Clustering Value-Added Trade: Structural and Policy Dimensions

Escaith, Hubert and Gaudin, Hadrien

World Trade Organization

10 July 2014

10 July 2014

Clustering Value-Added Trade: *Structural and Policy Dimensions*

Hubert Escaith[‡] and Hadrien Gaudin[‡‡]

‡: World Trade Organization, Geneva (CH)
‡‡ École Polytechnique, Paris (F)

 The authors wish to acknowledge the contribution of Amaury Decludt on initial drafts and the statistical support of the WTO Statistics Group. Any errors remain the fault of the authors.

Clustering Value-Added Trade:
Structural and Policy Dimensions

Hubert Escaith [‡] and Hadrien Gaudin [‡‡]
‡: World Trade Organization; ‡‡ École Polytechnique, Paris.

***Summary*:** The paper builds a typology of value-added traders according to their economic and trade policy characteristics. In the process, it defines clusters of countries according to the multi-dimensional criteria defined by value-added, economic and trade policy indicators. A second approach focuses on the relationships existing between the variables themselves, using multi-criteria and graph analysis. Natural resources endowments, on the one hand, and services orientation, on the other one, are among the most determinant variables for defining Trade in Value Added (TiVA) clusters. The level of economic development remains a crucial determinant of the TiVA profile as is the size of the economy, even if not as important as initially expected. Pro-active GVC up-grading strategies, such as investments in ICT and R&D tend to foster a higher foreign content in exports, compensating the lower domestic margin by higher volumes. Inward-oriented protectionist policies are not particularly successful in exporting higher share of domestic content, except in services exports; but in this case, export volumes remain marginal.

Key words: Trade in value-added, global value chains, trade policy, input-output analysis, effective protection rate, exploratory data analysis.

JEL codes: D57, F13, F14, F15, F23, O19, O24

Contents:

CLUSTERING VALUE-ADDED TRADE:
STRUCTURAL AND POLICY DIMENSIONS

1. INTRODUCTION

Thanks to an increasing international fragmentation of production networks, Global Value Chains (GVCs) have become a dominant feature of today's global economy. This phenomenon has variously been called fragmentation, unbundling, offshoring, vertical specialization, slicing-up of the value-added chain or trade in tasks (WTO, 2008). This new phase of the globalization process challenges conventional understanding on how to interpret trade statistics and, therefore, how to design trade policies. Some researchers even suggest GVCs, by undermining the old Ricardian law of comparative advantages, determined a paradigm change in international economics (Grossman and Rossi-Hansberg, 2006). Even if this remains an open question, the fact is that GVCs alter many of the stylised facts on which international economics models are based. Actually, GVCs impact on a wide range of policy domains, not just those related to trade policy. Amongst others, they have far reaching impacts on competitiveness, industrial policy, employment and labour skills, between and within countries' equity and income distribution, access to markets, etc.

Yet GVCs are still an unchartered territory from an empirical perspective. Up to the mid-2000s, anecdotic data were available through case studies but comprehensive aggregate level analyses were more limited. It is not before the 2000s that systemic efforts to produce internationally consistent estimators were put in place, first in the academia (e.g., Daudin et al., 2006, building on Hummels et al., 2001). After a pilot study realised with IDE-JETRO's Asian Input-output data (WTO and IDE-JETRO, 2011b), international efforts conducted to establishing a first global database built on official data released in 2012 by WIOD (a UE sponsored project); subsequently OECD and WTO released their Trade in Value-Added indicators in 2013 (*TiVA*) building on the OECD's expertise in harmonizing IO and linking individual countries' IO matrices with trade flows in intermediate goods and services (see OECD-WTO, 2012 for a background technical note).

Most empirical papers published on trade in value-added address specific statistical or economic issues; mapping the trade in value-added territory remains to be done. The present essay intends to start filling this gap by building a typology of value-added traders according to their structural economic characteristics and their trade policy options. In this perspective, it differs from, but complements, the research programmes aiming at mapping global value chains from the trade network geographical perspective "*who trades what with whom?*" as is most often intended (Ng and Yeats, 1999; Koopman et al., 2012; De Backer and Miroudot, 2013, to cite only a few).[1] In particular, we look into the determinants of vertical specialization and the domestic contents of sectoral exports, two of the most common measures of trade in value added.

Moving from the general to the particular, the paper starts by looking at the relationship between economic characteristics and trade in value-added, before defining clusters of countries according to multi-dimensional criteria defined by a series of *TiVA*, economic and trade policy indicators. Rather than focusing on the characterization of country profiles, the second approach concentrates on the relationships existing between the variables themselves. Multi-criteria and graph analysis are used to identify the degree of association between sub-sets of indicators belonging to those three variable spaces. This study being —to our knowledge— among the first ones of its kind, we limited our exploration to the main dimensions of the variable space and did not go further than scratching the surface of the rich information contained in the data sets. This process may obviously suffer from a selection bias in the choice of variables that requires further consideration. As extending the analysis to more variables is limited by the number of observations, the inclusion of more countries in the *TiVA* database in the next years will provide an opportunity to explore further dimensions.

Conclusions summarize the stylized facts that were identified, while highlighting the normative limits of the results obtained. As the readers will be reminded again and again in this essay, exploratory data analysis follows the "*estimate, don't test*" approach to statistics and does not

[1] For a comprehensive review of the GVC literature, see Park, Nayyar and Low (2013).

pretend identifying causalities and models. Its aim is helping the analysts and decision makers in organizing their empirical knowledge by pointing at underlying patterns and stylised facts.

2. GROSS VS. VALUE-ADDED INTERNATIONAL TRADE MEASUREMENT

Gross trade statistics derive usually from customs registers (merchandise trade statistics) and balance of payments (trade in services statistics). They measure the commercial value of the products that are exchanged between two countries. When all the production processes take place in one country and trade is in final products, they are also a good indicator of the economic value that is created by and retained by the exporting country. But today, trade is increasingly composed of intermediate products that are exchanged within production networks (global value chains) and traditional trade statistics suffer from a double counting bias: an input imbedded into goods for processing will cross several borders before reaching its final destination. In addition, gross exports may not reflect adequately the economic value that the exporter created, considering that the commercial valuation retained by customs administration includes the value of all the imported parts and components (including intermediate services) used in the production of this export.

Measuring trade in value-added is a way of correcting for double counting and estimating the various sources (by country and industry) that contributed to the value-added along the international supply chain.

2.1 Definition and first results

The Trade in Value-Added approach traces the value added by each industry and country in the production chain and allocates the value-added to these source industries and countries. Estimates of the value added content of trade rely typically on Leontief inverse matrices based on international input-output (I-IO) tables, which integrate national accounts and bilateral trade statistics. I-IO tables present the advantage to capture in a cost-effective manner not only direct linkages and exchanges between countries and sectors but, after applying standard Leontief transformation, also the indirect sectoral linkages (See Box 1). Even if *TiVA* goes up to year 2009 in its May 2013 version (to be updated end of 2014), the analysis focuses on 2008, as 2009 was affected by a deep recession and may not be representative.

Box 1 Introduction to the measure of trade in value-added.

Value-added reflects the value that is added by industries in producing goods and services. It follows the definition of value-added (in basic prices) used in the System of National Accounts (1993 SNA) and is equivalent to the difference between its output (in basic prices) and the sum of its intermediate inputs (in purchasers prices) of goods and services. It is equivalent to the compensation for labour (Compensation of Employees) and compensation for capital (Operating Surplus), but also includes a component for 'Other taxes on Production'. Input-output tables reflect both the interrelationships between domestic industries and between industries and final demand categories (households, government, investment and exports). Furthermore, they reflect how intermediate imports are used in producing goods and services, and how imports of final goods are consumed.

The basic idea behind measuring the value-added content in trade flows is relatively straightforward. Starting from the Leontief model, the total output of a (national or global) economy is given by the sum of intermediate consumption (inputs used for production) and final demand (consumption, investment, exports):

$$\mathbf{X = AX + Y} \qquad [1]$$

$$\mathbf{X = [I - A]^{-1} Y} \qquad [2]$$

Where

Y is the *n*x1 final demand vector and **X** is the *n*x1 vector of total production (*n* being the number of industries);

A is the matrix of technical coefficients, derived by normalizing the intermediate coefficients Z_{ij} by the value of total production ($a_{ij} = Z_{ij}/X_i$); where Z_{ij} is the intermediate consumption of products from sector *i* by *j* (*i* and *j* being possibly in different countries) and X_i is the total production of sector *i*.

$(\mathbf{I-A})^{-1}$ is known as the Leontief inverse matrix (**L**)

Considering **V** as the *nxn* diagonal matrix of value added coefficients, the total value added created in the economy (**VA,** a nx1 vector) is equal to:

$$\mathbf{VA_{nx1} = V\ L\ Y} \qquad \mathbf{[3]}$$

Under the hypothesis of homogeneity within the various components of the final demand Y, in particular that exported products are produced using the same production function (a_{ij}) as products destined to the domestic market, equation [3] can be used to measure the domestic value-added content [$\mathbf{VA^E}$, a nxn matrix] of gross exports [**E,** a nxn diagonal matrix based on the vector of gross exports].

$$\mathbf{VA^E_{nxn} = V\ L\ E} \qquad \mathbf{[4]}$$

In practice, extending X and L to cover many countries and sectors while maintaining the basic national accounts identities is a challenging statistical process. The measurement issues are also more complex because some of the exported value-added may return to the country of origin as imports of intermediate or final products (see Escaith, 2014, for a review of the measurement issues; Koopman, Powers, Zhi Wang and Shang-Jing Wei, 2014, for a detailed discussion). Eventually, the homogeneity assumption does not hold and may even become unrealistic in some cases (China, Mexico) where for some industries a large share of exports results from of deeply integrated global value chains relying much more on imported inputs than the rest of the economy.

A series of GVC indicators can be derived from equation [4].

Foreign content or Vertical Specialization (*VS*) is obtained by a column summation of the VA^E matrix (excluding domestic sectors) divided by gross exports [**E**] in each country, which yields a vector of VS shares, as defined by Hummels et al. (2001). Similarly, summing along rows (and excluding domestic sectors) and normalizing by [**E**] provides the share of domestic value added embodied in intermediates products that will be used as inputs by foreign countries.

The GVC participation index proposed by Koopman, Powers, Zhi Wang and Shang-Jing Wei (2010) adds the two calculations (columns and rows). It measures the share of foreign value-added embodied in gross exports and domestic contribution to the exports of third countries. Escaith (2014b) suggests excluding from the calculation of the second term the direct domestic value-added incorporated in the exports of primary commodities. The rationale for such exclusion is that commodities are undifferentiated products commonly traded on large spot or future markets. They do not always imply the kind of long term business-to-business relationship that characterizes international supply chains and may over-estimate the actual participation in GVCs. To account for this bias, one should consider only the direct and indirect exports of domestic value added originating from the secondary or tertiary sectors plus the indirect exports of embodied value-added from primary sectors (but not the direct ones).

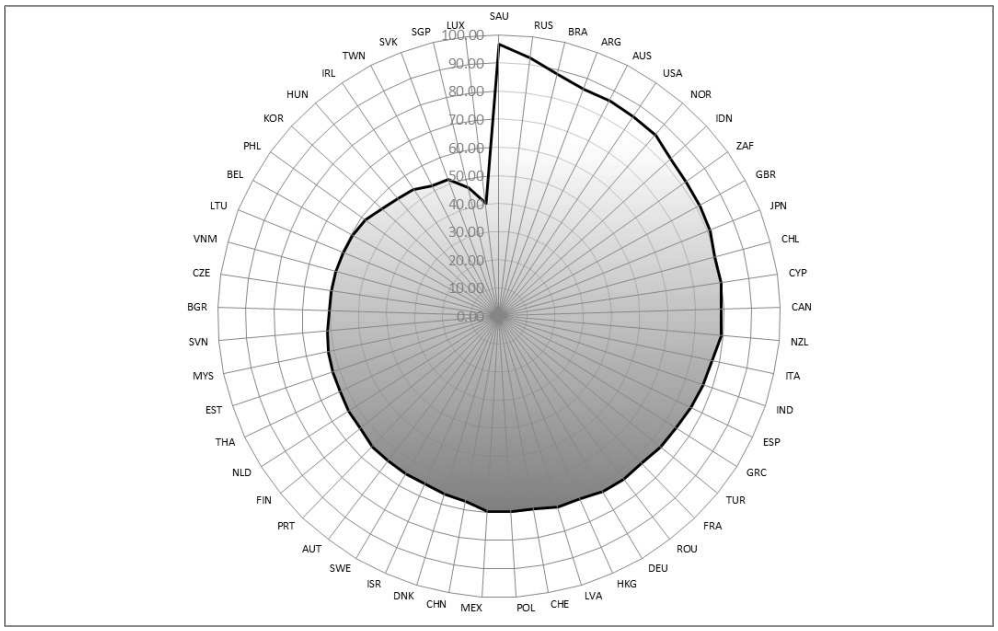
Deriving from the notion of backward and forward linkages, an additional indicator that can be derived is the average propagation length (APL), first introduced by Dietzenbacher and Romero (2007), and developed by Inomata (2008). Completing the measure of the strength of industrial linkages, APL allows estimating the length of supply chains, by simulating the propagation of supply or demand shocks through the vertical integration of production processes. For a review and application of some GVC indicators, see De Backer and Miroudot (2013), Escaith and Inomata (2013) and Zhi Wang, Shang-Jin Wei and Kunfu Zhu (2014).

The OECD-WTO TiVA database used in this paper measures trade in value-added by means of the global IO table elaborated by the OECD and covering 57 countries (May 2013 release) from 1995 to 2009. The sectorial level of detail used covers 37 industries. [2]

Figure 1 shows that the domestic value-added content embodied in gross total exports of goods and services varies widely from country to country. Saudi Arabia exports include almost 100% of domestic content while this share drops to 40% in the case of Luxemburg. In other words, Luxemburg relies on average for 60% on foreign inputs for her exports. Because the share of foreign content is one of the main indicators of participation in global value chains, one can infer that Luxembourg is better inserted in those value chains (more vertically specialized, to use the *VS* indicator proposed by Hummels et al. (2001).

[2] For further information on the methodology see OECD-WTO (2012) 'Trade in Value-Added: Concepts, Methodologies and Challenges'.

Figure 1 Share of domestic value-added in gross total exports, 2008



Source: Based on OECD-WTO *TiVA* database (May 2013 release).

The countries that rank highest on the domestic value-added content (and therefore lowest for the *VS* criterion) are exporters of upstream primary products (Saudi Arabia, Russian Federation, Brazil or Argentina); conversely, the countries where the foreign value-added content (*VS*) is highest are downstream service oriented economies (Luxemburg, Singapore).

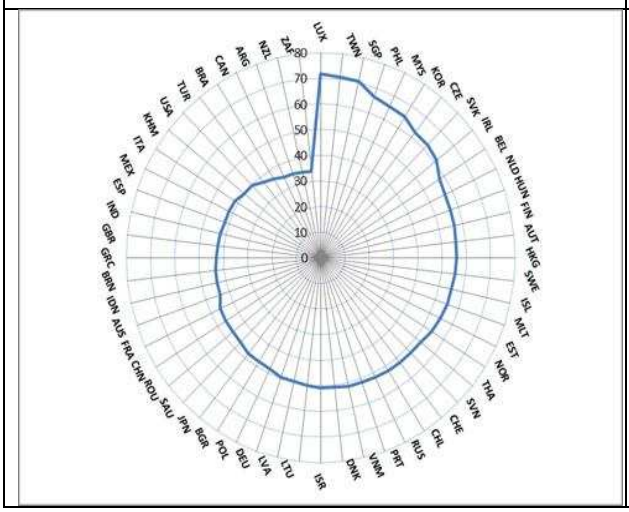
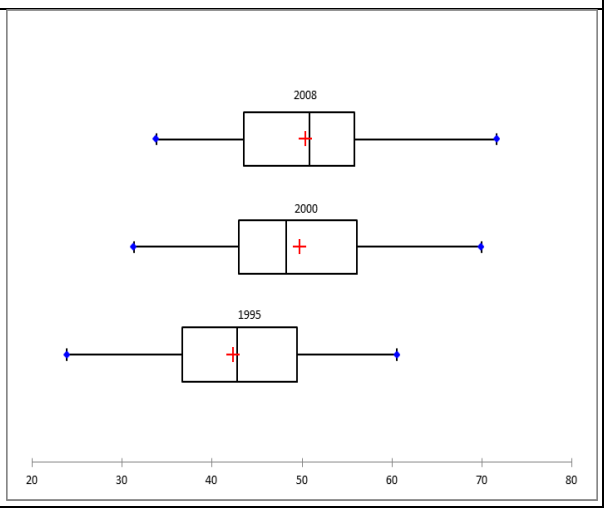
Yet *VS,* which measures the imported content in the composition of export, is not the sole indicator of value chains insertion, as it tends to be higher for countries specializing in downstream activities (closer to final demand) while not considering the contribution of upstream GVC sectors. Downstream sectors' exports increasingly rely on significant intermediate imports (and, so, value added created by industries in upstream countries). For example, Saudi Arabia is mainly exporting upstream products (fuels and oil derivatives) that are key inputs for downstream value chains. Even if her *VS* is low, her exports are key precursor inputs for many global value chains.

As mentioned in Box 1, a more comprehensive indicator is the GVC participation index. It adds the foreign value added contained in exports (roughly similar to the original *VS*) and the domestic value added that is exported to third countries in intermediate goods in order to be reprocessed. The higher the foreign value-added embodied in gross exports and the higher the value of intermediate goods exported to third countries and used as inputs to produce their exports, the higher the participation of a given country in the global value chain.

The GVC participation index is not a symmetric image of the domestic value added content in exports and there are some significant changes in the relative ranking of each economy according to the two indicators. When the downstream use of domestic value added for further processing in third countries is taken into account, natural resources exporters show much higher insertion in GVCs. If one ranks countries first by foreign value added content (data not shown here) then by exports for further processing, Saudi Arabia gains 13 places and South Africa 29. When some gain, others lose: in Asia, for example, Indonesia gains 20 places while Philippines, a downstream exporter oriented towards the production of goods and services for final demand, loses 19. [3]

[3] This index of GVC participation is purely quantitative and does not provide much indication on the quality of the GVC insertion. Producers of commodities that are traded on international markets have shallower business relationship with their clients than producers of specific intermediate products like automotive parts and components, which are made to meet the special requirements of a single customer. Escaith (2014b) corrects for this bias. The present paper uses the original calculation, as implemented in the OECD-WTO TiVA database (May 2013 release).

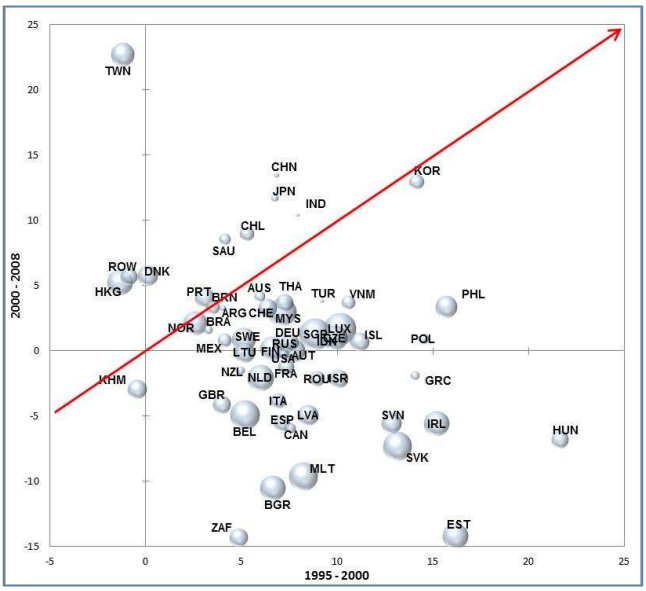
Figure 2. GVC participation index, 1995-2008

Participation index 2008	Boxplots 1995-2008
	

Note: In panel (b), each box represents the first and third quartiles of the distribution, the line inside the box indicates the median and the crux the mean (57 observations). Whiskers indicate extreme values.
Source: Based on OECD-WTO *TiVA* database.

Panel (b) also shows that countries integration in GVCs has increased rapidly between 1995 and 2000. The pace of progress has been slower afterwards; as shown in Figure 3, most countries lie below the 45° diagonal, evidencing a lower progression during the 8 years that followed 2000 than the 5 years that preceded it. A year to year average would even accentuate this difference. Most of the economies that stand above the 45° line are located in Asia, with the exception of Denmark, Portugal and Saudi Arabia.

Figure 3 GVC participation index, 1995-2008: 45° scatter plot

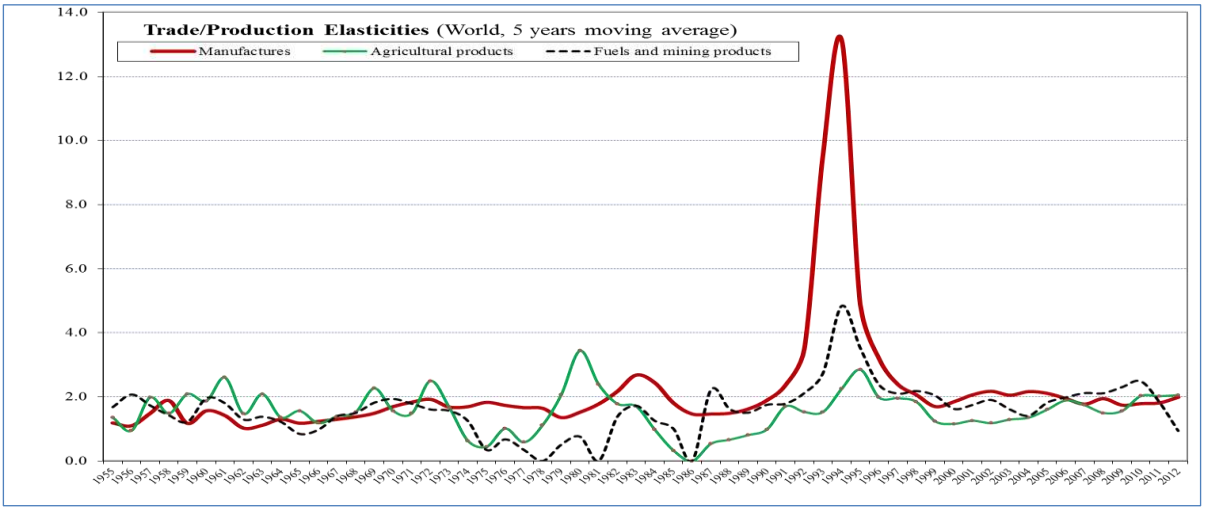


Note: The horizontal axis shows the evolution (in percentage points) of the index between 1995 and 2000 while the horizontal axis indicates the change between 2000 and 2008. Points lying on the 45° line indicate similar rates of growth. The size of the bubble refers to the value of the index in 1995.
Source: Based on OECD-WTO TiVA database.

This result is substantiated by the evolution of World Trade/Output elasticities (Figure 4). The peak period is centred on 1994-1995 for primary and industrial goods for both final and intermediate use, but the case for manufacture production is outstanding. From 1989 to 1994, world exports of manufacture increased 10 times more than the volume of output, a signal that manufacturing was

being fragmented between several different countries and that goods in process of production were now crossing several borders instead of being entirely manufactured locally. The new global production network spread relatively rapidly and elasticity returned to normal in the 2000s. After this date, the new global production model has stabilised at a much higher Trade/Output level and the growth of world trade tends to grow twice as much as the volume of world production. Only some countries have continued to engage into more diversified global value chains: those that stand above the 45° line in Figure 3, among others. [4]

Figure 4 Long term evolution of the Trade-Production elasticity, 1955-2012



Note: Rolling ratio of the five year growth rate in volume of Exports over growth of Production; trade and output include both intermediate and final products. The graph points correspond to the last year of the period, for example, 1955 refers to the 1950-55 trade/output elasticity.
Source: Elaborated on the basis of WTO ITS database.

2.2 Does size matter?

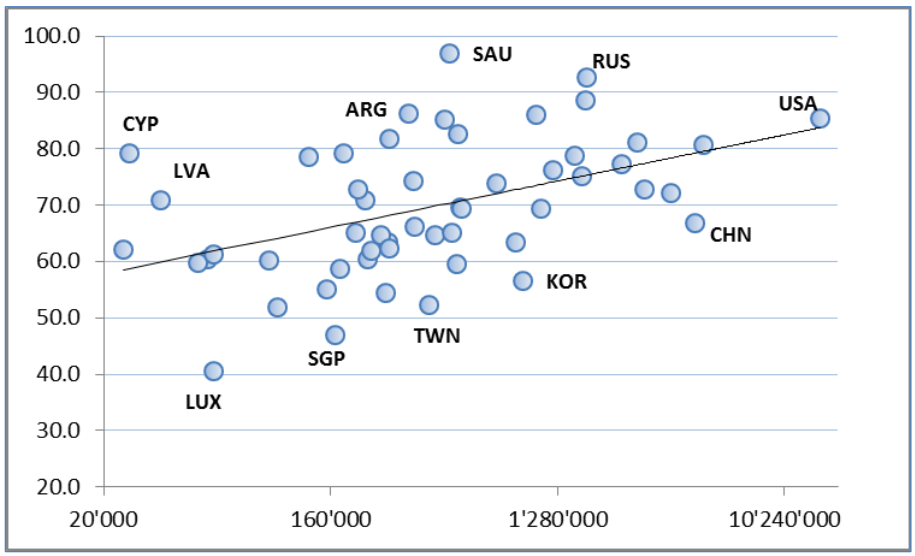
Upstreamness or downstreamness are not the sole factors affecting vertical specialization and the share of imported content in exports. United Kingdom is a service oriented economy but ranks just after South Africa in terms of domestic value-added content in her exports. Intuitively, the size of an economy is also an important factor: *ceteris paribus*, it will be much easier to find an adequate supplier of inputs in a large economy than in a small one. Reasoning *ad absurdum*, it is clear that the hypothesis that 'size does not matter' cannot hold. If all world economies but a small one were to confederate, the value of the large partner's total exports in commercial (customs) value would be very close to its trade-in-value added content, as all but a tiny part of the large confederation exports would be home based. [5]

Using firm-level surveys to analyze the relationship between the use of imported inputs by firms and country size, Amin and Islam (2014) determine that small countries rely disproportionately more on imported inputs than large countries do. Under the hypothesis of homogeneity of output for domestic and export use, the vertical specialization index of small countries should be higher, and the domestic value-added content should be lower. A first run at our data does not fully confirm this result. Crossing the domestic VA content of exports against size, proxied by the logarithm of GDP provides only with a loose fit (R^2=0.2).

[4] The calculation of trade/output elasticity in the WTO's ITS database builds on a larger number of countries than the *TiVA* database.

[5] The only possibility for a different outcome would be for the confederation's exports to be based almost exclusively on intermediate imports from the tiny Rest of the World country. A possible assumption, but hardly a plausible one.

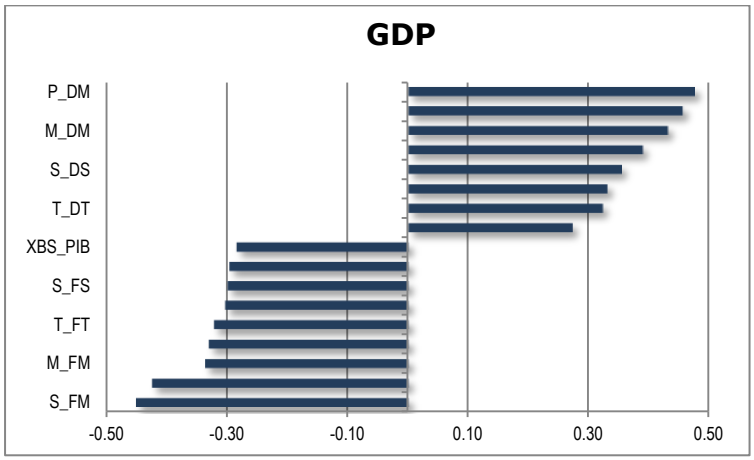
Figure 5 Trade in Value-Added: domestic content and GDP size, 2008



Note: Horizontal axis: GDP in million USD (log scale); vertical axis: domestic value-added content in total exports (%). The trend line is a log-curve but appears as a straight line due to the rescaling of the horizontal axis.
Sources: OECD-WTO's *TiVA* and World Bank's *WDI* databases.

Actually, GDP is fairly well correlated with a series of trade in value-added variables, as seen in Figure 2. On the negative correlation side, we find the total foreign content in total gross exports (T_FT) with an even stronger negative correlation for foreign manufacture value-added (VA) in services and manufacture exports (S_FM and M_FM, respectively). At the other side of the spectrum and as expected, one finds the total domestic content in total gross exports (T_DT), with the highest correlation coefficient when it comes to domestic manufacture VA in primary and manufacture exports (P_DM and M_DM).

Figure 6 Correlation of economic size with a selection of trade and *structural variables, 2008*

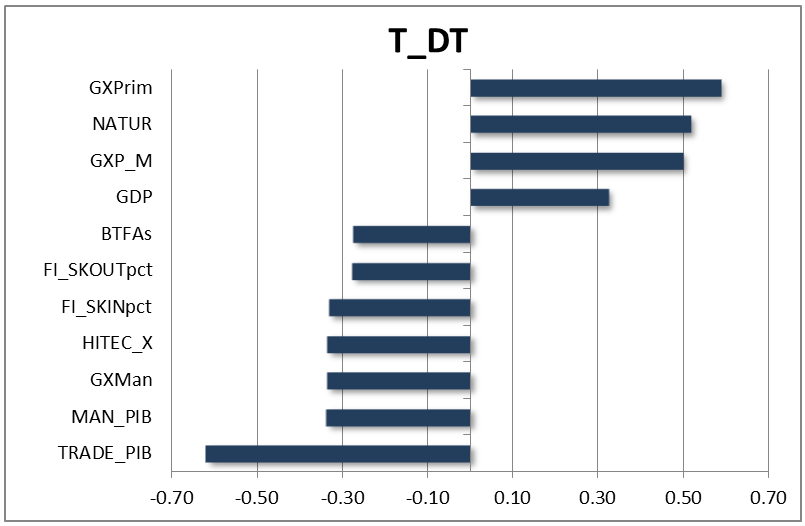


Notes: All coefficients are statistically different from 0 with a significance level alpha=0.05. For a dictionary of variables used in the analysis, see Annex 1.
Sources: Based on OECD-WTO's *TiVA* and World Bank's WDI databases.

Yet, even if pairwise correlations are highly significant, they are not very strong; the absolute value of the highest and lowest coefficients is lower than 0.5. While total GDP (size) is a potentially relevant variable in Figure 7 (R=0.32, significant at 1-alpha=0.95), it is not the most determinant one for determining domestic VA content in total exports. Natural resources endowment and a comparative advantage in primary products are much more influential than

economic size. Even if one may argue than economic size is negatively correlated (-0.30) with the trade coefficient (Trade_PIB), the high absolute value found in the (negative) correlation of the trade coefficient indicates that other qualitative factors — such as the degree of insertion in value chains — determine this result, rather than sheer economic size.

Figure 7 Correlation of total domestic value-added content in gross exports with GDP and a selection of other variables*, 2008*



Notes: All coefficients are statistically different from 0 with a significance level alpha=0.05. For a dictionary of variables used in the analysis, see Annex 1.
Sources: Based on OECD-WTO's *TiVA* and World Bank's WDI databases

3. MAPPING COUNTRIES ACCORDING TO THEIR TIVA PARAMETERS

The next move consists in applying Exploratory Data Analysis (EDA) techniques so as to analyse in more details how the countries are distributed in relation to the set of economic and trade indicators. To do so, we build a database composed of series of *a priori* relevant variables, adding to the set of TiVA variables a series of indicators providing information on the structural properties of the domestic economies, their trade policy or their attractiveness to foreign investors. In the following sections, we shall refer to these variables as part of the following three variable spaces: (i) TiVA, (ii) Structural, (iii) Trade Policy.

3.1 Initial mapping of the observations

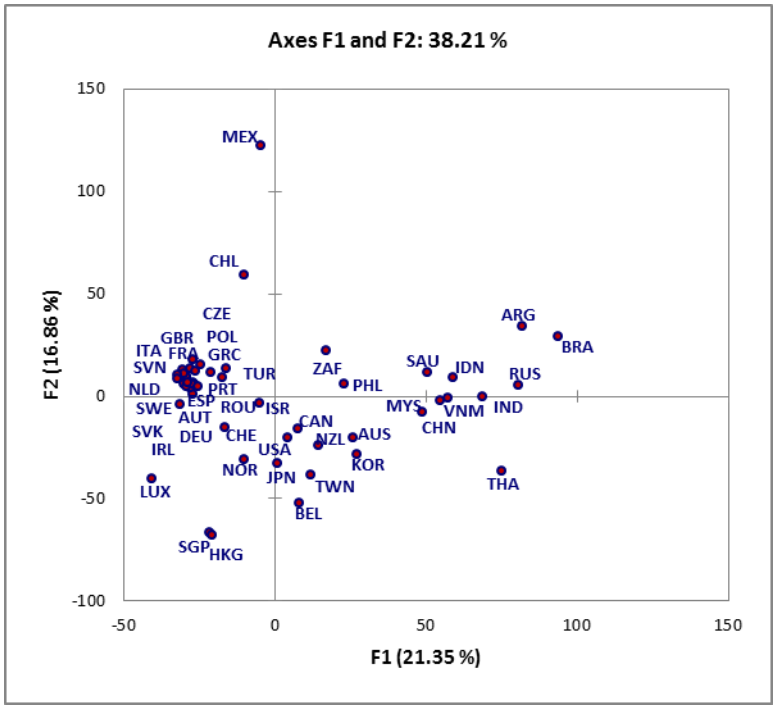
The first step of the EDA is a preliminary exploration on the data structure, applying principal component analysis (PCA) to the entire data set (172 variables in total). PCA projects observations (economies), thereby reducing a p-dimensional space (p initial variables) to a lower dimensional space while preserving as much information as possible. It is particularly apt at dealing with multi-criteria analysis with lot of collinearity. Actually, PCA reduces the numerous initial dimensions of the dataset to a few ones, putting all-together the most correlated variables and identifying new uncorrelated ones (principal components) that capture most of the information (*id est*, the variance) while being uncorrelated with other principal components.

There are, in theory, as many principal components as variables, but some components explain a much larger share of the total variance than others. The best situation is when the first two or three components "explain" about 80% of total variance; the worst case is when no component has better explanatory power than other ones (a totally randomly distributed dataset).

Our present case falls in-between. The first two principal components explain less than 40% of the variance. By including two other components, barely 52% of total information is accounted for. We would have to include 11 dimensions to "explain" 80% of the data-set variance. In a few words,

this is not a case for clear-cut analysis and we will probably need to look at multiple influences. Most probably, individual country's specificities, not reducible to a fixed combination of the selection of structural or policy variables included in the exploratory analysis, tend to explain a large share of the variance.

Figure 8 Initial Exploratory Analysis: Biplot of countries, first and second Principal Components (2008)



Note: PCA reduces a p-multiple dimensional space (p: number of initial variables, 172 in the present case) to a lower dimensional space, correlated with the initial dimensions (see Table 1) while preserving as much information (or variance) as possible. Here, the two components represent 38% of total variance. Some labels in the North-West panel corresponding to EU countries have been deleted or moved to improve graph readability.
Sources: See Annex 1.

Figure 8 presents the results of the projection of observations (countries) according to the first two principal components; to facilitate the interpretation of these two factorial axis, Table 1 displays the main correlations between the two first axis and the variables.

According to the table, size does not matter substantially: GDP does not qualify for the selection criteria (|R|>0.5); indeed its correlation with the first and second axis is close to zero (0.09 and 0.003, respectively). Per capita GDP is more significant, even if it did not pass the 0.5 "test": its correlation is -0.49 for the F1 (richer countries being at the left hand side, poorer on the right of the graph) and 0.34 for F2.

If we look at the variables ordering the horizontal axis F1, we note that, on the right (East) side, we find a strong correlation with high effective protection (at MFN), high share of agriculture in GDP and low imports of foreign services in primary production (i.e., low vertical specialization for the primary activities). [6] On the left hand (West) side of the horizontal axis, we find countries that signed relatively more RTAs, inducing significant drop in effective protection, and have a high share of services in their GDP. This general pattern is nevertheless fuzzy; there is no clear-cut ordering of the observations according to these criteria as the first axis "explains" only 20% of the total variance. Roughly, we expect to find highly-connected services oriented economies on the left-hand side and natural-resources rich ones on the right.

[6] We should keep in mind that the negative-positive, left-right or East-West orientation of the variables is interchangeable in a PCA; therefore, there is no normative ordering of the observations.

Table 1 Correlation of selected variables with the first two factorial components *(|R|>0.5)*

Variable	F1	F2	Variable	F1	F2	Variable	F1	F2
AEPR012	0.86	0.34	EPro006	0.68	0.29	NP018	0.45	0.60
NP012	0.86	0.37	S_DP	0.68	0.13	EPro011	0.60	0.34
EPro012	0.85	0.27	AEPR009_dP	-0.68	0.59	EPro013	0.60	0.34
BTFAs	-0.79	0.32	AEPR015	0.66	0.39	AEPR011_dP	-0.15	0.60
NP020	0.79	0.34	EPro009_dP	-0.66	0.60	EPro018	0.32	0.59
NP010	0.78	0.38	EPro008_dP	-0.65	0.65	AEPR004_dP	-0.55	0.59
AEPR010	0.78	0.40	NP011	0.65	0.39	AEPR008	0.34	0.59
NP015	0.77	0.47	NP015_dP	-0.62	0.65	AEPR017_dP	-0.32	0.59
NP006	0.77	0.33	EPro017_dP	-0.27	0.65	NP019_dP	-0.58	0.40
AEPR018_dP	-0.09	0.76	NP017_dP	-0.44	0.64	NP008	0.50	0.58
AEPR020	0.75	0.30	EPro003_dP	-0.64	0.22	EPro010_dP	-0.39	0.58
NP018_dP	-0.11	0.75	EPro020	0.64	0.23	EPro011_dP	-0.15	0.58
AEPR006	0.75	0.33	NP007_dP	-0.64	0.41	EPro009	0.58	0.45
EPro018_dP	-0.11	0.75	EPro015_dP	-0.50	0.64	NP005	0.57	0.49
EPro010	0.74	0.24	AEPR011	0.64	0.36	NP005_dP	-0.57	0.48
NP008_dP	-0.74	0.56	AGR_PIB	0.63	0.15	EPro004_dP	-0.54	0.57
NP016_dP	-0.73	0.41	NP011_dP	-0.22	0.63	AEPR013	0.57	0.42
NP013_dP	-0.55	0.72	AEPR018	0.36	0.63	P_FP	-0.27	-0.56
NP012_dP	-0.56	0.72	AEPR010_dP	-0.37	0.63	XBS_PIB	-0.22	-0.55
NP009_dP	-0.71	0.60	EPro008	0.35	0.62	NATUR	0.55	0.17
AEPR013_dP	-0.53	0.71	EPro016_dP	-0.62	0.28	EPro007_dP	-0.55	0.35
EPro012_dP	-0.53	0.71	M_DP	0.62	0.24	TRADE_PIB	-0.25	-0.55
NP003_dP	-0.70	0.33	S_DM	0.62	0.05	AEPR009	0.55	0.50
NP013	0.70	0.48	NP010_dP	-0.41	0.62	MBS_PIB	-0.28	-0.52
EPro013_dP	-0.47	0.69	AEPR015_dP	-0.56	0.62	FI_SKOUTpct	-0.29	-0.52
AEPR012_dP	-0.56	0.69	NP017	0.61	0.47	AEPR005_dP	-0.51	0.51
NP009	0.69	0.49	EPro015	0.61	0.46	T_DP	0.51	0.23
AEPR008_dP	-0.69	0.61	AEPR007_dP	-0.61	0.43	AEPR019_dP	-0.51	0.14
AEPR003_dP	-0.69	0.34	NP004_dP	-0.61	0.58	P_DP	0.39	0.51
AEPR016_dP	-0.68	0.38	SER_PIB	-0.60	-0.25	EPro005_dP	-0.49	0.51

Notes: Variables are selected when their correlation coefficient with F1 or F2 is greater than or equal to 0.5 in absolute value; all values are for 2008, see the dictionary of variables used in the analysis in Annex 1.

F2 is more correlated with trade policy variables, in particular the incidence of PTAs on nominal protections (tariff variables ending in "_dP"). Here, we expect to find on the upper panel countries with relatively high level of MFN tariffs but who entered into preferential trade agreements. At the other end of F2, we expect to find open countries on an MFN basis, with low natural resources base.

When looking at the countries (Figure 9), the first visual impressions are:

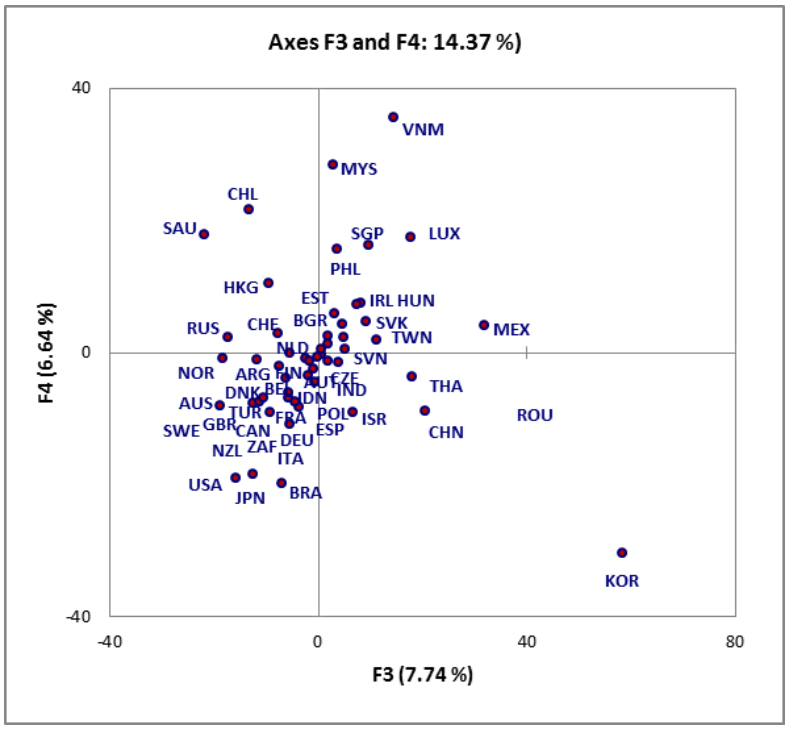
- the tight clustering of European economies observed in the lower North-East quarter, and
- an opposition between services oriented economies and natural resources exporters.

On the far left of the graph, we find services oriented Luxemburg, Singapore and Hong-Kong, while large Latin American developing countries stand on the right side. [7] But the distinction is not clear-cut: Norway – a resources rich country – stands on the left hand panel, reflecting her high per capita income. The other striking point – the tight clustering of most European countries (their contiguity on the graph shows similarity for the two principal components) – extends beyond the European borders, with the proximity of Near East countries such as Israel and Turkey. On the contrary, two European countries, Belgium and Luxemburg, stand somewhat apart from the EU27 block. Luxemburg is highly specialised in financial services exports and Belgium stands apart of other EU countries for – inter alia – the low incidence of domestic value-added in the exports of natural resources-based products.

If we wish to find a dimension where sheer economic size, measured as GDP, is determining, we have to go as far as the fourth principal component, which "explains" only 7% of total variance (Figure 9 and Table 2). Even in this case, GDP is not very significant for this component, as its correlation with the negative side of the axis is only 0.43 (larger economies being located in the lower part of the graph). The relative unimportance of economic size on production indicators, albeit unintuitive, has been documented in Rose (2006).

[7] Chile and Mexico – resources-rich countries that are relatively more inserted in GVCs and RTAs than the other large Latin American countries – stand closer to the vertical axis of the graph. Mexico appears as an outlier for the F2 axis (high MFN protection but strong incidence of RTAs).

Figure 9 Observations according to the third and fourth Principal Components (2008)



Note: Some observation labels close to the gravity centre have been deleted or moved to enhance graph readability.

Actually, the horizontal axis (F3) tends to distribute countries according to their GVC insertion, as measured by the foreign content in their exports (T_FT) on the left-hand side, vs. a higher reliance on domestic content (T_DT) on the right-hand side. On the right side, we also find countries with relatively high level of tariff protection in non-agricultural (NAMA) products. High level of MFN protection in agriculture is associated, on the contrary, with the left hand side of the graph. The vertical dimension, besides the economic-size aspects already mentioned, corresponds to economies with relatively high trade intensity and protection of the agricultural sector in the bottom-half of the graph, vs. NAMA protection in the upper part. But it is probable that Korea weighted disproportionally in the definition of the F4 dimension. [8]

Table 2 Correlation of variables with the third and fourth factorial components *(|R|>0.4)*

Variable	F3	F4	Variable	F3	F4	Variable	F3	F4
T_FT	0.70	0.34	M_FS	0.45	0.54	AEPR002	-0.10	0.46
T_DT	-0.70	-0.34	XBS_PIB	0.25	0.53	AEPR003	-0.45	0.26
EPro005	-0.67	0.34	P_DM	0.20	-0.53	P_FP	0.45	0.09
AEPR005	-0.66	0.32	TRADE_PIB	0.28	0.53	EPro019	0.07	0.44
NP001	0.65	-0.38	T_DM	0.19	-0.52	GDP	-0.19	-0.43
AEPR001	0.65	-0.39	M_DS	-0.52	-0.50	NP002	0.01	0.43
EPro001	0.65	-0.38	EPro003	-0.51	0.30	NP019_dP	-0.03	0.43
NP001_dP	0.63	-0.41	MBS_PIB	0.29	0.51	GXMan	0.43	-0.32
AEPR001_dP	0.62	-0.41	S_FM	0.28	0.50	GXPrim	-0.43	0.29
EPro001_dP	0.61	-0.41	S_DS	-0.49	-0.43	S_FS	0.28	0.41
T_FM	0.61	0.28	M_DM	-0.42	-0.48	M_FP	0.41	-0.06
M_FM	0.57	0.50	EPro002	-0.09	0.48	EPro017	0.41	-0.34
T_FP	0.56	-0.17	NP019	0.21	0.47	AEPR019	-0.01	0.41
S_FP	0.56	0.11	EPro014_dP	-0.47	-0.24	CONS_PIB	-0.16	-0.40
MAN_PIB	0.54	-0.08	T_FS	0.47	0.41			

Notes: Variables are selected when their correlation coefficient with F1 or F2 is greater or equal to 0.4 in absolute value; all values for 2008, see the dictionary of variables used in the analysis in Annex 1.

[8] A more thorough analysis would call for considering Korea as an outlier and remove her from the sample. But the low significance of F4 and the small size of the original sample do not support deeper an analysis, as any result is expected to be statistically fragile and sample-dependent.

3.2 Clustering economies according to their economic and trade profiles

Further exploratory data analysis is now performed in order to define with more accuracy the patterns that may help mapping the various dimensions lying behind the observed variations in trade in value-added. The techniques that are implemented aim at identifying clusters of countries according to their specificities. The first procedure adopted is Agglomerative Hierarchical Clustering, an iterative EDA technique used to build "homogeneous groups" of observations on the basis of their characteristics as given by a set of variables. The agglomerative approach successively unites pairs of individual observations and then sub-sets of observations, according to their similarities. Starting from as many clusters as observations in the sample, it ends up with merging all individual observations into a single class. Where to truncate the resulting tree between these two extremes for defining an optimal number of clusters can be determined by a combination of parametric methods building on variance decomposition and – as often in EDA – expert's judgement.

The method builds on a matrix describing the similarity or dissimilarity between the observations. They are successively applied to each of the three variable-spaces (i) Structural, (ii) TiVA and (iii) Trade Policy dimensions.

3.2.1 In relation to *TiVA* variables

Clustering analysis is applied to the sample of countries, taking into consideration for building the similarity matrix only the sub-set of TiVA variables, as defined in Annex 1. Using hierarchical clustering, the number of clusters was set *a priori* to 5 so as to obtain enough details. [9] Clustering results are always tentative and each one of the aggregative method has its strengths and weaknesses. We use Complete Linkage and Ward's Linkage to test the robustness of groups. Complete linkage, a hierarchical clustering method similar to average linkage, is less susceptible to be affected by random noise and the presence of outliers, but it can unnecessarily break large clusters as it favours compact shapes. Ward's agglomerative hierarchical clustering procedure method attempts to minimize the sum of the square distances of points from their cluster centroid and favours dense clusters.

Table 3 Hierarchical clustering of observations according to *TiVA* variables

Ward's method

Class	1 (CAN)	2 (SWE)	3 (SVN)	4 (KOR)	5 (SGP)
Objects	17	21	8	4	3
Within-class variance	1267.43	584.83	375.96	612.24	821.34
	ARG	AUT	BGR	CHN	IRL
	AUS	BEL	CZE	KOR	LUX
	BRA	CHE	EST	THA	SGP
	CAN	CYP	HUN	TWN	
	CHL	DEU	MYS		
	GBR	DNK	SVK		
	IDN	ESP	SVN		
	IND	FIN	VNM		
	JPN	FRA			
	MEX	GRC			
	NOR	HKG			
	NZL	ISR			
	RUS	ITA			
	SAU	LTU			
	TUR	LVA			
	USA	NLD			
	ZAF	PHL			
		POL			
		PRT			
		ROU			
		SWE			

Complete linkages

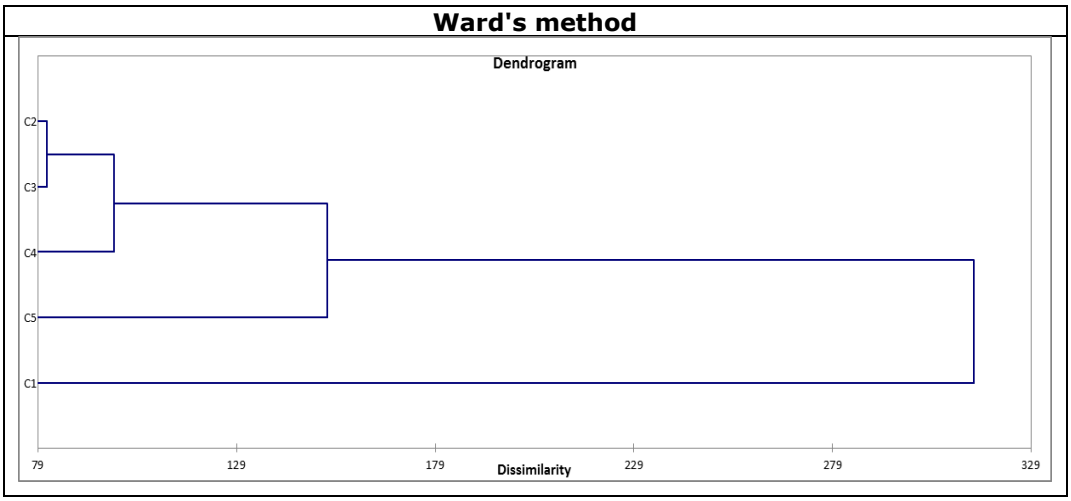
Class	1 (ESP)		2 (SVN)	3 (THA)	4 (IRL)	5 (SAU)
Objects	28		19	3	2	1
Within-class variance	930.16		472.70	859.02	1038.32	0.00
	ARG	IND	AUT	CHN	IRL	SAU
	AUS	ITA	BEL	IDN	LUX	
	BRA	JPN	BGR	THA		
	CAN	LVA	CZE			
	CHE	MEX	EST			
	CHL	NLD	FIN			
	CYP	NOR	HUN			
	DEU	NZL	ISR			
	DNK	POL	KOR			
	ESP	ROU	LTU			
	FRA	RUS	MYS			
	GBR	TUR	PHL			
	GRC	USA	PRT			
	HKG	ZAF	SGP			
			SVK			
			SVN			
			SWE			
			TWN			
			VNM			

[9] The optimal number of clusters defined on pure statistical grounds for was 4 after merging the two closest groups obtained as shown in the dendograms that illustrate the hierarchical tree in **Error! Reference source not found.**. We choose to keep the five clusters for illustrative purpose.

Groupings are first constituted according to Ward's method, less prone to be influenced by scale effect. The within-class variance provides an indication on the compactness of each cluster but should be evaluated in relation to the number of objects belonging to the cluster. For instance, the 2nd cluster exhibits more within-class variance than the 3rd but can still be considered as a more coherent construct considering that it includes almost three times as many members.

Group 5, centred on Singapore, hosts small and open service-oriented economies. Group 4 includes East-Asian developing economies, well inserted in international supply chains at the example of its most central 3 economy, Korea. Eastern European countries that form group 3 are also well inserted in EU supply chains; the presence of Vietnam in this group being somewhat surprising as it shares little with them, besides having also been part of the Soviet bloc. Group 2 is very close to Group 3 (see Figure 10**Error! Reference source not found.**) and gathers most other European countries, plus Hong Kong and Israel. Group 1 is a rather loose cluster (within-class variance is at its highest) which includes all remaining countries. Rather surprisingly, European countries such as Austria and UK are included here, rather than in Group 2 (the presence of Norway, an European oil-exporting country, is more understandable).

Figure 10 *TiVA* Variables: Simplified clustering dendogram



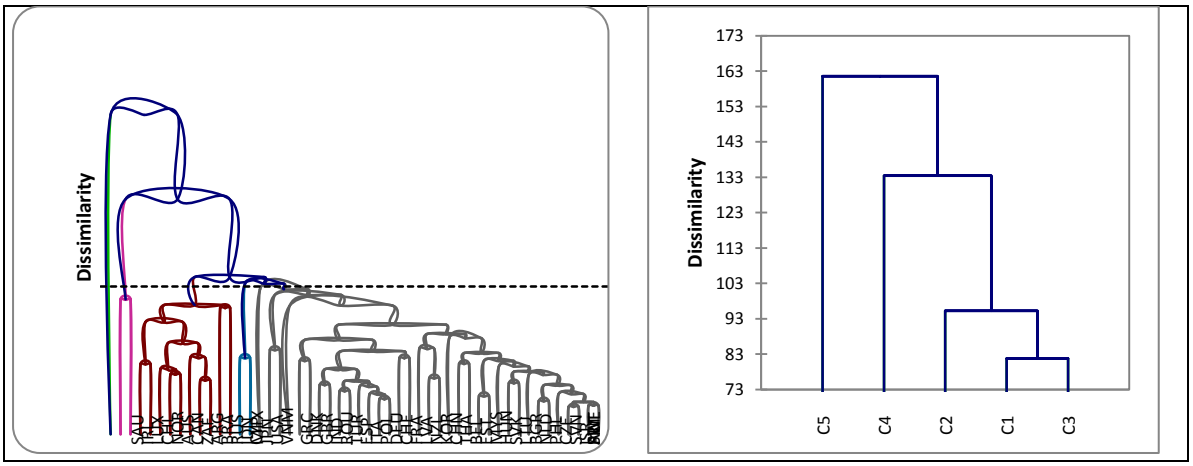
Note: Based on Ward's method.

But, as mentioned previously, this story-line is somewhat contingent to the choice of clustering method and the analysis should focus on the most robust clusters. To illustrate this, clustering according to the complete linkage method partially reshuffles the cards. Only Ireland and Luxembourg and, to a lesser extent, China and Thailand, keep on projecting a clear identity on their cluster. Saudi Arabia, which was before associated with other natural resources rich countries such as Russia or South Africa, appears now as a clear outlier. The two largest groups (1 and 2) are rather fuzzy and amorphous.

Another method (average linkage, Table 4) that provides a more balanced within class variance among clusters would point to three distinct groups on the extremes sides of the dendrogram (commodity exporters, manufacture exporters and small open economies). These groups have the lowest within-class variance. In contrast, a fuzzier group (Cyprus, Denmark, Greece, Japan, USA, Vietnam) stands in the middle of the spectrum close to the group 3 of service oriented "post-industrial" economies. The outlier corresponds to commodity-rich Saudi Arabia.

To sum up, the hierarchical clustering according to TiVA variables reveals a contrast between countries whose TiVA pattern is well identified and the others. The formers find themselves in the same group whatever the method employed, whereas the classification of the latter depends on the type of hierarchical clustering. Among the well-identified TiVA patterns, are the one at the extremes sides of the T_DT spectrum (i) manufacturing economies, (ii) primary good producers, and (iii) small open economies. The classification of other emerging countries and service oriented economies into one of these identified patterns or as outliers depends on the method employed.

Figure 11: Hierarchical dendrogram, average linkage



Note: Average linkage method.

Table 4 Alternative hierarchical clustering according to TiVA variables, average linkage (2008)

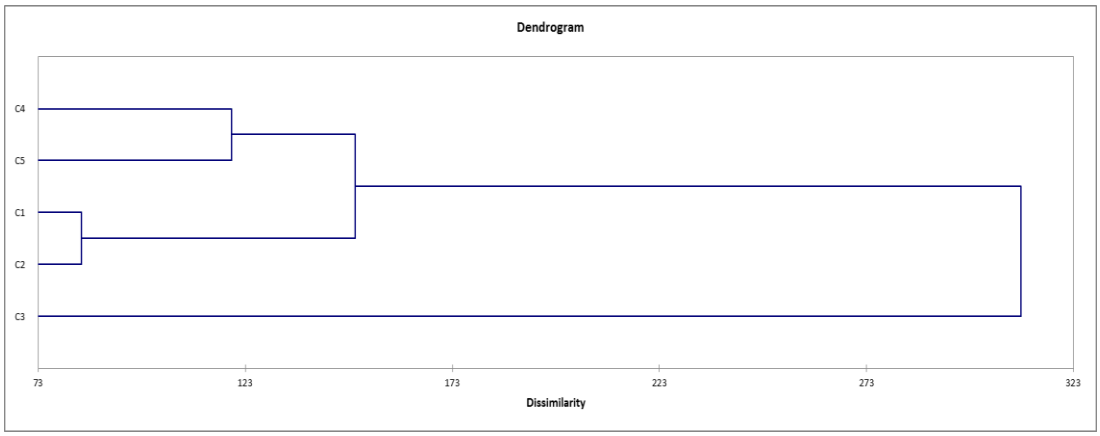
Average Linkage

Class	1 (AUT)	2 (ARG)	3 (CHE)	4 (IRL)	5 (SAU)
Objects	20	10	19	3	1
Within-class variance	401.65	702.70	835.22	821.34	0.00
	AUT	ARG	CHE	IRL	SAU
	BEL	AUS	CYP	LUX	
	BGR	BRA	DEU	SGP	
	CHN	CAN	DNK		
	CZE	CHL	ESP		
	EST	IDN	FRA		
	FIN	MEX	GBR		
	HUN	NOR	GRC		
	ISR	RUS	IND		
	KOR	ZAF	HKG		
	LTU		ITA		
	MYS		JPN		
	NLD		LVA		
	PHL		NZL		
	PRT		POL		
	SVK		ROU		
	SVN		TUR		
	SWE		USA		
	THA		VNM		
	TWN				

3.2.2 In relation to structural variables

Clustering analysis is now applied to the sub-set of structural economic variables, as defined in Annex 1. Using an optimal criterion for truncating the agglomerative tree leads to select 6 clusters, but Saudi Arabia formed a class by herself, so the final number was reduced to 5. Figure 12 shows the clustering tree, starting with the 5 classes described in Table 5 and converging into a single large cluster containing the entire sample.

Figure 12 Structural Variables: Simplified Clustering Dendogram



Note: Based on Ward's method.

Even after forcing a reduced number of classes, Saudi Arabia remains a class by herself (Table 5), leaving only two large clusters and two smaller ones.

Table 5 Hierarchical clustering of observations according to structural variables

(a) Ward's method

Class	1 (POL)	2 (ESP)	3 (HKG)	4 (IDN)	5 (SAU)
Objects	21	20	4	7	1
Within-class variance	3.10E+23	1.03E+25	3.72E+22	2.55E+24	0.00E+00
	ARG	AUT	BEL	CHN	SAU
	AUS	CHE	HKG	IDN	
	BGR	CYP	LUX	IND	
	BRA	DEU	SGP	MYS	
	CAN	DNK		PHL	
	CHL	ESP		THA	
	CZE	FIN		VNM	
	EST	FRA			
	HUN	GBR			
	LTU	GRC			
	LVA	IRL			
	MEX	ISR			
	NOR	ITA			
	NZL	JPN			
	POL	KOR			
	ROU	NLD			
	RUS	PRT			
	SVK	SWE			
	SVN	TWN			
	TUR	USA			
	ZAF				

(b) Complete linkage

Class	1 (POL)	2 (RUS)	3 (BEL)	4 (HKG)	5 (SAU)
Objects	23	25	2	2	1
Within-class variance	9.31E+23	8.32E+24	1.02E+23	1.38E+21	0.00E+00
	ARG	AUS	BEL	HKG	SAU
	BGR	AUT	LUX	SGP	
	BRA	CAN			
	CHN	CHE			
	CZE	CHL			
	EST	CYP			
	HUN	DEU			
	IDN	DNK			
	IND	ESP			
	LTU	FIN			
	LVA	FRA			
	MEX	GBR			
	MYS	GRC			
	NZL	IRL			
	PHL	ISR			
	POL	ITA			
	ROU	JPN			
	SVK	KOR			
	SVN	NLD			
	THA	NOR			
	TUR	PRT			
	VNM	RUS			
	ZAF	SWE			
		TWN			
		USA			

Note: the large value of variance is due to the inclusion of GDP as one of the variables. Ward's method aggregates two groups so that within-group inertia increases as little as possible to keep the clusters homogeneous; in complete linkages, agglomeration tends to dilate the data space and to generate compact clusters.

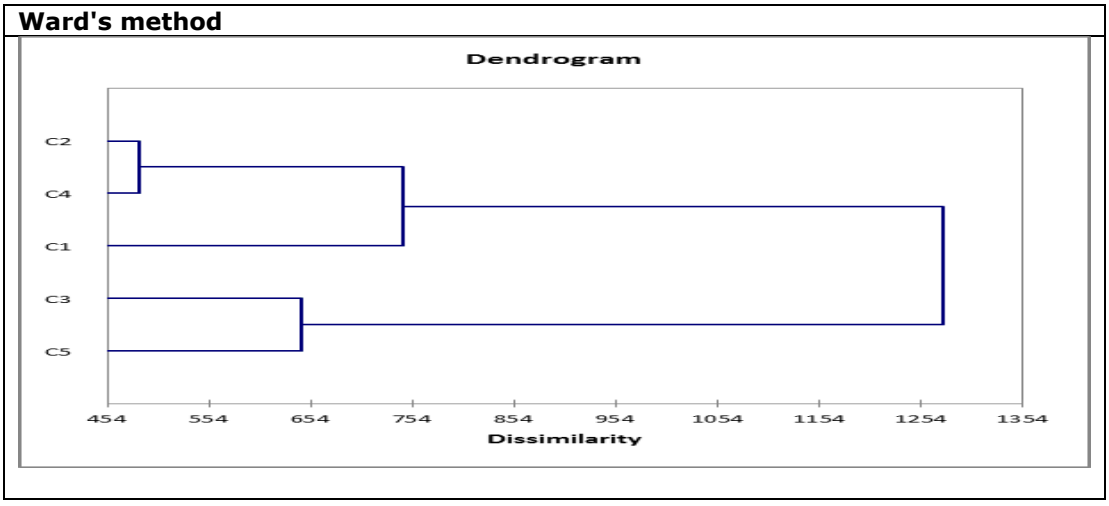
Within the large clusters, the first one (mainly Eastern European and natural resources rich countries) is the most compact on the basis of within-class variance despite including countries of very different economic profile. This anomalous result is due to the scale effect of GDP, which weights disproportionately on the variance. This cluster is distinct from the 2nd one, made mostly of Western European and Asian developed economies plus the USA. A few developing countries join this club of advanced economies; they are mainly small and relatively high income economies such as Chinese Taipei and Israel. The separate clustering of Eastern and Western European economies indicates that economic convergence within the EU region was still far from complete in 2008, on the eve of the global crisis. Cluster 3 is built around services-oriented small economies in Europe and Asia. The 4th cluster is made of Asian developing countries. Indonesia, despite her large endowment in natural resources, is nevertheless classified in this cluster: regional proximity may apparently supersede comparative advantages.

The complete linkage method changes somewhat the perspective while keeping the main characteristics. Saudi Arabia remains an outlier but the cluster of small services-oriented economies is now split in two, differentiating between Asian and European countries. The large Asian developing economies are now reclassified in the first cluster. In this new context, the Russian Federation joins the second cluster of more advanced economies and becomes her central point. Note that letting the algorithm choosing the optimal number of clusters, the USA would be in a single class. Hierarchical clustering is usually performed as the first step of a more detailed data exploration and further analysis should be performed to identify more precisely the variables that determine the closest associations or, on the contrary, isolate the outliers. But our objective here is only to provide a first view of the various facets of the multi-dimensional aspects of the country profiles.

3.2.3 In relation to Trade Policy variables

Similar feature of dense clusters coexisting with outliers is found when analyzing the trade policy space.

Figure 13 Trade Policy variables: Simplified dendogram of clustering



Note: Based on Ward's method.

Group 5 and 4 are outliers (Mexico and Korea, respectively) with respect to their tariff schedules, once all dimensions (nominal MFN, preferences and effective rates) are factored-in. The third and largest group (27 members) gathers most European countries, plus Turkey which is closely associated to this region. Two clusters are loose ones: Cluster 1 should be associated to commodity exporters and cluster 2 to services economies. But the split is not clear-cut: Cluster 1 also includes emerging countries and cluster 2 contains some manufacture exporters. In addition, Group 2 is made of a mix of developed and advanced developing countries while Group 1 (the

most loosely tight cluster) gathers the rest of the observations. Korea, classified as an outlier, is nevertheless close to Group 2 while Mexico shares similarities with Group 3 (Table 6). [10]

The complete linkage method provides additional information. In addition to Mexico and Korea, Chile can also be considered also as an outlier for the specificity of its tariff policy. While all participants to group 2 and some of group 1 merged with group 3 to form a single mega cluster of 42 members, the first cluster, still centred on Indonesia and prominently made of developing countries, confirm its specificity.

Table 6 Hierarchical clustering of observations according to Trade Policy variables

Ward's method

Class	1 (IDN)	2 (CAN)	3 (ESP)		4 (KOR)	5 (MEX)
Objects	12	12	27		1	1
Within-class variance	6405.50	4845.36	1439.16		0	0
	ARG	AUS	AUT	.../...	KOR	MEX
	BRA	BEL	BGR	IRL		
	CHN	CAN	CHL	ITA		
	IDN	CHE	CYP	LTU		
	IND	HKG	CZE	LUX		
	MYS	ISR	DEU	LVA		
	PHL	JPN	DNK	NLD		
	RUS	NOR	ESP	POL		
	SAU	NZL	EST	PRT		
	THA	SGP	FIN	ROU		
	VNM	TWN	FRA	SVK		
	ZAF	USA	GBR	SVN		
			GRC	SWE		
			HUN	TUR		

Complete linkages

Class	1 (IDN)	2 (EST)		3 (CHL)	4 (KOR)	5 (MEX)
Objects	8	42		1	1	1
Within-class variance	6543.53	3759.12		0	0	0
	ARG	AUS	JPN	CHL	KOR	MEX
	BRA	AUT	LTU			
	CHN	BEL	LUX			
	IDN	BGR	LVA			
	IND	CAN	NLD			
	MYS	CHE	NOR			
	RUS	CYP	NZL			
	THA	CZE	PHL			
		DEU	L			
		DNK	PRT			
		ESP	ROU			
		EST	SAU			
		FIN	SGP			
		FRA	SVK			
		GBR	SVN			
		GRC	SWE			
		HKG	TUR			
		HUN	TWN			
		IRL	USA			
		ISR	VNM			
		ITA	ZAF			

3.3 Identifying stable clusters

A clear-cut point of the analysis above is that the clustering method has a strong influence on the composition of clusters. On the other hand, some associations of countries appear more stable irrespective of the methodology used. It is natural, therefore, to expect that these stable clusters are actually built on robust economic characteristics. In order to identify such groupings, the following section investigates the robustness of groupings according to their stability relative to the various clustering methods implemented above.

Table 7 provides details on the most stable groups according to all linkage methods. The analysis will focus on the results obtained when using *TiVA* and Structural Variables databases only because trade policy database tends to be unstable (split, in the table terminology). These groups can be easily categorized along the exports category dimension. Groups 1 and 2 are commodities exporters (their share of commodities exports in total exports is the highest in the sample); groups 3 to 6 belong to manufacture products exporters and groups 7 and 8 display higher services exports.

For the sake of clarity, groups 3 to 6 were split according to their regional location, Europe or Asia. Manufacture exporters and commodities exporters can be further subdivided into developing and developed countries. Interestingly, this distinction by development or income level is no longer decisive when it comes to differentiating service-oriented economies.

[10] While the results do not explicitly indicate the source of similarities, one may draw similarities between the role of regional preferences within NAFTA for Mexico and similar regional arrangements for EU countries.

After gathering each dual group according to their export specificity (and geography, for manufacture exporters), we find more stability according to *TiVA* clustering than for structural variables. The latter constitute a proxy for the level of development and resources endowment. These observations support the idea that countries that export predominantly commodity products may share the same *TiVA* profile as other commodity exporters although their development levels are different. This is also verified for manufacture exporters (with one exception for Asian countries, with *TiVA*, complete linkage clustering). However, this does not apply to service oriented economies, as mentioned above, as if comparative advantages in services could not mature before the economy reached a certain level of structural development.

However, it should be reminded that, in a more general perspective, the level of development still remains a critical determinant of *TiVA* indicators because the likeliness for a country to export predominantly a specific category of products (whether commodities, manufacture or services) depends strongly on this dimension.

Table 7 Selected groups of countries that fall in the same group according to TiVA and structural data clustering.

	Database	**TiVA**			**Structural Data**		**Policy Data**	
Interpretation	*Countries \ Linkage Method*	*Ward's*	*Complete*	*Average*	*Ward's*	*Complete*	*Ward's*	*Complete*
1. Commodities Exporters Developing	ARG BRA CHL MEX ZAF	Stable	Stable	Stable	Stable	Stable	**Split**	**Split**
2. Commodities Exporters Developed	AUS CAN NOR RUS	Stable	Stable	Stable	Stable	Stable	**Split**	**Split**
Sum Commodities exporters [a]	ARG BRA CHL MEX ZAF AUS CAN NOR RUS	Stable	Stable	Stable	Stable	**Split**	**Split**	**Split**
3. Manufacturing EU12	BGR CZE EST HUN SVK SVN	Stable	Stable	Stable	Stable	Stable	Stable	Stable
4. Manufacturing Europe and Near East	AUT FIN ISR SWE	Stable	Stable	Stable	Stable	Stable	**Split***	Stable
Sum Manufacturing European Zone [ab]	BGR CZE EST HUN SVK SVN AUT FIN ISR SWE	Stable	Stable	Stable	**Split**	**Split**	**Split**	Stable
5. Manufacturing Asian developing	CHN THA	Stable	Stable	Stable	Stable	Stable	Stable	Stable
6. Manufacturing Asian developed	KOR TWN	Stable	Stable	Stable	Stable	Stable	**Split**	**Split**
Sum Manufacturing Asian [a]	CHN THA KOR TWN	Stable	**Split**	Stable	**Split**	**Split**	**Split**	**Split**
7. Large or service oriented European	CHE CYP DEU DNK ESP FRA GRC ITA	Stable	Stable	Stable	Stable	Stable	**Split**	Stable
8. Other large or service oriented	JPN USA	Stable	Stable	Stable	Stable	Stable	**Split**	Stable
Sum Service oriented economies [a]	CHE CYP DEU DNK ESP FRA GRC ITA JPN USA	**Split**	Stable	Stable	Stable	Stable	**Split**	Stable

Notes: [a] The sum of two stable sub-groups can be split if the two sub-groups do not belong to the same cluster.
[b] Israel is not part of Europe but was added to indicate strong similarity with some EU economies.

Table 8 is similar to Table 7 but further aggregates the results by extending the stability criterion to 2 clusters. This means that a pre-defined group of countries is still considered as stable even if the member countries are allocated into 2 different clusters for each hierarchical clustering method. According to the table, the most striking outcome is the stability of EU27 countries (excepting Belgium, Luxemburg and Ireland, which belong to small open economies, and Great Britain, which stands out of the group on *TiVA* data clustering using Ward's Linkage). Those

countries stand together whatever the clustering method employed, no matter what variable dimension is used, even the trade policy data.

Table 8 also shows that, like Europe, most Eastern Asian economies demonstrate a high degree of unity (although 4 countries in the sample had to be removed for being outliers respective to the cluster). However, those economies do not belong to the same RTAs, which explains the splitting into 3 or more clusters according to the trade policy criterion. Also worth noticing is the stability of the group of commodities exporters to *TiVA* variables and structural data, although the group splits on the trade policy criterion, evidencing the difference in trade policies between developed and developing commodities exporters.

Comparing the results with Table 7, one notes that the abundance of natural resources clusters non-EU developed economies with developing or transition countries. In other terms, natural resources endowment remains a key marker of *TiVA* profiles, even at different level of economic development.

Table 8 Selected group of countries and their relative stability according to the database and linkage method; 2 clusters

With consideration of 2 clusters	Database	TiVA			Structural Data		Policy Data	
	Countries \ Linkage Method	Ward's	Complete	Average	Ward's	Complete	Ward's	Complete
Commodities Exporters	ARG AUS BRA CHL MEX ZAF CAN NOR RUS IDN	Stable	Stable	Stable	Stable	Stable	**Split**	**Split**
Small Open Economies	LUX SGP IRL BEL	Stable	Stable	Stable	Stable	**Split**	Stable	Stable
Most EU members	EU [a]	Stable	Stable	Stable	Stable	Stable	Stable	Stable
Most East Asian	EAST ASIA [b]	Stable	Stable	Stable	Stable	Stable	**Split**	**Split**

Notes: a Excludes Ireland, Belgium, Luxemburg (small open economies) and Great-Britain
b In the sample. Excludes Japan, Hong Kong, the Philippines and Indonesia (commodity exporter)

4 THE DRIVERS OF *TIVA*: EXPLORATION OF THE VARIABLE SPACE [11]

The specialization of countries according to their economic and trade profile responds most probably to complex latent factors. The previous sections showed that capturing these specificities is elusive yet possible. Turning the table, the next step was to look at the variable spaces, instead of the countries themselves (the observations). Indeed, many EDA techniques can be applied either on the observations sample or on the variable space. The first step here is to look at similarities, highlighting pair-wise associations between *TiVA* variables and other structural or trade policy dimensions.

4.1 Identifying similarities

The hierarchical clustering process starts by calculating the dissimilarity between the objects to be clustered (here, variables). It is then relatively straightforward to derive a similarity (or associativity) index from the dissimilarity indicators generated during the data processing. [12] Table 9 shows the main bipolar associations (retaining only pairs of variables when the similarity index is equal to or higher than 40) between *TiVA* variables on the one hand and structural and trade policy variables on the other hand.

[11] The term "Drivers" is used here as a short-hand for "Underlying factors" and does not imply any causal relationship; as mentioned, EDA stops at highlighting associations and does not intent to test causalities.

[12] The first stage consists in deriving a dissimilarity index from the observed distances by normalizing them with respect to both the minimum and maximum values observed for the full sample (DI= 100.(Obs-Min)/(Max-Min), then to take the complement to 100 in order to obtain a similarity (associativity) indicator instead of a distance.

Table 9 *TiVA* variables: Main dipole associations with other (structural and trade policy) variables

TiVA	Other	Similarity	*TiVA*	Other	Similarity	*TiVA*	Other	Similarity	*TiVA*	Other	Similarity
M_DP	NATUR	66.0	M_FM	XBS_PIB	54.7	M_FP	MBS_PIB	45.8	P_FM	SER_PIB	42.0
T_DM	MAN_PIB	63.4	M_DP	NP020	54.2	P_FS	SER_PIB	45.1	P_DS	RD_PIB	41.9
S_FS	XBS_PIB	63.1	M_DP	NP015	53.5	M_FP	TRADE_PIB	45.1	M_DP	NP004	41.8
T_FS	MBS_PIB	62.7	P_DS	SER_PIB	52.9	M_DS	SER_PIB	44.9	P_DM	RD_PIB	41.7
T_FS	TRADE_PIB	62.4	P_FM	MBS_PIB	52.8	P_FP	RD_PIB	44.8	T_FM	TRADE_PIB	41.7
S_FS	TRADE_PIB	62.1	T_FM	MAN_PIB	52.8	M_FM	NP019_dP	44.7	M_FP	NP001_dP	41.6
T_FS	XBS_PIB	61.7	M_DP	NP013	52.2	P_FP	SER_PIB	44.7	P_FP	SHIP	41.5
S_FS	MBS_PIB	61.0	P_FM	TRADE_PIB	52.0	P_FS	PERCAP	44.5	T_DP	XBAL_PIB	41.3
M_FS	MBS_PIB	59.8	M_DP	NP012	51.5	S_FM	TRADE_PIB	44.2	M_DM	NP004	41.1
T_FT	MBS_PIB	58.5	M_DP	NP011	51.2	P_DM	SHIP	44.0	M_DM	NP015	41.0
M_FS	TRADE_PIB	58.4	M_DP	NP017	50.8	T_FT	MAN_PIB	43.6	P_FP	URB	41.0
T_FT	TRADE_PIB	57.9	P_FM	XBS_PIB	50.7	T_FP	MAN_PIB	43.5	M_DM	NP020	40.7
M_FS	XBS_PIB	56.8	P_DM	MAN_PIB	50.3	M_FP	XBS_PIB	43.3	M_FM	NP008_dP	40.6
S_DM	MAN_PIB	56.8	M_DP	NP006	49.7	M_DP	NP002	43.2	M_DM	NP005	40.6
P_FS	MBS_PIB	56.7	S_DP	MAN_PIB	47.8	P_DS	URB	43.0	M_FS	NP008_dP	40.6
P_FP	TRADE_PIB	56.7	S_FP	MAN_PIB	47.5	M_FP	SHIP	42.6	P_FM	PERCAP	40.5
P_FS	TRADE_PIB	56.1	S_FS	XBAL_PIB	47.4	M_FP	NP001	42.4	M_DM	NP008	40.4
T_FT	XBS_PIB	56.1	M_DP	NP010	46.8	S_FM	XBS_PIB	42.4	T_FM	XBS_PIB	40.4
M_FM	MBS_PIB	55.9	M_DP	NP014	46.7	M_FS	NP019_dP	42.1	MAN_PIB	NP020	40.3
M_FM	TRADE_PIB	55.8	M_DP	NP009	46.5	T_FM	MBS_PIB	42.1	S_FS	SER_PIB	40.3
P_FP	XBS_PIB	55.4	S_FM	MBS_PIB	46.2	M_DP	NP018	42.0	M_FM	NP005_dP	40.2
P_FP	MBS_PIB	55.3	M_FS	NP007_dP	45.8	M_DM	NATUR	42.0	M_FS	NP003_dP	40.1
P_FS	XBS_PIB	55.2									

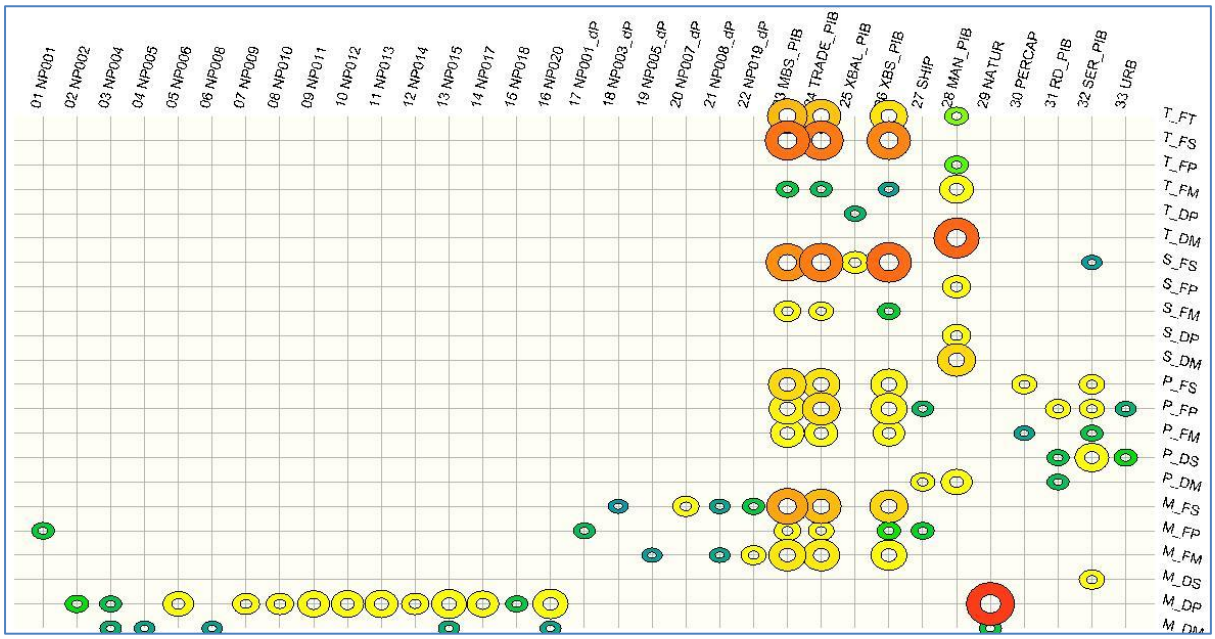
Note: The degree of association is based on a 0-100 index derived from the Euclidian distance between variables used in hierarchical clustering. Only cases were the index is at least equal to 40 were selected.

Figure 14 helps visualizing those results. The closest association is between natural resources endowments (NATUR) and the domestic primary content in manufacture exports (M_DP). This *TiVA* variable is also associated with a number of trade policy variables (variables 01 to 16, on the left-hand side of the horizontal axis), indicating high level of nominal protection. Conversely, a high foreign content in manufacture (*TiVA* variables coded M_FM, M_FP and M_FS) is associated with a strong incidence of preferential agreements on nominal protection (variables numbered 17 to 22) and high openness to trade (variables coded 23 to 26). Openness to trade is also strongly associated with *TiVA* variables related to foreign content in primary and tertiary exports (as well as in total export, coded by the variable T_FT).

Among the structural variables which were selected on the *TiVA*-associativity criteria, the share of manufacture in GDP (MANU_PIB) appears most frequently. It is closely associated with a high share of domestic manufacture value-added in total exports (T_DM) and, to a lesser extent, to domestic manufacture value-added in primary and tertiary exports (P_DM and S_DM). Curiously and illustratively, the share of manufacture in GDP is not associated with the share of domestic manufacture value-added in manufacture exports, as if manufacture exports often relied on a high foreign content, even in industrialized countries. This result coincides with one of the main conclusions of the new industrial revolution through global value chains: "imports make exports".

Research and Development, an important variable from a policy perspective, is only loosely associated with the domestic content of services and manufacture in exports of primary products. There is also no strong association of R&D with the domestic value-added content of manufacture or services exports. This result is not unexpected when considering that manufacture exports are increasingly done through GVC networks and the providers of R&D may not engage in actual manufacturing activities. As far as services exports are concerned, traditional balance of payments trade statistics may not capture some R&D flows, either because they are embodied into merchandises trade or because they are recorded in third countries for reasons of fiscal optimization.

Figure 14 Association maps of *TiVA* with other structural and trade policy variables (selected cases)



Note: The size and colour of a torus (from blue to red) indicate the strength of the associativity.
Source: Authors' elaboration on the basis of Table 9

No highly significant relationships with *TiVA* or structural variables were found when examining in details the trade policy space. The sole cases where correlations are suggestive (higher than 0.5 in absolute value) correspond to an inverted relationship between on the one hand nominal and effective protection rates and on the other hand the level of development of an economy (represented by their income per capita and the share of services in the GDP). This lack of strong economic determinism in the design of trade policy was also observed in Diakantoni and Escaith (2009).

4.2 Graph analysis

The next sections complement the previous results by investigating similarities between variable yet through a different approach, using tools derived from graph and network theories. The starting point is similar to clustering as it initiates from a matrix of Euclidian distances between variables (based on their association/correlation observed in the countries). After transforming these distances into similarity indexes and taking out the most tautological associations, a graph analysis was performed. [13] The resulting graphs are (loosely) associated with cluster methodologies: they were drawn using force-directed Fruchterman-Reingold algorithm using a relatively high repulsive coefficient in order to clearly separate the nodes that are poorly related. Under such high repulsive parameter, the algorithm tends to produce globular sub-graphs. The balance between attractive and repulsive force in the graph is analogically similar to the joint minimization of within-cluster distance and maximization of between-cluster distance of K-Mean clustering.

In order to maintain a fair level of readability, Figure 15 shows only the strongest linkages (strength of edges equal or higher than 60 on a 0 to 100 scale) involving *TiVA* variables. At this level of detail, four dipoles emerge (from North to South): {M_DM;T_DT}; {S_DM;S_DP}; {M_DP;NATUR}; and {T_DM;MAN_PIB}. Their interpretation is relatively straightforward:

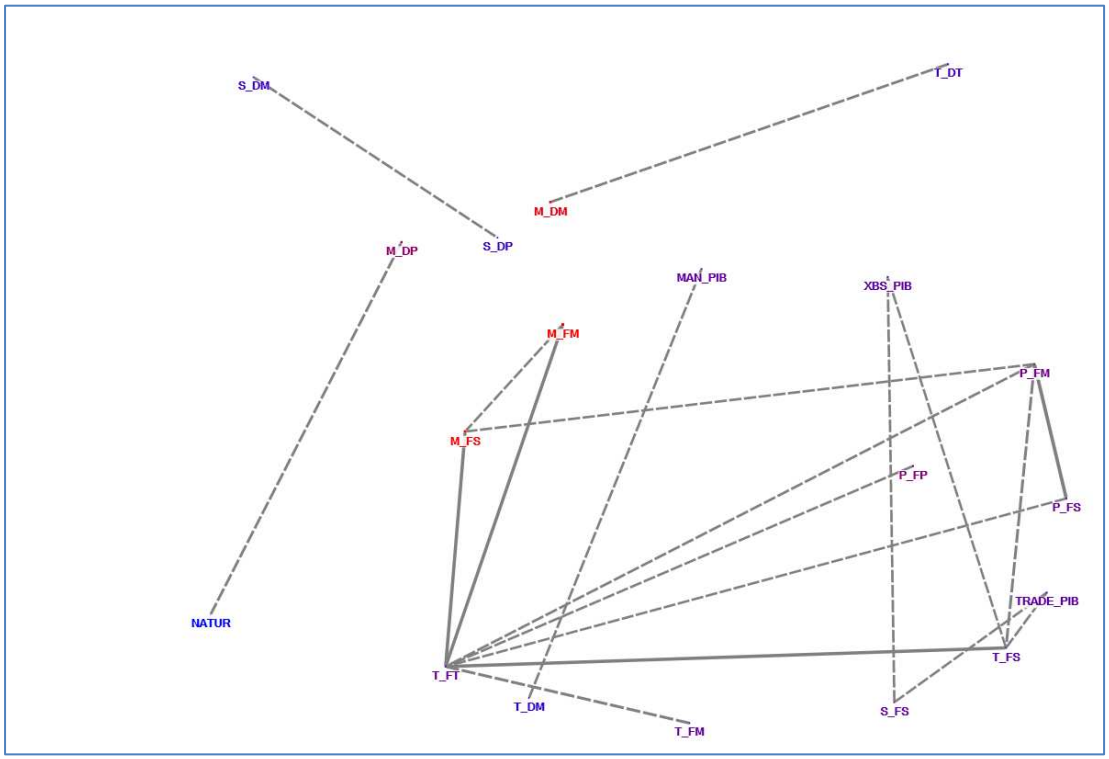
- a high share of domestic value-added in total exports is usually related to a high domestic contribution of manufacture in same sector exports;
- a high incidence of domestic manufacture value-added and of domestic primary value-added in total services exports are usually closely associated;

[13] The index ranges from 100 for the closest pair of variables to 0 for the farthest. Tautological associations were those linking nominal and effective rates of protection for the same product.

- a high domestic content of primary value-added in manufacture exports is related to natural resource endowments and, similarly,
- a high content of manufacture value-added in total exports is associated with a high share of manufacture in domestic GDP.

These four dipoles are complemented by a larger cluster which connects those *TiVA* variables that indicate a high level of vertical specialization (high contribution of foreign content in exports). All these variables are associated with export-oriented open economies (XBS_PIB and TRADE_PIB).

Figure 15 Graph analysis of the TiVA variable space: main relationships



Source: Drawn by the authors. Only close relationships (association index equal or higher than 60) are displayed. Most connected variables (including invisible edges, as long as the degree of association is higher than 30) are coloured in red; lesser connected ones are in blue.

5. VARIABLE REDUCTION AND ASSOCIATION

Our last section complements the previous one by applying principal component and cluster analysis to the complete variable space.

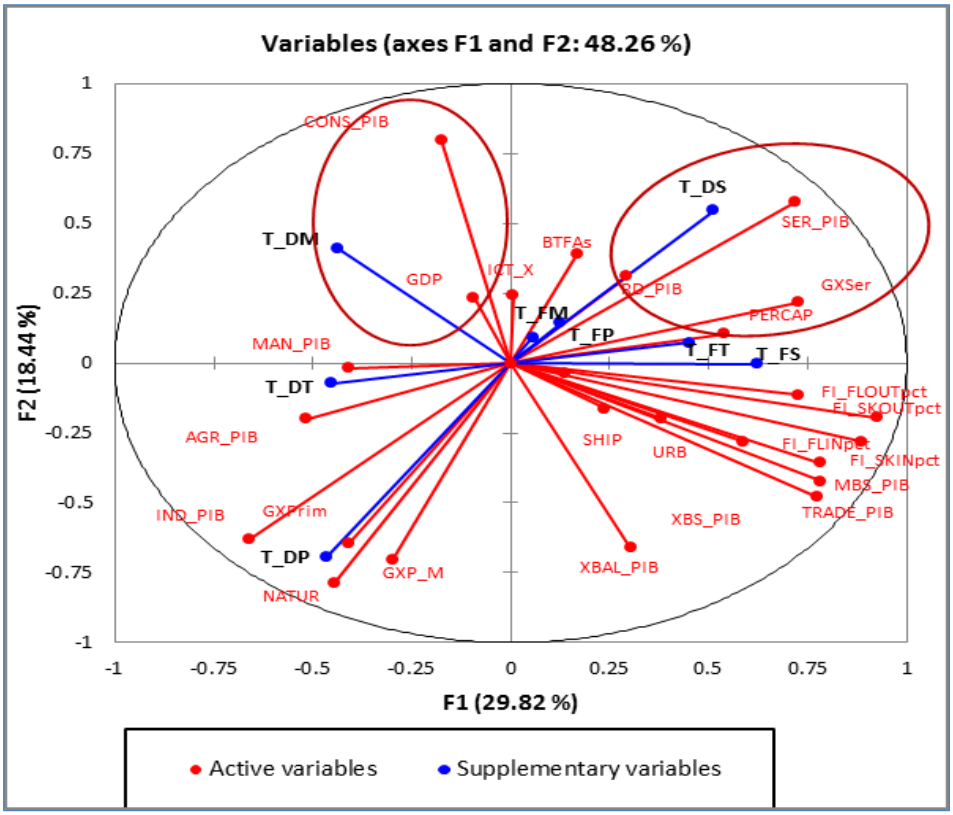
5.1 Projecting TiVA on the structural and trade policy space

Partial PCAs were conducted in order to visualise the relationship between *TiVA* indicators and each of the two other sets of variables; the first one includes only structural indicators and the second only trade policy indicators. In each case, *TiVA* variables were added as supplementary variables in the PCA calculation. Supplementary variables do not interfere with the calculation of principal components but their projection on the resulting variable space shows how these *TiVA* variables position themselves in relation to active variables. In order to simplify the graph, only the most aggregated *TiVA* variables (foreign and domestic value-added in total exports, by sectoral origin) were inserted in the calculation.

Figure 16 shows the results obtained with structural economic variables. The structural variables round up exporters according to their openness to trade and to foreign direct investment, as well as their specialization (ie, services vs. natural resources). On the right side of the horizontal axis (explaining 30% of total variance), we find trade related indicators: ratio trade over GDP,

incidence of FDI stock and flow. Vertical specialization variables T_FS and T_FT (services and total foreign content in total exports) tend to be associated with those variables. Moving towards the North-East, we observe that the domestic content of services VA in total exports (T_DS) is closely related to the weight of services in GDP. Conversely (South-West quadrant) and as expected, domestic primary sector content in total exports is related to natural resources endowment, the ratio of primary exports over manufacture exports and the weight of good-producing (non-services) sectors in the economy. Domestic manufacture content in total exports is more closely related to the ratio consumption/GDP (North-West quadrant).

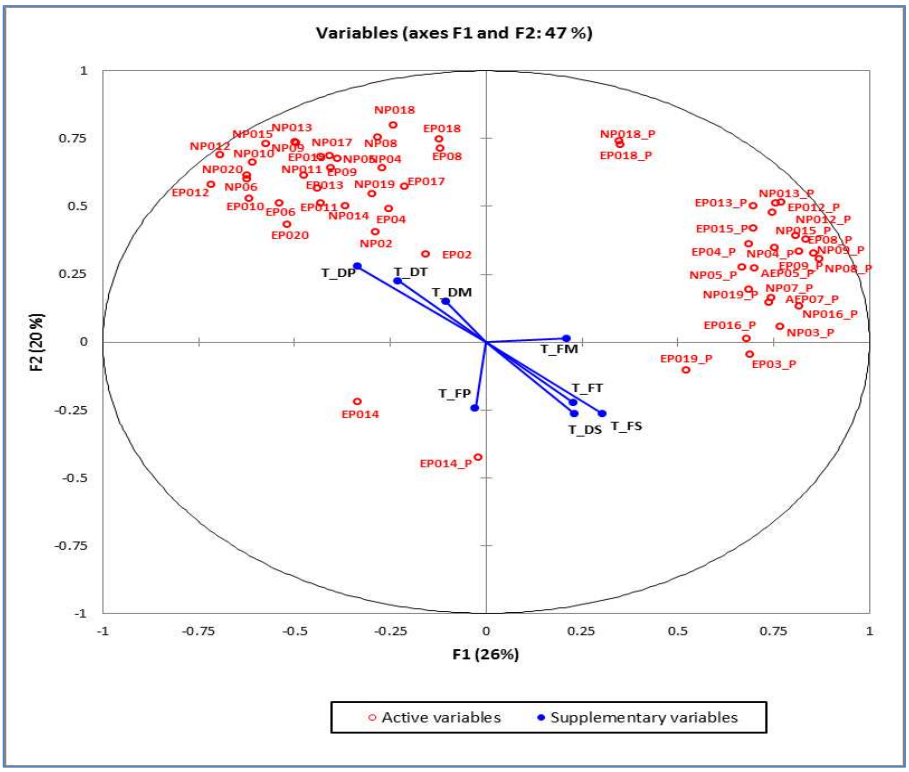
Figure 16 *TiVA* variables as supplement to a PCA on structural variables



The trade policy variables which help profiling *TiVA* variables along the first diagonal of Figure 17 in the North-West quadrant correspond principally to high MFN levels of nominal and effective protection (sector 14 of computation equipment, standing separately in the southern quadrant, probably for the role of the IT Agreement where duty free treatment on IT equipment is granted on an MFN basis: here, the impact of additional preferences is reduced and the two variables stands relatively close together). The trade policy variables capturing the effect of RTAs on this protection stand mainly in North-East quadrant and have less traction on the *TiVA* variables. High domestic content (all sectors, manufacture and services) tend to be associated with high MFN protection. Vertical Specialization and high domestic services content in total exports are, on the contrary, associated with low level of MFN protection.

The foreign content of primary sector VA in total exports (T_FP) and the foreign content of manufacture sector VA in total exports (T_FM) are responding to specific influences, more related in this case to the role of RTAs on applied nominal and effective tariffs (active variables terminating with the suffix "_P" in the graph). The effect is positive in the case of manufacture, indicating an influence of RTAs on vertical specialization. It is negative in the case of VA originating in foreign primary sector: RTAs seem not to induce more integration in primary sector value-chains. But the influence remains weak from a statistical point of view.

Figure 17 *TiVA* variables as supplement to a PCA on trade policy variables



Note: In order to simplify the graph, only active variables with a relatively large absolute correlation with the first two components are presented. Absolute effective protection variables (AEP) were omitted when the corresponding sectoral EPR variable was already selected.

5.2 Associations between variables in the TiVA, structural and trade policy spaces.

After joining the three sets of variables, a first hierarchical cluster analysis is applied to the resulting sample. A first run agglomerates the variables in three major clusters (Table 4). The third cluster is not of interest in our present case as it gathers only trade policy variables and is closely associated with the influence of preferential agreements: the presence of the variable indicating the number of bilateral and regional treaties (BTFAs) and the dominance of variables measuring preference margins variable code ending in "_P"). Unlike the formers, the first and second clusters, provide important information on the degree of association between *TiVA* variables and the other two sets of variables.

The first cluster is representative of open economies relying intensively on foreign value-added for their exports in almost all sectoral compartments. The sole sector where domestic value-added dominates is services, particularly indirectly imbedded in other exports (variable coded "T_DS"). Note that, even if those countries have a relatively high share of gross services in their exports and their domestic economy, they incorporate a large foreign content ("S_FS"), revealing a high degree of vertical specialization. Closely associated with deep vertical specialization and services orientation are high-tech products, R&D, urbanization, shipping lines, active foreign direct investment policy, and high per-capita income. Trade policy variables indicate also a close association with high nominal and effective protection in agriculture, although mitigated by preferential agreements.

The second cluster, by symmetry, is representative of countries that are less inserted in value-chains and rely mostly on their own domestic capacities for their exports. Even when they export services (not their main strength), those services are intensive in domestic value-added (variable coded "S_DS"). Actually, this variable is the closest to the cluster's centroid. But the main export specializations of those predominantly large economies are primary products and manufactures. Those large economies rely either on domestic consumption or on large endowment of natural resources. In close association with this cluster are high nominal and effective protections in all

sectors except agriculture and fuels. Food production is another sector which does not receive high nominal protection, but unlike agriculture and fuels, it benefits from a relatively high effective one.

Despite its high degree of protection against imports, the second group is not characterised by trade surpluses. Actually, the variable "XBAL_PIB" measuring the external balance on goods and services in relation to GDP is associated with the first cluster – the relatively more open ones, with respect to foreign value-added in particular. Albeit this result should not be interpreted from a causality perspective, this remains an interesting characteristic.

Table 10 Hierarchical clustering of all variables in three classes (Ward's criterion)

Class	1 (TRADE_PIB)	2 (S_DS)		3 (NP009_dP)	
Objects	36	77		59	
	P_FP	**P_DP**	.../...	BTFAs	.../...
	P_FM	**P_DM**	*AEPR009*	NP003	AEPR012_dP
	P_FS	**P_DS**	NP010	NP007	NP013_dP
	M_FP	**M_DP**	EPro010	EPro007	EPro013_dP
	M_FM	**M_DM**	AEPR010	AEPR007	AEPR013_dP
	M_FS	**M_DS**	NP011	NP002_dP	NP014_dP
	S_FP	**S_DP**	AEPR011	EPro002_dP	EPro014_dP
	S_FM	**S_DM**	NP012	AEPR002_dP	AEPR014_dP
	S_FS	**S_DS**	EPro012	NP003_dP	NP015_dP
	T_FP	**T_DP**	AEPR012	EPro003_dP	EPro015_dP
	T_FM	**T_DM**	NP013	AEPR003_dP	AEPR015_dP
	T_FS	**T_DT**	EPro013	NP004_dP	NP016_dP
	T_DS	*GXPrim*	AEPR013	EPro004_dP	EPro016_dP
	T_FT	*GXMan*	NP014	AEPR004_dP	AEPR016_dP
	GXSer	*GXP_M*	EPro014	NP005_dP	NP017_dP
	HITEC_X	*AGR_PIB*	AEPR014	EPro005_dP	EPro017_dP
	ICT_X	*CONS_PIB*	NP015	AEPR005_dP	AEPR017_dP
	MBS_PIB	*GDP*	EPro015	NP006_dP	NP019_dP
	PERCAP	*IND_PIB*	AEPR015	EPro006_dP	EPro019_dP
	RD_PIB	*MAN_PIB*	NP016	AEPR006_dP	AEPR019_dP
	SER_PIB	*NATUR*	EPro016	NP007_dP	NP020_dP
	SHIP	*NP002*	AEPR016	EPro007_dP	EPro020_dP
	TRADE_PIB	EPro002	NP017	AEPR007_dP	AEPR020_dP
	URB	AEPR002	EPro017	NP008_dP	
	XBAL_PIB	EPro003	AEPR017	EPro008_dP	
	XBS_PIB	AEPR003	NP018	AEPR008_dP	
	FI_SKINpct	NP004	EPro018	NP009_dP	
	FI_SKOUTpct	EPro004	AEPR018	EPro009_dP	
	FI_FLINpct	AEPR004	NP019	AEPR009_dP	
	FI_FLOUTpct	NP005	EPro019	NP010_dP	
	NP001	EPro005	AEPR019	EPro010_dP	
	EPro001	AEPR005	NP020	AEPR010_dP	
	AEPR001	NP006	EPro020	NP011_dP	
	NP001_dP	EPro006	AEPR020	EPro011_dP	
	EPro001_dP	AEPR006	NP018_dP	AEPR011_dP	
	AEPR001_dP	NP008	EPro018_dP	NP012_dP	
		EPro008	AEPR018_dP	EPro012_dP	
		AEPR008			
		NP009			
		EPro009			

A closer look suggests that three clusters is perhaps too aggregated a partition for adequately capturing the wealth of information included in the aggregated dataset. In order to obtain a more granular picture, we differentiate nine clusters (Table 11). Note that those nine classes are subsets of the three clusters discussed previously: C1 and C2 will merge to form the first cluster of Table 4; C8 and C9 and C5 will form cluster 3 and all 5 remaining classes gather in cluster 2.

Splitting the first cluster of vertically specialised economies into two classes does not bring much additional information, as the sole result is to separate the agricultural trade policy variables (C2) from the *TiVA* and economic criteria. Agricultural protectionism is not particularly associated with any of the other variables in particular. Similarly, the disaggregation of cluster 3 into two sub-sets is not particularly relevant in the present case, as no *TiVA* variables are associated with this grouping.

More interestingly, disaggregating the second cluster sheds additional light on the underlying patterns, particularly when looking at classes C3, C4 and C5:

- C3 characterises situations where exports are concentrated on primary products, thanks to the abundant natural resources endowment of the exporters. The associated *TiVA* variables indicate a large proportion of domestic value-added sourced from the primary sectors. The trade policy variables reveal a relatively high protection granted to sector 2 (Mining and quarrying).
- C4 deals with the larger countries of cluster 2 that exhibit a relatively high manufacture sector and/or are mainly driven by domestic consumption. *TiVA* variables associated with this class of observations reflect a strong reliance on domestic value added in manufacture and services. No specific trade policy is associated with this category.
- C5 characterises the countries in cluster 2 with a relative importance of agricultural production in their GDP. Those countries, probably the poorest of the sample, exhibit the strongest protectionist stance in their trade policy. *TiVA* variables relate to the domestic content in their services exports, but this should not be interpreted as a sign of comparative advantages in services export (services exporters are classified in cluster 1), rather a symptom of marginalisation (services exporters are vertically inserted in global value chains and tend to rely on foreign inputs to foster their exports).

Table 11 Hierarchical clustering of all variables in nine classes (Ward's criteria)

Class	C1 (TRADE_PIB)	C2 (NP001)	C3 (T_DP)	C4 (S_DS)	C5(AEPR012)	C6 (AEPR005)	C7 (NP018)	C8 (NP004_dP)	C9 (AEPR011_dP)
Objects	30	6	10	11	34	13	9	38	21
	P_FP	NP001	**P_DP**	**P_DM**	**S_DP**	EPro003	NP004	BTFAs	NP002_dP
	P_FM	EPro001	**M_DP**	**P_DS**	**S_DM**	AEPR003	EPro004	NP003	EPro002_dP
	P_FS	AEPR001	**T_DP**	**M_DM**	*AGR_PIB*	EPro005	AEPR004	NP007	AEPR002_dP
	M_FP	NP001_dP	*GXPrim*	**M_DS**	NP005	AEPR005	NP018	EPro007	NP005_dP
	M_FM	EPro001_dP	*GXP_M*	**S_DS**	NP006	NP014	EPro018	AEPR007	EPro005_dP
	M_FS	AEPR001_dP	*IND_PIB*	**T_DM**	EPro006	EPro014	AEPR018	NP003_dP	AEPR005_dP
	S_FP		*NATUR*	**T_DT**	AEPR006	AEPR014	NP018_dP	EPro003_dP	NP006_dP
	S_FM		NP002	*GXMan*	NP008	NP016	EPro018_dP	AEPR003_dP	EPro006_dP
	S_FS		EPro002	*CONS_PIB*	EPro008	EPro016	AEPR018_dP	NP004_dP	AEPR006_dP
	T_FP		AEPR002	*GDP*	AEPR008	AEPR016		EPro004_dP	NP011_dP
	T_FM			*MAN_PIB*	NP009	NP019		AEPR004_dP	EPro011_dP
	T_FS				EPro009	EPro019		NP007_dP	AEPR011_dP
	T_DS				AEPR009	AEPR019		EPro007_dP	NP014_dP
	T_FT				NP010			AEPR007_dP	EPro014_dP
	GXSer				EPro010			NP008_dP	AEPR014_dP
	HITEC_X				AEPR010			EPro008_dP	NP019_dP
	ICT_X				NP011			AEPR008_dP	EPro019_dP
	MBS_PIB				EPro011			NP009_dP	AEPR019_dP
	PERCAP				AEPR011			EPro009_dP	NP020_dP
	RD_PIB				NP012			AEPR009_dP	EPro020_dP
	SER_PIB				EPro012			NP010_dP	AEPR020_dP
	SHIP				AEPR012			EPro010_dP	
	TRADE_PIB				NP013			AEPR010_dP	
	URB				EPro013			NP012_dP	
	XBAL_PIB				AEPR013			EPro012_dP	
	XBS_PIB				NP015			AEPR012_dP	
	FI_SKINpct				EPro015			NP013_dP	
	FI_SKOUTpct				AEPR015			EPro013_dP	
	FI_FLINpct				NP017			AEPR013_dP	
	FI_FLOUTpct				EPro017			NP015_dP	
					AEPR017			EPro015_dP	
					NP020			AEPR015_dP	
					EPro020			NP016_dP	
					AEPR020			EPro016_dP	
								AEPR016_dP	
								NP017_dP	
								EPro017_dP	
								AEPR017_dP	

6. CONCLUSIONS

Building on a series of research initiatives, the release of the OECD-WTO *TiVA* database in January 2013 was the first attempt by international agencies to measure world trade in value-added. Combining the *TiVA* results with other economic and trade policy variables, the paper uses Exploratory Data Analysis techniques to identify the underlying patterns that characterize participation in global value chain trade. It should be noted, nevertheless, that the *TiVA* country coverage at the time of writing this report was limited to OECD countries and the main emerging economies. Most African and many Middle-East, Latin American or Caribbean countries are kept outside the reach of this analysis; these countries may have different characteristics from the ones in the sample with regard their mode of insertion in global value chains.

Due to the geographic fragmentation of production, trade in final goods has been complemented by trade in intermediate inputs of goods and services along global value chains. Albeit the supply chain concept was known many years ago, our results show that the rise of trade in tasks along those value chains became prevalent in the 1990s and reached maturity in the early 2000s. After this date, the dynamics of increasing fragmentation remained mainly concentrated in Asia.

The trade profile of the various economies in terms of their value-added composition reproduces a series of characteristics that still reflect the traditional comparative advantages of each country and its level of development, besides reflecting their openness to international trade. Natural resources endowments, on the one hand, and services orientation, on the other one, are among the most determinant variables for defining *TiVA* clusters. A more detailed analysis of the stability of groups of countries according to different methods reveals that once their predominant merchandise export category is defined (whether commodities or manufacture), countries with similar *TiVA* patterns can evidence diverse development levels. However this is not true for service exporters, which tend to be more homogeneous from an economic perspective. Thus, the level of economic development remains a crucial determinant of the TiVA profile.

The size of the economy is also a contributing factor, even if not as decisive as initially expected. Small economies tend to be more integrated into global value chains and exhibit higher content of imported content in their exports. But, despite this higher reliance on imports, they also tend to have competitive exports, leading to surplus in their trade balance. The data reviewed tend, therefore, to support the hypothesis that for export-led strategies, "imports create exports". Conversely, large inward-oriented economies, relying more on internal demand, are those evidencing the largest share of domestic value-added in their exports. An exception to the latter is represented by economies experiencing a high domestic indirect value-added content of goods-producing industries (primary or secondary sectors) in their exports of services. This particular *TiVA* profile reflects symptoms of underdevelopment (in particular, a high share of agriculture in total GDP).

What are the policy variables that could influence the *TiVA* profiles? Investments in ICT and R&D, development of international transport logistics (shipping lines), active foreign direct investment policy: all those value chain upgrading variables are related with a ***high*** foreign content by unit of export. *Id est*, the results do not support the mercantilist objective of relying less on imports.

Similarly, maintaining high levels of nominal and effective protection are not convincing policies. A closer analysis shows that the countries that enforce such high level of protection are not particularly successful in exporting high shares of domestic content, except in services export. But exports of services in this group of countries are not particularly dynamic and do not represent, for the countries involved in protectionist policies, a high share in their total sales to the rest of the world.

Regional trade agreements and active foreign investment policies tend to foster vertical specialization, *id est*, promote a higher foreign content in exports. This result is consistent with the conclusions of WTO (2011) that regional trade agreements are primarily geared at facilitating trade and investment interactions for closer GVC integration. RTAs appear, nevertheless, not very successful in promoting value-chains based on primary sectors. This may be due to the fact that exporting commodities is relatively less difficult and does not require the additional trade enabling effect of joining GVCs. Nevertheless, this result indicates that the up-grading potential offered by GVCs remains largely untapped.

A limit of the analysis – and a serious shortcoming of protectionist trade policies – is that the *TiVA* data refer to the distribution of value added between its foreign and domestic sectoral contents. As other indicators of "global value chain upgrading" such as the "Smiley Curve" demonstrate, focusing on shares tends to obscure a basic law of business: high volumes can compensate for small margins. Thus, the economies most open to imports of intermediate products are also those which were able to export more and record a trade surplus.

*****

REFERENCES

Amin, M. and A. Islam (2014) 'Imports of Intermediate Inputs and Country Size', World Bank Policy Research Working Paper No. 6758.

Chris Ding and Xiaofeng He (2004) 'K-means Clustering via Principal Component Analysis' Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley.

Daudin, G., P. Monperrus-Veroni, Ch. Rifflart and D. Schweisguth (2006) 'Le commerce extérieur en valeur ajoutée' Revue de l'OFCE 98 pp. 129-165.

De Backer, Koen and Sébastien Miroudot, 2013. "Mapping Global Value Chains," OECD Trade Policy Papers 159, OECD Publishing.

De Sousa, J (2012) "The currency union effect on trade is decreasing over time," Economics Letters, Elsevier, vol. 117(3), pages 917-920.

Degain, Ch., H. Escaith and A. Maurer (2013) 'Comparison of methodologies to estimate trade in value added terms', MIWI, WTO mimeo September 2012.

Diakantoni, A. and H. Escaith (2009) 'Mapping the Tariff Waters' WTO Staff Working Paper ERSD-2009-13.

Diakantoni, A. and H. Escaith (2014) ´Trade in Tasks, Tariff Policy and Effective Protection Rates', forthcoming, WTO Research Papers.

Dietzenbacher, E.; Romero, I. (2007) 'Production Chains in an Interregional Framework: Identification by Means of Average Propagation Lengths' International Regional Science Review vol. 30-4 pp.362-383.

Escaith, H. (2014) 'Mapping Global Value Chains and Measuring Trade in Tasks', Asian Development Bank, forthcoming.

Escaith, H (2014b) 'The Policy Space Dimensions of Trade in Value-Added', conference paper, 22nd International Input-Output Conference, Lisbon.

Grossman, G. M. and Rossi-Hansberg, E. (2006) 'Trading tasks: A simple theory of offshoring', NBER Working Paper No. 12721.

Hummels, D., Jun Ishii, and Kei-Mu Yi (2001) 'The Nature and Growth of Vertical Specialization in World Trade' Journal of International Economics, 54(1): 75-96.

Koopman, R, Z Wang, and S-J Wei (2014), "Tracing Value-added and Double Counting in Gross Exports", American Economic Review, 104(2): 459–494. Also available as NBER Working Paper 18579, 2012.

Koopman, R., W. Powers, Z. Wang and Shang-Jing Wei (2010). "Give credit to where credit is due:
tracing value added in global production chains", NBER Working Papers Series 16426,
September 2010.

Ng, F., and A. Yeats (1999) 'Production Sharing in East Asia: Who Does What for Whom, and Why?' Policy Research Working Paper Series 2197, The World Bank, Washington, DC.

OECD-WTO (2011) 'Trade in Value-Added: Concepts, Methodologies and Challenges (Joint Concept Note, MIWI e-document).

OECD-WTO (2012) 'Note on Measuring Trade in Value Added' e-document (WTO-MIWI website).

Park, A., G. Nayyar and P. Low (2013) 'Supply Chains Perspective and Issues: A Literature Review', Fung Global Institute and World Trade Organization.

Rose, A. (2006), "Size Really Doesn't Matter: In Search for a National Scale Effect," Journal of the Japanese and International Economies, 20(4): 482-507.

WTO (2011) 'World Trade Report 2011: The WTO and preferential trade agreements, from co-existence to coherence', Geneva.

WTO and IDE-JETRO (2011b) 'Trade patterns and global value chains in East Asia: From Trade in Goods to Trade in Tasks', Geneva and Tokyo.

WTO (2008) 'World Trade Report 2008: Trade in a Globalizing World', Geneva.

Zhi Wang, Shang-Jin Wei and Kunfu Zhu (2014) 'Gross trade accounting: A transparent method to discover global value chain-related information behind official trade data: Part 2', Vox EU 16 April.

Annex 1. *TiVA* and other variables used in the analysis

The Trade in Value-Added initiative launched by OECD and WTO in 2012 attempts to account for the implicit double counting in current gross flows of trade, and measures trade flows according to where the value is added (labour compensation, taxes and profits) by industrial sectors and countries in the production of any good or service that is exported. This requires a full set of inter-country I-O tables, where all bilateral exchanges of intermediate goods and services are accounted for: in other words an international input-output table.

Identifying backwards linkages from those export oriented sectors producing tradable goods (agriculture, manufacture) allows mapping where the domestic value added was created, either domestically or internationally. The break-up of domestic content by direct and indirect sectoral value added reveals that a large chunk of the value originates indirectly from service sectors. This break-down is particularly important for identifying up-stream sectors (typically, services) which are not considered as exporters by traditional statistics. [14]

Using the *TiVA* database first released in January 2013 then updated in May 2013, the paper identified a set of variables defining the national (home vs. foreign) and sectoral (primary, secondary and tertiary) origin of the value-added imbedded in sectoral exports. For example, in the case of the manufacture sector, the following *TiVA* indicators are computed (Table 12); similar indicators are computed for the primary (agriculture and mining) and tertiary (services) sectors.

Table 12 Example of *TiVA* indicator for the Manufacture sector

Manufacture export, % ***domestic*** VA from Manufacture
Manufacture export, % ***domestic*** VA from Primary
Manufacture export, % ***domestic*** VA from Services
Manufacture export, % ***foreign*** VA from Manufacture
Manufacture export, % ***foreign*** VA from Primary
Manufacture export, % ***foreign*** VA from Services

Note: domestic VA includes both direct and indirect sectoral contributions to the total value of the output. Indirect contribution refers to a situation when the industry (in this case, manufacture) is a supplier to other exporting sectors (primary goods or services).

Other variables describe the economic structure of the exporters (GDP and its composition, per capita income, intensity of R&D; incidence of foreign direct investment, etc.) and are sourced mainly from the World Bank (World Development Indicators). Trade policy indicators (nominal and effective protection by sector, incidence of preferential regimes on MFN treatment) are derived from WTO and OECD database, using Diakantoni and Escaith (2014). All values refer to 2008 or closest year; whenever possible, missing data were imputed using other sources or interpolation.

Table 13 Dictionary of variables utilised in the analysis.

Indicator description	Indicator code	Source
Agriculture, value added (% of GDP)	AGR_PIB	WDI, World Bank
Final consumption expenditure, etc. (% of GDP)	CONS_PIB	WDI, World Bank
High-technology exports (% of manufactured exports)	HITEC_X	WDI, World Bank
ICT service exports (% of service exports, BoP)	ICT_X	WDI, World Bank
Industry, value added (% of GDP)	IND_PIB	WDI, World Bank
Manufacturing, value added (% of GDP)	MAN_PIB	WDI, World Bank
Imports of goods and services (% of GDP)	MBS_PIB	WDI, World Bank
GDP per capita (current US$)	PERCAP	WDI, World Bank
Research and development expenditure (% of GDP)	RD_PIB	WDI, World Bank
Services, etc., value added (% of GDP)	SER_PIB	WDI, World Bank
Trade (% of GDP)	TRADE_PIB	WDI, World Bank
External balance on goods and services (% of GDP)	XBAL_PIB	WDI, World Bank
Exports of goods and services (% of GDP)	XBS_PIB	WDI, World Bank
GDP (current US$)	GDP	WDI, World Bank
Liner shipping connectivity index (maximum value in 2004 = 100)	SHIP	WDI, World Bank

[14] OECD-WTO (2011).

Population in urban agglomerations of more than 1 million (% of total population)	URB	WDI, World Bank
Total natural resources rents (% of GDP)	NATUR	WDI, World Bank
Labour participation rate, total (% of total population ages 15+)	LAB_PART	WDI, World Bank
Agricultural land (% of land area)	AGR_LAND	WDI, World Bank
Cost to export (US$ per container)	COS_EXP	WDI, World Bank
Cost to import (US$ per container)	COS_IMP	WDI, World Bank
Current account balance (% of GDP)	CA_BAL	WDI, World Bank
Employment in agriculture (% of total employment)	AGR_EMP	WDI, World Bank
Employment in industry (% of total employment)	IND_EMP	WDI, World Bank
Employment in services (% of total employment)	SER_EMP	WDI, World Bank
Employment to population ratio, 15+, total (%)	EMP_POP	WDI, World Bank
Gross national expenditure (% of GDP)	GRO_EXP	WDI, World Bank
Gross savings (% of GDP)	GRO_SAV	WDI, World Bank
Inflation, consumer prices (annual %)	INF_CPI	WDI, World Bank
International tourism, receipts (% of total exports)	INT_TOUR	WDI, World Bank
Internet users (per 100 people)	INT_USER	WDI, World Bank
Labour force participation rate, total (% of total population ages 15-64)	LAB1564	WDI, World Bank
Labour force with tertiary education (% of total)	EDUC_TER	WDI, World Bank
Land area (sq. km)	LAND	WDI, World Bank
New businesses registered (number)	NEWBIZ	WDI, World Bank
Oil rents (% of GDP)	OIL	WDI, World Bank
Passenger cars (per 1,000 people)	PAS_CAR	WDI, World Bank
Population ages 15-64 (% of total)	AGE_WORK	WDI, World Bank
Population, total	POP	WDI, World Bank
Public spending on education, total (% of GDP)	SPE_EDUC	WDI, World Bank
Rural population (% of total population)	RURAL	WDI, World Bank
Time required to start a business (days)	TIMBIZ	WDI, World Bank
Time to export (days)	TIM_EXP	WDI, World Bank
Number of Free Trade Agreements enforced	BTFAs	de Sousa, J.[15]
Foreign Direct Investment, Inward Stock (USD Million)	FDI_SK_IN	UNCTAD
FDI Outward Stock (USD Million)	FDI_SK_OUT	UNCTAD
FDI Inward Flow (USD Million)	FDI_FL_IN	UNCTAD
FDI Outward Flow (USD Million)	FDI_FL_OUT	UNCTAD
FDI Inward Stock (percent GDP)	FI_SKINpct	UNCTAD
FDI Outward Stock (percent GDP)	FI_SKOUTpct	UNCTAD
FDI Inward Flow (percent GDP)	FI_FLINpct	UNCTAD
FDI Outward Flow (percent GDP)	FI_FLOUTpct	UNCTAD
Primary exports (Gross, % total)	GXPrim	*TiVA*
Manufacture exports (Gross, % total)	GXMan	*TiVA*
Services exports (Gross, % total)	GXSer	*TiVA*
Ratio Primary Exports / Manufacture Exports (Gross, %)	GXP_M	*TiVA*
Manufacture export, % domestic Value-Added from Manufacture	M_DM	*TiVA*
Manufacture export, % domestic VA from Primary	M_DP	*TiVA*
Manufacture export, % domestic VA from Services	M_DS	*TiVA*
Manufacture export, % foreign VA from Manufacture	M_FM	*TiVA*
Manufacture export, % foreign VA from Primary	M_FP	*TiVA*
Manufacture export, % foreign VA from Services	M_FS	*TiVA*
Primary export, % domestic VA from Manufacture	P_DM	*TiVA*
Primary export, % domestic VA from Primary	P_DP	*TiVA*
Primary export, % domestic VA from Services	P_DS	*TiVA*
Primary export, % foreign VA from Manufacture	P_FM	*TiVA*
Primary export, % foreign VA from Primary	P_FP	*TiVA*
Primary export, % foreign VA from Services	P_FS	*TiVA*
Services export, % domestic VA from Manufacture	S_DM	*TiVA*
Services export, % domestic VA from Primary	S_DP	*TiVA*
Services export, % domestic VA from Services	S_DS	*TiVA*
Services export, % foreign VA from Manufacture	S_FM	*TiVA*
Services export, % foreign VA from Primary	S_FP	*TiVA*
Services export, % foreign VA from Services	S_FS	*TiVA*
Total export, % domestic VA from Manufacture	T_DM	*TiVA*
Total export, % domestic VA from Primary	T_DP	*TiVA*
Total export, % domestic VA from Services	T_DS	*TiVA*
Total export, total % domestic VA from all sectors	T_DT	*TiVA*
Total export, % foreign VA from Manufacture	T_FM	*TiVA*
Total export, % foreign VA from Primary	T_FP	*TiVA*
Total export, % foreign VA from Services	T_FS	*TiVA*

[15] De Sousa, José(2012), pages 917-920.

Total export, total % foreign VA from all sectors	T_FT	*TiVA*
Nominal Protection at Most Favoured Nation, including Ad Valorem Equivalents, for each good producing sector of *TiVA*	NP	WTO IDB
Effective Protection Rate (including AVEs), for each good producing sector of *TiVA*	EPro	WTO IDB
Absolute Effective Protection (numerator of the EPR, including AVEs), for each good producing sector of *TiVA*	AEPR	WTO IDB
Difference between "NP at MFN" and "NP including preferences", for each good producing sector of *TiVA* (here, sector 001)	NP001_dP	WTO IDB
Difference between Effective Protection Rate at MFN and including preferences (here, sector 002)	EPro002_dP	WTO IDB
Difference between Absolute Effective Protection at MFN and including preferences (here, sector 003)	AEPR003_dP	WTO IDB

Table 14 Dictionary of disaggregated sector reference numbers listed in the WTO IDB database.

Sector codes/names
001 - Agriculture, hunting, forestry and fishing
002 - Mining and quarrying
003 - Food products, beverages and tobacco
004 - Textiles, textile products, leather and footwear
005 - Wood and products of wood and cork
006 - Pulp, paper, paper products, printing and publishing
007 - Coke, refined petroleum products and nuclear fuel
008 - Chemicals
009 - Rubber & plastics products
010 - Other non-metallic mineral products
011 - Basic metals
012 - Fabricated metal products, except machinery & equipment
013 - Machinery & equipment, nec
014 - Office, accounting & computing machinery
015 - Electrical machinery & apparatus, nec
016 - Radio, television & communication equipment
017 - Medical, precision & optical instruments
018 - Motor vehicles, trailers & semi-trailers
019 - Other transport equipment
020 - Manufacturing nec; recycling (include Furniture)

Table 15 Sample of 53 economies covered in the analysis and their ISO codes

Name	ISO3	Name	ISO3	Name	ISO3	Name	ISO3
Argentina	ARG	Finland	FIN	Latvia	LVA	Saudi Arabia	SAU
Australia	AUS	France	FRA	Lithuania	LTU	Singapore	SGP
Austria	AUT	Germany	DEU	Luxembourg	LUX	Slovak Rep.	SVK
Belgium	BEL	Greece	GRC	Malaysia	MYS	Slovenia	SVN
Brazil	BRA	Hong Kong SAR	HKG	Mexico	MEX	South Africa	ZAF
Bulgaria	BGR	Hungary	HUN	Netherlands	NLD	Spain	ESP
Canada	CAN	India	IND	New Zealand	NZL	Sweden	SWE
Chile	CHL	Indonesia	IDN	Norway	NOR	Switzerland	CHE
China	CHN	Ireland	IRL	Philippines	PHL	Chinese Taipei	TWN
Cyprus	CYP	Israel	ISR	Poland	POL	Thailand	THA
Czech Rep.	CZE	Italy	ITA	Portugal	PRT	Turkey	TUR
Denmark	DNK	Japan	JPN	Romania	ROU	United Kingdom	GBR
Estonia	EST	Korea. Rep.	KOR	Russian Fed.	RUS	United States	USA
						Vietnam	VNM