Web Appendix for "Vertical Trade Specialization and the Formation of North-South PTAs"

This appendix presents a short discussion of the methodology in the paper, summary statistics for the main variables and a number of additional robustness checks and comparisons.

1 The Panel Bootstrap compared with other approaches

Bootstrap methods perform well when the sample is a good approximation of the actual population, which is likely to be the case here with 9×148 countries and the corresponding dyads in the sample. Using a panel bootstrap and resampling from N panels with replacement also addresses the additional problems of time dependence in a short panel, the presence of slow-moving variables, and of heteroskedastic residuals in the first stage. All three can obviously lead to biased results. Recent work shows that panel resampling performs best when $N > T$, since bootstrapping relies on asymptotics in N.[1] This is precisely the most common panel setup in international relations applications, and optimal for the present problem where N is approximately 1100 panels but T only 12 years. Specifically for the case of a binary dependent variable, detailed Monte Carlo studies show that the bootstrap performs much better than common parametric methods in the presence of rarely changing variables and temporal dependence.[2] The panel bootstrap is the only method that largely avoids false positives, while standard methods can lead to misleading inferences. In Monte Carlo "placebo" studies, error rates are 3-5 times higher for time dummies and Taylor series approximations[3] and 4-6 higher when using cubic splines.[4] By contrast, the panel bootstrap comes within 1-2% of true error rates regardless of the time dependence. In other words, a 5% significance level is really what it is. This should give greater confidence in the results obtained with this method: a significant coefficient is far less likely to be uncorrelated with the outcome, positives are more likely to be positives, and the inference is generally more conservative.

[1] Kapetanios 2008.

[2] Bischof 2009.

[3] For this approach see Carter and Signorino 2010.

[4] This standard procedure is described in Beck, Katz and Tucker 1998, perhaps the most widely used method in political science when faced with duration-dependent data and a binary dependent variable.

2 Summary Statistics

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Distance	9.052	0.543	5.756	9.891	12623
(ln) Product of GDPs	37.586	2.459	30.392	46.024	12623
(ln) Product of GDP/cap	18.597	1.293	14.603	24.505	12623
(ln) Product of FDI	19.359	2.515	10.566	27.814	12623
(ln) GDP/cap North	10.478	0.812	9.744	13.275	12623
(ln) GDP/cap South	8.119	1.025	4.859	11.23	12623
No. WTO members	137.58	7.21	128	148	12623
MTN round ongoing	0.49	0.5	0	1	12623
Alliance	0.044	0.206	0	1	12623
Percent. share of VIIT, 25% uv diff.	5.592	11.423	0	100	12623
Percent. share of VIIT, 15% uv diff.	6.336	12.772	0	100	12339
Percent. share of VIIT, 35% uv diff.	5.736	11.374	0	99.143	12339
Democracy North	8.641	3.801	-2	10	12623
Democracy South	1.642	6.305	-10	10	12623
PTA coverage North	0.059	0.146	0	0.9	12623
PTA coverage South	0.161	0.218	0	1	12623
(ln) Othertrade	10.605	3.091	-4.509	19.298	12338

Table 1 displays a variety of summary statistics. The share of VIIT in the bilateral trade reaches a maximum of 100% because of three observations relating to Norway and Liberia. The trade in question consists of container ships built in Norway and registered in Liberia, and "second-hand" ships that are bought by Norwegian shipping companies from Liberian-registered carriers. These values and observations, however, bias the analysis *against* the central hypothesis of the paper, since Norway and Liberia evidently did not sign a PTA and are unlikely to do so in the foreseeable future. Arbitrarily dropping these observations merely overstates the effect of VIIT on PTA formation, so that they are left in the dataset for the analysis in the paper.

3 Additional Robustness Checks

Several additional robustness checks are shown in tables 2 to 4. First, it could be argued that treating the European Union as one actor is incorrect, or that the EU has been particularly important in promoting North-South regionalism. Column (8) in table 2 show the results when dyads involving the EU are dropped from the sample. The substantive conclusions are unaffected. Likewise, the shift of the US towards bilateralism has been

cited as important in motivating other countries to move towards PTAs.[5] Column (9) shows the estimation results with the US dropped from the sample. Again, the role of vertical IIT appears unchanged.

Second, I exclude developed countries that primarily export natural resources and agricultural products to their developing-country partners—Australia, Canada, New Zealand and Norway. Dropping these from the analysis does not appreciably change the results, as shown in column (10) of table 2.

Next, it is clear that Korea cannot be classified as a typical developing country by any reasonable standard. Similarly, Singapore may be an outlier because it is a small city-state with a high technological base, but more focused on attracting FDI than exporting it. Dropping these two countries from the analysis does not change the results, as shown in column (11) in table 2.

The WTO classification of "developing countries" is often of more historical than practical relevance. As a general test, I have therefore re-classified developing countries based on their capital-per-worker ratio in the year 2000.[6] I use a cut-off of below 5,000 to exclude countries that evidently do not have much of an industrial base at all, and a cut-off of 50,000 to exclude countries that may be called developing at the WTO, but actually have a higher capital concentration than many developed economies this applies to the major oil and gas exporters, among others. These results are shown column (12) in table 3. Nothing changes in the substantive conclusions.

It is also worthwhile to explore whether the trade relationship is primarily determined by its nature or its volume, and whether horizontal intra-industry trade and inter-industry trade are determinants of the formation of North-South PTAs. It is not possible to include the measured or predicted values of the volume of VIIT trade and the total trade into the same equation, as these are too highly correlated (0.85), as is to be expected when VIIT makes up a considerable share of North-South trade. We can, however, include both the VIIT share and the value of the remaining trade (in constant 2000 USD, logged) in the equation. The remaining trade consists of horizontal intra-industry trade and trade based on comparative advantage. As show in column (13), this does not change anything about the substantive result; moreover, the value of the remaining trade is not found to be statistically significant. Note that this does *not* mean that studies are wrong to include the total volume of trade in equations to investigate determinants of PTA formation, but the composition of this trade and the underlying political economy of trade liberalization will be different, as argued in this paper.

The results do not substantially change either if we consider the year in which a PTA is signed instead of the year in which a PTA enters into force, as shown in column (14) in table 3. The substantive effect is slightly smaller, and the overall model fit as suggested by the Akaike information criterion is worse (AIC 721.11 for the "year of signature" model

[5] Aggarwal 2009, 11.

[6] The data is from Baier, Dwyer and Tamura 2005.

versus 490.86 for the "year of entry into force").

The most commonly used estimation technique for binary dependent variables in the presence of duration dependence is to include natural cubic splines (Beck, Katz and Tucker 1998). While this approach obviously does not address the selection effect in the current model, it is useful to compare the results with those in the paper. Column (15) in table 3 show these results. The coefficient estimate for the share of VIIT is clearly significant, but given the selection effect, its estimation is biased. Unfortunately it is not possible to include time dummies in the second stage because of separation problems (Carter and Signorino 2010, 275-278) that cause too many observations to be dropped.

As argued in the paper, there are theoretical reasons to suspect a selection effect. The main problem with any instrumental variable or two-stage approach is the "exclusion restriction," i.e. the need to ensure that the errors in the first equation are uncorrelated with the unobserved component of the outcome of the second equation. To test whether this is a problem, we can include the residuals from the first stage in the second stage and see if they are statistically significant. Column (16) shows these results. When included in the second stage, the residuals are not statistically significant and of substantively negligible magnitude.

Furthermore, we can also include country dummies in the first stage of the regression (recall that DISTANCE otherwise functions as dyad fixed effect), dyad-dummies, or use the within estimator in the first stage. The estimations results for both stages are shown in columns labeled (17), (18) and (19) respectively in table 4. While the coefficients are somewhat smaller, they clearly remain statistically significant, and the conclusions are unchanged.

Table 2: Robustness checks: Subsamples

	(8)		(9)		(10)		(11)	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Predicted VIIT share	.160***	.047	.208***	.055	.196***	.036	.130**	.046
Distance	–1.149***	.307	–1.145***	.250	–1.120***	.251	–1.176***	.259
Democracy North	–.160***	.041	–.160***	.042	–.207***	.046	—	
Democracy South	.135***	.034	.158***	.033	.107**	.035	.159***	.036
Alliance	.333	.478	.383	.585	.541	.631	.280	.451
No. WTO members	–.191**	.058	–.276***	.065	–.229***	.068	–.256***	.068
MTN round ongoing	2.263**	.807	2.689**	.994	2.253*	1.034	2.742**	1.000
GDP/cap North	–1.978	1.876	–.056	.171	–.057	.182	.078	.162
GDP/cap South	.078	.225	.340	.182	.366	.222	.568*	.222
PTA coverage North, $t-1$	2.204***	.665	1.916***	.577	1.670*	.792	1.958***	.568
PTA coverage South, $t-1$	2.660**	.947	2.872***	.802	3.804***	.850	2.572**	.836
Constant	48.209*	20.175	37.480***	9.621	30.777***	9.295	30.722**	10.380
No. of observations	11387		11188		6929		11195	
R^2								
Pseudo-R^2	.36		.43		.39		.40	

$^{*}p < 0.05,^{**}p < 0.01,^{***}p < 0.001$. Two-tailed tests are conducted for all estimates.

[a]Standard errors (SE) are bootstrapped by resampling from panels with 1000 repetitions.

Table 3: Robustness checks, continued

	(12)		(13)		(14)		(15)	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Predicted VIIT share	.143***	.033	.152**	.057	.163***	.040		
OtherTrade			.076	.093				
ln VIIT share							.016*	.008
Distance	–.985***	.255	–1.182***	.230	–1.169***	.213	–1.654***	.198
Democracy North	–.195***	.046	–.152***	.040	–.143***	.038	–.114***	.032
Democracy South	.115***	.034	.135***	.029	.127***	.026	.127***	.024
Alliance	.543	.533	.238	.449	.536	.376	.704	.393
No. WTO members	–.189**	.061	–.241***	.060	–.241***	.053	–.894***	.238
MTN round ongoing	1.712	.910	2.540**	.871	2.655***	.770	3.450***	1.044
GDP/cap North	.007	.199	–.119	.164	–.057	.149	.378**	.130
GDP/cap South	.204	.249	.290	.185	.381*	.160	.480**	.167
PTA coverage North, $t-1$	1.934**	.704	1.949**	.604	1.876***	.534	2.564***	.466
PTA coverage South, $t-1$	3.687***	.804	2.887***	.765	2.808***	.726	1.537*	.658
Constant	25.504**	8.937	33.591***	8.687	32.975***	7.460	116.073***	30.523
No. of observations	6472		12338		12570		12623	
R^2								
Pseudo-R^2	.36		.39		.37		.40	

$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$. Two-tailed tests are conducted for all estimates. Standard errors (SE) are bootstrapped by resampling from panels with 1000 repetitions, except for (15) which provides robust standard errors.

Table 4: Robustness checks, continued

	(16)		(17)		(18)		(19)	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Predicted VIIT share	.169***	.043	.098***	.024	.036***	.011	.072*	.034
Stage 1 residuals	.010	.009						
Distance	–1.202***	.239	–1.278***	.238	–1.457***	.248	–1.643***	.230
Democracy North	–.155***	.042	–.124**	.046	–.130**	.043	–.154***	.042
Democracy South	.136***	.031	.114***	.032	.122***	.031	.141***	.030
Alliance	.332	.464	.844	.468	.723	.474	.449	.451
No. WTO members	–.182***	.044	–.144**	.044	–.120**	.046	–.118*	.046
MTN round ongoing	1.592**	.579	1.312*	.576	1.119	.575	1.032	.580
GDP/cap North	–.033	.163	.367**	.129	.364**	.132	–.067	.225
GDP/cap South	.276	.185	.073	.180	.324	.169	.256	.200
PTA coverage North, $t-1$	1.995***	.591	2.440***	.540	2.454***	.551	2.039***	.596
PTA coverage South, $t-1$	3.002***	.775	3.066***	.840	2.392**	.827	2.435**	.899
Constant	25.830***	6.716	19.038**	6.780	16.107*	7.096	22.561**	7.246
No. of observations	12623		12623		12623		12623	
R^2								
Pseudo-R^2	.38		.38		.37		.36	

$^{*}p < 0.05,^{**}p < 0.01,^{***}p < 0.001$. Two-tailed tests are conducted for all estimates.
Standard errors (SE) are bootstrapped by resampling from panels with 1000 repetitions.

References

Aggarwal, Vinod K. 2009. "Reluctance to Lead: U.S. Trade Policy in Flux." *Business and Politics* 11(3):1–21.

Baier, Scott L., Gerald P. Dwyer and Robert Tamura. 2005. Factor Returns, Institutions, and Geography: A View From Trade. Technical report Atlanta, GA: .

Beck, Nathaniel, Jonathan N. Katz and Richard Tucker. 1998. "Taking time seriously: Time-series-cross-section analysis with a binary dependent variable." *American Journal of Political Science* 42(4):1260–1288.

Bischof, Jonathan M. 2009. "A Bootstrap Approach to Time Invariance in Panel Data." Paper presented at the American Political Science Association (APSA) Annual conference in Chicago, September 1-4, 2009.

Carter, David B. and Curtis S. Signorino. 2010. "Back to the Future: Modeling Time Dependence in Binary Data." *Political Analysis* 18(3):271–292.
URL: *http://pan.oxfordjournals.org/cgi/content/abstract/18/3/271*

Kapetanios, George. 2008. "A bootstrap procedure for panel data sets with many cross-sectional units." *Econometrics Journal* 11(2):377–395.