Online Appendix

Figure 1: Evolution of PTA depth and breadth
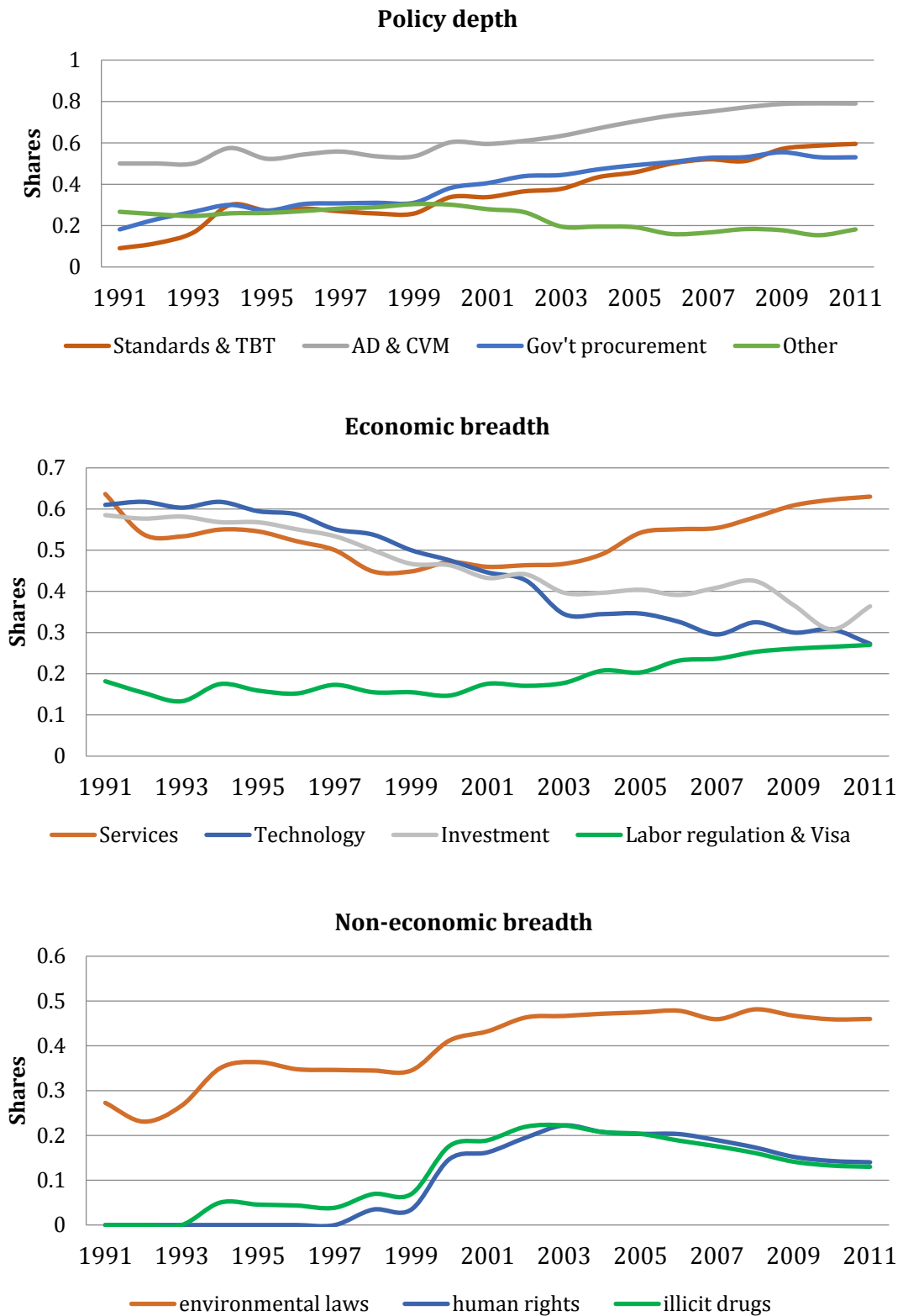
Policy depth

Economic breadth

Non-economic breadth

Table 1: Regression Samples

Master	PTA	PTA, GDP	PTA, Tariff	Master	PTA	PTA, GDP	PTA, Tariff	Master	PTA	PTA, GDP	PTA, Tariff
Afghanistan	X		X	Gambia, The	X	X	X	Norfolk Island			
Albania	X	X	X	Georgia	X	X	X	Northern Mariana Islands			
Algeria	X	X	X	Germany	X	X	X	Norway	X	X	X
Andorra				Ghana	X	X	X	Oman	X	X	X
Angola	X	X	X	Gibraltar				Pakistan	X	X	X
Anguila				Greece	X	X	X	Palau	X	X	X
Antigua and Barbuda	X	X	X	Greenland	X	X		Panama	X	X	X
Argentina	X	X	X	Grenada	X	X	X	Papua New Guinea	X	X	X
Armenia	X	X	X	Guadeloupe				Paraguay	X	X	X
Aruba	X	X		Guatemala	X	X	X	Peru	X	X	X
Australia	X	X	X	Guinea	X	X	X	Philippines	X	X	X
Austria	X	X	X	Guinea-Bissau	X	X	X	Pitcairn			
Azerbaijan	X	X	X	Guyana	X	X	X	Poland	X	X	X
Bahamas, The	X	X	X	Haiti	X	X	X	Portugal	X	X	X
Bahrain	X	X	X	Honduras	X	X	X	Qatar	X	X	X
Bangladesh	X	X	X	Hong Kong, China	X	X	X	Reunion			
Barbados	X	X	X	Hungary	X	X	X	Romania	X	X	X
Belarus	X	X	X	Iceland	X	X	X	Russian Federation	X	X	X
Belgium	X	X	X	India	X	X	X	Rwanda	X	X	X
Belize	X	X	X	Indonesia	X	X	X	Saint Helena			
Benin	X	X	X	Iran, Islamic Rep.	X	X	X	Saint Pierre and Miquelon			
Bermuda	X	X	X	Iraq	X	X		Samoa	X	X	
Bhutan	X	X	X	Ireland	X	X	X	San Marino	X	X	
Bolivia	X	X	X	Israel	X	X	X	Sao Tome and Principe	X	X	
Bosnia, Herzegovina	X	X	X	Italy	X	X	X	Saudi Arabia	X	X	X
Botswana	X	X	X	Jamaica	X	X	X	Senegal	X	X	X
Brazil	X	X	X	Japan	X	X	X	Seychelles	X	X	X
British Virgin Islands				Jordan	X	X	X	Sierra Leone	X	X	X
Brunei	X	X	X	Kazakhstan	X	X	X	Singapore	X	X	X
Bulgaria	X	X	X	Kenya	X	X	X	Slovak Republic	X	X	
Burkina Faso	X	X	X	Kiribati	X	X		Slovenia	X	X	
Burundi	X	X	X	Korea, Dem. Rep.				Solomon Islands	X	X	
Cambodia	X	X	X	Korea, Rep.	X	X	X	Somalia	X		
Cameroon	X	X	X	Kuwait	X	X	X	South Africa	X	X	X
Canada	X	X	X	Kyrgyz Republic	X	X	X	Spain	X	X	X
Cape Verde	X	X	X	Lao PDR	X	X	X	Sri Lanka	X	X	X
Cayman Islands	X			Latvia	X	X		St. Kitts and Nevis	X	X	X
Central Afr. Republic	X	X	X	Lebanon	X	X	X	St. Lucia	X	X	X
Chad	X	X	X	Lesotho	X	X	X	St. Vincent & Grenadines	X	X	X
Chile	X	X	X	Liberia	X	X		Sudan	X		

Country	PTA	GDP	tariff	Country	PTA	GDP	tariff	Country	PTA	GDP	tariff
China	X	X	X	Libya	X	X		Suriname	X	X	X
Christmas Island				Lithuania	X		X	Swaziland	X	X	X
Cocos Islands				Luxembourg	X	X	X	Sweden	X	X	
Colombia	X	X	X	Macao	X	X	X	Switzerland	X	X	X
Comoros	X	X	X	Macedonia, FYR	X	X	X	Syrian Arab Republic	X	X	X
Congo, Dem. Rep.				Madagascar	X	X	X	Tajikistan	X	X	X
Congo, Rep.	X	X	X	Malawi	X	X	X	Tanzania	X	X	X
Cook Islands				Malaysia	X	X	X	Thailand	X	X	X
Costa Rica	X	X	X	Maldives	X		X	Togo	X	X	X
Cote d'Ivoire	X	X	X	Mali	X	X	X	Tokelau			
Croatia	X	X	X	Malta	X	X	X	Tonga	X	X	X
Cuba	X	X	X	Marshall Islands	X	X		Trinidad and Tobago	X	X	X
Cyprus	X	X	X	Martinique				Tunisia	X	X	X
Czech Republic	X	X		Mauritania	X	X	X	Turkey	X	X	X
Denmark	X	X	X	Mauritius	X	X	X	Turkmenistan	X	X	
Djibouti	X	X	X	Mexico	X	X	X	Turks and Caicos Isl.			
Dominica	X	X	X	Micronesia	X	X		Tuvalu			
Dominican Republic	X	X	X	Moldova	X	X	X	Uganda	X	X	X
East Timor				Mongolia	X	X	X	Ukraine	X	X	X
Ecuador	X	X	X	Montserrat				United Arab Emirates	X	X	X
Egypt, Arab Rep.	X	X	X	Morocco	X	X	X	United Kingdom	X	X	X
El Salvador	X	X	X	Mozambique	X	X	X	United States	X	X	X
Equatorial Guinea	X	X		Myanmar	X		X	Uruguay	X	X	X
Eritrea	X	X		Namibia	X	X	X	Uzbekistan	X	X	
Estonia	X	X		Nauru				Vanuatu	X	X	X
Ethiopia	X	X	X	Nepal	X	X	X	Venezuela	X	X	X
Faeroe Islands	X			Netherlands	X	X	X	Vietnam	X	X	X
Falkland Island				Netherlands Antill.	X			Wallis and Futura Isl.			
Fiji	X	X	X	New Caledonia	X	X		Western Sahara			
Finland	X	X		New Zealand	X	X	X	Yemen	X	X	X
France	X	X	X	Nicaragua	X	X	X	Yugoslavia,FR Serbia/Monte.			
French Guiana				Niger	X	X	X	Zambia	X	X	X
French Polynesia				Nigeria	X	X	X	Zimbabwe	X	X	X
Gabon	X	X	X	Niue							

Notes: the first column lists the 221 countries in the *master* sample, this is the sample that results from merging trade flows and geographic variables. The second column marks with an "X" those countries that also belong to the PTA sample. The third columns marks with an "X" those countries that, in addition, belong to the GDP sample (those countries for which nominal GDP and GDP deflators are non-missing). The third column marks with an "X" those countries that belong to the PTA and tariff samples.

Data

The data and code to replicate the results in the chapter are available at http://terpconnect.umd.edu/~limao/handbook_pta. The use of either should be cited as follows:

Limão, Nuno. 2016. "Preferential Trade Agreements" in: *Handbook of Commercial Policy*, edited by K. Bagwell and R. Staiger, Elsevier.

I.   Sources

1.   Bilateral trade flows
- o   Source: UN Comtrade database, available from WITS https://wits.worldbank.org/
- o   Description: annual bilateral imports and exports since 1962 for more than 200 countries

2.   Preferential Agreements
- o   Source: Database on Economic Integration Agreements, by Baier and Bergstrand, Version September 30, 2015 (last downloaded in January 2016) http://www3.nd.edu/~jbergstr/DataEIAsSeptember2015/EIA%20Database%20September%2030,%202015.zip
- o   Description: records the economic integration of bilateral country pairings for 195 countries annually from 1950 through 2011. Depending on the level of economic integration, a country pairing was assigned a number code from 0, representing no trade agreement, to 6, representing an Economic Union.
- o   For figure 1 only we use the figures reported by WTO (2011).

3.   WTO membership
- o   Source: WTO, https://www.wto.org/english/thewto_e/whatis_e/tif_e/org6_e.htm
- o   Description: dates of WTO membership and pre-WTO GATT membership (where applicable) for its 162 members as of November 2015

4.   Breadth and depth of agreements
- o   Source: Updated dataset on the Content of PTAs, Version April 2013, WTO, https://www.wto.org/english/res_e/publications_e/wtr11_dataset_e.htm
- o   Description: The dataset is an extension of the data provided in Horn et al. (2010). The agreements covered represent almost 90 per cent of world trade and cover most regions of the world: the United States, the European Union, Africa, the Middle East, Oceania, Asia, and Central and South America. The dataset includes PTAs between WTO members as well as agreements between WTO members and non-members. This updated version consists of a comprehensive breakdown of 100 PTAs signed between 1958 and 2011. For the methodology, see Section D.2 of the 2011 World Trade Report.

5.   Bilateral tariffs
- o   Source: UNCTAD Trade Analysis Information System (TRAINS), available from WITS https://wits.worldbank.org/
- o   Description: provides data on a wide set of trade control measures at the HS-based tariff line level (HS 6-digit). These data are reported by 150 countries. Depending on the country, the data are available from 1988 onwards.
- o   For figure 1 only we use the country groupings defined by the World Bank (via the WDI).

6.   Geographic variables
- o   Source: GeoDist database by CEPII, see Mayer and Zignano (2011) for documentation, http://www.cepii.fr/cepii/en/bdd_modele/presentation.asp?id=6
- o   Description: different measures of bilateral distances available for 225 countries

7.   National data
- o   Source: World Development Indicators database by the World Bank
- o   Description: annual nominal GDP in US dollars and GDP deflators since 1960.

II. Construction

Using the data sources described in the previous section we construct an (unbalanced) panel for the period 1962-2010. As is customary in the literature, we use data in five-year intervals starting in 1965. Since the trade data starts in 1962, we also add that year to the panel and use it only if we need to create a lagged PTA variable for 1965.

The regression analysis is carried using different samples of countries depending on the specification since country coverage varies across datasets. When combining the different datasets we used the following methodology. First, we merged bilateral trade flows with geographic variables and we kept only the dyads that belonged to both datasets. Consequently, we discarded the dyads that belonged to only one of the aforementioned datasets. Let us called the resulting dataset, the *master* dataset. Second, we sequentially merged the rest of the datasets, the *using* datasets (tariffs, national data, etc), and we only kept the observations that belong to the *master* dataset. Therefore, it is possible that a pair of countries in the final dataset has missing values for tariffs, national, or agreements data.

In the following paragraphs we describe specific changes we applied to the original datasets.

1. Bilateral trade flows

- Following Feenstra et al. (2005) we use import flows as reported by the importing country since they tend to be reported with more accuracy than export flows as reported by the exporting country. However, if the import flow is missing and the reverse export flow is not, we impute this last value as an import. For example, if country A does not report imports from country B, but country B reports exports to A, we assign the value of exports from B to A as an import by A from B. We identify imputed observations with a dummy variable called *flag* that equals 1 if the value of the import flow is imputed.

- We rename some countries that split/unify in order to have consistent country names through time. We rename Belgium-Luxembourg (1962-1998) as Belgium, Former Ethiopia as Ethiopia, Former Sudan as Sudan, and the Soviet Union as Russian Federation.

2. Breadth and depth of agreements

- The original WTO agreement data has 100 agreements between 116 countries classified in 52 categories. We aggregate these categories in four (mutually-exclusive) groups according to the depth of the agreements and five (mutually-exclusive) groups according to their breadth.

3. Tariffs

- We use Most Favored Nation, Preferential, and Effectively Applied bilateral tariffs in their simple average and weighted average form as calculated by TRAINS. Both types of averages are calculated over all 6 digit lines imported by the reporter country from a given partner. Weighted averages are weighted by the value of the import flow of each tariff line.

- This dataset aggregates the countries belonging to the European Union as a single reporter. Since our dataset is at the country level, we imputed to each EU country the tariffs applied by the Union to third countries, taking into account its date of accession. We also imputed tariffs equal to zero between EU member countries.

- In this dataset missing data is particularly severe. To deal with this issue, we resort to missing-values imputation. If the bilateral tariff in year t is missing, we impute the average of years t-1 and t+1. If the value in t-1 is missing, we impute the average of t-2 and t+1. If the

value in t+1 is missing, we impute the average of t-1 and t+2. If both years t-1 and t+1 are missing, we impute the average of year t-2 and t+2. Finally if all years from t-2 to t+2 are missing, we leave the observation as a missing value.

4. Geographic Variables

o The original CEPII dataset has four measures of distances and two measures of common language. We use as a measure of distance between two countries the simple distance between the most populated cities and the criterion for common language is the official or primary language.

5. GDP Deflators

o The base year of the GDP deflators in the original WDI dataset vary across countries. We normalize GDP deflators so that the value in the year 2000 equals 100 for all countries.