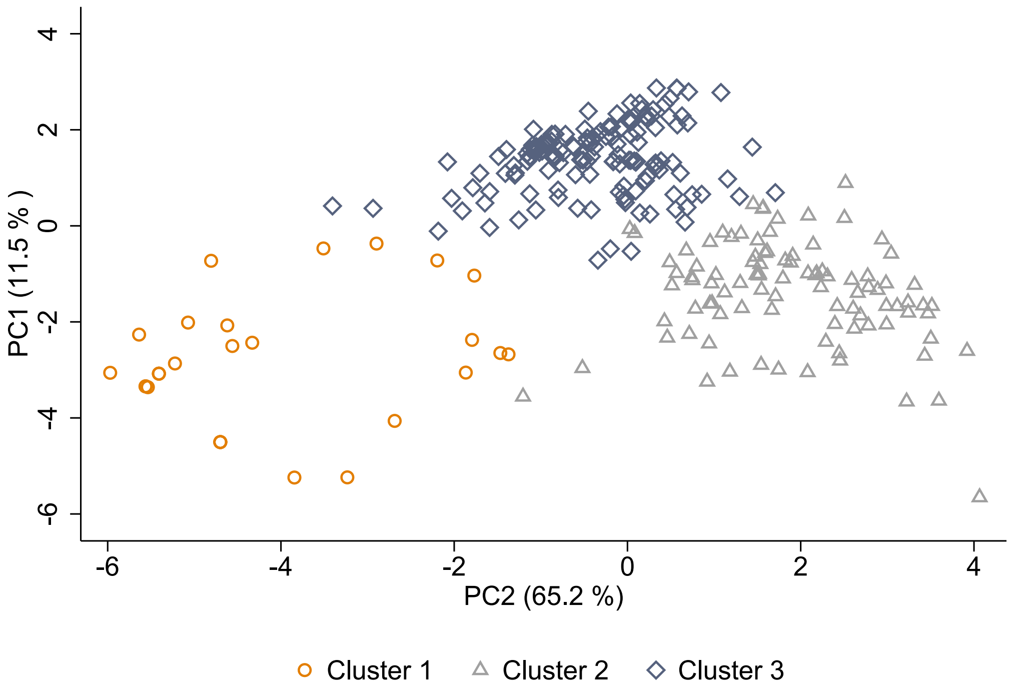
[20] The five additional PTAs in cluster #1, compared to the genetic algorithm, are: EFTA-Serbia, EFTA-Ukraine, EFTA-Montenegro, EFTA-Bosnia and Herzegovina, Peru-Mexico.

Table 5: PPML: Gravity Estimations of the elasticity of trade to PTAs by Alternative Cluster Definitions

Dep Var:	X_{ijt}/X_{jt}					
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{PTA}^{K\#1}_{ij,t}$	0.525	0.542	0.500	0.525	0.482	0.501
	(0.041)	(0.053)	(0.037)	(0.048)	(0.038)	(0.048)
$\text{PTA}^{K\#2}_{ij,t}$	0.186	0.219	0.189	0.220	0.185	0.215
	(0.031)	(0.039)	(0.029)	(0.036)	(0.029)	(0.036)
$\text{PTA}^{K\#3}_{ij,t}$	0.112	0.127	0.088	0.095	0.105	0.117
	(0.038)	(0.046)	(0.042)	(0.049)	(0.040)	(0.047)
Transitory $\text{PTAs}_{ij,t}$	0.198	0.224	0.207	0.237	0.200	0.228
	(0.037)	(0.046)	(0.039)	(0.048)	(0.038)	(0.047)
INTL BRDR∗1980	-1.077	-1.063	-1.076	-1.061	-1.078	-1.064
	(0.048)	(0.076)	(0.048)	(0.076)	(0.048)	(0.076)
INTL BRDR∗1990	-0.793	-0.789	-0.790	-0.785	-0.794	-0.790
	(0.038)	(0.056)	(0.037)	(0.055)	(0.037)	(0.055)
INTL BRDR∗2000	-0.262	-0.257	-0.265	-0.261	-0.266	-0.262
	(0.028)	(0.043)	(0.028)	(0.043)	(0.028)	(0.043)
Cluster definition	PAM		Hierarchical		Reclassify "borderline" PTAs	
Period	1978-2018	1978-2018	1978-2018	1978-2018	1978-2018	1978-2018
N. Country ID	136	136	136	136	136	136
Data	Raw	Raw	Raw	Raw	Raw	Raw
Observations	338,685	338,685	338,685	338,685	338,685	338,685
FEs	it, jt, ij	it, jt, ij three-way correction	it, jt, ij	it, jt, ij three-way correction	it, jt, ij	it, jt, ij three-way correction

Notes: Exporter-time (*it*), Importer-time (*jt*) and Exporter-Importer (*ij*) fixed effects are always included. Standard errors in parentheses are clustered by country-pair. In column 2, column 4 and column 6 both standard errors and point estimates are corrected using the Weidner & Zylkin (2021) procedure, implemented in Stata with *ppml_fe_bias*. The Genetic algorithm in column 1 and column 2 classifies: 24 PTAs in cluster #1; 113 PTAs in cluster # 2 and the remaining 141 PTAs in cluster # 3. The Hierarchical algorithm in column 3 and 4 classifies: 33 PTAs in cluster #1; 103 PTAs in cluster # 2 and the remaining 142 PTAs in cluster # 3. The manual reclassification in column 5 and 6 implies: 34 PTAs in cluster #1; 94 PTAs in cluster # 2 and the remaining 150 PTAs in cluster # 3. In column 2, 4 and 6 both standard errors and point estimates are corrected using the Weidner & Zylkin (2021) procedure, implemented in Stata with *ppml_fe_bias*.

Figure 3: Cluster Space



Note: spatial representation of the 3 clusters. Each point represents a trade agreement; x-axis and y-axis are defined using the first two principal components of the 18 features used by the clustering algorithm, centered around zero.

3 Addressing the endogenity of the content of PTAs

As discussed in the Introduction, Blanga-Gubbay et al. (2020) find evidence that big firms in the United States tend to lobby for specific provisions to be included in PTAs. We consequently need to control also for the intensity of bilateral (time-varying) economic interests that may encourage lobbing activity. Baier & Bergstrand (2007) show that dyadic fixed effects are required to control for self-selection of country pairs into PTAs. We already introduced these fixed effects in the estimated equations, in this section we go beyond this approach to investigate endogeneity concerns. As cross-country data on political economy incentives to lobbying – generally used as instrumental variables – are not available, we proceed in three steps.

First, we reallocate randomly country pairs into groups of RTAs, keeping the number of groups and the number of pairs in each group constant.

Second, we introduce a control for the bilateral intensity of GVC income by country pair. This measure traces the income of domestic factors (capital and labor) generated through foreign production chains, thus capturing the intensity of bilateral economic interests (Johnson 2018). Importantly, there is also a potential issue of endogeneity in such metric, as bilateral trade flows are partly determined by GVCs. This is why we rely on the extraction method to filter the contribution of direct trade to ensure we only use "Indirect" GVC income flows.[21]

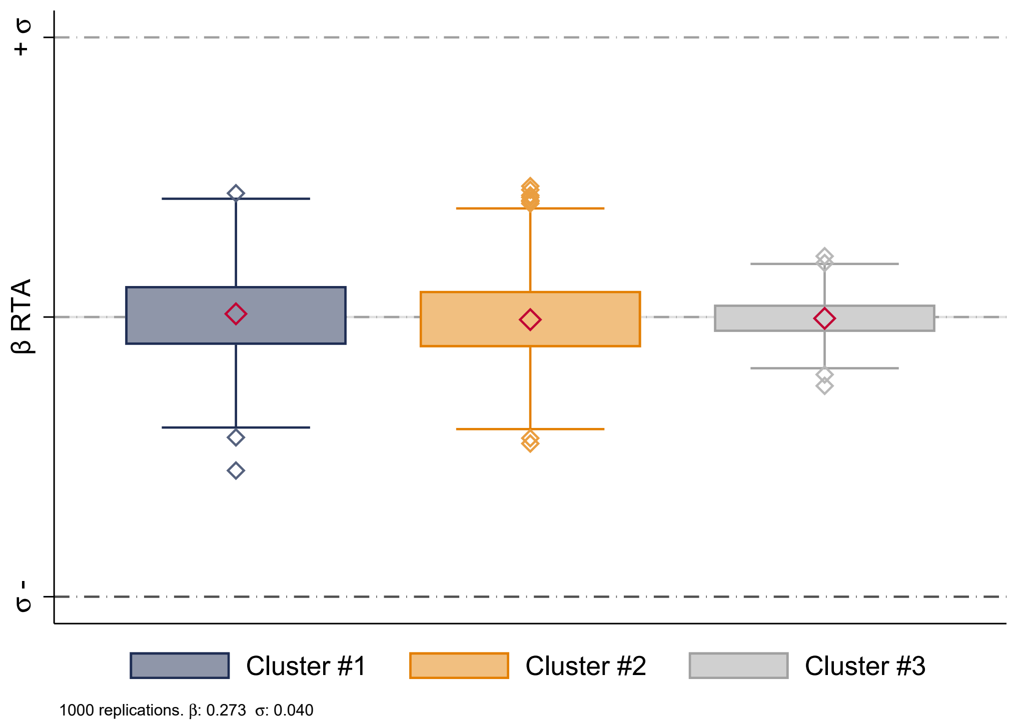
Third, we control for the "dependency" of the importing country vis-a-vis its partner country by normalizing trade flows by destination country total absorption.

[21] See Borin & Mancini (2019) for a presentation of the extraction method as well as other insightful decompositions.

3.1 Randomization of the grouping of country pairs

Figure 4 reports the distribution of the estimated coefficients when the country-pairs sharing a trade agreement in the estimation sample are randomly assigned to a given cluster. After 1,000 replications the estimated coefficient for each cluster is not statistically different from the main PTAs coefficient reported in column 3 of Table 4. If the different impact of PTAs grouped in different clusters is not an artefact, then randomly allocating PTAs to three clusters (and keeping the number of PTAs in each cluster unchanged) should end up in all clusters having the same impact on trade.

Figure 4: Distribution of estimated coefficients for randomly defined clusters (1,000 replications)



Note: The boxes plot the distribution of the estimated partial trade effects of randomly defined clusters. In each of the 1,000 replications we randomly assign country pairs with an PTA to a given cluster.

3.2 Controlling for indirect GVC participation and alternative timing

In order to control for the intensity of country pair production linkages, which may encourage lobbying for selected provisions, we include as a control variable bilateral (indirect) GVC income flows. As the measure is built with an extraction method, it excludes income generated through bilateral trade flows between countries i and j and captures only the bilateral income generated throughout production linkages with third countries. We rely on the EORA MRIO database to trace income flows through production chains; as the data cover the period 1990-2015 we restrict the estimation sample accordingly to 1993-2018, leaving a 3-year lag between the GVC participation and trade variables.[22] Results of this new specification are shown in Table 6. Column 1 replicates the baseline specification (i.e. column 6 of Table 4) but on a shorter period, due to the availability of the multi-region input-output database needed to compute bilateral GVC income flows. In columns 2 to 4 we include, as a control for the intensity of bilateral production linkages, a dummy variable taking the value of 1 if GVC income flows are above the 99^th, 90^th and 75^th percentile respectively. Finally, in column 5 we introduce in the estimated equation an hyperbolic transformation of the indirect GVC income variable instead of a dummy; while column 6 replicates the last estimation using the correction for incidental parameter problem from Weidner & Zylkin (2021).

In Table 7, we test across periods the robustness of our main finding, i.e. preferential trade agreements in cluster # 1 are more effective in promoting trade between members. To proceed we first report in column 1 the same specification as in column 3 of Table 6, and then replicate the same specification over shorter and more recent time periods: 1998-2018 (column 2), 2003-2018 (column 3), and 2008-2018 (column 4). As in the previous tables, column 5 of Table 7 replicates the last estimation using the correction for incidental parameter problem.

Reassuringly, neither the inclusion of the indirect GVC intensity nor the change in time horizons affects the relative magnitude of the estimated coefficients for the different clusters.

[22] The EORA database provides sectoral input-output (IO) tables for 189 countries (and a Rest of the World aggregate) https://worldmrio.com/eora26/.

Table 6: Elasticity of trade to PTAs by Cluster controlling for GVC^{Indirect} income flows

Dep Var:	X_{ijt}/X_{jt}					
	(1)	(2)	(3)	(4)	(5)	(6)
PTA$^{K\#1}_{ij,t}$	0.447	0.447	0.446	0.446	0.440	0.454
	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.048)
PTA$^{K\#2}_{ij,t}$	0.139	0.139	0.138	0.139	0.135	0.152
	(0.028)	(0.028)	(0.028)	(0.028)	(0.029)	(0.033)
PTA$^{K\#3}_{ij,t}$	0.050	0.050	0.050	0.050	0.050	0.073
	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.058)
GVC$^{99pc}_{ij,t-3}$		-0.002				
		(0.035)				
GVC$^{90pc}_{ij,t-3}$			0.032			
			(0.027)			
GVC$^{75pc}_{ij,t-3}$				0.042		
				(0.031)		
$IHS(GVC_{ij,t-3})$					0.026	0.034
					(0.017)	(0.024)
Period	1993-2018	1993-2018	1993-2018	1993-2018	1993-2018	1993-2018
N. Country ID	120	120	120	120	120	120
Observations	230,059	230,059	230,059	230,059	230,059	230,059
FEs	it, jt, ij	it, jt, ij	it, jt, ij	it, jt, ij	it, jt, ij	it, jt, ij three-way correction

Exporter-time (it), Importer-time (jt), Exporter-Importer (ij) FEs, International Border and Transitory PTAs always included. Standard errors in parentheses clustered by country-pair. $GVC_{ij,t-3}$) measures the income generated through *indirect* supply chain linkages between country i and j, excluding any direct effect of trade in intermediate goods between country pair by extraction method. In column 2 to 4 *indirect* $GVC_{ij,t-3}$) income is included as a dummy variable if bilateral flows are above the 95th, 90th or 75th percentile respectively. In column 5 and 6 $IHS(GVC_{ij,t-2})$ refers to inverse hyperbolic sine function (IHS). In column 6 both standard errors and point estimates are corrected using the Weidner & Zylkin (2021) procedure, implemented in Stata with *ppml_fe_bias*.

Table 7: Elasticity of trade to PTAs by Cluster and Period, controlling for GVC^{Indirect} income flows

Dep Var:	X_{ijt}/X_{jt}				
	(1)	(2)	(3)	(4)	(5)
$\text{PTA}^{K\#1}_{ij,t}$	0.446	0.388	0.341	0.396	0.387
	(0.039)	(0.043)	(0.046)	(0.100)	(0.117)
$\text{PTA}^{K\#2}_{ij,t}$	0.138	0.142	0.135	0.122	0.137
	(0.028)	(0.030)	(0.029)	(0.033)	(0.037)
$\text{PTA}^{K\#3}_{ij,t}$	0.050	-0.047	0.046	0.082	0.075
	(0.048)	(0.050)	(0.039)	(0.064)	(0.070)
Period	1993-2018	1998-2018	2003-2018	2008-2018	2008-2018
N. Country ID	120	111	110	99	99
Observations	230,059	177,745	127,580	77,849	77,849
FEs	it, jt, ij	it, jt, ij	it, jt, ij	it, jt, ij	it, jt, ij three-way correction

Exporter-time (it), Importer-time (jt), Exporter-Importer (ij) FEs, International Border and Transitory PTAs always included. Standard errors in parentheses clustered by country-pair. $\text{GVC}^{90pc}_{ij,t-3}$ is included in all the regressions. $\text{GVC}^{90pc}_{ij,t-3}$ is a dummy variable taking the value of 1 if bilateral GVC income flows is higher than the 90^{th} percentile. GVC measures the income generated through *indirect* supply chain linkages between country i and j, excluding any direct effect of trade in intermediate goods between country pair by extraction method.

4 General equilibrium gravity and counterfactuals

Relying on the analysis of the impact of different clusters of signed PTAs, in this section we assess the economic consequences of deepening existing trade agreements.

4.1 Background

A large body of literature focusing on PTAs and trade relies on the assessment of the *partial* impact of agreements on trade within the countries that are signatories of such agreements (see Limão (2016) for a survey). However, PTAs affect the global matrix of relative trade costs between country pairs and the prices faced by exporters and importers in any country through general equilibrium effects. Thus, quantifying the trade effects of PTAs of uneven ambition can be done in calibrated Computable General Equilibrium (CGE) models or in estimated structural gravity models. Notwithstanding the drawback to rely on elasticities estimated outside the model, CGE models offer a flexible tool to assess the sectoral impact of detailed tariff shocks. However, when the information on trade cost reduction is not sector-specific (as it would be with tariffs or with the sector level estimation of the trade impeding impact of regulations), but origin-destination specific, with no sectoral dimension, as in the case of the used database on PTAs, the advantage of the large sectoral decomposition of these models vanishes, making the structural gravity approach more appealing.

Against this background a recent strand of literature is using estimated models (or a combination of estimation and calibration) inspired by the structural gravity literature initiated by Anderson & Van Wincoop (2003), to assess the GE effects of shocks to the matrix of trade costs (see Yotov et al. (2017) for a didactic presentation). A first intrinsic advantage of these models is to have the trade elasticity estimated with the data used for the counterfactual exercise. The second advantage is to be rather agnostic in terms of the trade effects of provisions of PTAs going beyond the phasing out of tariffs among signatory countries.[23] A natural extension of such modeling approach, in line with the spirit of this paper, is to assess the uneven impacts of PTAs (Baier et al. 2019), provided that the "ambition" of signed agreements differs strongly.

4.2 Quantification strategy

We rely on a general equilibrium gravity model for and endowment economy and quantify the *intensive* margin of regional integration, whereby the countries increase the ambition of existing agreements instead of signing

[23] The last generation of these models relies on the properties of the Poisson Pseudo Maximum Likelihood (PPML) estimator (Silva & Tenreyro 2006) demonstrated by Fally (2015): the solution of the GE system of equations derived from a gravity model can indifferently be estimated (Anderson, Larch & Yotov 2018, Fontagné & Santoni 2021) or computed with a solver Head & Mayer (2014). And when the error term is in a multiplicative form (Anderson et al. 2018) this is equivalent to the so-called "hat algebra" resolution in line for instance with the approach coined as "trade theory with numbers" (Arkolakis, Costinot & Rodríguez-Clare 2012) – and thus not fundamentally different from what a resolution of a CGE implies.

new PTAs.[24] The first consequence of deepening agreements, as evidenced above by our gravity estimation, is to modify the overall structure of trade costs. The bilateral trade adjustment, induced by the change in the policy, manifests through two channels: (i) a direct effect driven by the estimated parameters $\beta_z, z = (1, ...3)$ in Equation 1; (ii) an indirect (general equilibrium) effect induced by third countries adjustments. In our case, the typical example is the impact on trade between the United States and Brazil of a simulated deepening of MERCOSUR. The usual trade diversion effect will show up, which here depends on the content of the agreement. The multilateral resistance terms (MRT thereafter) *à la* Anderson & Van Wincoop (2003) act as general equilibrium trade cost indices transmitting local policy shocks to the overall matrix of trade frictions (these effects are formally described using the standard gravity system of equations reported in Appendix A). The inward MRT P_j on the importer side accounts for the impact on consumers and the outward MRT Π_i for the impact on producers in the exporter country. Ultimately, the effects spill also on the price of the exported variety (by the representative producer) and on the expenditure in the importing country. This corresponds to the *general equilibrium effects* for an endowment economy (Head & Mayer 2014).

After switching the PTA dummies to their new values reflecting the design of the counterfactual, in each simulation X_{ijt} is predicted using the new matrix of PTAs while constraining the coefficients β_z, β_T and the μ_{ij} of equation 1 to their initial value, to obtain counterfactual values for the MRTs and eventually solving for the associated general-equilibrium effects in an endowment economy.[25]

Starting from the baseline trade costs matrix, $t_{ij,t}$:

$$\widehat{t_{ij,t}^{1-\sigma}} = exp(\widehat{\mu_{ij}} + \sum_{z=1}^{3} \widehat{\beta_z} PTA_{ij,t}^{z} + \widehat{\beta}\text{Transitory}PTA_{ij,t} + \widehat{\beta_T} INTL\ BRDR_{ij} * T + ln(X_{ij,t}/\widehat{X_{ij,t}}))$$

where the inclusion of the ratio between observed and predicted trade from Equation 1, $ln(X_{ij,t}/\widehat{X_{ij,t}})$, ensures a perfect fit for the observed trade flows; the equilibrium in each counterfactual derives directly from the adjustments in our system of equations induced by the change in the trade costs vector, $t_{ij,t}$.

In order to trace these effects we follow Yotov et al. (2017) and Fontagné & Santoni (2021), using the following notations[26]: Y_i the value of production in the exporting country, E_j the expenditure at destination, and Y the value of world output. Q_i is the endowment (the quantity produced) by the exporter country, p_i the factory-gate price of the exporter and ϕ_i is related to the trade balance. The direct effect of a change in trade costs on trade flows between exporter i and importer j, X_{ij}, can be inferred from the estimated coefficients

[24] For comparison purposes, we provide in the 6 first columns of Table A4 in the Appendix a sense of the impact of the extensive margin of regional integration, whereby EAP countries sign new PTAs of different level of ambition with all countries with which they do not yet have an agreement.

[25] In year 2018, 0.06% of dyadic fixed effects cannot be identified due to separation: this is equivalent to 80 observations out of 12,544 (i.e. 112*112). As we need a value for the dyadic fixed effect for the general equilibrium exercise, as a measure of trade easiness between country-pairs, we assume that the bilateral trade easiness value is equal to 0.

[26] See Appendix A for a more detailed presentation

of the structural gravity equation, holding the MRTs Π_i and P_j constant. In turn, MRTs are impacted by the change in trade costs implied by our counterfactuals, because deepening an existing PTA or signing a new between countries i and j will affect the overall matrix of trade costs and thus the structure of relative prices. These indirect effects as well as their feedbacks on exporter and importer countries relative prices concur in determining the final GE effects for an endowment economy.

As the series of bilateral estimated fixed effects are the counterparts of the MRTs when relying on a PPML estimator (Fally 2015), we follow Anderson et al. (2018), Yotov et al. (2017) and solve our system of equations accordingly. The same approach pertains to our counterfactuals, whereby the system is solved with the alternative trade frictions derived from signing missing PTAs among countries in a region or alternatively deepening the already signed ones. We firstly recover β_z (the average trade cost elasticity over the period considered) and μ_{ij} (the bilateral fixed effects) from the baseline gravity Equation (1) including both intra-national and inter-national yearly trade flows covering the period 1978-2018; then we solve the counterfactual gravity system and compute the associated general equilibrium indices. We solve the model using the "estibration" procedure (Anderson et al. 2018, Yotov et al. 2017) which gives a solution identical to the "exact hat" algebra (Dekle, Eaton & Kortum 2007).[27]

4.3 Deepening existing or new PTAs

We proceed in this section in four steps. Results are given in terms of relative variation, i.e. in percentage deviation in the counterfactual relative to the baseline scenario. Computations are executed on trade flows for year 2018.

In a first counterfactual, we switch all the *existing* agreements from their current level of ambition to the highest level of ambition and compute the change in countries' total exports and GDP. Individual effects are weighted by country GDP or export and add up to a 3.9 percent increase in exports and 0.9 percent increase in GDP for the world economy as a whole. The results by country of such generalized shift to deep trade agreements, hence at the intensive margin, are shown in Figure 5. Few countries are negatively affected by the global deepening of trade cooperation but overall gains are significantly positive for most of the countries, including for many developing economies.

In a second counterfactual, we shift *region by region* all the existing agreements of the region from their current level of ambition to the highest level of ambition and again compute the change in total exports and GDP for the countries in each region separately before aggregating these results at the region level. Results for the different regions are reported in column (1) and (2) of Table 8. As an illustration, deepening all existing

[27] Multilateral resistance terms and factory-gate prices are indexes subject to a normalisation. This imposes to choose a reference country not directly impacted by the counterfactual and to restrict the exercise to cross-sectional comparisons across countries. See Chapter 2 of Yotov et al. (2017) for a detailed discussion.

agreements among countries in the East Asia and Pacific region would lead to, on average, a 5.99 percent increase in their exports and a 1.05 percent in their GDP, compared to a baseline where these agreements have the current level of depth. This average increase is done in general equilibrium but at the country level before aggregation, and the detailed effects at the country level are shown in the two first columns in Table A3 in the Appendix for this region.

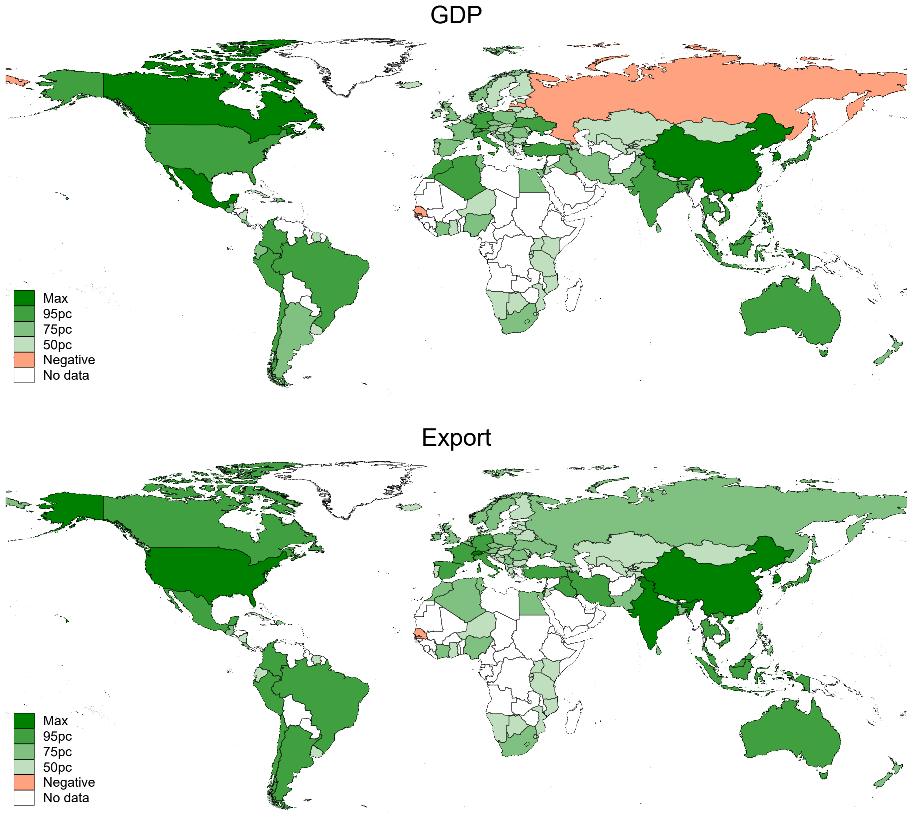
The third step addresses a different question: is it preferable for a given region to deepen the PTAs signed within other countries in the region? Or alternatively to deepen the PTAs signed with countries external to the region? Specifically, in columns (3) and (4) of Table 8, we report the results for trade and GDP when we switch all the existing agreements *within* each region of the world economy from their current level of ambition to the highest level of ambition, and keep constant the content of all current extra-regional PTAs. In contrast, columns (5) and (6) present the results when we switch all the existing agreements of each region with partners *outside* of the region from their current level to the highest level of ambition while keeping unaltered the content of PTAs with other regional partners. As before, we compute the change in total exports and GDP of each region to illustrate the results and we report the detail for the EAP region in columns (3) to (6) of Table A3 in the Appendix.

This decomposition provides interesting new insights and helps explaining the mechanisms at play. The economic effect of deepening existing trade agreements within regions is driven by two factors. First, the actual level of ambition of existing agreements within the different regions, as opposed to between regions. Second, the economic size of trading partners with whom the deepening of an existing agreement is envisaged. Given the geography and depth of the existing agreements, the economic effects are very different across regions. On one extreme, we have a situation where most of the potential gains come from deepening intra-regional PTAs, as in the East-Asia and Pacific case where agreements in the baseline are relatively shallow and the economic size of partners is significant. On the other extreme, we have a situation where all the gains come from deepening agreements with partners outside the region. This is the case for a region like the Middle East and North Africa where larger markets are predominantly extra-regional. In between these situations, Latin America, Africa, South Asia and Europe and Central Asia would have similar gains in deepening trade agreements within or outside the region.

We ultimately turn to a different question. Instead of deepening existing agreements, we are signing new agreements with a different level of ambition, allowing us to address the impact of the extensive margin of regional integration in a fourth step. Columns (1) to (6) in Table A4 in the Appendix show the trade and GDP impact of EAP countries signing all missing agreements with or without the region. The results are detailed by country in the region and aggregated in the penultimate row of the table. The last row shows the impact on countries out of the region. On the extensive margin as well signing a deep agreement makes a big difference:

the average impact on exports of EAP countries is 17.8% and the impact on GDP 1.8%, to be compared with respectively 4.7% and 0.4% would new shallow agreements be signed. A side result is illustrated by comparing columns (3) and (4) with columns (7) and (8) respectively.

Figure 5: General equilibrium effects of deepening existing PTAs



Note: the map above shows the contribution of each country to the "World" effect as reported in column 1 and 2 of Table 8. Five countries report a moderate or insignificant contraction in GDP following a global deepening of trade cooperation: Russia (0.063 percent), Kuwait (0.006), Senegal (0.003), Palestine and Latvia (both < 0.002).

Table 8: General-Equilibrium effects of deeper trade integration by Region, year 2018.

Country Name	Iso3	*All* PTA^{Deep} Δ Export (1)	Δ GDP (2)	*Within Region* PTA^{Deep} Δ Export (3)	Δ GDP (4)	*With RoW* PTA^{Deep} Δ Export (5)	Δ GDP (6)
East Asia & Pacific	EAP	5.99	1.05	4.66	0.86	1.24	0.05
Europe & Central Asia	ECA	1.66	0.72	1.14	0.55	0.50	0.08
Latin America & Caribbean	LAC	8.74	3.83	2.79	0.73	4.68	1.45
Middle East & North Africa	MENA	8.13	0.37	0.87	0.04	4.76	0.17
North America	NA	6.35	1.06	2.91	0.65	2.15	0.23
South Asia	SA	15.61	1.68	1.96	0.17	5.05	0.45
Sub-Saharan Africa	SSA	3.83	0.80	0.39	0.10	1.93	0.37

Note: percent change compared to the baseline in total exports and GDP of participating countries. We assume $\sigma = 5$. The reference country for the normalization is South Africa. The list of PTAs for the reference country remains at the baseline. For each region, all the existing agreements of the region are shifted from their current level of ambition to the highest level of ambition. The variation in total exports and GDP for the countries in each region were computed separately and aggregated using GDP or exports weights.

5 Conclusion

This paper uses new data on the content of trade agreements and a structural gravity general equilibrium model to quantitatively assess the economic impacts of deepening trade agreements. Based on a clustering of 278 PTAs, comprising 910 provisions grouped in 18 policy areas, we have shown that PTAs of higher ambition are associated with a much larger trade elasticity to PTAs. This finding confirms that using an average effect of PTAs, disregarding the depth and content of trade agreements, is misleading. We then simulated a series of full general equilibrium counterfactual situations for endowment economies, revealing the economic impacts of deepening all existing PTAs, and of deepening trade agreements within regions and between regions. These exercises focusing on the intensive margin of PTAs suggest that deepening trade agreements can have significant effects on trade and GDP. The effects of deepening trade agreements within and between regions vary across countries and world regions depending on the initial depth of trade agreements and on the size of regional markets, but are largely positive. Potential differences in collective preferences between regions, which make negotiating deep trade agreements more difficult, suggest that for most countries the low-hanging fruits may be in deepening trade agreements within regions.

Bibliography

Anderson, J. E., Larch, M. & Yotov, Y. V. (2018), 'GEPPML: General equilibrium analysis with PPML', *The World Economy* **41**(10), 2750–2782.

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

Anderson, J. E. & Yotov, Y. V. (2016), 'Terms of trade and global efficiency effects of free trade agreements, 1990–2002', *Journal of International Economics* **99**, 279–298.

Arkolakis, C., Costinot, A. & Rodríguez-Clare, A. (2012), 'New trade models, same old gains?', *American Economic Review* **102**(1), 94–130.

Arthur, D. & Vassilvitskii, S. (2007), K-means++: The advantages of careful seeding, *in* 'Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms', SODA '07, Society for Industrial and Applied Mathematics, USA, p. 1027–1035.

Baier, S. L. & Bergstrand, J. H. (2004), 'Economic determinants of free trade agreements', *Journal of International Economics* **64**(1), 29–63.

Baier, S. L. & Bergstrand, J. H. (2007), 'Do free trade agreements actually increase members' international trade?', *Journal of international Economics* **71**(1), 72–95.

Baier, S. L., Yotov, Y. V. & Zylkin, T. (2019), 'On the widely differing effects of free trade agreements: Lessons from twenty years of trade integration', *Journal of International Economics* **116**, 206–226.

Blanchard, E. J., Bown, C. P. & Johnson, R. C. (2016), Global value chains and trade policy, Technical report, Working paper NBER N.21883.

Blanga-Gubbay, M., Conconi, P. & Parenti, M. (2020), 'Lobbying for globalization'.

Borin, A. & Mancini, M. (2019), Measuring What Matters in Global Value Chains and Value-Added Trade, Policy Research Working Paper Series 8804, The World Bank.
URL: *https://ideas.repec.org/p/wbk/wbrwps/8804.html*

Bown, C. P., Conconi, P., Erbahar, A. & Trimarchi, L. (2021), 'Trade protection along supply chains'.

Bown, C. P., Erbahar, A. & Zanardi, M. (2021), 'Global value chains and the removal of trade protection', *European Economic Review* **140**, 103937.

Dekle, R., Eaton, J. & Kortum, S. (2007), 'Unbalanced trade', *American Economic Review* **97**(2), 351–355.
URL: *http://www.aeaweb.org/articles?id=10.1257/aer.97.2.351*

Eaton, J., Kortum, S. & Sotelo, S. (2013), *International Trade: Linking Micro and Macro*, Vol. 2 of *Econometric Society Monographs*, Cambridge University Press, p. 329–370.

Egger, P. H., Larch, M. & Yotov, Y. (2022), 'Gravity estimations with interval data: Revisiting the impact of free trade agreements', *Economica* **89**(353), 44–61.

Egger, P. H. & Nigai, S. (2015), 'Structural gravity with dummies only: Constrained anova-type estimation of gravity models', *Journal of International Economics* **97**(1), 86–99.

Fally, T. (2015), 'Structural gravity and fixed effects', *Journal of International Economics* **97**(1), 76 – 85.

Fontagné, L., Guimbard, H. & Orefice, G. (2022), 'Tariff-based product-level trade elasticities', *Journal of International Economics* **137**, 103593.
URL: *https://www.sciencedirect.com/science/article/pii/S0022199622000253*

Fontagné, L. & Santoni, G. (2021), 'Gvcs and the endogenous geography of rtas', *European Economic Review* **132**, 103656.
URL: *https://www.sciencedirect.com/science/article/pii/S001429212100009X*

Head, K. & Mayer, T. (2014), *Gravity Equations: Workhorse, Toolkit, and Cookbook*, Vol. 4 of *Handbook of International Economics*, Elsevier, chapter 3, pp. 131–195.

Head, K. & Mayer, T. (2021), 'The united states of Europe: A gravity model evaluation of the four freedoms', *Journal of Economic Perspectives* **35**(2), 23–48.
URL: *https://www.aeaweb.org/articles?id=10.1257/jep.35.2.23*

Johnson, R. C. (2018), 'Measuring Global Value Chains', *Annual Review of Economics* **10**(1), 207–236.
URL: *https://ideas.repec.org/a/anr/reveco/v10y2018p207-236.html*

Kharel, P. (2019), 'The effect of free trade agreements revisited: Does residual trade cost bias matter?', *Review of International Economics* **27**(1), 367–389.

Limão, N. (2016), Preferential trade agreements, *in* 'Handbook of Commercial Policy', Vol. 1, Elsevier, pp. 279–367.

Mattoo, A., Mulabdic, A. & Ruta, M. (2017), Trade creation and trade diversion in deep agreements, Technical report, Policy Research Working Paper 8206.

Mattoo, A., Rocha, N. & Ruta, M. (2020), *Handbook of Deep Trade Agreements*, number 34055 *in* 'World Bank Publications', The World Bank.
URL: *https://ideas.repec.org/b/wbk/wbpubs/34055.html*

Rousseeuw, P. J. (1987), 'Silhouettes: a graphical aid to the interpretation and validation of cluster analysis', *Journal of computational and applied mathematics* **20**, 53–65.

Silva, J. S. & Tenreyro, S. (2006), 'The log of gravity', *Review of Economics and Statistics* **88**(4), 641–658.

Sotelo, S. (2019), Practical Aspects of Implementing the MultinomialPML Estimator, mimeo.

Weidner, M. & Zylkin, T. (2021), 'Bias and consistency in three-way gravity models', *Journal of International Economics* **132**, 103513.
URL: *https://www.sciencedirect.com/science/article/pii/S0022199621000933*

Yotov, Y. V. (2012), 'A simple solution to the distance puzzle in international trade', *Economics Letters* **117**(3), 794–798.

Yotov, Y. V. (2021), 'The variation of gravity within countries (or 15 reasons why gravity should be estimated with domestic trade flows)'.

Yotov, Y. V., Piermartini, R., Monteiro, J.-A. & Larch, M. (2017), *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*, United Nations publication, World Trade Organization.

Appendix

A The Calculation of the General-Equilibrium Baseline and Counterfactual

Partial- and general equilibrium effects can be described using a standard gravity system reported in Equations A1 to A3, following typical notation: t_{ij} is the vector of trade frictions, Y_i the value of production in the exporting country, E_j the expenditure at destination, and Y the value of World output. Q_i is the endowment (the quantity produced) in the exporting country, p_i the factory-gate price, and ϕ_i is capturing trade balance. The direct effect of a change in trade costs on trade flows between exporter i and importer j, X_{ij}, can be inferred by the estimated coefficients from Equation A1, holding the MRTs Π_i and P_j constant. However, MRTs are not constant, as changing bilateral trade costs between any i and j affects the world structure of trade frictions and relative prices. The indirect (general equilibrium) effect on i and j can accordingly be derived from Equation A2, which is the solution of the conditional General-Equilibrium effects. As a PTA between i and j becomes deeper, this reduces the MRTs for this country pair but increases them for third countries. The feedback on i and j of these changes is given by the General-Equilibrium effect described in terms of the price of the representative firm and expenditure in i (Equation A3).

$$\text{Partial Equilibrium}\begin{cases} X_{ij} = & (\frac{t_{ij}}{\Pi_i P_j})^{1-\sigma}\frac{Y_i E_j}{Y} \end{cases} \tag{A1}$$

$$\text{Conditional GE}\begin{cases} \Pi_i^{1-\sigma} = & \sum_j(\frac{t_{ij}}{P_j})^{1-\sigma}\frac{E_j}{Y} \\ P_j^{1-\sigma} = & \sum_i(\frac{t_{ij}}{\Pi_i})^{1-\sigma}\frac{Y_i}{Y} \end{cases} \tag{A2}$$

$$\text{GE}\begin{cases} p_i = & \frac{(Y_i/Y)^{\frac{1}{1-\sigma}}}{\alpha_i \Pi_i} \\ E_i = & = \phi_i Y_i = \phi_i Q_i p_i \end{cases} \tag{A3}$$

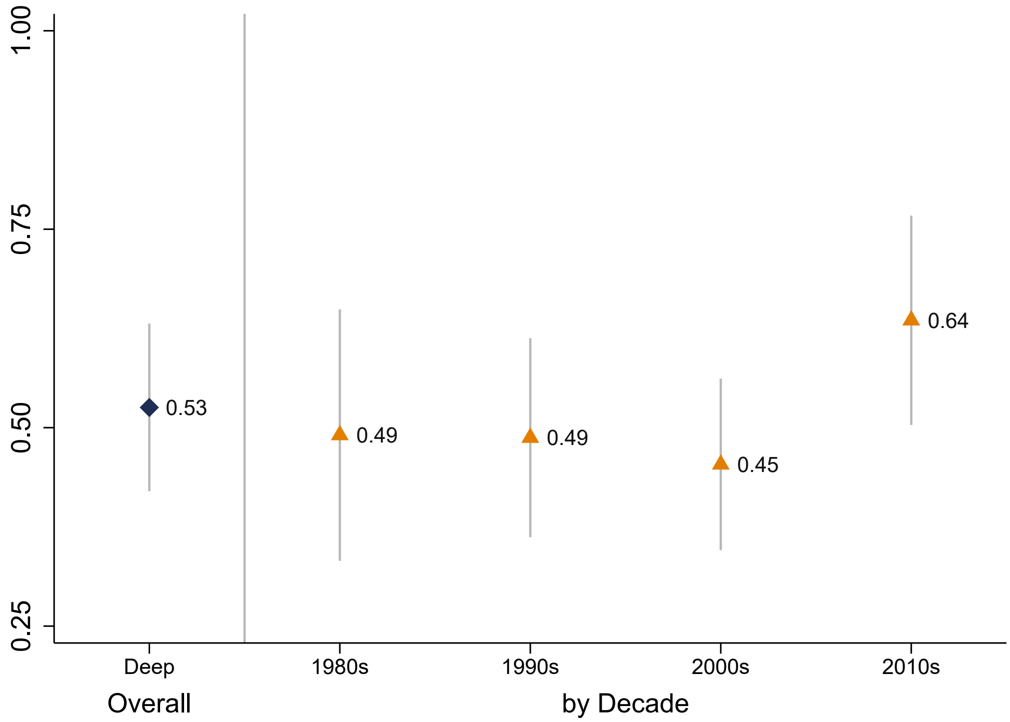
B Additional tables and figures

Table A1: Actual values of the raw and normalized matrix for the area "Services"

Policy Area	Provision	Cluster # 1 EFTA	...	Cluster # 2 PAN-SGP	CHL-JPN	...	Cluster # 3 APTA	TUR-GEO	...	Average
Raw Matrix										
Services	836	0	...	1	1	...	0	0	...	0.29
Services	837	1	...	1	0	...	0	0	...	0.27
Services	850	5	...	1	2	...	1	0	...	1.68
Services	...	...	...	...	...	...	...	...	...	...
Provisions by PTA	62	22	...	50	48	...	17	0	...	
Coverage by cluster		**0.289**		**0.460**			**0.267**			
Normalized Matrix										
Services	836	0.00	...	3.43	3.43	...	0.00	0.00	...	
Services	837	3.71	...	3.71	0.00	...	0.00	0.00	...	
Services	850	2.97	...	0.59	1.19	...	0.59	0.00	...	
Services	...	...	...	...	...	...	...	...	...	...
Score by PTA		0.952	...	2.483	2.901	...	1.084	0.000	...	
Score by cluster		**0.840**		**1.36**			**0.846**			

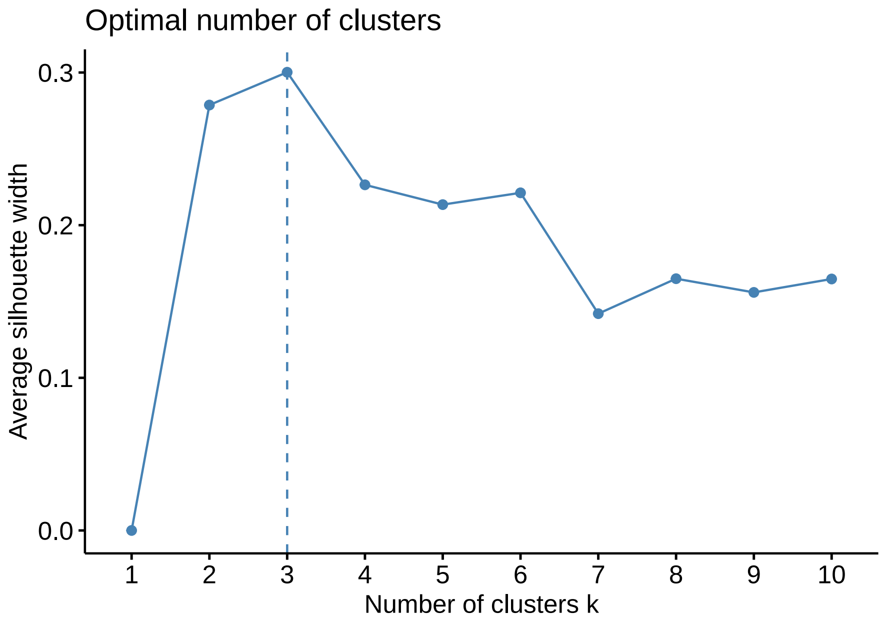
Notes: APTA Asia Pacific Trade Agreement, EFTA European Free Trade Association. Provisions: # 836, Are there sector-specific chapters (e.g. financial services, telecommunications)? Code: $Services - struc_{chapt}$. # 837, Are there sector-specific provisions in an annex to a chapter on investment or CBTS (such as express delivery as an annex to CBTS chapter)? Code: $Services - struc_{chapt1}$. # 850, To be considered a service supplier of a party to the agreement, in the case of the supply of services through commercial presences, does a juridical person have to: "being incorporated under the domestic law of the party and have substantive business operations in the territory of a member" (coded with value 1, most restrictive), ... , "being owned or controlled by natural persons of the other party" (coded with value 6, most liberal). Code: $Services - roo_{m3}$.

Figure A1: PPML: Gravity Estimations of the elasticity of trade to Deep PTAs, by decade



Note: baseline specification refers to column 5 of Table 4; the effect by decade is computed using an interaction between $\text{PTA}_{ij,t}^{K\#1}$ and a dummy variable for each decade, including controls and fixed effects as in the baseline regression.

Figure A2: Average Silhouette Value



Note: Optimal number of clusters of Regional Trade Agreements over 18 provision areas.

Table A2: PPML: Gravity Estimations of the elasticity of trade to PTAs by Cluster with symmetric country-pair FEs

Dep Var:	X_{ijt}				X_{ijt}/X_{jt}	
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{PTA}_{ij,t}$	0.559	0.578	0.268			
	(0.048)	(0.054)	(0.041)			
$\text{PTA}_{ij,t}^{K\#1}$				0.613	0.527	0.484
				(0.047)	(0.041)	(0.072)
$\text{PTA}_{ij,t}^{K\#2}$				0.219	0.197	0.252
				(0.046)	(0.030)	(0.034)
$\text{PTA}_{ij,t}^{K\#3}$				0.108	0.099	0.135
				(0.055)	(0.040)	(0.040)
Transitory $\text{PTAs}_{ij,t}$		0.078	0.102	0.204	0.210	0.131
		(0.051)	(0.047)	(0.055)	(0.038)	(0.043)
INTL BRDR∗1980			-0.852	-0.840	-1.067	-0.990
			(0.044)	(0.043)	(0.049)	(0.063)
INTL BRDR∗1990			-0.522	-0.524	-0.782	-0.773
			(0.035)	(0.034)	(0.038)	(0.045)
INTL BRDR∗2000			-0.164	-0.166	-0.253	-0.314
			(0.027)	(0.027)	(0.028)	(0.033)
Intra-National flows	Raw	Raw	Raw	Raw	Raw	Extrapolated
Period	1978-2018	1978-2018	1978-2018	1978-2018	1978-2018	1978-2018
Frequency	Yearly	Yearly	Yearly	Yearly	Yearly	Yearly
N. Country ID	133	133	133	133	133	142
Observations	343,280	343,280	343,280	343,280	343,280	597,180
FEs	it, jt, Sym-ij	it, jt, Sym-ij	it, jt, Sym-ij	it, jt, Sym-ij	it, jt, Sym-ij	it, jt, Sym-ij

Notes: Exporter-time (it), Importer-time (jt) and Exporter-Importer (ij) fixed effects are always included. Standard errors in parentheses are clustered by country-pair. In column 6 missing values in domestic sales are linearly interpolated and the remaining missing values are extrapolated using the evolution of a country total exports.

Table A3: General-Equilibrium effects of deeper trade integration for EAP countries, year 2018.

Country Name	Iso3	*All* PTA^{Deep}		*Within Region* PTA^{Deep}		*With RoW* PTA^{Deep}	
		Δ Export	Δ GDP	Δ Export	Δ GDP	Δ Export	Δ GDP
		(1)	(2)	(3)	(4)	(5)	(6)
Australia	AUS	5.28	5.90	4.91	5.25	0.69	0.17
China	CHN	4.93	0.31	4.17	0.26	0.70	-0.01
Fiji	FJI	7.83	1.86	5.32	1.34	2.07	0.19
Hong Kong	HKG	-4.20	-0.46	-3.03	-0.33	-1.68	-0.24
Indonesia	IDN	14.41	4.64	12.64	4.13	1.92	0.20
Japan	JPN	3.82	0.39	2.81	0.29	1.14	0.03
Korea	KOR	12.95	2.68	7.04	1.44	4.38	0.45
Malaysia	MYS	3.55	8.26	3.10	7.52	0.91	0.31
Mongolia	MNG	-0.14	0.25	-0.26	0.22	0.44	0.14
New Zealand	NZL	6.68	3.18	6.72	3.22	0.36	-0.03
Philippines	PHL	4.08	9.93	3.83	9.69	0.69	0.05
Singapore	SGP	4.08	8.47	2.76	6.04	1.66	1.06
Thailand	THA	11.20	5.42	10.02	4.93	1.39	0.21
Viet Nam	VNM	4.39	9.08	3.87	8.47	0.77	0.06
EAP		5.99	1.05	4.66	0.86	1.24	0.05
RoW		0.37	0.09	-0.37	-0.07	0.16	0.10

Note: percent change compared to the baseline in total exports and GDP of participating countries. We assume $\sigma = 5$. The reference country for the normalization is South Africa. EAP: East Asia & Pacific. In the counterfactual the level of ambition of each PTA already signed by an EAP country is shifted to the highest level of ambition. This is done for PTAS within the region in columns (3 and (4), for PTAs with the Rest of the World in columns (5) and (6) and for all agreements signed by countries in the EAP region in columns (1) and (2).

Table A4: General-Equilibrium effects of deeper integration for EAP countries, year 2018, *Extensive Margin.*

Country Name	Iso3	*New* PTA^{Deep} Δ Export (1)	Δ GDP (2)	*New* PTA^{Medium} Δ Export (3)	Δ GDP (4)	*New* $PTA^{Shallow}$ Δ Export (5)	Δ GDP (6)	*New* PTA^{Dummy} Δ Export (7)	Δ GDP (8)
Australia	AUS	4.56	4.17	2.30	2.04	1.31	0.98	2.29	2.03
China	CHN	23.36	1.34	11.19	0.65	6.16	0.29	11.12	0.65
Fiji	FJI	28.15	6.65	14.44	3.24	7.30	1.59	14.35	3.22
Hong Kong	HKG	43.52	5.07	21.73	2.43	10.99	1.17	21.60	2.42
Indonesia	IDN	8.41	1.84	4.12	0.88	2.19	0.41	4.10	0.88
Japan	JPN	31.48	2.95	15.41	1.43	8.11	0.66	15.32	1.42
Korea	KOR	5.45	0.86	2.70	0.43	1.51	0.20	2.69	0.43
Malaysia	MYS	4.64	5.23	2.36	2.55	1.27	1.24	2.35	2.54
Mongolia	MNG	12.32	22.68	6.44	11.67	3.49	5.85	6.41	11.60
New Zealand	NZL	9.82	4.13	4.90	2.00	2.58	0.96	4.87	1.98
Philippines	PHL	5.23	5.85	2.67	2.88	1.43	1.41	2.66	2.87
Singapore	SGP	3.93	3.93	2.00	1.92	1.09	0.93	1.98	1.91
Thailand	THA	8.74	2.78	4.34	1.34	2.28	0.64	4.31	1.33
Viet Nam	VNM	6.29	5.09	3.20	2.50	1.66	1.23	3.18	2.49
EAP		17.79	1.76	8.63	0.86	4.68	0.39	8.58	0.86
RoW		5.39	0.95	2.62	0.45	1.22	0.23	2.61	0.45

Note: percent change compared to the baseline in total exports and GDP of participating countries. We assume $\sigma = 5$. The reference country for the normalization is South Africa. EAP: East Asia & Pacific. The PTAs elasticity by depth are the same as in the baseline, i.e. column 6 of Table 4. The estimated elasticity for the dummy approach to PTA, used in column 7 and column 8, takes the value of: 0.258 with a standard error clustered by country-pair of 0.034. All the estimated equations include exporter-time, importer-time and exporter-importer fixed effects as well as controls for international border and transitory PTAs.

Table A5: Country coverage for the counterfactual exercise, year 2018

Countries with non-missing domestic sales		Additional countries using extrapolated domestic sales	
Argentina	Italy	Albania	Mongolia
Armenia	Jordan	Azerbaijan	Mozambique
Australia	Japan	Burundi	Malaysia
Austria	Kenya	Belgium	Namibia
Bulgaria	Kyrgyzstan	Bangladesh	Nigeria
Bahrain	Korea	Bermuda	Nicaragua
Bosnia and Herzegovina	Sri Lanka	Barbados	Netherlands
Belarus	Morocco	Botswana	Norway
Brazil	Mauritius	Chile	Nepal
Canada	Niger	Cote d'Ivoire	Pakistan
Switzerland	New Zealand	Cuba	Rwanda
China	Palestine	Cyprus	Senegal
Colombia	Peru	Dominican Republic	Singapore
Costa Rica	Philippines	Algeria	Suriname
Czech Republic	Poland	Estonia	Slovenia
Germany	Portugal	Ghana	Swaziland
Denmark	Romania	Guatemala	Trinidad and Tobago
Ecuador	Russian Federation	Hong Kong	Tunisia
Egypt	Slovakia	Hungary	Uganda
Spain	Sweden	Indonesia	Uruguay
Finland	Thailand	Iraq	
Fiji	Tajikistan	Iceland	
France	Turkey	Kazakstan	
United Kingdom	Tanzania, United Rep. of	Kuwait	
Georgia	Ukraine	Lithuania	
Greece	United States of America	Latvia	
Croatia	Uzbekistan	Moldova	
India	Viet Nam	Maldives	
Ireland	Serbia and Montenegro	Mexico	
Iran	South Africa	Macedonia	
Israel		Malta	

Notes: The overall number of countries is 112; for 51 of them domestic sales have being extrapolated using total exports.