Maastricht University

UNITED NATIONS UNIVERSITY

UNU-MERIT

Working Paper Series

#2017-038

On the relationship between the breadth of PTAs and trade flows

Rod Falvey and Neil Foster-McGregor

Maastricht Economic and social Research institute on Innovation and Technology (UNU-MERIT)
email: info@merit.unu.edu | website: http://www.merit.unu.edu

Maastricht Graduate School of Governance (MGSoG)
email: info-governance@maastrichtuniversity.nl | website: http://www.maastrichtuniversity.nl/governance

Boschstraat 24, 6211 AX Maastricht, The Netherlands
Tel: (31) (43) 388 44 00

UNU-MERIT Working Papers
ISSN 1871-9872

Maastricht Economic and social Research Institute on Innovation and Technology
UNU-MERIT

Maastricht Graduate School of Governance
MGSoG

UNU-MERIT Working Papers intend to disseminate preliminary results of research carried out at UNU-MERIT and MGSoG to stimulate discussion on the issues raised.

September 2017

On the Relationship between the Breadth of PTAs and Trade Flows

Rod Falvey[♦]

Neil Foster-McGregor[*]

Abstract

This paper uses matching econometrics to extend the literature investigating the impact of Preferential Trading Arrangements (PTAs) on goods trade flows. Heterogeneity in PTAs is accounted for through a 'provision count index' derived from data provided in a recent World Bank study (Hofmann et al, 2017). PTA formation now involves two separate, sequential decisions – first, whether two trading partners should form a PTA and, second, if they do, how broad that agreement should be. We find that our explanatory variables are significant for both decisions, but often have opposing effects on each. Using our matched PTA and non-PTA groups of country-pairs, we estimate a dose response function which indicates that arrangements with few provisions and arrangements with many provisions do not appear to have a significant impact on goods trade flows between their members. PTAs in an intermediate range are shown to have a significant positive effect. We then relate these outcomes to the actual content of the PTAs using the concept of 'provision intensity'.

Keywords: Preferential Trade Agreements, Trade Flows, PTA Breadth.

JEL Classification: F10, F15

[♦] Rod Falvey, Bond Business School, Bond University, Australia. Email: rfalvey@bond.edu.au

[*] Neil Foster-McGregor, UNU-MERIT, Boschstraat 24, 6211AX, Maastricht, The Netherlands. Email: foster@merit.unu.edu

1. Introduction

Preferential Trading Arrangements (PTAs) have proliferated, particularly in the last twenty years, as shown in Figure 1 which plots the number of country-pairs in our sample with a PTA along with the number of country-pairs that formed a PTA in each year from 1970 to 2015. The share of world trade between PTA members has increased from 22% in 1965 to 60% in 2010, while the share of World Trade Organisation (WTO) country pairs with PTAs rose from 2% in 1965 to more than 25% in 2010 and their corresponding trade share within the WTO rose from 30% to 60% (Limao, 2016). With no progress in multilateral liberalisation in recent years and its immediate prospects looking particularly grim, PTAs have been and are likely to continue to be the most important source of trade policy reform for most countries. While average tariffs, and hence the preferential margins, have been falling, there remain a wide range of non-tariff barriers to be negotiated away.

The quantitative analysis of PTAs has also proliferated. The bulk of this analysis has employed some variant of the 'gravity equation', interpreted as a reduced form equation which may arise from a variety of trade models explaining bilateral trade flows. The typical approach has been to include a dummy variable indicating whether or not a bilateral trade flow was covered by a PTA and to interpret the estimated coefficient on this dummy as indicating the average effect of a PTA. Results from this literature tend to suggest that PTAs have a positive impact upon trade flows, but there is a great deal of heterogeneity in the estimated effects.[1] This heterogeneity takes many forms with PTA effects varying across time (Baier and Bergstrand, 2009), across agreements (Eicher and Henn, 2011), and by trade partner (Eicher and Henn, 2011; Cheong et al, 2014).

In this paper we extend a recent literature interested in identifying and explaining the heterogeneous effects of PTAs (e.g. Baier et al, 2016; Cipollina and Salvatici, 2010; Eicher and Henn, 2011). The starting point for our analysis is the observation that PTAs differ quite widely in terms of the characteristics of the country-pairs that form them, and that these differences have tended to change over time. To illustrate this Table 1 reports means and standard deviations of selected characteristics of the country pairs that formed PTAs in each of the four decades between 1970 and 2010. The largest average PTAs (as measured by the sum of their

[1] See Cipollina and Salvatici (2010) and World Bank (2005) for meta-analyses of the trade effects of PTAs. World Bank (2005), for example, considers 362 estimates of a PTA dummy from 17 studies that cover different PTAs, time periods and equation specifications. One-third of the estimates are statistically insignificant, over 10% are negative and significant, and only just over 50% are positive and significant. The mean estimate is 0.79 but the standard error is 1.3.

GDPs) were signed in the 1990s and the smallest in the 2000s. In the most recent period we also observe that PTAs have been signed largely between country-pairs that do not share a common language or a common border. We further observe a tendency for PTA partners to become increasingly distant over the period.[2]

Figure 1: Number of New Country-Pairs with PTAs by Year

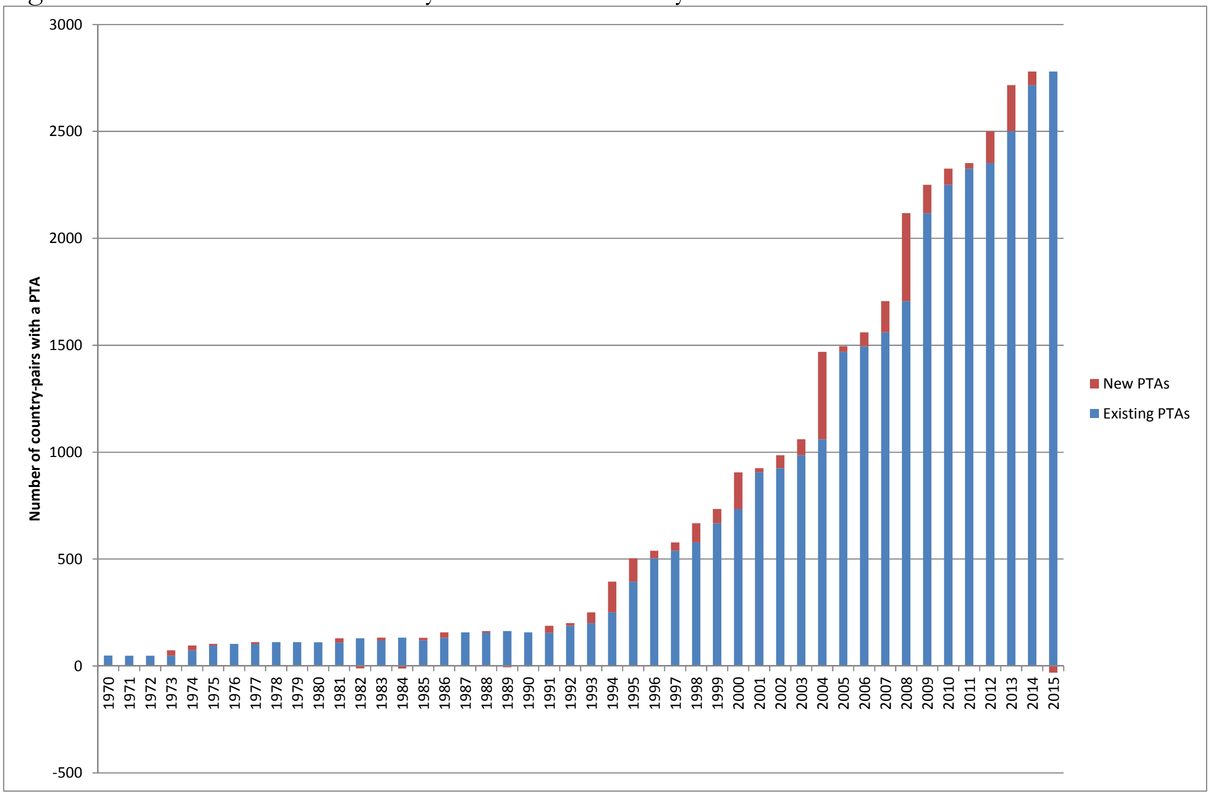


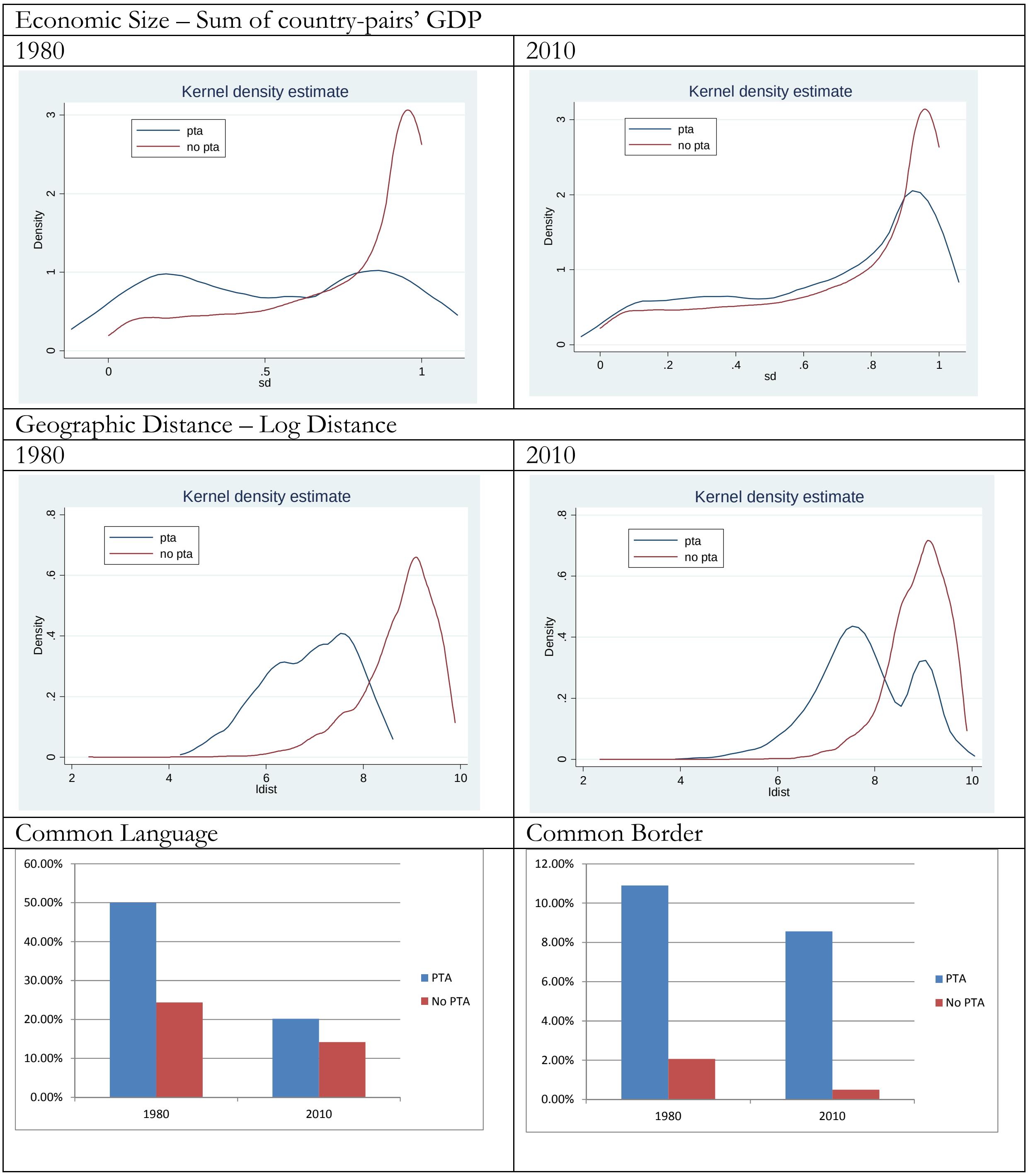
Table 1: Characteristics of new PTA Country Pairs (by decade)

	PTAs signed in:			
	1980s	1990s	2000s	2010s
GDP Sum	48.907	50.101	47.233	49.457
	(4.933)	(4.350)	(3.762)	(3.054)
Dist	0.069	0.107	0.099	0.238
	(0.052)	(0.063)	(0.065)	(0.199)
Lang (%)	0.581	0.434	0.449	0.137
	(0.497)	(0.500)	(0.498)	(0.344)
Adj (%)	0.048	0.132	0.189	0.040
	(0.216)	(0.342)	(0.392)	(0.197)

Notes: Standard deviations in parentheses. GDP Sum is the log of the sum of the real GDPs of the two PTA members; Distance is the distance between the two PTA partners divided by the maximum distance in the sample; Lang is the percentage of new PTA pairs that share a common language; and Adj is the percentage of new PTA pairs that share a common land border.

[2] Other figures (not reported) indicate that PTAs have been signed by country-pairs that are increasingly different in both their levels of GDP and GDP per capita, further suggesting the extension of PTA relations to increasingly distant – both economically and geographically – country-pairs.

Figure 2: Differences in Characteristics between PTA Members and Non-Members



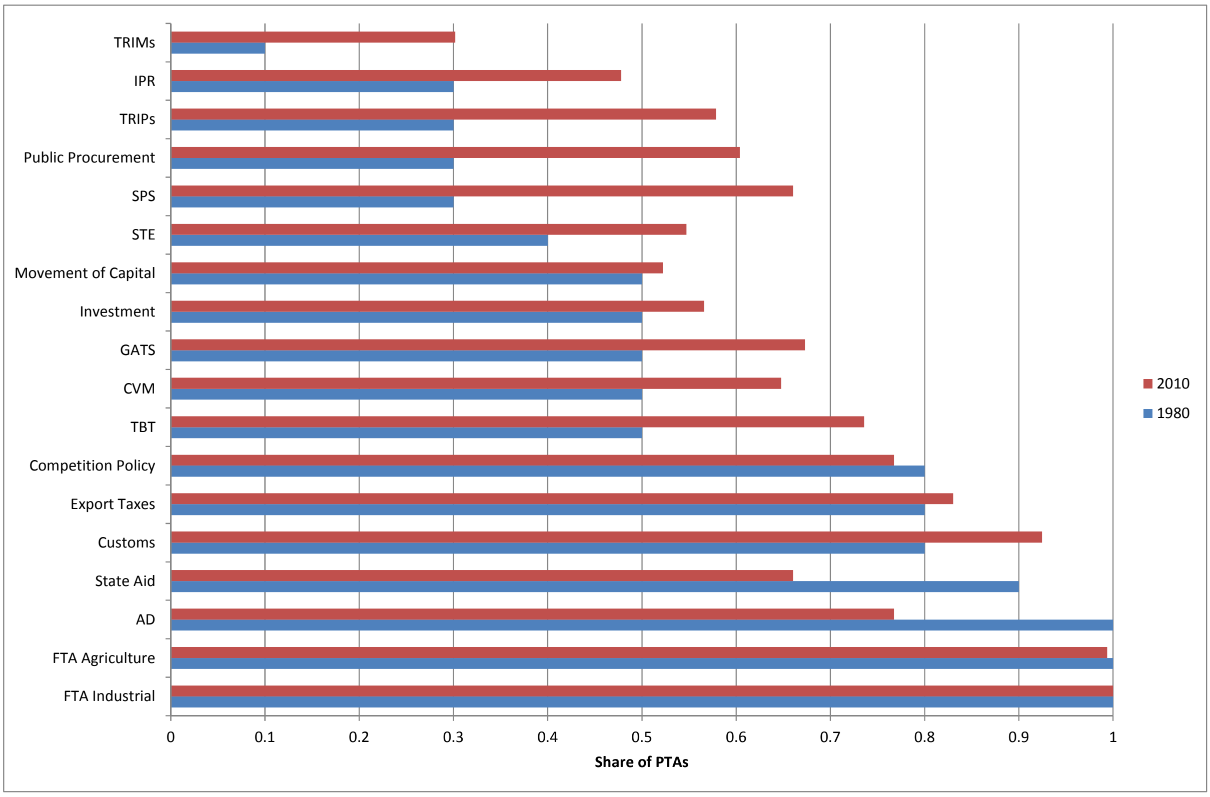
The differences between PTA members and non-PTA participants are further illustrated in Figure 2, which reports the distributions of economic size (sum of GDPs) and distance (log of distance between capital cities) for country-pairs in a PTA and country-pairs that were not in a PTA for the years 1980 and 2010. In both years PTA members have tended to be more similar in economic size than randomly-paired non-members, but this difference has weakened over time and by 2010 it was confined to the upper tail of the distribution. PTA members have always tended to be closer geographically than non-members – hence their alternative label as Regional Trading Arrangements. For the year 2010 however we observe the emergence of a bimodal distribution, with the second peak in the distribution being at a similar value for non-members, an outcome suggestive of a movement in recent times to the signing of PTAs with more distant partners. This figure also reports information on the share of PTA and non-PTA country-pairs that share a common language or common border in 1980 and 2010. These shares tend to be much larger for PTA country-pairs in both years, but have also tended to decline over time (and in the case of common language at least have converged towards the value for non-country-pairs).

In addition to changes in the characteristics of country-pairs that enter into PTAs, there have also been changes in the characteristics of the agreements signed, and an alternative explanation for why PTAs might appear to have heterogeneous effects is that the arrangements themselves are heterogeneous. The existence of differences in the depth and breadth of PTAs has long been recognised as evidenced by the distinctions drawn between Free Trade Areas, Customs Unions, Common Markets and Economic Unions.[3] If all these different types of arrangements are included in a gravity equation with a single PTA dummy, the estimated coefficient on that dummy will provide an average effect of trade agreements, with the effect being averaged across both PTAs with heterogeneous membership and PTAs with heterogeneous provisions. Controlling for the latter requires a variable that captures the breadth of PTAs in terms of their provisions. A recent World Bank study (Hoffman et al., 2017) develops a database that lists the coverage and legal enforceability of the provisions included in the entire set of PTAs in force and notified to the World Trade Organization as of 2015 (i.e. 279 agreements signed by 189 countries between 1958 and 2015). This database contains information on the inclusion of up to 52 policy areas, which the authors categorise in various ways. For our purposes the most useful category is that containing the 18 'Core' provisions which the authors describe as those "that the literature identifies as more meaningful from an economic point of view." (Hoffman et al., 2017, p3.). The subject areas of these core provisions are listed in Figure 3 (and in Section 4 below) which also indicates changes in the breadth of coverage by listing the fraction of the agreements that contain each core provision in 1980 and 2010. All provisions appear in a higher proportion of

[3] Vicard (2009) distinguishes between some of these types of PTA, finding a significant effect of the four different types of PTA that he considers.

PTAs in 2010 than in 1980, except those relating to competition policy, state aid and antidumping.[4]

Figure 3: Proportion of PTAs including Core Provisions



Notes: This figure reports the share of all PTAs that include the relevant provision as a legally enforceable commitment in the text of the agreement. TRIMs refer to trade related investment measures; IPR to intellectual property rights; TRIPs to the trade related aspects of intellectual property rights agreement; SPS to sanitary and phytosanitary standards; STE to state trading enterprises; GATs to the general agreement on trade in services; CVM to counter-veiling measures; and TBT to technical barriers to trade; and AD to anti-dumping duties.

These changes in the characteristics of both PTA participants and the PTAs themselves will be reflected in the effects of PTAs on bilateral trade flows and may be at least partly responsible for the heterogeneity in estimated effects noted earlier. Investigating this proposition is our objective in this paper, where we employ a matching econometrics approach. Using this method, we create both a 'treatment' (joined in a PTA) and a 'control' (not joined in a PTA) group of country-pairs of trading partners by selecting on observable characteristics and comparing observations drawn from the 'same distribution'. Thus for each trading pair in a PTA, we construct a control group of trading pairs with nearly identical economic characteristics but no PTA, with the average difference in trade flows between the pairs with PTAs and their control groups providing an estimate of the effects of the PTA treatment. Baier and Bergstrand (2009) [BB (2009), hereafter]

[4] Hoffman et al. (2017) also draw a distinction between those provisions which are 'legally enforceable' and those which are not. Here we restrict attention to legally enforceable provisions.

argue for using matching as a non-parametric benchmark for PTA treatment effects, rather than the more popular gravity equation estimation. The log-linear gravity equation does not capture non-linearities in the effects of PTAs on trade flows, and while it can be modified by introducing interaction effects, matching does not require the assumption of any specific functional relationship. BB (2004) show that selection into PTAs is non-random, depending on some of the same variables that determine bilateral trade flows. The combination of non-random selection into PTAs and omitted non-linearities can bias OLS estimates of PTA effects, and BB (2009) argue that that their matching estimates of PTA treatment effects are much more stable and economically plausible than the average treatment effect estimates using typical cross-section OLS gravity equations (with or without country fixed effects).[5]

In this paper we extend the BB (2009) matching approach to take some account of heterogeneity in PTA provisions. In the next two sections we review their methodology and apply it to our (larger and longer) sample. The results are qualitatively equivalent as expected. We then extend the analysis to encompass 'PTA breadth' by using a 'provision count index' constructed from the (Hoffman et al., 2017) data discussed above. We note that PTA formation now involves two separate, sequential decisions – first, whether two trading partners should form a PTA and, second, if they do, how broad that agreement should be. We find that our explanatory variables are significant for both decisions. Section 5 then investigates how heterogeneous PTAs affect bilateral trade flows. Using our matched PTA and non-PTA groups of country-pairs, we estimate a dose response function which demonstrates that PTAs with few provisions and PTAs with many provisions do not appear to have a significant impact on goods trade flows between their members. PTAs in an intermediate range are shown to have a significant positive effect. We then relate these outcomes to the actual content of the PTAs using the concept of 'provision intensity'. Section 6 provides a summary and conclusions.

2. Methodology and Data

As noted, our starting point is BB (2009) who employ the Anderson and van Wincoop (2003) estimating equation for the sum of the bilateral trade flows between country i and country j (T_{ij}):

$$\ln\left[\frac{T_{ij}}{(GDP_i GDP_j)}\right] = \alpha_0 + \alpha_1 \ln Dist_{ij} + \alpha_2 Adj_{ij} + \alpha_3 Lang_{ij} + \alpha_4 PTA_{ij} - \ln P_i^{1-\sigma} - \ln P_j^{1-\sigma} + \varepsilon_{ij} \quad (1)$$

[5] In the case of panel data the possibility of using country-pair fixed effects to control for unobserved heterogeneity and self-selection into PTAs is a standard response to this endogeneity problem.

Where GDP_i (GDP_j) is GDP in exporter i (importer j); $Dist_{ij}$ is the distance between i and j; Adj_{ij} is a dummy variable taking the value 1 if i and j share a land border; $Lang_{ij}$ is a dummy variable taking the value 1 if i and j share an official language; and PTA_{ij} is a dummy variable taking the value 1 if i and j are in a PTA. $P_i^{1-\sigma}$ and $P_j^{1-\sigma}$ are multilateral resistance (MR) terms whose (implicit) solutions can be obtained from the N nonlinear market-equilibrium conditions:

$$P_i^{1-\sigma} = \sum_{k=1}^{N} P_k^{\sigma-1}(GDP_k/GDP^T)e^{\alpha_1 \ln Dist_{ki} + \alpha_2 Adj_{ki} + \alpha_3 Lang_{ki} + \alpha_4 PTA_{ki}} \quad (2)$$

where GDP^T is world income. Unbiased estimates of α_0 to α_4 can be obtained by estimating (1) including country fixed effects for the MR terms.[6]

The results of estimating (1) for our sample, including and excluding country fixed effects, are shown in Table 2. Those estimates which make no allowance for multilateral resistance, present coefficients with the expected signs and that are generally statistically significant for all the explanatory variables, including the PTA dummy. Those results which do control for multilateral resistance using country fixed effects, present a similar outcome for the control variables, but for the PTA dummy we now find no significant coefficients in the years before 2000, but a positive and significant effect thereafter.

Table 2: Gravity Regression Results (with and without partner fixed effects)

	1980	1980	1990	1990	2000	2000	2010	2010
LnGDPsum	0.990***	1.105***	1.007***	0.817***	1.098***	0.928***	1.139***	0.860***
	(0.0105)	(0.140)	(0.00898)	(0.108)	(0.00693)	(0.0520)	(0.00712)	(0.0544)
LnDist	-1.265***	-1.559***	-1.295***	-1.571***	-1.142***	-1.518***	-1.239***	-1.675***
	(0.0449)	(0.0481)	(0.0384)	(0.0419)	(0.0274)	(0.0312)	(0.0307)	(0.0347)
Lang	0.951***	0.866***	1.012***	0.841***	1.102***	0.849***	1.181***	1.028***
	(0.0828)	(0.0864)	(0.0726)	(0.0741)	(0.0585)	(0.0611)	(0.0598)	(0.0633)
Adj	-0.0265	0.0731	0.200	0.270	0.787***	0.653***	0.590***	0.366**
	(0.225)	(0.243)	(0.192)	(0.196)	(0.129)	(0.145)	(0.151)	(0.172)
PTA	1.373***	0.175	0.809***	0.0208	0.999***	0.739***	0.717***	0.296***
	(0.214)	(0.255)	(0.162)	(0.190)	(0.0838)	(0.0834)	(0.0557)	(0.0626)
Constant	-22.23***	-25.97***	-22.91***	-13.71***	-29.26***	-15.80***	-30.25***	-10.49***
	(0.617)	(6.187)	(0.572)	(5.147)	(0.428)	(2.502)	(0.439)	(2.632)
Reporter/Partner FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	4,674	4,674	5,760	5,760	11,349	11,349	12,730	12,730
R-squared	0.637	0.747	0.681	0.785	0.709	0.784	0.714	0.794
F-Stat	1883***	63.50***	3017***	81.75***	5811***	123.7***	6282***	114.1***

Notes: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

[6] With panel data it is common to include time-varying country fixed effects to account for the fact that these MR terms need not be constant over time.

Before we conclude that the average PTA had no significant impact on bilateral trade between its members until after the WTO came into effect in the mid-1990s, we should recall that log-linear gravity regressions may provide inaccurate estimates of the effects of a PTA if (nonlinear) interactions between the PTA dummy and the trade-flow determinants are significant or if trading partners' selection into PTAs is non-random. Evidence of significant differences between PTA pairs and countries which did not form a PTA were shown in Figure 2 above. There is some evidence in the literature that the effects of PTAs depend on bilateral distance (Frankel et al, 1997) and on GDPs and populations (BB, 2004). In these circumstances parametric estimates of treatment effects will be biased and for this reason BB (2009) suggest employing a matching estimator.

As noted earlier, the key to the matching estimator lies in selecting treated and control groups of country pairs that are virtually identical in all other respects so as to simulate a random assignment into treatment and control. BB (2009) note that three conditions should be satisfied.[7] The first condition is known as *'ignorability of treatment'* or *'selection on observables'*. Its purpose is to ensure that the assignment to treatment is random and is achieved by choosing for each treated pair a control group that is closely matched in terms of all relevant variables influencing trade (except the PTA). BB (2006) show that applying a first order log-linear Taylor-series expansion (around a symmetric equilibrium) to system (2) gives a reduced-form function of linear combinations of the exogenous variables as follows:

$$\ln T_{ij} = \beta_0 + \beta_1 \ln(GDP_i GDP_j) + \beta_2 BVDist_{ij} + \beta_3 BVAdj_{ij} + \beta_4 BVLang_{ij} + \beta_5 BVPTA_{ij} + \varepsilon_{ij} \quad (3)$$

where $BVDist_{ij} = \ln Dist_{ij} - \left(\frac{1}{N}\right)\sum_{k=1}^{N} \ln Dist_{ik} - \left(\frac{1}{N}\right)\sum_{k=1}^{N} \ln Dist_{kj} + \left(\frac{1}{N^2}\right)\sum_{k=1}^{N}\sum_{h=1}^{N} \ln Dist_{kh}$

$BVAdj_{ij} = Adj_{ij} - \left(\frac{1}{N}\right)\sum_{j=1}^{N} Adj_{ik} - \left(\frac{1}{N}\right)\sum_{i=1}^{N} Adj_{kj} + \left(\frac{1}{N^2}\right)\sum_{i=1}^{N}\sum_{j=1}^{N} Adj_{kh}$

$BVLang_{ij} = Lang_{ij} - \left(\frac{1}{N}\right)\sum_{k=1}^{N} Lang_{ik} - \left(\frac{1}{N}\right)\sum_{i=1}^{N} Lang_{kj} + \left(\frac{1}{N^2}\right)\sum_{k=1}^{N}\sum_{h=1}^{N} Lang_{kh}$; and

$BVPTA_{ij} = PTA_{ij} - \left(\frac{1}{N}\right)\sum_{j=1}^{N} PTA_{ij} - \left(\frac{1}{N}\right)\sum_{i=1}^{N} PTA_{ij} + \left(\frac{1}{N^2}\right)\sum_{i=1}^{N}\sum_{j=1}^{N} PTA_{ij}$

In (3) the 'raw' explanatory variables have been 'adjusted' to take account of how the same variables affect the other trade flows for these trading partners. Thus in the case of $BVDist_{ij}$ the distance between i and j is now measured relative to the average distance of both i and j from their trading partners (the final term in each expression is independent of i and j). Similar interpretations apply to the other explanatory variables. Estimates of (3) are shown in Table 3,

[7] See BB09 for a more detailed explanation and a more formal presentation of these conditions.

where we find results similar to those correcting for MR in Table 2 in terms of sign and significance of the non-PTA variables. The PTA variable is significant for 1980 in Table 3, but not in Table 2.

Table 3: Trade Gravity Regression Results - Equation (3)

	1980	1990	2000	2010
LnGDPsum	0.911***	0.956***	1.063***	1.126***
	(0.0104)	(0.00877)	(0.00659)	(0.00667)
BVDist	-1.539***	-1.561***	-1.524***	-1.685***
	(0.0533)	(0.0455)	(0.0337)	(0.0379)
BVAdj	0.0591	0.299	0.672***	0.358**
	(0.236)	(0.200)	(0.133)	(0.160)
BVLang	0.760***	0.763***	0.806***	0.979***
	(0.0995)	(0.0865)	(0.0684)	(0.0711)
BVPTA	0.404*	0.0785	0.711***	0.298***
	(0.244)	(0.187)	(0.0890)	(0.0653)
Constant	-42.42***	-45.03***	-50.17***	-54.52***
	(0.687)	(0.565)	(0.431)	(0.475)
Country FE	No	No	No	No
Observations	4,674	5,760	11,349	12,730
R-squared	0.639	0.684	0.717	0.720
F-Stat	2038***	3229***	6320***	6723***

Notes: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

This set of explanatory variables has been derived from a well specified trade model, and to the extent that (increased) trade among its members is the objective of a PTA, then these variables should also be relevant to the PTA formation decision.[8] Based on this, BB (2009) include $BVDist_{ij}$, $BVAdj_{ij}$, $BVLang_{ij}$ and $\ln GDPsum_{ij}$ as the explanatory variables in their matching regression. An important consideration for successful matching is that the distributions of these covariates for the treated and matched untreated pairs be virtually indistinguishable. We present evidence in Appendix 1 that provides some support for this assumption.

The second requirement that should be satisfied is the '*overlap*' condition. This requires that for each possible set of observations on the explanatory variables there are both PTA flows and non-PTA flows. BB (2009) argue that this is satisfied because of the large number of PTAs. The third requirement has two parts and satisfying each of them is a little more problematic. The first

[8] Other sets of variables have been proposed to explain PTA formation (see, for example, BB, 2004 and Bergstrand and Egger, 2013). As a robustness exercise we also matched on an alternative set of variables extracted from the literature. We added income dissimilarity (*ID*), measured as the absolute difference in the GDPs per capita of the two countries divided by the sum of their GDPs per capita, and size dissimilarity (*SD*), measured as the absolute difference in the real GDPs of the two countries divided by the sum of their GDPs, to *Dist, Lang, Adj* and *Gdpsum*. We found that the qualitative outcomes were the same for both sets of variables and so do not report the alternatives in the interest of brevity. They are available on request.

part, that the '*treatment is unique*', ensures that the treatment is identical for each treated observation. This condition will only be satisfied if all PTAs have the same effects on the bilateral trade flows of their members, which is the standard assumption that we make when we introduce a single PTA dummy into a gravity equation. We will discuss this further below. The second part of this condition is '*non-interference*', which ensures that the treatment does not influence the non-treated observations. BB (2009) note that by selecting on $BVDist_{ij}$, $BVAdj_{ij}$, $BVLang_{ij}$ they account for the general equilibrium interactions, other than PTAs, among the non-PTA pairs in calculating the treatment effects. Potential trade diversion from the PTAs remains a concern, however.

Data on trade flows are collected from UN COMTRADE. In our analysis we use the mirror flow to construct bilateral exports (i.e. we use imports into the partner country to measure exports from the reporter) when it is available. When the mirror flow is not available we use the raw export data to fill in the gaps when these data are available. We consider the bi-directional trade flow using the sum of imports and exports for each country-pair in our dataset.[9]

Finally, let PT_{ij} and NT_{ij} denote the total goods trade between countries i and j with and without a PTA respectively. Only one of PT_{ij} and NT_{ij} will be observable for each pair, and the matching estimator imputes the missing values using average outcomes for pairs with similar values for the covariates. Suppose we use the m closest matches to each PTA country-pair, and let $I(ij)$ denote the set of indices of these matches for pair ij. Then the trade values used are:

$$NT_{ij}^* = \begin{cases} NT_{ij} & PTA_{ij} = 0 \\ \frac{1}{M}\sum_{kh\in \mathrm{I(ij)}} NT_{kh} & PTA_{ij} = 1 \end{cases} \quad \text{and} \quad PT_{ij}^* = \begin{cases} \frac{1}{M}\sum_{kh\in \mathrm{I(ij)}} PT_{kh} & PTA_{ij} = 0 \\ PT_{ij} & PTA_{ij} = 1 \end{cases},$$

and the estimator for the average treatment effect for the sample is:

$$ATE = \frac{1}{N}\sum_{ij}[PT_{ij}^* - NT_{ij}^*]$$

while the average treatment of the treated (those countries that joined a PTA) is:

$$ATT = \frac{1}{N_T}\sum_{PTA_{ij}=1}[PT_{ij} - NT_{ij}^*]$$

where N is the total number of country-pairs in the sample, and N_T is the number of country-pairs in a PTA.

[9] Data on other variables are from two main sources. Data on distance, common language and adjacency are from CEPII, while data on GDP and GDP per capita are from the World Bank's World Development Indicators. As already mentioned, data on PTAs is from the paper of Hofmann et al (2017).

3. Matching Results

Table 4 shows the results from estimating the BB (2009) probit equation for selection into a PTA on our sample – i.e.

$$PTA_{ij} = \gamma_0 + \gamma_1 \ln\left(GDP_i GDP_j\right) + \gamma_2 BVDist_{ij} + \gamma_3 BVAdj_{ij} + \gamma_4 BVLang_{ij} + \varepsilon_{ij} \qquad (6)$$

The coefficient on the log product of the two countries GDPs is positive as expected and significant except in 1980. The other (composite) variables require more careful interpretation. For example the distance variable (BVDist) is the (log) distance between the exporter and importer minus the sum of the average (log) distance of the exporter and the importer from all their potential trading partners (the other term is a constant). So this variable will be large when the two parties concerned are distant from each other relative to their average trading partners. Its negative and significant coefficient indicates that relatively distant countries are less likely to form a PTA. Likewise BVAdj (BVLang) will be negative for two countries that do not share a common border (language). It will be positive if they do but will also be decreasing in the number of other common borders (languages) possessed by either of them. Interestingly, adjacent countries appear to have been significantly less likely to be in a PTA last century, but are significantly more likely to be PTA partners in the present century. Countries sharing a common language have always been more likely to form a PTA. These results, except for the changing sign on the BVAdj variable, are as expected.

The matching outcomes are shown in the bottom panel, where we have provided the corresponding estimates from BB (2009) for comparison.[10] Our results show a declining impact of PTAs on trade flows from 1980 to 2000, but in 2010 the result is closer to 1990 levels.

4. Heterogeneous PTAs

As noted earlier, one explanation for why PTAs might appear to have heterogeneous effects (besides the heterogeneity of membership) is that the arrangements themselves are heterogeneous. In this section and the next we consider two issues related to the breadth of the agreements. Firstly, how do the determinants of entering into a PTA impact on its breadth (this section). Secondly how are trade flows related to the breadth of trade agreements (next section).

---

[10] Note that BB (2009) use 96 potential trading partners while we have up to 183.

Table 4: Probit Matching Equations (3 nearest neighbours)

Equation (6)	1980	1990	2000	2010
LnGDPsum	0.00884	0.0635***	0.0112*	0.0437***
	(0.0140)	(0.0126)	(0.00646)	(0.00468)
BVDist	-0.934***	-0.948***	-0.872***	-0.766***
	(0.0667)	(0.0584)	(0.0276)	(0.0211)
BVAdj	-0.702***	-0.522***	0.283***	0.435***
	(0.218)	(0.185)	(0.0993)	(0.104)
BVLang	0.238*	0.428***	0.533***	0.176***
	(0.139)	(0.126)	(0.0590)	(0.0464)
% Correctly Predicted				
All	74.14	86.12	87.68	80.90
PTA pairs	16.36	16.77	27.70	29.66
Non-PTA pairs	75.18	87.80	92.54	91.43
ATT(6)	1.515***	0.972***	0.448***	1.085***
	(0.239)	(0.077)	(0.125)	(0.082)
BB (2009)	0.83**	1.00**	0.59**	

Notes: Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Our data on PTAs and the 'breadth' of PTAs is from the recent work of Hofmann et al (2017). This database contains information on the inclusion of 52 policy areas and their legal enforceability for 279 trade agreements signed between 189 countries in the period 1958 to 2009. In our analysis we focus on the 18 provisions considered the core provisions by Hofmann et al (2017). These 18 provisions relate to: (i) free trade in industrial goods; (ii) free trade in agricultural goods; (iii) Customs administration; (iv) Export taxes; (v) Anti-dumping measures; (vi) Competition policy; (vii) Technical barriers to trade; (viii) Sanitary and phytosanitary standards (SPS); (ix) State aid; (x) GATS; (xi) Counter-veiling measures; (xii) TRIPs; (xiii) Public procurement; (xiv) Investment measures; (xv) Movement of capital; (xvi) State trading enterprises; (xvii) IPRs; (xviii) TRIMs. To capture the depth of PTAs, we calculate for each country-pair and at each point in time the proportion of these 18 core provisions that are included in a PTA.[11] Thus 0 indicates no PTA and 1 indicates a PTA that includes all of the core provisions. We label this indicator PTAB.

Taking account of differences in the breadth of PTAs means that country-pairs contemplating bilateral trade liberalisation now have two decisions to make - whether to enter into a PTA and, if they do, how broad that PTA should be. Country-pairs self-select into PTAs before agreeing upon the breadth of any such agreement, and ignoring this self-selection issue when considering the determinants of the breadth of PTAs could lead to biased estimates, as illustrated below. To

[11] This 'provision count' measure effectively weights each type of provision equally in the index. Other divisions and subdivisions of these provisions have been proposed. For example, we can subdivide the core provisions into 'border' and 'behind the border' provisions, depending on whether the policies that the provision regulates are applied at the border or not. Limao (2016) proposes several divisions based on alternative criteria.

deal with this we adopt the Heckman Two-Stage Correction for selection bias. In the first stage we model the decision for a country-pair to enter into a PTA using a probit equation. From these first stage results we calculate the inverse Mills ratio and include this variable in the second stage fractional probit equation[12] of the breadth of PTAs to control for self-selection.

Table 5 shows the results from directly estimating a fractional probit equation of the breadth of PTA choice. As we noted earlier these estimates combine two selection effects – selection into PTA (or not) and selection of breadth of PTA. Where the explanatory variables have different impacts on these two decisions, the estimates below will confound their two effects, which can lead to errors when interpreting the outcomes. Thus the results in Table 5 appear to suggest that larger countries sharing a common language are more likely to choose broader PTAs, while more distant countries choose narrower PTAs (if any). Interestingly countries that are contiguous are also likely to choose narrower PTAs.

Table 5: Fractional Probit Results (all observations) for PTA Breadth

	1980	1990	2000	2010
Ln GDPsum	0.00633	0.0578***	0.0292***	0.0530***
	(0.0181)	(0.0157)	(0.00717)	(0.00430)
BVDist	-0.829***	-0.867***	-0.764***	-0.619***
	(0.0641)	(0.0567)	(0.0243)	(0.0206)
BVAdj	-0.629***	-0.537***	-0.0107	0.0283
	(0.211)	(0.180)	(0.0873)	(0.0756)
BVLang	0.199**	0.273***	0.105*	-0.0738*
	(0.0789)	(0.0806)	(0.0555)	(0.0402)
Constant	-10.13***	-13.00***	-10.00***	-9.243***
	(1.238)	(1.116)	(0.443)	(0.321)
Observations	4,674	5,760	11,349	12,730

Notes: *** p<0.01, ** p<0.05, * p<0.1

These, in some cases puzzling, outcomes may be a consequence of failing to account for the different roles of these explanatory variables in the overall PTA decision. To illustrate this we present the results from the Heckman procedure in Table 6.[13] Since the first stage equation should include at least one variable that is not included in the second stage, and we don't wish to 'drop' any of our current explanatory variables, we construct $OPTA$, a variable capturing the

[12] The fractional probit equation is similar to the standard probit model, but allows one to estimate models where the dependent variable is a fractional variable (i.e. ranging between 0 and 1).

[13] Normally, the second stage regression model is estimated using OLS. But since we measure breadth as a proportion, we use the fractional probit model in the second stage.

number of additional PTAs that the country-pair has (with third countries) and include it in the first stage regression. This variable is lagged 5 years to avoid simultaneity issues and is not included in the second stage.

A comparison of the Stage 1 probit results in Table 6 with the probit results in Table 4 indicates the influence of including OPTA. The coefficient on OPTA itself is positive and significant in all years. The greater the number of other PTAs with third parties that these trading partners have the more likely they are to be in one together. There are minor changes in the significance of a few coefficients on the other explanatory variables, but these are not inconsistent with the outcomes for the same variable in other years. None of our conclusions on the variables that influence the PTA decision are affected by the addition of OPTA.

Table 6 Heckman Two-Step Correction

	Stage 1: Formation - Probit Results				Stage 2: Breadth – Fractional Probit Results			
	1980	1990	2000	2010	1980	1990	2000	2010
LnGDPsum	0.138***	0.253***	0.0703***	0.00391	-0.0129	-0.0213	0.0389***	0.0243***
	(0.0226)	(0.0260)	(0.00894)	(0.00567)	(0.0141)	(0.0144)	(0.00602)	(0.00462)
BVDist	-0.864***	-1.117***	-0.940***	-0.926***	0.215**	0.137	0.519***	0.256***
	(0.0844)	(0.0884)	(0.0356)	(0.0252)	(0.0988)	(0.0887)	(0.0581)	(0.0306)
BVAdj	-0.450	-0.117	0.314***	0.556***	0.279	0.203	-0.246***	-0.159**
	(0.284)	(0.259)	(0.118)	(0.113)	(0.231)	(0.175)	(0.0744)	(0.0708)
BVLang	0.680***	1.278***	0.898***	0.205***	-0.360	-0.760***	-0.778***	-0.679***
	(0.211)	(0.218)	(0.0730)	(0.0524)	(0.318)	(0.260)	(0.0562)	(0.0587)
OPTA	19.11***	34.62***	16.50***	11.85***				
	(1.498)	(2.441)	(0.776)	(0.315)				
Constant	-17.60***	-26.63***	-14.45***	-10.82***	3.492***	3.208***	4.682***	2.318***
	(1.535)	(1.905)	(0.633)	(0.373)	(0.896)	(1.160)	(0.755)	(0.449)
Obs	3,978	4,920	8,587	11,568	94	138	636	2,090

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Next we look at the patterns of coefficients on the explanatory variables across the two stages. The first observation to be made is that few of the explanatory variables are significant in the Stage 2 equation in 1980 (only distance) and 1990 (only common language and income differences). This may be a consequence of the small number of PTAs in our sample in these years. Somewhat unexpectedly, only for LnGDPsum are the estimated coefficients consistently of the same sign (+) when significant across the two equations. The others all have opposite signs when they are significant. This has interesting implications. For example, we see that more distant countries are less likely to form a PTA, but if they do it is likely to be broader. This feature cannot be picked up by a 'combined' estimation as in Table 5, where distance appears to have a negative effect on PTA breadth. Similar interpretations apply to the other variables. Thus sharing a common language or a common border makes a PTA more likely, but it will be narrower in scope on average. The direct probit results for PTA breadth in Table 5 essentially mimic the first stage probit results in Table 6, leading to potentially erroneous inferences about the role that these explanatory variables play in the choice of PTA breadth.

5. Heterogeneous PTAs and Trade Flows

In this section we use our results to investigate the effects of PTA breadth on bilateral trade flows. Table 7 presents estimates of gravity equations including PTAB as an explanatory variable. Comparing these results with those in Table 3, we see that the pattern of signs and significance of all the other explanatory variables is the same in the two tables. Furthermore, we have equivalent results on the PTA variables in the two Tables. The estimated coefficients on BVPTA in Table 3 and BVPTAB in Table 7 are positive in all years, and are statistically significant in all years except 1990. But these estimates may be biased because, as before, they do not take account of country-pair self-selection into PTAs or the endogeneity of the breadth of the PTAs they choose.

Allowing for heterogeneous PTAs gives us two groups of country pairs: those who do not form a PTA (and whose treatment (PTAB) is zero) and those who do. But the latter can receive different levels of treatment ($1 \geq PTAB > 0$), and to take account of this we are interested in estimating the *dose response function* (DRF) which gives the causal effect of the different intensities of treatment (i.e. different values of PTAB) on (the log of) bilateral trade within the observed sample. Cerulli (2015) develops a model to estimate this DRF that doesn't rely upon any assumption of normality and that can be used when many pairs have a treatment level of zero.[14] The approach is applicable when treatment is exogenous (i.e. selection into treatment depends only on factors that are observable) or endogenous (selection into treatment depends on both observable and unobservable factors). In this analysis we adopt the version that relies on exogenous treatment.[15]

[14] It thus has some advantages over the approach of Bia and Mattei (2008). In the case where many units are not subject to treatment we will observe a non-zero probability mass at a zero value of treatment, meaning that the assumption of a normal distribution is untenable.
[15] Further details are provided in Appendix 2.

Table 7: Gravity Regressions of Trade Flows on the Depth of PTAs

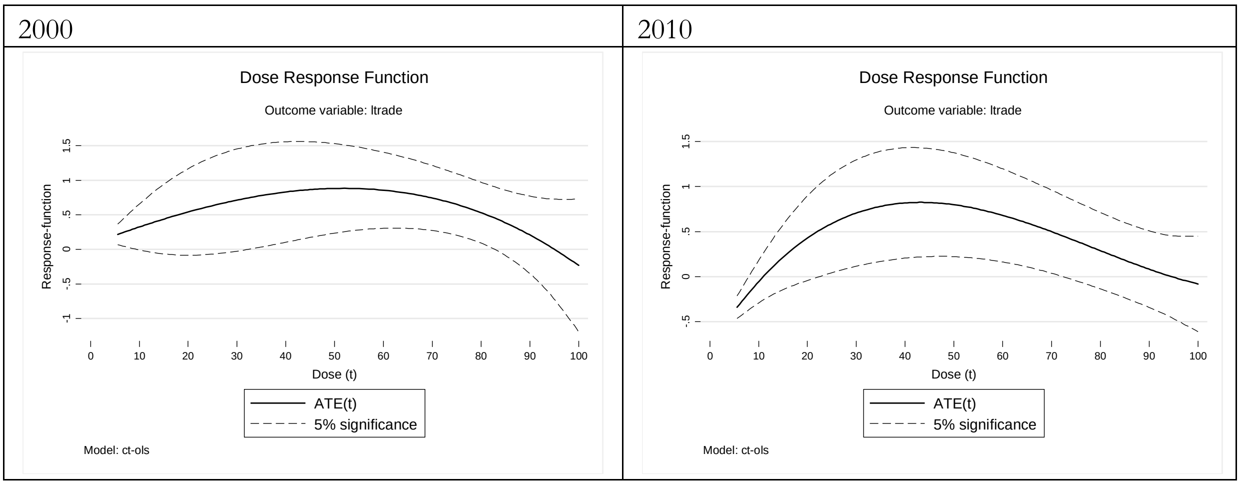
	1980	1990	2000	2010
LnGDPsum	0.911***	0.956***	1.062***	1.126***
	(0.0104)	(0.00879)	(0.00662)	(0.00668)
BVDist	-1.541***	-1.566***	-1.565***	-1.724***
	(0.0526)	(0.0451)	(0.0329)	(0.0377)
BVAdj	0.0603	0.301	0.768***	0.402**
	(0.236)	(0.199)	(0.134)	(0.161)
BVLang	0.760***	0.765***	0.850***	0.996***
	(0.0995)	(0.0865)	(0.0681)	(0.0711)
BVPTAB	0.540*	0.00338	0.679***	0.165**
	(0.319)	(0.245)	(0.122)	(0.0790)
Constant	-42.43***	-45.08***	-50.51***	-54.89***
	(0.682)	(0.564)	(0.431)	(0.474)
Observations	4,674	5,760	11,349	12,730
R-squared	0.639	0.684	0.717	0.720
F-Stat	2034	3245	6299	6695

Notes: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

We deal with the problem of self-selection into a PTA by using the same matched sample of countries identified in Section 3.[16] The results are illustrated in Figure 4. Given the relatively small number of PTA country-pairs in 1980 and 1990 we find estimates of the DRF that are very imprecisely estimated, and we choose not to report these results in the main text (the figures are reported in Appendix 3), concentrating instead on the results for 2000 and 2010. Figure 4 shows that results for these two years are quite similar. In particular, they suggest that the effect of a PTA is initially increasing in its breadth, with the effect being low (and often insignificant) or negative for the narrowest PTAs. The impact of PTAs then increases and becomes significant for country-pairs with a PTA breadth in the middle of the distribution, with the maximum effect being found for values of breadth between 0.4 and 0.6. As PTA breadth increases beyond these levels the impact of PTAs begins to decline, such that for the highest values of breadth the impact of PTAs is again insignificant.

[16] In a further robustness test we instead use the full sample of observations and include the inverse Mills ratio calculated for the Heckman procedure to control for selection into PTAs. Results using this approach are qualitatively very similar to those found using the matched sample.

Figure 4: Dose Response Functions



This apparently counterintuitive result that a broader PTA can in fact result in a smaller net trade increase (or none at all) might be explained by the diverse nature of these PTA provisions, even those in the economic core. Provisions related to the protection of intellectual property rights, the movement of capital and other investment-related measures may be aimed at inducing increased economic interaction that does not necessarily result in increased international trade in goods. Indeed, with the potential substitutability between goods trade and capital movements and the possibility that a stronger IPR regime could discourage some trade through a market power effect, we see that a PTA that emphasises these types of provisions may actually reduce trade flows.[17]

With these thoughts in mind we construct Figure 5. Here we used the 2010 data to separate our PTA country pairs into 4 strata based on PTAB. Strata characteristics are shown in Table 8. We then calculate the expected probability of each provision appearing in a PTA between a country pair located in each stratum and subtract it from the actual frequency with which that provision appears in that stratum. This gives us a measure of ‘provision intensity’ for each of the strata. If the actual frequency exceeds the expected then the PTA country pairs in the corresponding stratum are ‘intensive’ in that provision. These intensities are illustrated in Figure 5. Note, first, that the lowest stratum only includes four (different) provisions and is intensive in agricultural and industrial tariff liberalisation. This ‘basic’ level of PTA breadth is shown to produce no significant increases in trade flows (in 2010) by the DRF estimated above. Second, the highest stratum, which is the other range where the DRF estimates no significant trade

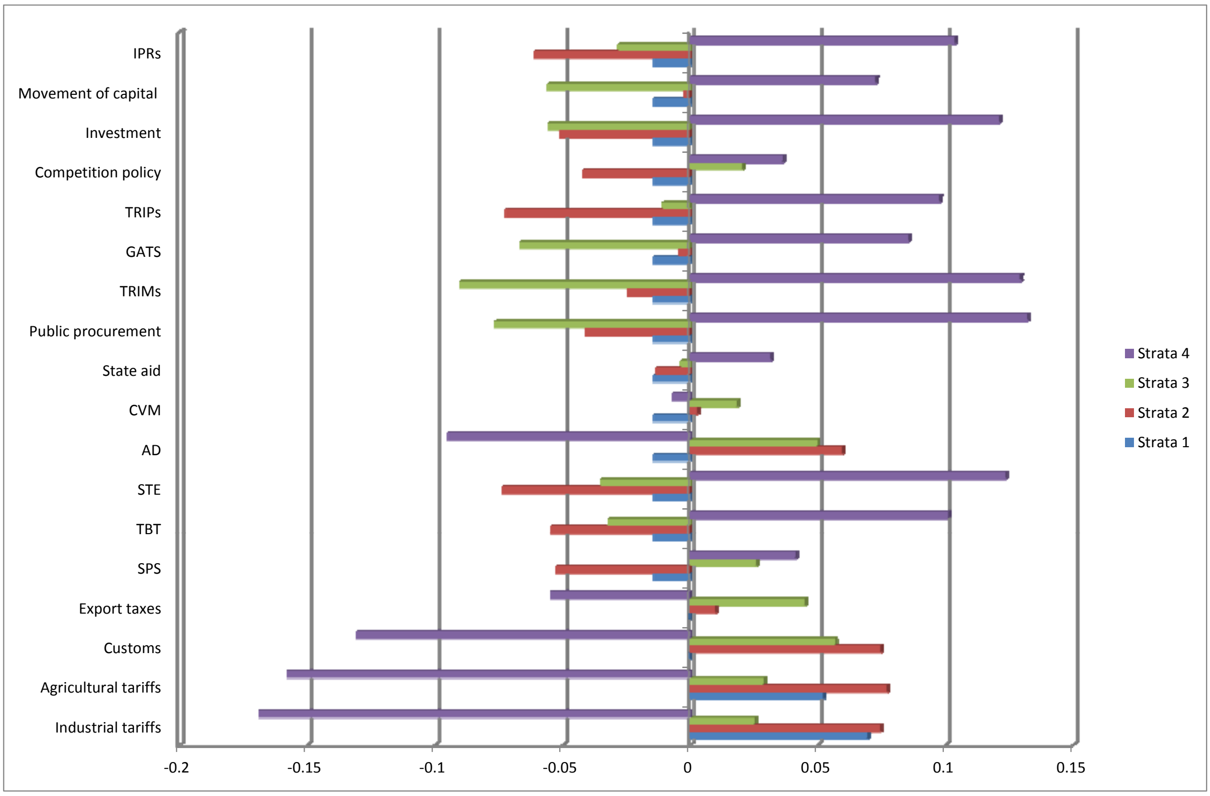
[17] On the possible substitutability or complementarity between goods trade and factor movements see for example Fontagne (1999). On the relationship between goods trade and the strength of IPR regimes see Smith (1999) and Falvey et al (2009).

increase, is the only stratum intensive in provisions relating to IPRs, the movement of capital, investment, TRIPS, GATS, TRIMs, state aid, state enterprises and technical barriers to trade. This list includes, inter alia, those provisions noted above that may be trade inhibiting rather than trade enhancing. Finally, the middle strata where the positive effects for trade are found, are the only strata intensive in provisions relating to countervailing measures, antidumping, export taxes and customs. These are provisions directly relating to goods trade flows.

Table 8: PTAB Strata Characteristics.

Stratum	PTAB range	# PTA country pairs	Percent of PTA pairs	Total Provisions Included
1	0-0.25	194	8.34	399
2	0.26-0.50	354	15.22	2045
3	0.51-0.75	436	18.74	4525
4	0.76-1.00	1342	57.70	20759
Total		2326	100	27728

Figure 5: Provision Intensities of PTAB Strata



6. Summary and Conclusions

Given the recent proliferation of preferential trading arrangements and their status as one of the few WTO-consistent ways in which countries might engage in 'multilateral' trade liberalisation in the near future, the importance of having some knowledge of their effects on trade flows seems obvious. Most estimates in the literature have been of the average effect of preferential trading arrangements on bilateral trade flows – i.e. the average effect of existing arrangements that differ quite widely in terms of the economic characteristics of their members and the provisions of the arrangements themselves. Little wonder then that these estimates have been sensitive to the composition of their sample, both in terms of the time period and the trading partners that are included.

In this study we have extended this analysis to allow for heterogeneity in the arrangements. In doing this we took advantage of a recently published data set that lists the provisions and assesses their 'enforceability' of those preferential arrangements reported to the WTO. From this we constructed a simple 'count measure' for each arrangement. Once we allowed that PTAs could be heterogeneous we were careful to model PTA formation as a two stage process. In the first stage countries choose their PTA partners. In the second stage they decide on the breadth of the PTA to form with those partners. While the same economic variables may influence each stage of the process, the directions of that influence could be opposite in the two cases. Thus we found that while more distant countries are less likely to form a PTA, if they do it is likely to be broader. Similarly, countries that share a common language or a common border are more likely to form a PTA, but it will be narrower in scope. Large trading partners are both more likely to form a PTA and more likely to choose a broad PTA.

We estimated the effects of PTA breadth on bilateral trade by first using a matching equation to create for each PTA country-pair a control group of untreated country pairs and using the differences in their bilateral trade flows to estimate a dose response function that related the trade change due to PTA treatment to the breadth of the PTA adopted. The estimated dose response function exhibited an inverted u-shape, with insignificant effects for both relatively narrow and relatively broad PTAs. To investigate further we divided PTAs into four strata based on their breadth. We found that those PTAs in the bottom strata, where the trade effects estimated by the dose response function are not significant, all consisted of the same four provisions relating to industrial and agricultural tariffs, customs and export taxes. Our evidence suggests that, in the WTO era at least, a basic PTA of this type has no significant effects on bilateral aggregate trade flows. PTAs in the top stratum, which also result in no significant trade increase according to our estimated dose response function, were shown to be intensive in a range of provisions dealing with economic interactions beyond trade flows in goods (e.g. services, foreign investment, and intellectual property rights). The implication here was that these interactions act as a substitute for goods trade; a possibility established in the literature and worthy of further exploration in this context. Those PTAs in the middle two strata, where we estimate a positive effect of PTA breadth on goods' trade flows, are the only PTAs intensive in provisions relating to contingent protection (counter-veiling measures and anti-dumping), export taxes and customs.

While our results on patent breadth should best be treated as exploratory, they do highlight the important point that the analysis of PTA provisions should take into account that the objectives of PTAs may extend beyond greater market access for goods exports. The apparent inability of the PTAs adopted by the nearly 60% of country pairs in the top stratum to generate a significant increase in bilateral goods trade flows does not necessarily imply that the majority of PTAs have failed. They may simply have broader objectives.

References

Anderson, J E., and E. van Wincoop, 2003. Gravity with gravitas: a solution to the border puzzle. *American Economic Review* 93 (no. 1), 170–192 March.

Baier, S. and J. Bergstrand, 2002. Economic determinants of free trade agreements. *Journal of International Economics*, 64(1), 29-63.

Baier, S. L. and J. H. Bergstrand 2004. Economic determinants of free trade agreements. *Journal of International Economics*, 64, 29-63.

Baier, S. L. and J. H. Bergstrand, 2009. Estimating the effects of free trade agreements on international trade flows using matching econometrics. *Journal of International Economics*, 77, 63-76.

Baier, S., Yotov, Y. and T. Zylkin, 2016. On the widely differing effects of free trade agreements: Lessons from twenty years of trade integration. CESifo Working Paper no. 6174, CESifo, Munich.

Bergstrand, J. and P. Egger, 2013. What determines BITs? *Journal of International Economics*, 90, 107-122.

Bia, M., and A. Mattei, 2008. A Stata package for the estimation of the dose–response function through adjustment for the generalized propensity score. *Stata Journal*, 8, 354–373.

Cerulli, G., 2015 ctreatreg: Command for fitting dose–response models under exogenous and endogenous treatment. *Stata Journal*, 15(4), 1019-1045.

Cheong, J., Kwak D.W. and K.K. Tang, 2014. Heterogeneous effects of preferential trade agreements: How does partner similarity matter? *World Development*, 66, 222-236.

Cipollina, M. and L. Salvatici, 2010. Reciprocal trade agreements in gravity models: a meta-analysis. *Revue of International Economics*, 18(1), 63–80.

Eicher, T.S. and C. Henn, 2011. In search of WTO trade effects: preferential trade agreements promote trade strongly, but unevenly. *Journal of International Economics*, 83(2), 137–153.

Falvey, R. E., Foster, N. and D. Greenaway, 2009. Trade, imitative ability and intellectual property rights. *Review of World Economics*, 145, 373-404.

Fontagne, L., 1999. Foreign direct investment and international trade: Complements or substitutes. OECD Working Paper 1999/03, OECD.

Frankel, J.A., E. Stein, and S-J Wei, 1997. Regional trading blocs in the world economic system. Peterson Institute for International Economics, Washington, DC.

Hirano, K., and G. W. Imbens. 2004. The propensity score with continuous treatments. In *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, ed. A. Gelman and X.-L. Meng, 73–84. Chichester, UK: Wiley.

Hofmann, C., Osnago, A. and M. Ruta, 2017. Horizontal depth: A new database on the content of preferential trade agreements. World Bank Policy Research Working Paper 7981, The World Bank.

Limao, N., 2016. Preferential trade agreements. Chapter 6 in Kyle Bagwell and Robert W. Staiger (eds.), *Handbook of Commercial Policy* Vol 1B, Elsevier, pp. 279-367.

Rubin, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66,688–701.

Smith, P. J. (1999). Are weak patent rights a barrier to US exports? *Journal of International Economics*, 48(1), 151–177.

Vicard, V, 2009. On trade creation and regional trade agreements: Does depth matter? *Review of World Economics*, 145(2), 167-187.

World Bank, 2005. Global economic prospects 2005: Trade, Regionalism, and Development. World Bank, Washington, DC.

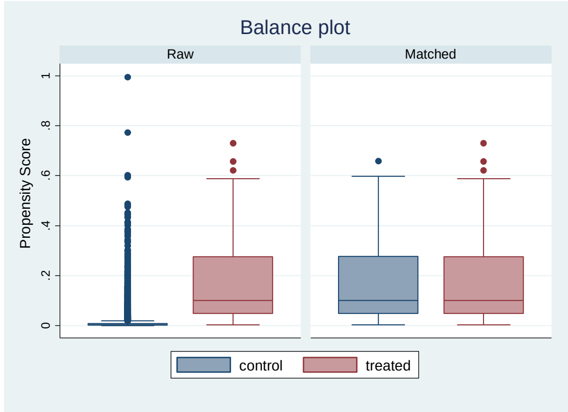
Appendix 1: Balancing Property of the Matched Sample

In this appendix we present some additional results aimed at testing whether the balancing property holds for the sample of PTA country-pairs and the matched sample of non-PTA country-pairs. We do this in two ways. Firstly, we present information on the standardised differences and the variance ratio of our matching variables for both the full sample of observations (raw) and the matched sample. For the balancing property to be satisfied we would expect to see that the standardised differences for the matched sample become close to zero (implying that the mean values of the matching variables for the PTA and matched non-PTA country-pairs are similar) and that the variance ratio moves towards a value of one (implying that the variance of our matching variables for the PTA and matched non-PTA samples are similar). For each year, results of these tests are reported in the left-hand column in Table A1 below. Secondly, we report information on the distribution of the constructed propensity score for both the raw and matched samples. For the balancing property to hold we would expect to see that the distribution of the propensity score for the matched non-PTA (i.e. untreated) sample resembles that of the PTA (i.e. treated) sample. Results in Table A1 tend to provide some support for the balancing property being met. Standardised differences in the matched sample tend to be smaller than in the raw sample, and in most cases are close to zero, while in most cases the variance ratio also moves towards one (the exception tending to be for the log of the sum of GDP). In terms of the balance plots the distributions for the matched non-PTA and PTA samples tend to be very similar, despite quite different distributions for the untreated non-PTA sample as a whole.

Table A1: Results of Balancing Property Tests

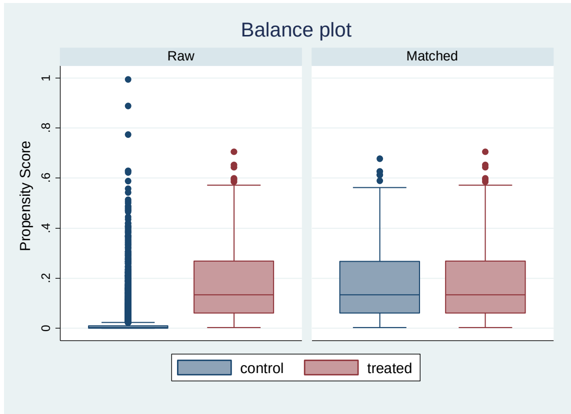
1980

	Standardized Differences		Variance Ratio	
	Raw	Matched	Raw	Matched
lgdpsum	-0.134	0.029	3.647	2.339
bvdist	-2.187	0.003	1.414	0.874
bvadj	0.501	0.003	5.425	0.882
bvlang	0.822	0.049	0.807	0.475



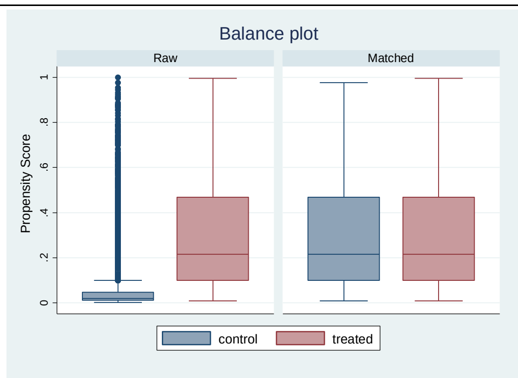
1990

	Standardized Differences		Variance Ratio	
	Raw	Matched	Raw	Matched
lgdpsum	0.105	-0.020	3.198	1.828
bvdist	-2.151	-0.015	1.397	0.965
bvadj	0.540	0.035	7.702	0.982
bvlang	0.873	0.009	1.064	0.695



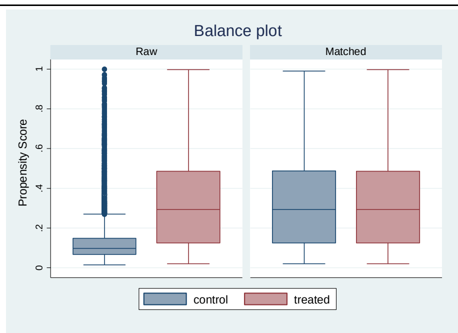
2000

	Standardized Differences		Variance Ratio	
	Raw	Matched	Raw	Matched
lgdpsum	-0.074	-0.090	2.149	1.686
bvdist	-1.648	0.025	2.214	0.931
bvadj	0.602	0.155	15.693	1.412
bvlang	0.684	0.043	2.108	1.267



2010

	Standardized Differences		Variance Ratio	
	Raw	Matched	Raw	Matched
lgdpsum	0.185	0.072	1.374	1.244
bvdist	-1.040	0.025	3.368	0.984
bvadj	0.415	0.055	14.365	1.171
bvlang	0.386	0.010	1.547	0.870



Appendix 2: Estimating the Dose Response Function

The potential trade outcomes are given by PT_{ij} for country pair ij when treated and NT_{ij} for the same pair when untreated. The treatment indicator, w_{ij}, takes the value 1 for treated and 0 for untreated units, while the variable $PTAB_{ij}$ takes values within the range [0,100].[18] The row vector $\mathbf{x_{ij}} = (LnGDPsum_{ij}, BVDist_{ij}, BVAdj_{ij}, BVLang_{ij})$ comprises the set of exogenous and observable characteristics (confounders). Let N be the total number of country pairs, with N_T being the number of PTA pairs and N_0 the number of non-PTA pairs. We then define two functions $g_T(x_{ij})$ and $g_0(x_{ij})$ as pair ij's responses to the vector of confounding variables when the pair is treated and untreated respectively.

Given the above, a specific population generating process for the two exclusive potential outcomes is assumed:

$$\begin{cases} w = 1: PT = \mu_T + g_T(\mathbf{x}) + h(PTAB) + \epsilon_T \\ w = 0: NT = \mu_0 + g_0(\mathbf{x}) + \epsilon_0 \end{cases} \quad \text{(A1)}$$

Where μ_T and μ_0 are scalars, ϵ_T and ϵ_0 are random variables with mean 0 and constant variance, and $h(PTAB)$ is a general derivable function of $PTAB$ that differs from zero only in the treated status. The causal parameters of interest are the population ATEs conditional on $\mathbf{x}$ and $PTAB$, i.e.

$$\begin{aligned} ATE(\mathbf{x}, PTAB) &= E(PT - NT|\mathbf{x}, PTAB) \\ ATT(\mathbf{x}, PTAB > 0) &= E(PT - NT|\mathbf{x}, PTAB > 0) \\ ATN(\mathbf{x}, PTAB = 0) &= E(PT - NT|\mathbf{x}, PTAB = 0) \end{aligned} \quad \text{(A2)}$$

By assuming that $g_0(\mathbf{x}) = \mathbf{x}\boldsymbol{\delta_0}$ and $g_T(\mathbf{x}) = \mathbf{x}\boldsymbol{\delta_T}$ are linear in parameters, the ATE conditional on $\mathbf{x}$ and $PTAB$ becomes:

$$ATE(\mathbf{x}, PTAB, w) = w \times \{\mu + \mathbf{x}\boldsymbol{\delta} + h(PTAB)\} + (1 - w) \times (\mu + \mathbf{x}\boldsymbol{\delta})$$

where $\mu = (\mu_T - \mu_0)$ and $\delta = (\delta_T - \delta_0)$. We have

$$ATE = p(w = 1) \times ATT + p(w = 0) \times ATN$$

where $p(\cdot)$ is a probability. This gives us

$$\begin{aligned} ATE: &= p(w = 1) \times (\mu + \bar{\mathbf{x}}_{PTAB>0}\delta + \bar{h}_{PTAB>0}) + p(w = 0) \times (\mu + \bar{\mathbf{x}}_{PTAB=0}\delta) \\ ATT: &= \mu + \bar{\mathbf{x}}_{PTAB>0}\delta + \bar{h}_{PTAB>0} \\ ATN: &= \mu + \bar{\mathbf{x}}_{PTAB=0}\delta \end{aligned} \quad \text{(A3)}$$

where $\bar{h}_{PTAB>0}$ is the average response function taken over $PTAB > 0$.

The Dose Response Function (DRF) is a function of the treatment intensity $PTAB$ and is given by averaging $ATE(\mathbf{x}, PTAB)$ over $\mathbf{x}$:

---

[18] The appropriate Stata command (ctreatreg) requires that the treatment variable lies between 0 and 100. For this analysis therefore we had to rescale PTAB to take values between 0 and 100.

$$ATE(PTAB) = \begin{cases} ATT + \{h(PTAB) - \bar{h}_{PTAB>0}\} & \text{if } PTAB > 0 \\ ATN & \text{if } PTAB = 0 \end{cases} \quad \text{(A4)}$$

Before estimating the DRF we need consistent estimates of the parameters of the potential outcomes in (A1). This can be achieved through the following baseline random coefficients regression model:

$$T_{ij} = \mu_0 + w_{ij} \times ATE + \mathbf{x_{ij}}\boldsymbol{\delta_0} + w_{ij} \times (\mathbf{x_{ij}} - \bar{\mathbf{x}})\delta + w_{ij} \times \{h(PTAB_{ij}) - \bar{h}\} + \eta_{ij} \quad \text{(A5)}$$

Under the assumption of unconfoundedness[19] this model can be estimated by OLS to obtain estimates of the ATE and the DRF. The only other requirement is a form for $h(PTAB)$, which is assumed to be a polynommial parametric form of degree m, i.e.:

$$h(PTAB_{ij}) = \lambda_1 PTAB_{ij} + \lambda_2 PTAB_{ij}^2 + \lambda_3 PTAB_{ij}^3 + \cdots + \lambda_m PTAB_{ij}^m$$

where $\lambda_i (i = 1, \ldots, m)$ are among the parameters to be estimated in (A5). The DRF is equal to the average treatment effect (ATE) given the level of treatment $PTAB$. Plotting this across the support of $PTAB$ allows one to observe the pattern of the DRF.

For completeness we report the dose response functions for 1980 and 1990. As mentioned in the text, for these years the pattern tends to be the same with the largest effects found for intermediate levels of PTA depth. Confidence intervals are particularly large however, probably reflecting the relatively small number of treated country-pairs in these two years.

[19] Note that in an alternative version of the model the unconfoundedness assumption can be relaxed through the use of instrumental variables. In our analysis we employ the matched sample of PTA and non-PTA observations to alleviate any concerns about this assumption.

Appendix 3: Additional Dose Response Function Results

Figure A1: Dose Response Function for 1980

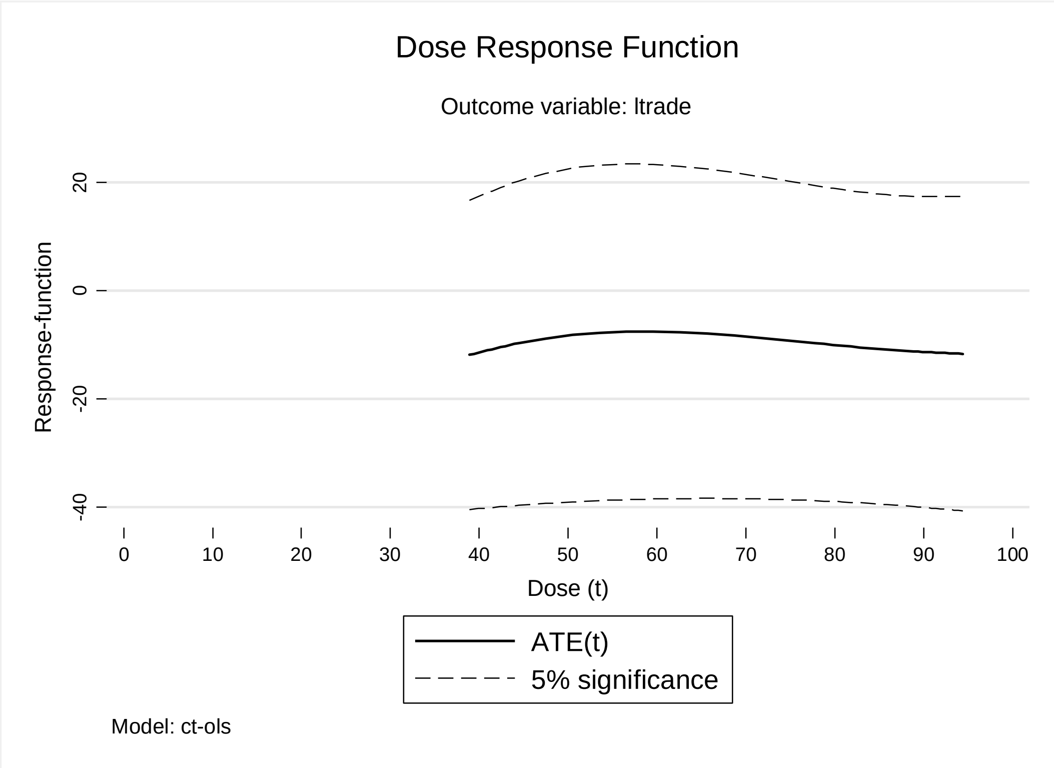
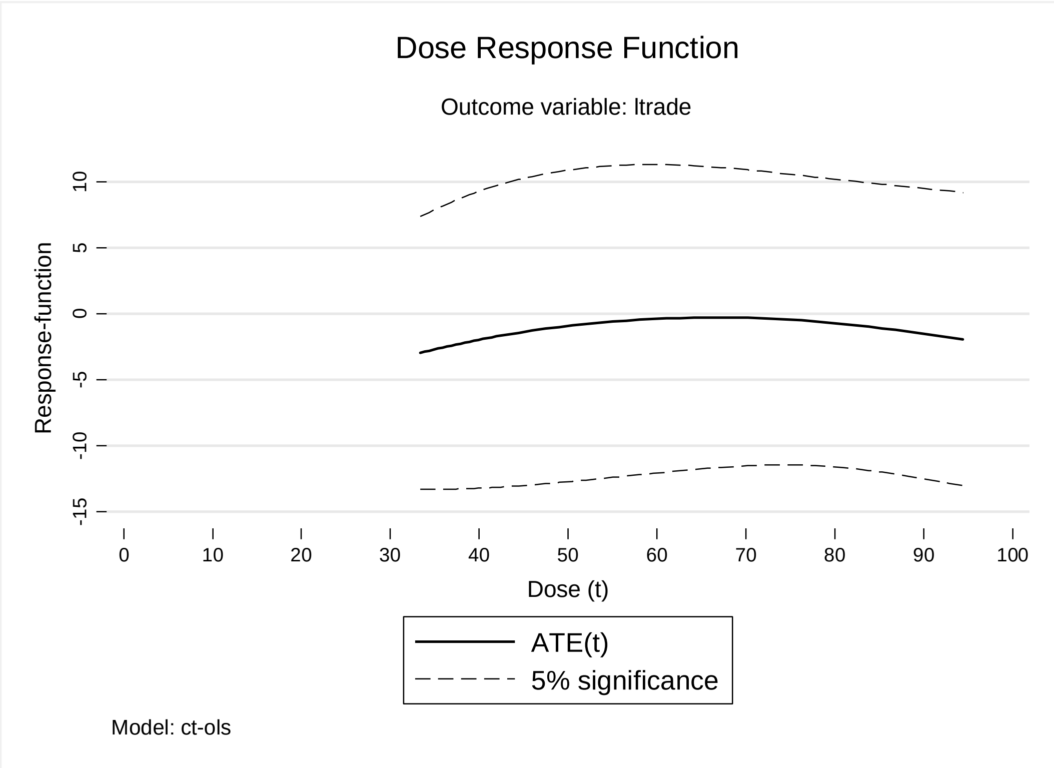


Figure A2: Dose Response Function for 1990



The UNU-MERIT WORKING Paper Series

2017-01 *The economic impact of East-West migration on the European Union* by Martin Kahanec and Mariola Pytliková
2017-02 *Fostering social mobility: The case of the 'Bono de Desarrollo Humano' in Ecuador* by Andrés Mideros and Franziska Gassmann
2017-03 *Impact of the Great Recession on industry unemployment: a 1976-2011 comparison* by Yelena Takhtamanova and Eva Sierminska
2017-04 *Labour mobility through business visits as a way to foster productivity* by Mariacristina Piva, Massimiliano Tani and Marco Vivarelli
2017-05 *Country risk, FDI flows and convergence trends in the context of the Investment Development Path* by Jonas Hub Frenken and Dorcas Mbuvi
2017-06 *How development aid explains (or not) the rise and fall of insurgent attacks in Iraq* by Pui-Hang Wong
2017-07 *Productivity and household welfare impact of technology adoption: Micro-level evidence from rural Ethiopia* by Tigist Mekonnen
2017-08 *Impact of agricultural technology adoption on market participation in the rural social network system* by Tigist Mekonnen
2017-09 *Financing rural households and its impact: Evidence from randomized field experiment data* by Tigist Mekonnen
2017-10 *U.S. and Soviet foreign aid during the Cold War: A case study of Ethiopia* by Tobias Broich
2017-11 *Do authoritarian regimes receive more Chinese development finance than democratic ones? Empirical evidence for Africa* by Tobias Broich
2017-12 *Pathways for capacity building in heterogeneous value chains: Evidence from the case of IT-enabled services in South Africa* by Charlotte Keijser and Michiko Iizuka
2017-13 *Is innovation destroying jobs? Firm-level evidence from the EU* by Mariacristina Piva and Marco Vivarelli
2017-14 *Transition from civil war to peace: The role of the United Nations and international community in Mozambique* by Ayokunu Adedokun
2017-15 *Emerging challenges to long-term peace and security in Mozambique* by Ayokunu Adedokun
2017-16 *Post-conflict peacebuilding: A critical survey of the literature and avenues for future research* by Ayokunu Adedokun
2017-17 *Effects of health insurance on labour supply: A systematic review* by Nga Le, Wim Groot, Sonila M. Tomini and Florian Tomini\
2017-18 *Challenged by migration: Europe's options* by Amelie F. Constant and Klaus F. Zimmermann
2017-19 *Innovation policy & labour productivity growth: Education, research & development, government effectiveness and business policy* by Mueid Al Raee, Jo Ritzen and Denis de Crombrugghe
2017-20 *Role of WASH and Agency in Health: A study of isolated rural communities in Nilgiris and Jalpaiguri* by Shyama V. Ramani
2017-21 *The productivity effect of public R&D in the Netherlands* by Luc Soete, Bart Verspagen and Thomas Ziesemer
2017-22 *The role of migration-specific and migration-relevant policies in migrant decision-making in transit* by Katie Kuschminder and Khalid Koser

2017-23 *Regional analysis of sanitation performance in India* by Debasree Bose and Arijita Dutta

2017-24 *Estimating the impact of sericulture adoption on farmer income in Rwanda: an application of propensity score matching* by Alexis Habiyaremye

2017-25 *Indigenous knowledge for sustainable livelihoods: Lessons from ecological pest control and post-harvest techniques of Baduy (West Java) and Nguni (Southern Africa)* by Leeja C Korina and Alexis Habiyaremye

2017-26 *Sanitation challenges of the poor in urban and rural settings: Case studies of Bengaluru City and rural North Karnataka* by Manasi Seshaiah, Latha Nagesh and Hemalatha Ramesh

2017-27 *Heterogeneous effects of bilateral investment treaties* by Rod Falvey and Neil Foster-McGregor

2017-28 *Willingness to pay for agricultural risk insurance as a strategy to adapt climate change* by Tigist Mekonnen

2017-29 *Social protection investments, human capital, and income growth: Simulating the returns to social cash transfers in Uganda* by Stephan Dietrich, Daniele Malerba, Armando Barrientos, Franziska Gassmann, Pierre Mohnen, Nyasha Tirivayi, Susan Kavuma and Fred Matovu

2017-30 *Opening and linking up: Firms, global value chains and productivity in Latin America* by Pierluigi Montalbano, Silvia Nenci and Carlo Pietrobelli

2017-31 *Husbands' return migration and wives' occupational choices* by Clotilde Mahé

2017-32 *The need to customise innovation indicators in developing countries* by Michiko Iizuka and Hugo Hollanders

2017-33 *Economic diversification: Explaining the pattern of diversification in the global economy and its implications for fostering diversification in poorer countries* by Clovis Freire

2017-34 *How inequality hurts growth: Revisiting the Galor-Zeira model through a Korean case* by Bogang Jun, Mary Kaltenberg and Won-Sik Hwang

2017-35 *Is the demand-pull driver equally crucial for product vs process innovation?* by Herbert Dawid, Gabriele Pellegrino and Marco Vivarelli

2017-36 *Testing linear growth rate formulas of non-scale endogenous growth models* by Thomas HW Ziesemer

2017-37 *Promoting structural transformation: Strategic diversification vs laissez-faire approach* by Clovis Freire

2017-38 *On the relationship between the breadth of PTAs and trade flows* by Rod Falvey and Neil Foster-McGregor