Mapping the Tariff Waters

Diakantoni, Antonia and Escaith, Hubert
WTO, DEFI

01. December 2009

Mapping the Tariff Waters

Antonia Diakantoni * and Hubert Escaith ** [1]
25 November 2009

Abstract

Tariff water –the difference between bound and applied duties– provides relevant information on domestic trade policy and WTO trade negotiations. This paper examines the general and sectoral tariff structure of 120 economies, using exploratory data analysis.

Keywords: Commercial Policy, Taxation, Tariff Duty, GATT-WTO, International Trade Agreements, MFN, Bound Tariff, Tariff Water.

JEL: F13, H20

Acknowledgements: The authors acknowledge helpful comments and suggestions by Jean-Michel Pasteels and technical support from Joaquin Montes and Stefan Froschl. Any remaining errors and analytical shortcomings remain the sole responsibility of the authors.

* Economic Research and Statistics Division, WTO. ** ERSD-WTO; DEFI

1. This article represents the opinions of the authors, and it is not meant to represent the position or opinions of the WTO or its Members, nor the official position of any staff members.

I. INTRODUCTION

Governments negotiate at the WTO the bound tariffs levels, i.e. the maximum tariffs they are authorised to charge on other WTO members. Countries differences in their bound levels as well as in applied tariffs under the most favoured nation principle (MFN) are most likely related to both their economic development and their national policies towards domestic protection and trade openness. To the extend that trade negotiations are reflected in these policies, analysing the existing tariff structures should expose any underlying negotiation strategies.

The present study examines the hypothesis that the particular structure of bound and applied tariffs of a given country, reflects a series of trade policy decisions that are related to its economic development and reveals national authorities' protection standpoint towards the various economic sectors. The analysis identifies the countries having similar overall structures across sectors, then analyses the specific patterns for each particular product sector. Finally, the degree of influence of countries' socio-economic characteristics on their tariff structure is evaluated. In doing so, the analysis takes into account that recently acceded members went through an accession process that was more demanding than previous GATT practices.

The relative situation of applied and bound tariffs provides also information on the negotiation margins of each economy, as open economies (low applied duties) with low bounds have less margin at the negotiation table than open economies with high bounds. Economies with high applied duties are in an intermediate situation as any reduction in the bound could cut the MFN applied and this may be resisted by domestic pressure groups. In addition, some members are also allowed to keep some of their products unbound while others have fixed bound duties at economic pointless levels i.e., much above the prohibitive tariff (Foletti, Fugazza, Nicita and Olarreaga 2009).
Thus, a better understanding of the tariff structure can shed additional light on the strategies behind trade negotiations and the building up of coalitions (Costantini, Crescenzi, De Filippis and Salvatici 2007).

Mapping the tariff waters has gained an additional interest after September 2008, when the global crisis severely affected the international economy. [2] In such a crisis situation, it is feared that governments will look for uncooperative exit solutions typical of the "Prisoner Dilemma", running unilaterally their applied tariff within the bound commitments at the expense not only of their partners, but eventually of themselves and the international governance.

After introducing the data and the statistical methodology, the study proceeds by exploring the tariff data from a general to a sectoral perspective. The fourth part explores the relationship between the tariffs structure and relevant socio-economic variables. Finally, the conclusion summarises the main results.

[2] In this paper, the term "tariff water" is used for the binding overhang, namely the difference between bound and MFN applied tariffs. From a historical context, the expression was used, in a somewhat more restrictive connection, in the Uruguay Round negotiations when countries offered tariff equivalents that were obviously inflated (Goode 2003). It should not be confused neither with the notion of "policy space" as many trade partners have concluded a series of regional or preferential trade agreements (PTA and RTA, respectively) that legally restrict their margin of manoeuvre.

II. THE TARIFF DATA

A. SOURCES OF DATA AND PROCESSING

The analysis covers 120 economies with the European Union 27 countries counting as one and Switzerland including Liechtenstein. While only WTO Members are considered, some had to be excluded from the analysis due to lack of information. The applied and bound duties are sourced from the WTO Integrated Data Base (IDB) and the WTO Consolidated Tariff Schedules database (CTS). The 2006 applied duties are considered; for 17 countries where this information is missing, the latest available year is considered. Missing applied duties and *ad valorem* equivalents are sourced from the UNCTAD's TRAINS database and ITC's MAcMap database, discarding the outliers (e.g. *ad valorem* equivalents higher than 1000%).

In order to obtain cross-country comparability, national tariff line duties are first aggregated up to the level of 6-digit HS sub-headings by calculating the simple average of included tariff lines. The water between the bound and applied duty is calculated at the HS 6-digit level; for products where the applied duty is not available (0.4% of all observations), the water has been left missing. Negative values of the water (3.5% of all observations) due to reporting errors, binding violations or inconsistencies in the calculation of *ad valorem* equivalents, are set to zero. Unbound products [3] (21.3 % of the observations) are considered sensitive for the economy; thus, the imputed water (at 6-digits) is set equal to the 3rd quartile of the water observed for similar products (same HS 4-digit products, or same HS 2-digit products when no 4-digit water is available). For products where the water at the HS 2-digit level is not available, the value has been left missing.

To provide economically meaningful information, applied duties and water have been aggregated by sector by using simple averages, according to UNCTAD's [4] economic classification based on factor intensity. This nomenclature regroups all HS products into 6 groups at the HS 4-digit level (

Table 1).

Table 1: Product sectors by factor intensity

Sector	Description	Typical products
A	Non-fuel primary commodities	Agricultural products (fish included)
B	Resource-intensive manufactures	Raw materials (aluminium, paper, leather, silk, etc.)
C	Low skill- and technology intensive manufactures	Textiles and clothing
D	Medium skill- and technology intensive manufactures	Organic and inorganic chemicals, rubber, machinery, cars
E	High skill- and technology intensive manufactures	Pharmaceuticals and hi-tech products
F	Mineral fuels	Coal, petroleum, other energy.

Source: UNCTAD's economic classification based on factor intensity

B. DATA ANALYSIS METHODOLOGY

In order to identify underlying patterns behind the large volume of data, a series of bottom-up procedures inspired by Exploratory Data Analysis (EDA) are applied. EDA seeks "connections" among complex multivariate data sets, without pretence of causality, that may help to formulate

[3] and not available bound tariffs

[4] The original classification follows the HS96 nomenclature; it was also transposed into the HS92 and in HS02 in order to match all countries applied and bound duties.

statistical models to test. Because bound, applied and binding overhang are related with a strict identity, the analysis focuses on applied tariff (the present situation) and the binding overhang (the tariff water). [5] The category F (Fuels) has been excluded, because it is (i) usually unrepresentative of national trade policy, (ii) particularly prone to measurement errors in notifications and interpretation, and (iii) plagued by many missing observations.

A series of statistical techniques are implemented in the study. Underlying similarities and dissimilarities between observations and variables are explored using clustering techniques in order to (i) develop a typology of objects; (ii) understand the conceptual classificatory function behind the typology; and (iii) formulate hypothesis about the data generation process. [6] Principal Component Analysis (PCA) is used to identify the underlying multidimensional data structure. PCA displays the data in a substantially reduced subspace of uncorrelated variables, that best preserves the statistical information (variance) contained in the original data. [7] Discriminatory analysis is applied in the end, with a view to verify whether the composition of the identified clusters during the analytical process could be linked to any trade policy or economic performance variables.

III. EXPLORING THE TARIFF WATERS

Adopting the time honoured "general to specific" statistical analysis, this section initiates with a global perspective, factoring-in all the tariff dimensions (sectors) into a common pattern, representative –at least in the tariff space– of each country's trade policy. The second part is successively focusing on each one of the tariff policy dimensions (sectoral tariffs) to scale more precisely the tariff profile of each individual member.

A. THE TARIFFS FROM A GLOBAL PERSPECTIVE

The entire set of variables and observations is explored first, in order to provide some general descriptive information of the dataset and to identify any general underlying patterns across both products and countries.

1. Descriptive statistics on tariffs

Table 2 presents a set of descriptive statistics concerning the observed distribution of applied tariffs and the binding overhang. Not surprisingly, the A category of agricultural products emerges almost immediately as an outlier: it spreads across a large interval and has the highest variance in both dimensions of applied and water.

A box plot of agricultural data [8] reveals the presence of outliers for applied tariff duties, with the maximum lying far away from the normally expected range. Aside from agriculture, the group D, of medium skill- and technology intensive manufactures (chemicals and cars), shows also an interesting distribution. At the contrary of A, it is a very compact group, distributed much sharply and more concentrated than a normal Gauss distribution. The distribution of applied tariffs in group C of low skills- and technology intensive manufactures (textiles and clothing) is the closest to normality, in our sample. It is relatively symmetric, albeit flatter than the normal distribution. Groups B and E come in-between the previous cases.

[5] Dropping one of the variables is not only an editorial choice, but a mathematical necessity. Because the three components are linearly dependent, they cannot be analysed jointly through most statistical procedures based on linear algebra.

[6] Testing those hypothesis belongs to "confirmatory data analysis", not covered here.

[7] PCA is similar to factor analysis, but it is more general (it works on the total variance of the observation and does not imply any underlying statistical model).

[8] Not published here; all the same, distribution patterns for individual variables can easily be inferred from Table 2 and Figure 1.

Table 2: Descriptive statistics on applied MFN and binding overhang

Variables: [a]										
Statistics:	App_A	App_B	App_C	App_D	App_E	Dif_A	Dif_B	Dif_C	Dif_D	Dif_E
No. of observations	120	120	120	120	120	120	120	120	120	120
No. of missing values	0	0	0	0	0	0	0	0	0	0
Minimum	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Maximum	54.2	30.0	29.3	30.5	31.2	148.2	115.7	115.0	119.5	116.8
Freq. of minimum	2.0	4.0	3.0	3.0	5.0	2.0	3.0	3.0	3.0	3.0
Freq. of maximum	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Range	54.2	30.0	29.3	30.5	31.2	148.2	115.7	115.0	119.5	116.8
1st Quartile	8.7	4.9	6.8	3.6	3.6	15.3	7.5	6.6	7.0	4.6
Median	12.9	8.3	11.6	6.5	8.1	26.5	22.6	17.3	21.2	17.4
3rd Quartile	17.1	12.6	16.3	8.5	11.1	62.0	40.5	36.9	41.9	36.4
Mean	14.3	9.1	12.4	6.9	7.8	38.8	28.5	25.9	26.9	23.5
Variance (n)	83.8	33.4	43.4	20.6	29.2	1077.1	724.4	666.0	617.9	528.3
Standard deviation (n)	9.2	5.8	6.6	4.5	5.4	32.8	26.9	25.8	24.9	23.0
Variation coefficient	0.6	0.6	0.5	0.7	0.7	0.8	0.9	1.0	0.9	1.0

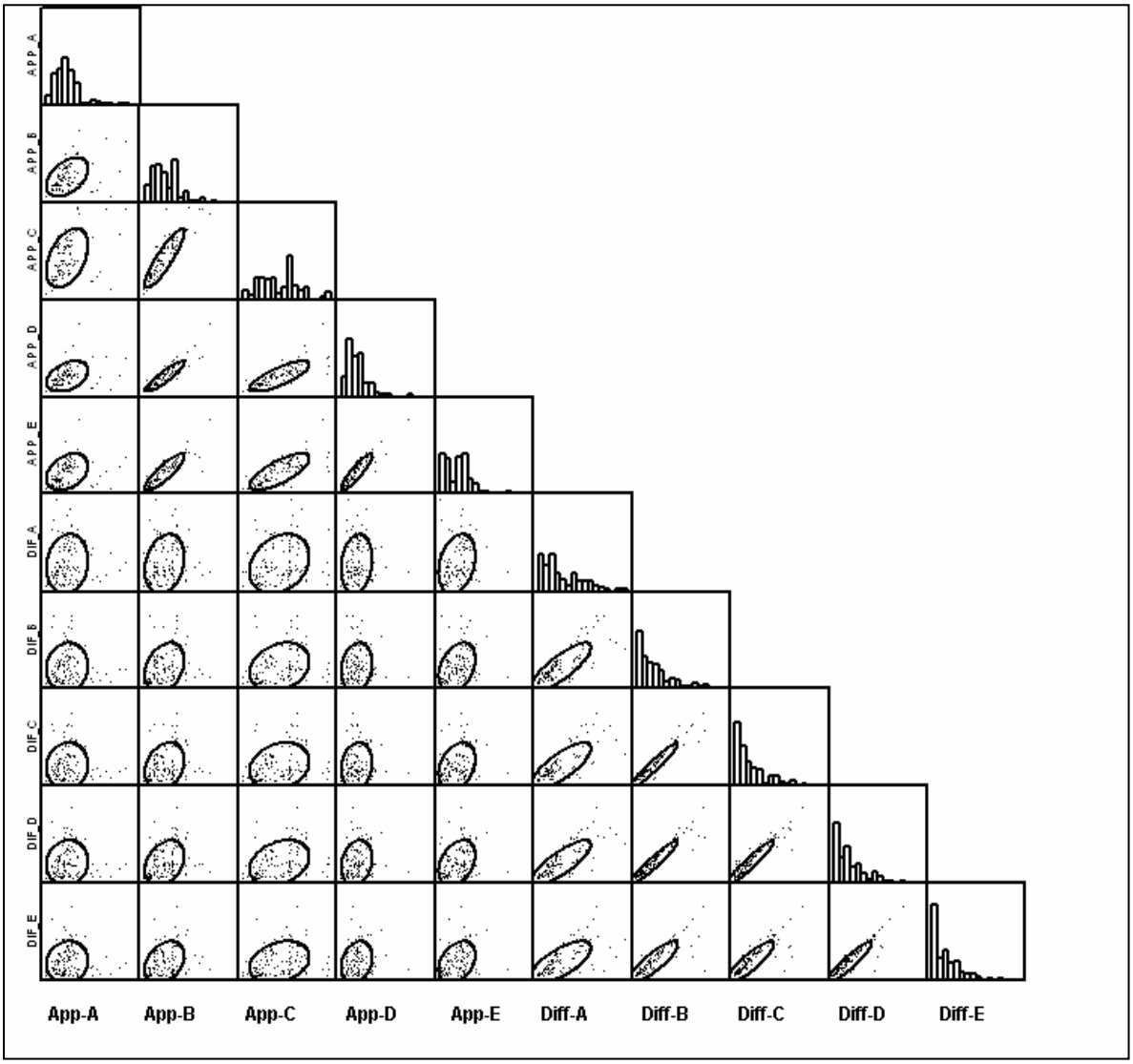
Notes: a/ Variable codes used along the paper are constructed using a prefix plus a suffix. Prefix "App_" stands for applied MFN, "Dif_" for difference between applied and bound. Suffix refer to product groups: A for Non-fuel primary commodities; B for Resource-intensive manufactures; C for Low skill- and technology intensive manufactures; D for Medium skill- and technology intensive manufactures; E for High skill and technology intensive manufactures (see

Table 1). All statistics are based on un-weighted average.
Source: Authors' calculation on the basis of WTO tariff data.

Applied tariffs are usually more regularly distributed than the binding overhang. The distribution of water is characterised by its concentration towards the lower end of the spectrum, together with a large variance of the observations. Here again, agriculture is an extreme case; all remaining groups —the non agricultural products, or NAMA products— share apparently more or less the same distribution.

The scatter-plot matrix (Figure 1) allows to visualize the individual and the joint distribution of applied tariff levels and tariff water, offering a fairly comprehensive view of the data. The diagonal shows the distribution histogram of each individual variable and complements the statistics presented in Table 2. Scatter-plots outside the diagonal illustrate, for each cell or "panel" formed by the intersection of a row and column, the bivariate relationship between the variables appearing on the horizontal and vertical axis. To facilitate the visualization of the bivariate relationship, a confidence interval ellipse is centred on the sample means of the respective variables. It indicates the sample covariance between the two variables (a measure of correlation) and its orientation. Highly correlated variables will exhibit an elongated ellipse along the first or second diagonal of the graph (positive or negative correlation).

Figure 1: Scatter-plot matrix of applied tariffs and difference



Source: Authors' calculation on the basis of WTO tariff data.

The histograms on the diagonal of the scatter-plot matrix show a clear distinction between the distribution of applied tariffs and water. The tariff waters share a common profile with a concentration of observations on low values and with a regular decrease of frequency when moving to higher values for the water. This patter is steeper for higher technology contents (categories D and E) and more diffuse and irregular for agriculture. Applied MFN show a wider difference in histograms, from log-normal (groups A, B and D) or asymmetric (E) to a more normal distribution around the average (C).

There is usually a strong bilateral interrelation between tariff waters, as shown by the elongated feature of the covariance ellipse. The relationship is strong for NAMA but, once again, water in agriculture is less closely related to water in other categories of goods (i.e., some countries may have high water for agriculture and low water for NAMA, or conversely). The relationship among applied tariffs is also present, albeit weaker, between B and the other NAMA categories C, D and E, and between D and E. The covariance between {C ; D}, and {C ; E} exists, but is weaker than for the other NAMA categories. As far as agriculture is concerned, no clear bivariate relationship appears from the graphs, as evidenced by the rounder ellipses. Similarly, there is no direct relationship between applied and differential tariffs.

The degree of similarity/dissimilarity among the product groups has been also be more formally measured by testing the following hypothesis: do the 5 product sectors belong to the same statistical population as far as the applied duties and water are concerned? For the applied duties the hypothesis was rejected, confirming that each group receives a specific tariff treatment and validating that the five sectors can be considered as separate entities. Even when the comparison is limited to NAMA groups (B, C, D and E), the hypothesis is rejected.

The situation for the water is somehow similar as the hypothesis has been also rejected, but the risk of wrongly rejecting the hypothesis is about 0.2, while for the applied tariffs was almost nil. Moreover, when the sample is reduced to the three NAMA groups B, C, E, the hypothesis of same origin is accepted: water is sufficiently similar across these three sub-group to consider them as homogeneous for this criteria.

Thus, from an analytical perspective, applied tariffs and binding overhang seem to follow a different logic. Agriculture follows a clearly specific internal logic (*sui generis*) in terms of applied duties and water. Moreover, the MFN applied levels in the NAMA sector are clearly distinct but the water looks rather homogeneous across NAMA products.

2. Geographical mapping of tariffs

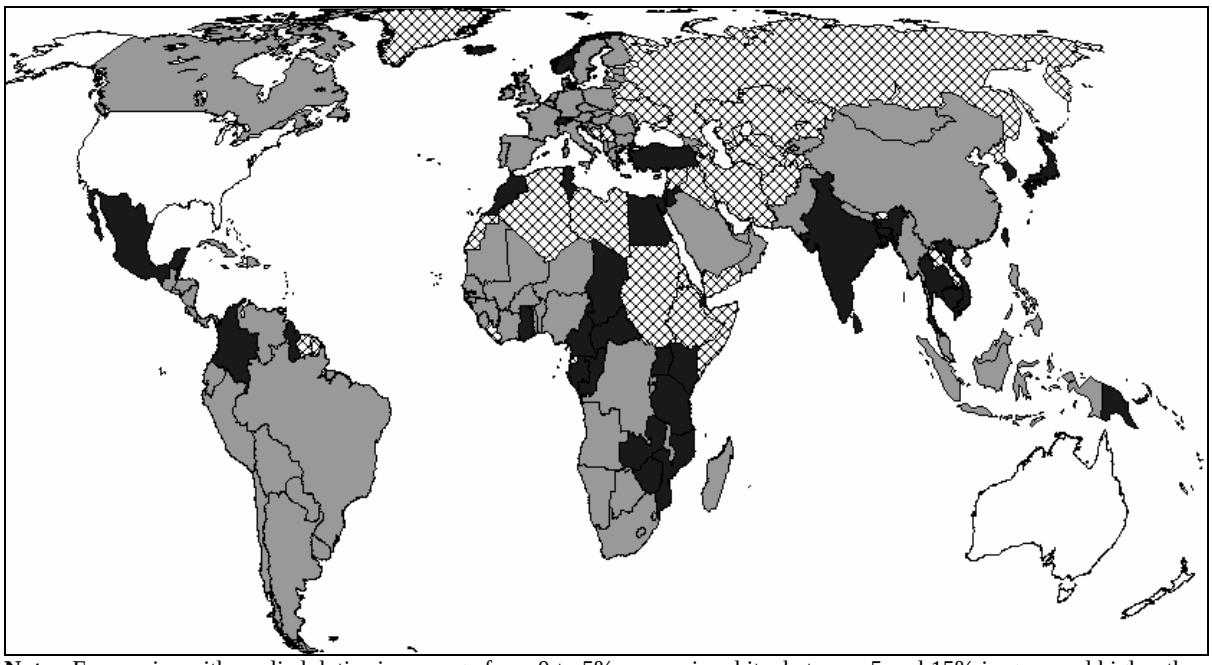
By categorizing tariffs (0 to 5%; between 5 and 15%; higher than 15%) and by colouring high values with dark colours, an illustration of the average applied tariffs in agriculture and in NAMA for the 120 economies is provided in the world maps 1 and 2 respectively. [9] The word maps 3 and 4 provide an overview of the average tariff water for agricultural and NAMA products by using the same scale and colours.

Countries seem to apply in average higher duties in agriculture than in NAMA, as shown in maps 1 and 2. For instance, in the Americas Mexico and Colombia look much darker in the agricultural sector, like Egypt in Africa and Norway, Switzerland and Iceland in Europe. Similarly Japan in Asia, Turkey and Saudi Arabia in Middle East apply higher duties in agriculture and appear darker in map 1. Nevertheless, a few countries applying lower duties in agriculture can be identified like Rwanda and Burundi in Africa, as they appear lighter in the agriculture map.

Globally, in agriculture the average applied duty is 14% and the water 39%; in NAMA the corresponding values are 9% and 26% and as expected, are lower than in agriculture. It has to be reminded that averages include AVEs which can fluctuate significantly from one year to another and the average NAMA water includes a high number of estimates for the unbound products. [10]

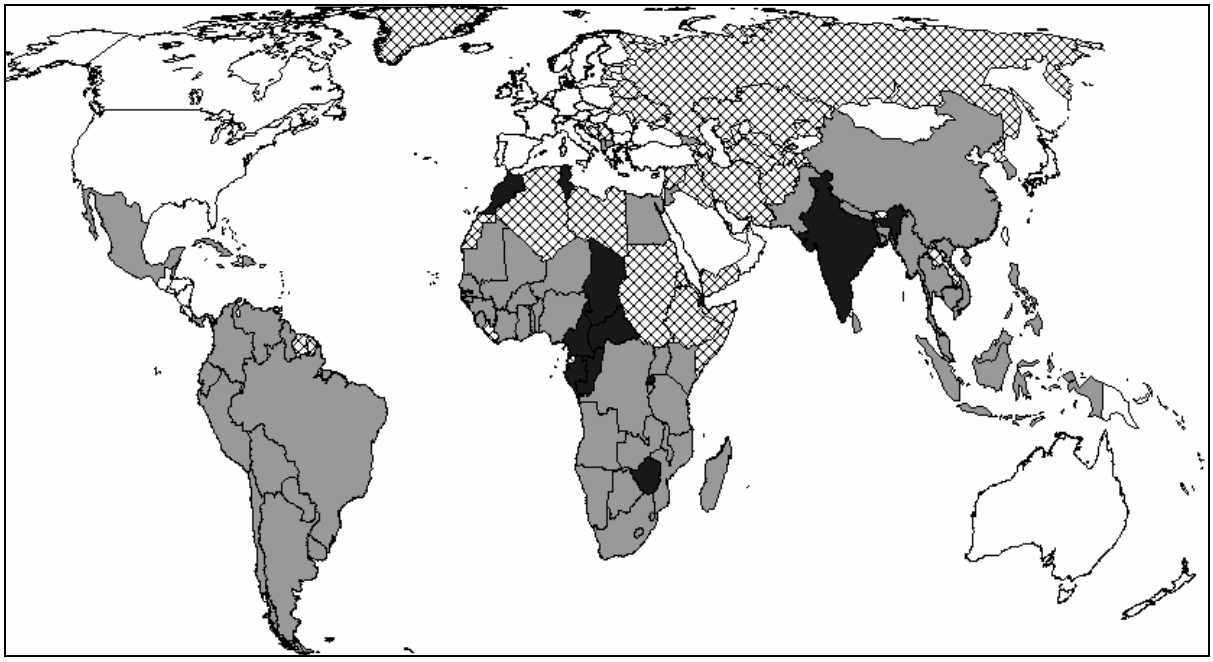
[9] NAMA products values are calculated by summing up sectors B, C, D and E.
[10] The NAMA groups, B, C, D and E include equivalent numbers of unbound products.

Map 1: Applied duties for agricultural products



Note: Economies with applied duties in average from 0 to 5% appear in white, between 5 and 15% in grey and higher than 15% in black. Crossed diagonals indicate that data are not available for the country.
Source: Authors' calculation on the basis of WTO tariff data.

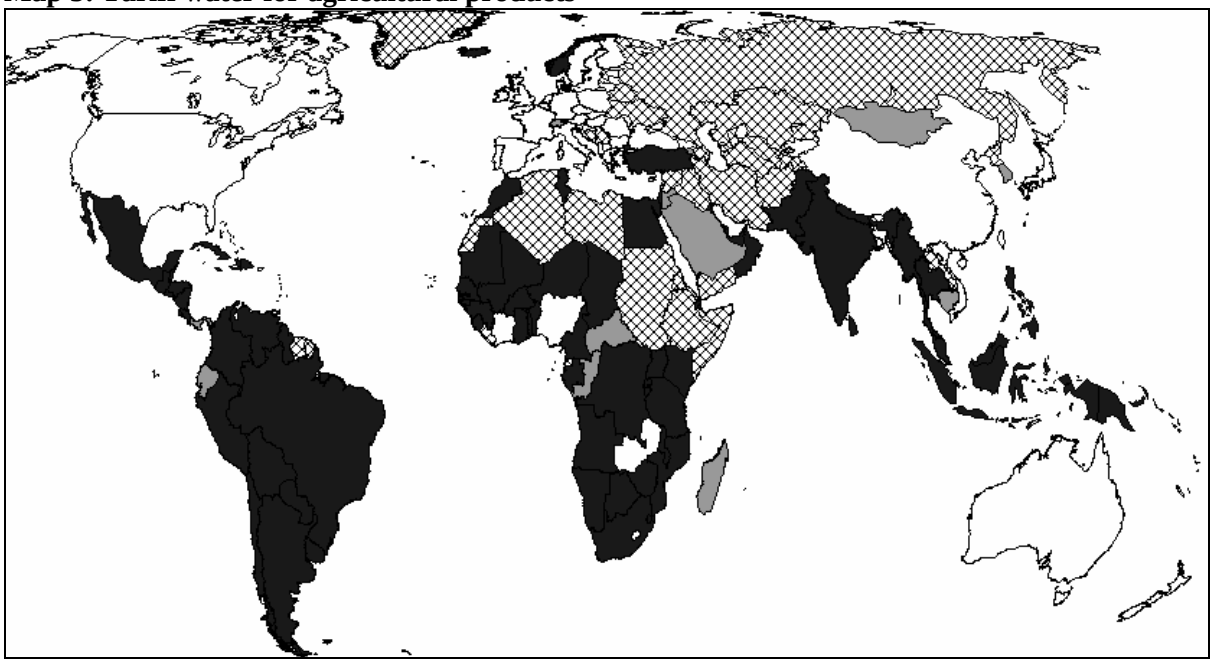
Map 2: Applied duties for NAMA products



Note: Economies with applied duties in average from 0 to 5% appear in white, between 5 and 15% in grey and higher than 15% in black. Crossed diagonals indicate that data are not available for the country.
Source: Authors' calculation on the basis of WTO tariff data.

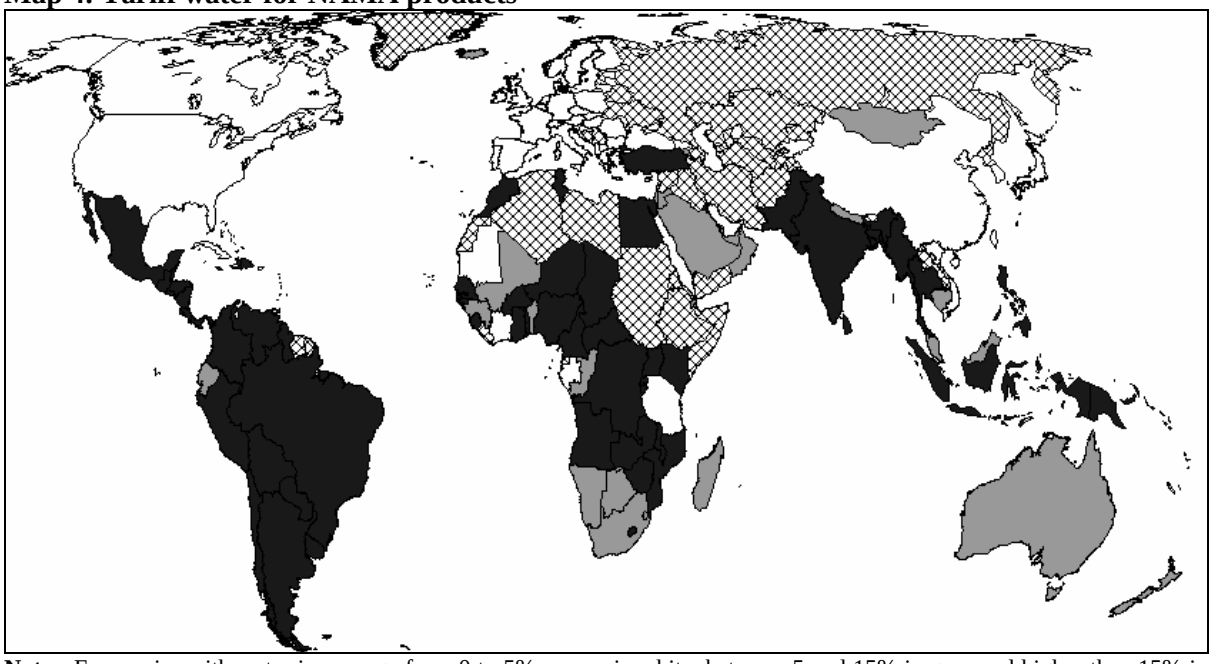
The visual exploration of the water in agriculture, see map 3, indicates that Latin America economies are characterised by deep water.

Map 3: Tariff water for agricultural products



Note: Economies with water in average from 0 to 5% appear in white, between 5 and 15% in grey and higher than 15% in black. Crossed diagonals indicate that data are not available for the country.
Source: Authors' calculation on the basis of WTO tariff data.

Map 4: Tariff water for NAMA products



Note: Economies with water in average from 0 to 5% appear in white, between 5 and 15% in grey and higher than 15% in black. Crossed diagonals indicate that data are not available for the country.
Source: Authors' calculation on the basis of WTO tariff data.

In most African and in some Asian countries, the margins in agriculture look also very high. Developed countries show rather low margins regardless their geographical location.

For several countries, the tariff water in NAMA (map 4) is lower than in agriculture; it is not the case for Australia, New Zealand and some African countries like Central African Republic, Guinea Bissau, Sierra Leone, Angola, etc. Apart from Latin American countries, there is no evidence that the water in developing countries is related to the geographical region. The visual exploration shows also the

diversity of situations regarding tariff schedules across countries and between agriculture and NAMA sectors.

3. Tariffs by stages of processing

Tariff policy is defined not only by the absolute level of market protection provided by applied and bound tariff, by also by the "effective protection" or "effective rate of assistance" provided by tariff escalation. Tariff escalation results in progressively higher import duties on semi-processed products than on raw materials, upwards to finished products. As mentioned by the Dictionary of Trade Policy Terms (Goode, 2003), this practice protects domestic processing industries and discourages the development of processing activity in the countries where the raw materials originate.

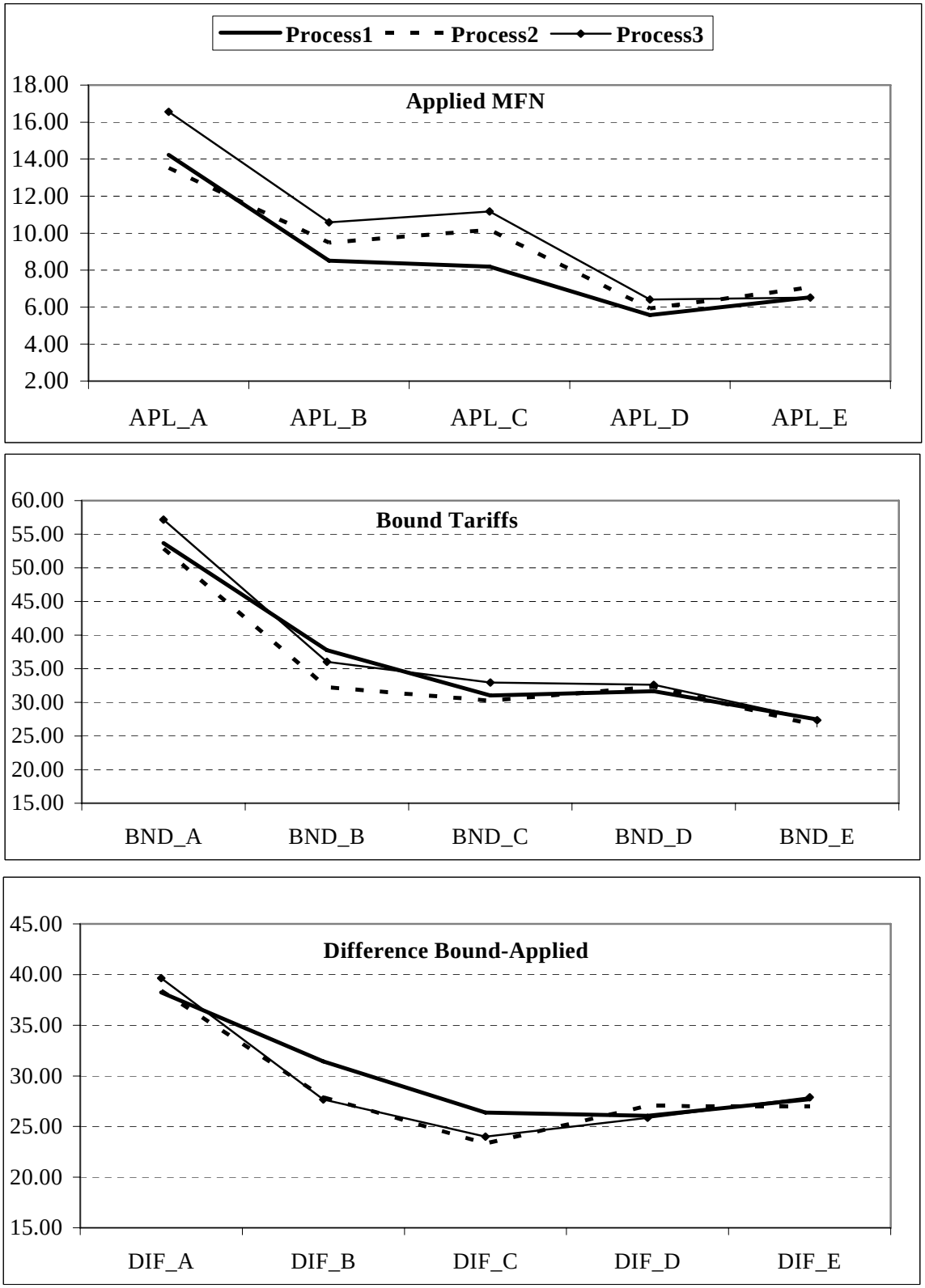
The extend of tariff escalation is explored by classifying products according to their stages of processing, from raw (stage 1) to partially processed (stage 2) and processed (stage 3) and computing the corresponding indicators of applied MFN, bound tariffs and water. The results presented in , show a complex pattern of escalation within product groups A, B, C and D in terms of applied duties, and within product groups A, C and D in terms of bound duties. A de-escalation pattern is noticeable across product groups for both applied and bound tariffs as duties decrease with the complexity and intensity of factoring (or, conversely, increase with the intensity of low skilled labour). [11] This pattern is more consistent with the protection of domestic labour than the promotion of industrialization through effective assistance. It should also be noted than those patterns observed for the sample of 120 members hide large variances across countries, and the coefficient of variation is always comprised between 0.7 and 1.

As shown in the first panel of **Figure 2** within each of the NAMA product groups, applied tariffs tend to be higher for the processed products (stages 2 and 3) belonging to group C (low skill and low technology manufactures). These labour intensive products, such as textile and clothing, receive a more significant protection than other manufactures in terms of tariff escalation. The tariff escalation pattern for bound averages is slightly different than for applied duties. Within each product group, countries tend to protect more the processed products (process3) but unprocessed goods tend to be more protected than semi-processed. Group B –raw materials- makes the exception, as the bound protection is higher for un-processed products (stage 1).

It should be also noted that the higher protection observed in panel 1 (applied tariffs) for labour intensive manufactures results more from an evolution of practices than an a priori policy, as the escalation graph for the bound tariffs (second panel) does not show significant differences between groups B, C and D.

[11] Comparing product groups, the relatively more technical intensive groups D and E have lower applied and bound tariffs than the other product groups; similarly, agriculture is always more protected than manufactured products. It should be noted that, when performing "within group" analysis, tariff escalation within the products "E" (high skill- and technology intensive manufactures) is not meaningful under the present classification in three stages, as this category includes practically no unprocessed articles.

Figure 2: Tariffs By Stage of Processing and Product Group, 2006



Notes: Three stages of processing, from low (Process 1) to high (Process 3). Product groups: A: Non-fuel primary commodities; B: Resource-intensive manufactures; C: Low skill and technology intensive manufactures; D: Medium skill and technology intensive manufactures; E: High skill and technology intensive manufactures.
Source: Authors' calculation on the basis of WTO tariff data.

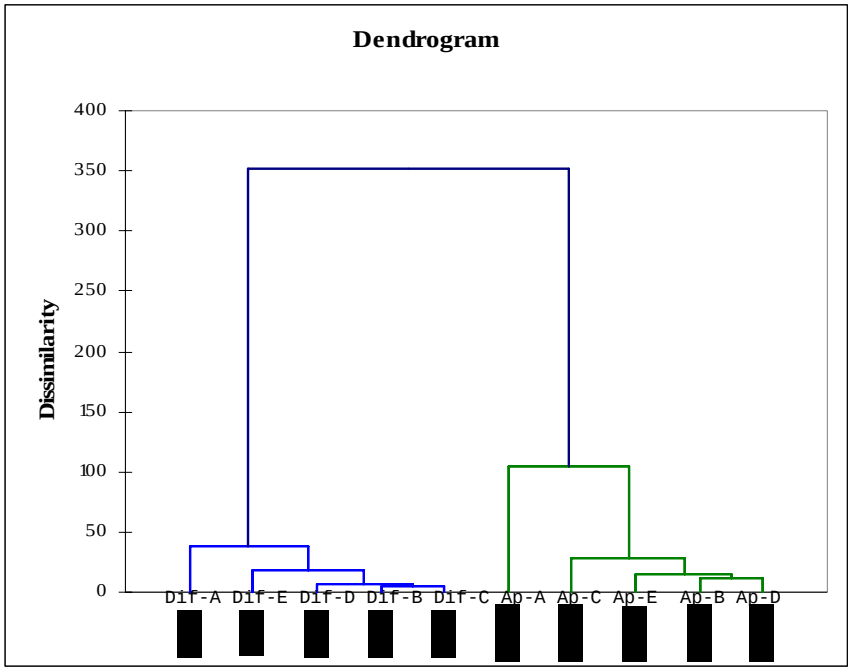
4. Looking for tariff policy patterns

After this first examination of the data, the present section is using EDA techniques to identify potential underlying structures in the data. The research is conducted in two directions: the variables (tariff duties) and the observations (countries). Structuring the variable space means looking for common tariff patterns across countries, while studying the observation space provides information on economies, based on their use of applied tariffs and binding overhang.

(a) Cluster analysis of tariff variables

The Agglomerative Hierarchical Clustering (AHC) of tariff variables is straightforward, as showed in Figure 3, and confirms the initial results of previous sections; the dendrogram should be read bottom-up by considering that the most similar objects are paired first. The tariff water presented at the left panel of the graph, clusters more rapidly than average tariffs, indicating more homogeneity than the applied MFN tariffs.

Figure 3: Hierarchical clustering of tariff variables



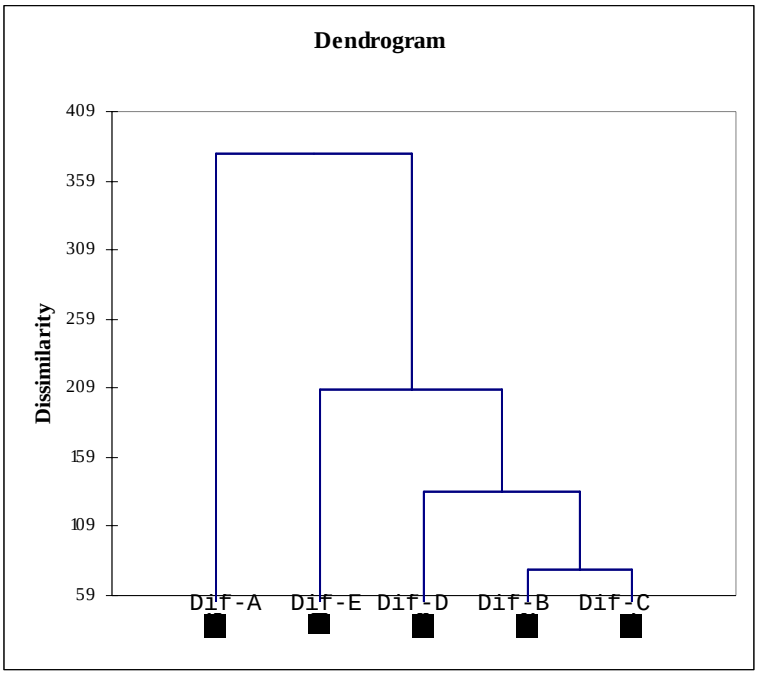
Notes: a/ Product groups: A, Non-fuel primary commodities; B, Resource-intensive manufactures; C, Low skill- and technology intensive manufactures; D, Medium skill- and technology intensive manufactures; E, High skill- and technology intensive manufactures
Source: Authors' calculation on the basis of WTO tariff data.

For both applied and water, agriculture stands out by its dissimilarity to other product groups. Its specificity measured by the dissimilarity index on the graph, is larger in the "applied tariff" dimension than in the binding overhang. Variables related to groups B, C and D are close in the water dimension, but less in the applied tariffs. Group C of low skill manufactures (textiles and clothing) differs from the other NAMA products, in the applied tariffs dimension.

(b) Cluster analysis of the economies

The analysis of the 120 "observations", included in our sample by visual examining a series of AHC experiments, indicated that countries could be "appropriately" grouped into 5 clusters (Figure 4). [12]

Figure 4: Hierarchical clustering of observations



Source: Authors' calculation on the basis of WTO tariff data.

The composition of the 5 clusters is defined by using the K-means clustering.[13] The classification as showed in Table 3, produces two large clusters, 1 and 2, including respectively 41 and 34 countries, and three medium sized clusters of about 15 economies. Clusters 1 and 4 are the most "compact", with the lowest within-class variance. Cluster 5 is the "loosest" one, with large within-class variance and high average distance to centroid. In other words, the typical tariff profile of classes 1 and 4 is shared by most of its members, while the association of class 5 participants with a "typical pattern" is fuzzier.

All clusters demonstrate relatively higher applied tariffs for group A (agriculture) and C (low skill/capital manufactures), although there is more dispersion concerning the treatment of agriculture. Product groups D and E (medium and high skill manufactures) tend to be generally less protected.

[12] Appropriate means that this number provided for reasonably balanced sub-samples, while maintaining significant distances between each cluster. As often in EDA, it was the result of a subjective judgement based on the interpretation of graphs. The following sections will show that this decision can also be based on quantitative information criteria.

[13] The procedure divides the dataset by minimising the within-group sum of squared errors in terms of the Euclidean distance from the group centroid. – i.e. the mean of variables included in the cluster. Because K-means is sensitive to initial partitioning, 10 simulations were run using different random initial partitions, and the best case was selected on the basis of minimum within-class variance, i.e., the partition giving the most compact clustering.

Table 3: Classification of economies in five clusters.

Cluster	Number of economies	Composition	Maximum distance to centroid
1	41	Albania, Armenia, Australia, Botswana, Côte d'Ivoire, Cambia, Canada, China, Congo, Croatia, Cuba, Ecuador, EC, FYROM Macedonia, Gabon, Georgia, Guinea, Hong-Kong, Japan, Jordan, Korea Rep., Kyrgyz, Macao, Madagascar, Mauritania, Moldova, Mongolia, Namibia, Nepal, New Zealand, Oman, Qatar, Saudi Arabia, Sing, South Africa, Swaziland, Switzerland, Taipei Chinese, UAE, USA, Viet Nam.	325
2	34	Argentina, Bahrain, Venezuela, Bolivia, Brazil, Brunei, C. African Rep., Chile, Costa Rica, Djibouti, Dominican Rep., Egypt, El Salvador, Fiji, Guatemala, Guinea Bissau, Haiti, Honduras, Indonesia, Maldives, Mexico, Morocco, Nicaragua, Panama, Papua NG, Paraguay, Peru, Philippines, Senegal, Sierra Leone, Sri Lanka, Thailand, Turkey, Urugay.	564
3	14	Benin Burkina Faso, Burundi, Colombia, Iceland, India, Israel, Malaysia, Mali, Myanmar, Niger, Norway, Tunisia, Zimbabwe.	1079
4	14	Angola, Antigua & Barbuda, Belize, Dominica, Ghana, Grenada, Guyana, Jamaica, Malawi, Pakistan, St. Lucia, St. Vincent, Trinidad & Tobago, Uganda.	456
5	17	Bangladesh, Barbados, Cameroon, Chad, Dem. Rep. Congo, Kenya, Kuwait, Lesotho, Mauritius, Mozambique, Nigeria, Rwanda, St, Kitts & Nevis, Solomon Isl., Tanzania, Togo, Zambia.	2434

Note: Classification based on applied tariffs and binding overhang in the five product groups A to E (see

Table 1). The maximum distance to centroid indicates possible outliers.
Source: Authors' calculation on the basis of WTO tariff data.

When contrasted with other groups, cluster 1 is characterized by comparatively low values for applied tariffs and water for all product groups (
Table 4). At the other extreme, cluster 5 is made of countries having both high applied tariffs in average (except in agriculture), and high binding overhangs. Class 4 is somewhat similar to class 5 albeit much more homogeneous, as seen above, but with lower values in both tariffs and waters.

Table 4: Tariff profile of the five clusters

Statistics:[a]	Average		Std. Deviation		Coefficient of variation	
Variables:	Applied	Water	Applied	Water	Applied	Water
1	7.7	5.7	2.7	2.4	0.4	0.4
2	10.5	24.4	3.1	1.2	0.3	0.0
3	12.6	25.4	5.4	19.8	0.4	0.8
4	10.8	51.2	3.3	11.5	0.3	0.2
5	12.5	76.9	2.9	10.1	0.2	0.1

Notes: a/ computed on the values of each class centroid for the five categories of products (A to E).
Source: Authors' calculation on the basis of WTO tariff data.

Compared with 4 and 5, clusters 2 and 3 shows similar high applied rates in all product groups, (and somewhat higher for agriculture) but lower water (except for agriculture). The difference between clusters 2 and 3 themselves is more subtle, and relates mostly to the dispersion of the tariff water rather than the applied level. Cluster 2 is characterized by more homogenous pattern of applied tariffs and water across the five product groups, while the third cluster shows higher variation, especially in tariff water.

The last step of the EDA on the general tariff profiles applies a principal component analysis to the variables and observations. Most of the variance (82%) is explained by the first two components

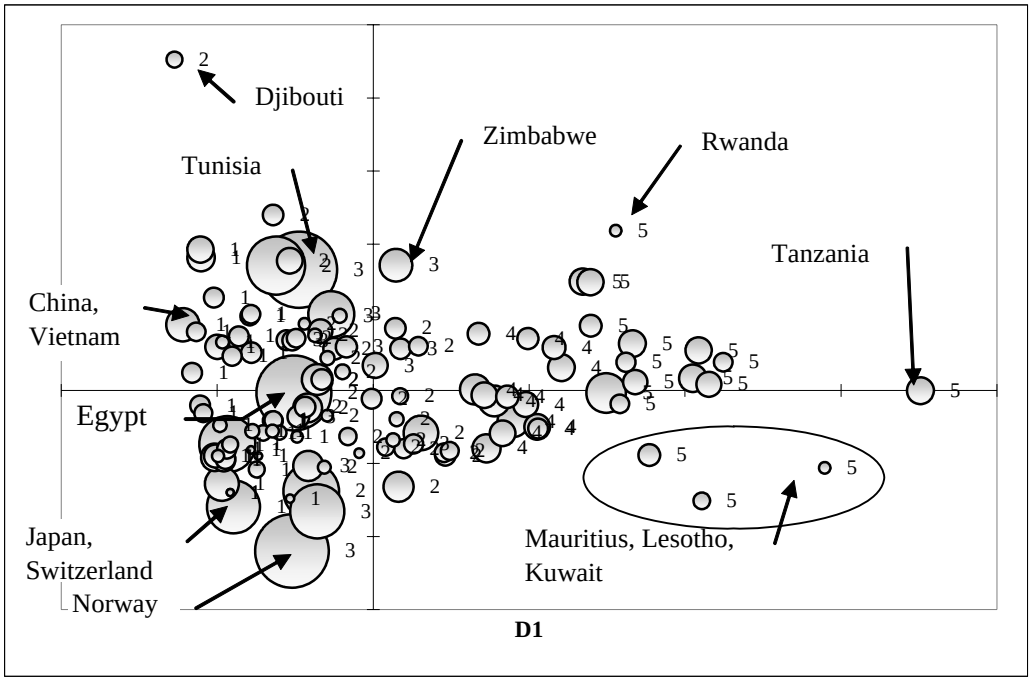
while the third explains only 8% of the variance. The first factor (53% of total variance) organises the observations by the level of binding overhang (see Table 5) and the second by the applied tariffs for non agricultural goods. Finally, the third dimension is determined by the level of applied tariffs in agriculture. Figure Figure 5 provides a projection of the 120 observations on a two dimensional graph, defined by the first two components as axis, and with the third component indicated by the size of the point.

Table 5: PCA factor loading

	D1	D2	D3
Applied:			
A	0.0	0.3	1.0
B	0.2	0.9	0.2
C	0.1	0.9	0.2
D	0.1	1.0	0.1
E	0.2	0.9	0.1
Water:			
A	0.9	0.1	0.1
B	1.0	0.1	0.0
C	1.0	0.1	0.0
D	1.0	0.2	0.0
E	0.9	0.2	0.0

Note: Factor loading after applying varimax rotation.
Source: Authors' calculation on the basis of WTO tariff data.

Figure 5: Principal Component Analysis of clusters



Notes: After varimax rotation. The numbers refer to the clusters, the size of the bubble indicates the score on D3.
Source: Authors' calculation on the basis of WTO' tariff data.

The ordering from left to right of the clusters according to their score in terms of binding overhang appears clearly on the graph. At the contrary, the second dimension of NAMA applied tariffs, does not discriminate between clusters: for each cluster, there are observations scattered across a large range of applied tariffs. The exception is cluster 4, which shows a clear concentration around the

average value. It appears also that clusters 2 and 3 are overlapping and could be merged without loosing too much information, at least according to the two principal components. Observation scoring relatively high on D3 (high applied in agriculture) tend to agglomerate on low value of D1.

From the projection, Djibouti appears as a clear outlier in the sample; at the contrary, Tanzania looks more as a extreme value for group 5, rather than a real outlier.[14] Mauritius, Lesotho and Kuwait are exocentric with respect to group 5, which visibly has the highest within-group variance (Table 3). Finally, Norway stands out for its low tariffs in non-agricultural products (second axis) and its high level of agricultural protection (third factor, shown by the size of the point in the graph).

B. TARIFFS BY SECTOR

Moving from general to specific, in this section the applied tariffs and water are explored by product sector. The research looks in every product sector for similarities and dissimilarities between the economies and investigates how the observations can be best regrouped (or split).

At first, the most evocative results obtained with clustering techniques are presented. Then, the negotiation margins are illustrated by following a simpler and more intuitive approach based on the relative position of each economy in relation to applied duties and water. The results for only two products are presented in this paper:

- group A -agricultural products-, for its specific role in partitioning the sample, and
- group C -textiles and clothing-, a symbolic NAMA product.

1. K-means cluster analysis for groups A and C

The k-means cluster analysis method is used again to partition the observations. The optimal number of clusters is defined on the basis of the Calinski - Harabasz pseudo-F index [15]. For group A and C the optimal number of clusters obtained is two and seven (see also footnote 13). [16]

In figure 6 are projected the 120 economies' average values of applied duties (horizontal axis) and binding overhang (vertical axis) in the agricultural sector. The cluster analysis regroups low overhang economies in cluster 1 and high overhang economies in cluster 2. The applied tariffs factor does not discriminate between clusters. It's worth mentioning that developed economies, with the exception of Iceland and Norway, are all included in cluster 1 which counts 80 economies.

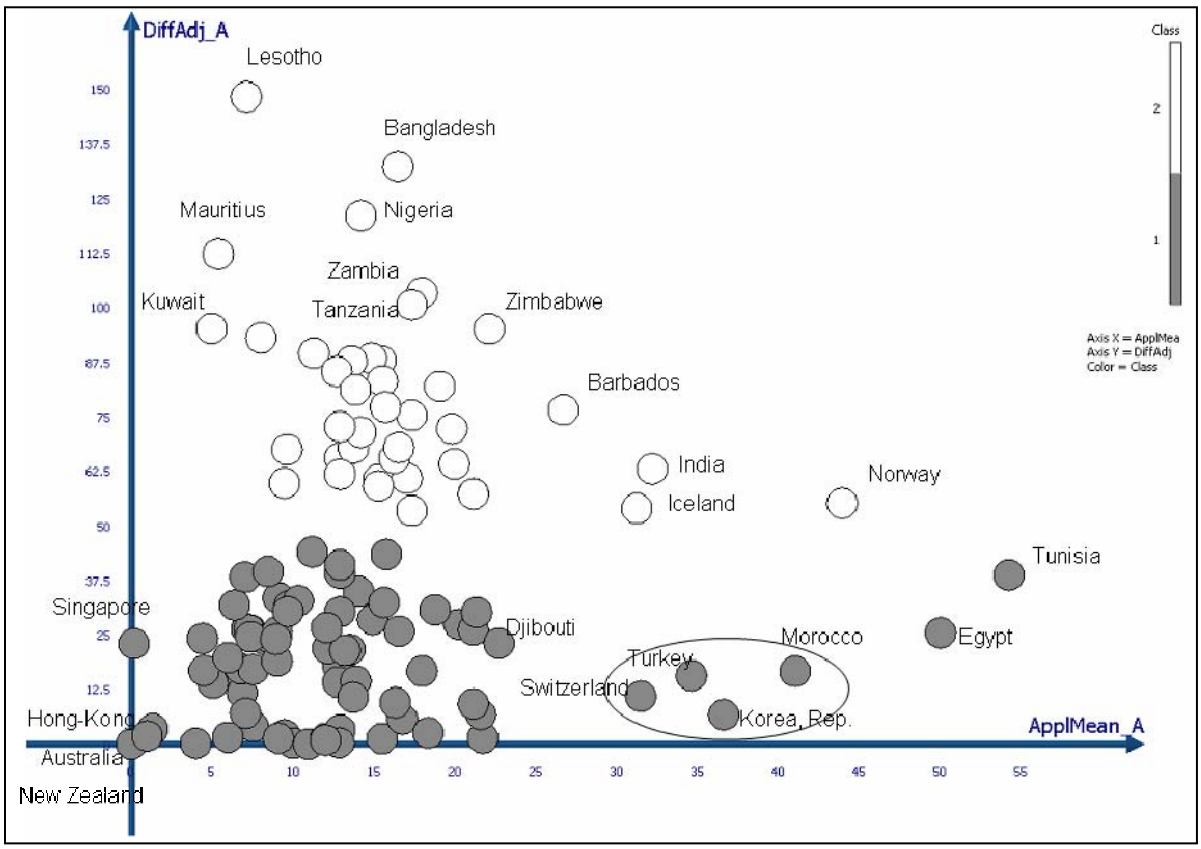
Extreme values lay at the top-left and bottom-right of the plot opposing Lesotho, Bangladesh, Nigeria and Mauritius to Tunisia, Egypt, Norway and Morocco. Djibouti and Tanzania depicted as extreme values/ outliers in the previous section are now laying near to the centres of clusters 1 and 2 and have rather a 'normal' tariff profile in this sector. Consequently, their agricultural tariff policy had to significant influence in their position in the global analysis.

[14] Reprocessing the data without Djibouti did not change substantially the results.

[15] The larger the index, the better the partition provided by the corresponding number of clusters. In the previous chapter, a more intuitive visual examination based on an initial hierarchical clustering was used to define this optimal number. The present approach is more objective, and easier to apply when many sub-groups have to be treated successively.

[16] For the remaining product groups, the obtained optimal number of clusters was 5 for group B, 4 for group D , 8 for both E and F.

Figure 6: K-means cluster analysis for agricultural products, product group A



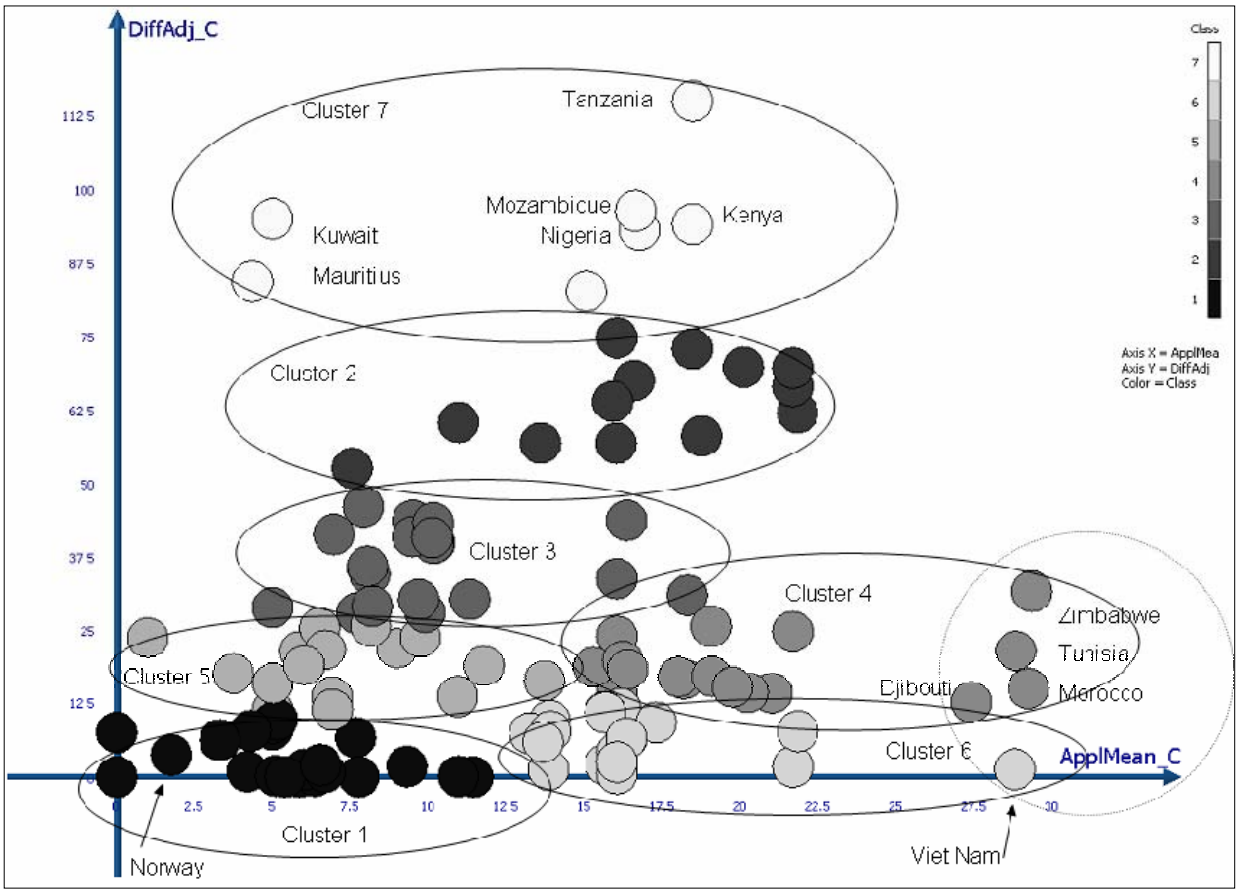
Source: Authors' calculation on the basis of WTO's tariff data.

The scatter diagram for the textiles and clothing in figure 7, gives a different picture of the 120 economies. Many data points have moved to the centre of the diagram compared to agriculture. The k-means cluster analysis regroups the data points into 7 clusters.

The clusters are well balanced and include equivalent number of countries. The smaller cluster is cluster 7, which includes only 7 economies and has the highest dispersion. Again, the binding overhang is the factor that basically discriminate the economies and forms the clusters.

Tanzania is an extreme value in this group of products, and has the highest overhang. It is included in cluster number 7 together with Kuwait and Mauritius. Djibouti, revealed as outlier in the global analysis shown in the previous section, lays at the bottom-right of the scatter together with Zimbabwe, Morocco and Viet Nam. Lesotho, almost an outlier for agriculture, has a 'normal' profile for textiles and clothing, and belongs to cluster 3. Norway has also a completely different behaviour in this product group, and is now part of cluster 1.

Figure 7: K-means cluster analysis for textiles and clothing, product group C



Source: Authors' calculation on the basis of WTO tariff data.

It is now clear that economies behave differently in their tariff policy according to the product sectors. In agriculture, economies with similar tariff policies form basically two groups; whereas in textiles and clothing sector, the situation is more complex: a larger variety of tariff policies is observed, leading to establish seven "policy" clusters.

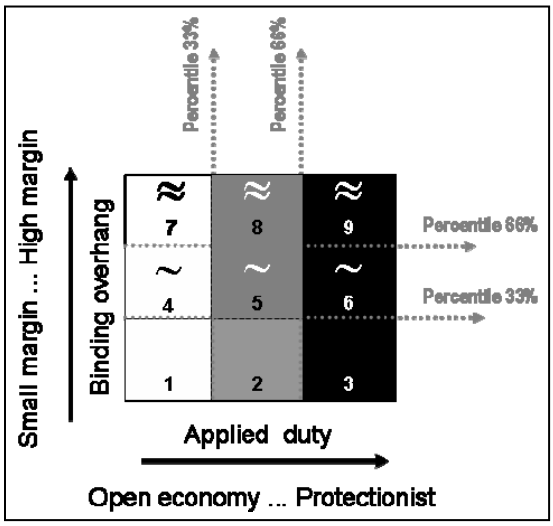
However, cluster analysis is a technique best used for dealing with multiple criteria; it is probably too sophisticated when dealing with only two dimensions, i.e. applied and water for a single product group. It is also sensitive to the presence of outliers. The following section offers a more intuitive and robust approach.

2. Illustrating the negotiation margins: groupings based on percentiles

In this section a different way to regroup countries is proposed: straightforward categories based on applied duties and tariff water percentiles are used to partition the data points. The 33rd and 66th percentiles [17] are used to regroup economies according to three categories (low / middle/ high) of applied duties and three categories (low / middle/ high) of binding overhang. These 3x3 categories split the scatter diagram of the 120 economies in 9 distinct groups (table 6) for every product group.

[17] Defined as the values below which stand 33% and 66% of all observations, respectively.

Table 6. Visual presentation of the percentile approach



Group 1 includes very open economies applying relatively low duties (below the 33rd percentile of applied duties of the sample) and having very shallow water (below the 33rd percentile of the binding overhang of the sample). Conversely, group 9 gathers very protectionist economies with very high applied duties (above the 66th percentile of applied duties of the sample) and very high margins (above the 66th of the binding overhang of the sample). Compared to cluster analysis, where groupings are based on abstract topological considerations, the percentile approach provides a more intuitive interpretation of the groupings and allows a cross-products / cross-countries comparison.

Table 7 shows the 9 percentile classes for agriculture. Classes 1 and 2 include the majority of developed countries and confirm, for example, the openness of the USA, Australia and New Zealand markets in the agricultural sector. EC and Canada are included in the second group as they apply in average, higher duties in agriculture than Class1 countries. It's interesting to observe that among developed countries the most protective in agriculture are Japan and Switzerland (class 3).

Table 7: The percentiles approach in product group A (Agriculture) - the composition of classes

	Members of the class	*Number*
CLASS 1	Albania, Armenia, Australia, Croatia, Cuba, Haiti, Hong-Kong China, Kyrgyz Rep., Macao China, Moldova, Mongolia, New Zealand, Oman, Qatar, Saudi Arabia, United Arab Emirates, United States	17
CLASS 2	Côte d'Ivoire, Canada, China, Ecuador, European Communities, FYROM Macedonia, Georgia, Madagascar, Panama, Peru, Senegal, Chinese Taipei	12
CLASS 3	Cambodia, Central African Republic, Congo, Japan, Jordan, Korea Rep., Morocco, Switzerland, Thailand, Turkey, Viet Nam	11
CLASS 4	Argentina, Bahrain, Botswana, Brazil, Brunei Darussalam, Chile, Guatemala, Honduras, Indonesia, Namibia, Nicaragua, Paraguay, Philippine, Singapore, South Africa, Swaziland, Uruguay	17
CLASS 5	Angola, Benin, Venezuela, Bolivia, Costa Rica, Dominican Rep., El Salvador, Guinea, Guinea Bissau, Mali, Mauritania, Nepal, Sierra Leone	13
CLASS 6	Djibouti, Egypt, Fiji, Gabon, Israel, Maldives, Mexico, Papua New Guinea, Sri Lanka, Tunisia	10
CLASS 7	Burundi, Kuwait, Lesotho, Malaysia, Mauritius, Myanmar	6
CLASS 8	Antigua and Barbuda, Burkina Faso, Colombia, Dem. Rep. of Congo, Malawi, Mozambique, Niger, Nigeria, Pakistan, Rwanda, St. Kitts and Nevis, St. Lucia, St. Vincent and Grenadines, Togo, Trinidad and Tobago	15
CLASS 9	Bangladesh, Barbados, Belize, Cameroon, Chad, Dominica, Ghana, Grenada, Guyana, Iceland, India, Jamaica, Kenya, Norway, Solomon Isl., Tanzania, Uganda, Zambia, Zimbabwe	19

Source: Authors' calculation on the basis of WTO tariff data.

It is also interesting to notice that Madagascar and Senegal are part of class 2, with very low water. The majority of developing countries are expected to be included in upper groups, where margins are higher. What is most probably unexpected is to find in the class 9 of very protective economies with very high margins Iceland, India and Norway. In the textiles and clothing sector, the majority of developed countries are, as expected, in class 1, except Australia and New Zealand which are in class 4. Class 1 is the most numerous and also includes many developing countries. [18]

[18] Albeit the uni-dimensional clustering is done using percentiles and is by definition balanced, the intersection of two uni-dimensional clusters needs not to be balanced (indeed, the intersection can be empty).

Table 8: The percentiles approach in product group C (Textiles and Clothing)- the composition of classes

	Members of the class	*number*
CLASS1	Armenia, Canada, Croatia, European Communities, Georgia, Hong-Kong China, Iceland, Japan, Korea Rep., Kyrgyz Rep., Macao China, Moldova, Norway, Oman, Saudi Arabia, Singapore, Switzerland, Chinese Taipei, United States	19
CLASS2	Albania, Cambodia, China, Cuba, FYROM Macedonia, Guinea, Jordan, Malaysia, Mauritania	9
CLASS3	Benin, Botswana, Burkina Faso, Burundi, Côte d'Ivoire, Congo, Gabon, Mali, Namibia, South Africa, Swaziland, Viet Nam	12
CLASS4	Australia, Brunei Darussalam, Chile, Haiti, Israel, Mongolia, Myanmar, New Zealand, Panama, Papua New Guinea, Qatar, Turkey, United Arab Emirates	13
CLASS5	Dominican Rep., Ecuador, Honduras, Indonesia, Madagascar, Paraguay, Peru, Philippine, Sri Lanka, Thailand, Uruguay	11
CLASS6	Argentina, Venezuela, Brazil, Central African Republic, Colombia, Djibouti, Egypt, India, Maldives, Mexico, Morocco, Nepal, Niger, Pakistan, Senegal, Tunisia	16
CLASS7	Angola, Bahrain, Costa Rica, Jamaica, Kuwait, Mauritius, Nicaragua, Trinidad and Tobago	8
CLASS8	Antigua and Barbuda, Barbados, Belize, Bolivia, Dem. Rep. of Congo, Dominica, El Salvador, Fiji, Ghana, Grenada, Guatemala, Guyana, St. Kitts and Nevis, St. Lucia, St. Vincent and Grenadines, Solomon Isl..	16
CLASS9	Bangladesh, Cameroon, Chad, Guinea Bissau, Kenya, Lesotho, Malawi, Mozambique, Nigeria, Rwanda, Sierra Leone, Tanzania, Togo, Uganda, Zambia, Zimbabwe	16

Source: Authors' calculation on the basis of WTO tariff data.

It's worth mention that some Least Developed Countries (LDC) appear in classes 2 and 3 - like Benin, Burkina Faso, Burundi, Cambodia, Guinea, Mali and Mauritania-, an unexpected result considering that these classes have very shallow water. This tends to indicate that tariff policies are not always directly related to the development factor, an hypothesis that will be further tested in the next section of the paper.

3. Negotiation margins: A cross-country and cross-product comparison

The percentiles approach divides in every sector the data points into 9 groups and consequently allows a cross-product / cross-countries comparison.

Table 9 shows the results for selected countries (all developed economies and selected CIS, developing countries and LDCs). For each country and group of product the applied duties (high, middle and low) are illustrated with colours (black, grey and white) and the water (large, middle and low) with waves (≈,∼,).

Developed countries are expected to naturally have low values in the water and therefore to be included in classes 1, 2 and 3. Conversely, LDCs are expected to have high values in the water and to a certain extent be included in classes 7, 8 and 9.

Surprisingly some of the developed countries have high applied duties (coloured black or grey) for some products –agriculture, textiles and clothing, pharmaceuticals- and have even high water (two waves). CIS countries show a 'developed economy' profile and appear as open as developed countries without benefiting from water in their tariffs.

In the other extreme, LDCs at the bottom-left of table 9 show a variety of profiles but globally apply high duties (with some exceptions) and do not always benefit from high margins in the binding overhang.

Developing countries, in the right part of

Table 9, have very different types of profiles, from 'developed country' profiles —Albania, Croatia, Hong-Kong China, Macao China, Singapore etc.— to 'LDC country' profiles —Antigua and

Barbuda, Bolivia, Dominica, Nigeria, St. Kitts and Nevis, St. Lucia, St. Vincent, Barbados, Belize, Grenada, Guyana, Kenya, Pakistan, Cameroon, Ghana and Zimbabwe.

Table 9: Negotiation margins for selected countries by product group

Developed economies and selected CIS and LDC countries						Selected developing economies					
Product group	A	B	C	D	E	Product group	A	B	C	D	E
Australia			~			Albania					
Canada						Argentina	~	~	~	~	~
European Communities						Brazil	~	~	~	~	~
Iceland	≈				~	Cameroon	≈	≈	≈	≈	≈
Japan						China					
New Zealand			~			Costa Rica	~	≈	≈	≈	≈
Norway	≈					Côte d'Ivoire					
Switzerland						Cuba					
United States						Ecuador		~	~		~
Armenia						Egypt	~	~	~	~	~
Georgia						Gabon	~				
Kyrgyz Rep.						Ghana	≈	≈	≈	≈	≈
Angola	~	≈	≈	≈	≈	Hong-Kong, China					
Bangladesh	≈	≈	≈	≈	≈	India	≈	~	~	~	~
Burkina Faso	≈	~		~		Israel	~		~		~
Burundi	≈	~		~	~	Jamaica	≈	≈	≈	~	≈
Cambodia		~				Jordan					
Central African Rep.		~	~	~	~	Kenya	≈	≈	≈	≈	≈
Djibouti	~		~			Mauritius	≈	≈	≈	≈	
Guinea	~	~		~		Mexico	~	~	~	~	~
Haiti		~	~	~	~	Morocco		~	~	~	≈
Madagascar			~	~	~	Nigeria	≈	≈	≈	≈	≈
Maldives	~		~	~	≈	Pakistan	≈	≈	~	≈	≈
Mali	~	~		~		Peru		~	~	~	~
Mauritania	~					Philippines	~	~	~	~	~
Myanmar	≈	~	~	~	~	Saudi Arabia					
Nepal	~	~	~	~		Singapore	~				
Niger	≈	≈	~	~	≈	South Africa	~				~
Senegal		~	~	~	~	Taipei,Chinese					
Sierra Leone	~	≈	≈	≈	≈	Thailand		~	~	~	~
Solomon Isl.	≈	≈	≈	≈	≈	Tunisia	~	~	~	~	~
Tanzania	≈	≈	≈	≈	≈	Viet Nam					
Togo	≈	≈	≈	≈	≈	Zimbabwe	≈	≈	≈	≈	~

Note: The level of applied duties, high, middle and low, are illustrated with black, grey and white and the level of water, large, middle and low, with waves (≈,~,).
Source: Authors' calculation on the basis of WTO tariff data.

IV. EXPLORING THE UNDERLYING ECONOMIC PATTERNS

This inventory of selected economies so called "*negotiation margins*" presented in the previous section tends to indicate that the variety of tariff policies followed by the countries are rather product-specific, and only loosely related to the economic development dimension. The present sections explore more thoroughly the relationship between trade policy and economic factors. Two complementary analyses are undertaken: the relationship between socio-economic variables and tariffs is explored initially at the product group level and then at the global level. Lastly, the socio-economic variables are used to determine the tariff policy and therefore, endorse any obtained analytical conclusions.

A. THE SOCIO-ECONOMIC DATA

A series of structural socio-economic variables and macroeconomic indicators are associated to the initial tariff indicators including the size of the economy in terms of population and GDP, the average per capita income, and the source of value added (agriculture, industry, services). Trade indicators cover openness, the relative share of trade in goods and services, and the main categories of imports and exports of merchandises (food, fuels and manufactures). Macroeconomic variables include inflation, current account balance, government expenditure and taxes. Other qualitative variables influencing the trade negotiations, such as the date of WTO membership, the level of development and the geographical location are also added. In total, 75 socio-economic variables are included. Most of the quantitative variables are grouped in a triplet including the 1995 and the 2006 levels, the variation from 1995 to 2000, and from 2000 to 2006. The socio-economic data are sourced from the World Bank 's World Development Indicators.

1. Exploring the socio-economic variables

Many of the socio-economic variables exhibit strong co-linearity, either because the indicators are partially redundant or because they are linked through a functional relationship.. Principal component analysis is applied to the initial dataset to reduce multi co-linearity and isolate, out of the 74 variables, the relevant variables to be retained. In order to avoid endogeneity issues, variables in level are measured in 1995, at the conclusion of the UR that fixed the bound tariffs, and values for 2006 are reflected through two sets of rates of growth (from 1995 to 2000, and from 2000 to 2006). A second analysis is performed with the 2006 level, in order to check the robustness of results (results did not differ substantially, except for the ordering of the components).

The first result is disappointing as no clear pattern can be established in the extended dataset. The main component, supposed to capture most of the information, explains less than 10% of the total variance. Any conclusion or interpretation of the results would have a very weak explanatory power: the four major principal components accounted for only 30% of the variance, and one would have to include up to 22 dimensions to explain at least 80% of the total variance.
With these caveats in mind, the socio-economic data set is structured by the following factors: (1) overall economic development dimension, (2) small oil and service oriented economies; (3) trade openness and macroeconomic variables.

- The first component F1 explains 10% of total variance. It is associated with rural countries on its positive segment, with a high share of agriculture in GDP, and low manufacturing and services contribution. This rural orientation remains extensive, with a low use of fertilizers: those are predominantly poor countries, as measured by the per capita income, and have a high population growth and a high share of food imports and fuel imports; They are

characterized by a high weight of inelastic imports, corresponding to a low share of manufactures in the imports. The governmental consumption has also a low incidence in GDP. As for their external sector, the countries with high scores for this dimension were highly indebted countries in 1995, but experienced a rapid decrease of their external indebtedness between 2000 and 2006. Because their large current account deficit has not improved markedly during the period, the reduction in external debt may be due to the impact of official assistance, such as the initiatives adopted in the framework of the Millennium Development Goals. Their exports are mainly agricultural products and low in manufactures. This dimension is closely associated with high levels of applied MFN (except in agriculture) and deep water (Table 10). Sub-Saharan Africa, as a region, is closely associated with positive values of F1.

- Countries with a high score for the second component F2 (capturing 8% of total variance) are associated with high inflation levels over the entire period, albeit decreasing between 2000 and 2006. They also tend to have high military spending in 1995, decreasing between 1995 and 2000. The share of industries predominantly extractive, representing in 1995 a large proportion of GDP, has further increased between 1995 and 2000. Correspondingly, the share of services in the economy has decreased markedly since 1995, and particularly between 1995 and 2000. Albeit a low share of agriculture in GDP in 1995, it has increased during the reporting period. This latter characteristic is not found in other components. Their external sector dominated by exports of fuels, has strengthened their current account position between 1995 and 2000. Despite registering no particular current account deficit in the base year, these countries had accumulated an important stock of external debts. Indebtedness decreased quickly between 1995-2000 and also, but less so, after 2000. Regional specificities are rather diffuse: Middle-East and African countries tend to be somewhat associated with the positive segment of F2, while CIS countries are closer to the negative values.

- The contribution of the two other axes in explaining the data structure is even lower (about 6% of total variance for each component) but reveals some additional discriminating factors. F3 shares some of the characteristics of F2 in the macroeconomic side (like high inflation) but is defined by larger countries (in population and total GDP), slightly poorer than the average in 1995 but which registered a high increase in the income per capita over the entire 1995-2006 period. The countries ranking high on this component tend to have a low participation of services in GDP, and no particular specialization in agriculture or in manufactures, at the difference of F1 (agriculture) and F2 (extractive industries) countries. External sector variables indicate that high scores on F3 are associated with relatively closed economies, with low trade coefficients in goods and, particularly, in services. They also register low values of foreign direct investment, which further decreased after 2000. F4 countries tend to suffer from moderate inflation, slowing down after 2000. At the difference of F3, those are small countries in terms of population and GDP. Despite some increase in the share of rural population, these countries show a sharp reduction in both demographic growth and share of agriculture in GDP, especially after 2000. Open to trade, they received high FDI flows. CIS countries tend to be associated with high scores on F4. No specific regional pattern has been identified for F3.

2. Associating tariff policy and socio-economic variables

(a) EDA on tariffs and socio-economic data sets.

The exploratory procedure combines the results of the tariff policy exploration (applied tariffs and water, clusters and scores) to the principal component analysis applied to the socio-economic variables. The additional variables are simply projected on the resulting principal components, but do

not interact with their computation. Thus these components are strictly determined by the socio-economic variables, as described in the previous section.

(i) Tariff by products and socio economic variables

At first, the applied tariffs and water for each product group are included as additional variables in the socio-economic Principal Component Analysis. As shown in Table 10, the only clear relation between tariff indicators (excluding agriculture) and socio-economic variables is the level of development, measured by the first component F1. But this remains a loose relationship as F1 explains only 10 per cent of the total variance. In addition, the correlation of NAMA tariff variables with the development variable F1 remains low (between 0.3 and 0.6). In addition, this closer relationship is limited to the applied level, and no strong systematic link exists between water and development. Finally, agriculture behave differently from the other product groups.

Table 10: Tariff variables and Socio-Economic Principal Components

(percentages and correlation coefficients)

	F1	F2	F3	F4
Variability (%)	9.9	8.4	6.3	5.6
Cumulative (%)	9.9	18.4	24.6	30.2
Factor loading:				
App_A	-0.02	-0.02	-0.08	-0.08
App_B	0.5	0.1	-0.1	-0.2
App_C	0.5	0	0	-0.3
App_D	0.4	0	-0.1	-0.2
App_E	0.6	0	-0.2	-0.2
Diff_A	0.3	0.1	-0.2	0
Diff_B	0.4	0.2	-0.2	-0.1
Diff_C	0.3	0.2	-0.2	-0.1
Diff_D	0.4	0.2	-0.2	-0.1
Diff_E	0.3	0.2	-0.2	-0.1

Notes: F1 to F4 are the four principal components extracted from socio-economic variables; applied tariffs and differences appear as supplementary variables. Factor loadings are similar to correlation coefficients in a normalized PCA.
Source: Authors' calculation on the basis of WTO tariff data.

The second socio-economic factor, F2 explaining 8% of total socio-economic variance, differentiates oil exporters from the other economies. Its impact on tariff structure is almost nil, as it is also the case for the remaining two principal components F3 and F4.

(ii) Overall tariff pattern and socio-economic variables

The previous section explored the relationship between socio-economic structure and tariffs for each individual product group. But a country's trade policy is defined by considering the entirety of products. This holistic approach led to 5 tariff policy clusters, as seen in previous sections of this paper. The relationship between these 5 clusters and the principal components representing the socio-economic variables is even more diffuse than for the product groups, with a correlation between 0 and 0.3 in absolute value (Table 11). Therefore, no clear and strong relationship can be identified between the overall tariff policy and broad socio-economic variables.

Table 11: Tariff policy clusters and Socio-Economic Principal Components

(percentages and correlation coefficients)

	PCA on Tariffs:			Tariff Policy Clusters:				
Factors:	**D1**	**D2**	**D3**	**1**	**2**	**3**	**4**	**5**
F1	0.3	0.5	-0.2	0.0	-0.3	0.2	0.0	0.1
F2	0.2	0.0	0.0	-0.1	-0.1	0.1	0.0	0.1
F3	-0.2	0.0	-0.1	0.0	0.2	-0.1	0.0	-0.1
F4	0.0	-0.2	0.0	-0.1	0.1	-0.1	-0.2	0.3
Selected variables:								
Fertilizers 2006	0.1	-0.4	0.0	0.0	-0.1	0.2	-0.1	0.0
GNI per cap. 1995	-0.1	-0.5	0.1	0.2	-0.2	0.1	-0.1	0.0
Trade services/GDP1995	0.3	0.0	0.1	-0.2	-0.1	-0.1	0.3	0.1
Export Food 1995	0.3	0.1	-0.1	-0.3	0.1	-0.1	0.4	0.0
Agricultural GDP 1995	0.1	0.4	-0.1	-0.1	-0.1	0.1	0.0	0.1
Ext. Debt 1995	0.2	0.3	-0.1	-0.1	0.0	-0.1	0.1	0.1
Manufacture GDP 1995	-0.3	-0.3	0.1	0.1	0.2	0.0	-0.3	-0.2

Notes: F1 to F4 are the four principal components extracted from socio-economic variables; applied tariffs and differences appear as supplementary variables. Factor loadings are similar to correlation coefficients in a normalized PCA.
Source: Annex 1.

The strongest relationship observed between clusters and individual socio-economic variables never exceeds 0.5 in absolute value (Annex 1). In addition, the robustness of the results was checked by substituting the 1995 levels by the final 2006 levels, endorsing the weakness of the conclusions: we are more in a situation of nuances than of contrasts. With these caveats in mind, the few significant relations between tariff policy clusters and socio-economic variables that can be mentioned refer to Cluster 1. This subset consists of economies with low applied duties and water in all products, and regroups recently acceded WTO members and developed countries characterized by a urban population, mature demography, low inflation, a high per capita GNI and a low weight of agriculture in both production and trade. Even though no clear-cut regional identification can be associated, these countries are mainly located in Europe, including the CIS.

Characterizing other clusters is much more tentative. Cluster 2 is somehow symmetric to cluster 1 from the socio-economic variables perspective; the correlation between the two clusters is -0.5. It contains relatively poor countries, with low incidence of public services, located in Latin America or the Caribbean. Cluster 3 economies distinguish themselves from the previous cluster countries for their higher governmental consumption, an increase of agriculture share in GDP and more intensive use of fertilizers. This is the only cluster associated with high applied tariffs in agriculture. Cluster 4 economies having high tariffs and deep water in both agriculture and NAMA, appear to be relatively poor rural developing countries, mainly food exporters suffering from inflation and external debt. These economies have nevertheless benefited from FDI flows. Cluster 5 economies are also poor countries, predominantly rural and very similar to cluster 4 economies with high water in both agricultural and NAMA sectors. They are still far from achieving demographic maturity, with high population growth and have benefited from a reduction in their external debt.

Once again, these correlations are weak, and do not define clear-cut socio-economic pattern. It is clear that the national socio-economic dimensions captured by the principal components do not improve significantly our understanding of national tariff policies. But, as often in statistics, the null hypothesis is not deprived of significance and despite the apparent weak outcome, this is an important result. There is no over-determination on the overall tariff policy by broad economic considerations: the decision-making process governing the definition of national tariff policy is the result of more complex interactions.

(b) Discriminating tariff patterns by using economic variables

The final step of the analysis consists in investigating whether socio-economic variables, despite the low individual correlation observed in the previous sections, can still reliably determine the five tariff policy profiles. As a first step, a discriminant analysis investigates for each country, if its initial clustering (based on purely tariff information) could have been inferred on the basis of socio-economic data. The confusion matrix, obtained by comparing the prior classification (initial clusters built on purely tariff data) to the results deduced from the socio-economic variables, provides some interesting information, especially on countries where the classification based on tariff structure appears unusual considering the economic situation.

Table 12: Confusion matrix for the five clusters (discriminant analysis based on socio-economic variables)

1. Full sample

from \ to	1	2	3	4	5	Total	% correct
1	32	1	1	0	0	34	94.1%
2	0	30	0	0	1	31	96.8%
3	1	1	9	0	1	12	75.0%
4	0	1	1	11	0	13	84.6%
5	1	0	0	0	13	14	92.9%
Total	34	33	11	11	15	104	91.3%

2. Cross validation

from \ to	1	2	3	4	5	Total	% correct
1	25	3	4	0	2	34	73.5%
2	0	26	2	2	1	31	83.9%
3	2	3	7	0	0	12	58.3%
4	0	3	2	8	0	13	61.5%
5	2	2	2	1	7	14	50.0%
Total	29	37	17	11	10	104	70.2%

Source: Authors' calculation.

Overall, most countries are found to be properly classified in the 5 clusters, except for cluster 3 where 25% of the countries seem misplaced according to their socio-economic characteristics (see part 1 of table 8).

To cross validate the results and verify the robustness of the initial classification, each country has been re-classified, after removing it from the sample (see part 2 of **Table** 12). Thanks to this "jacknifing" process, the discriminant function is computed independently of the particular country being classified. [19] The difference between the full sample analysis results and the cross validation results is inversely dependent on the size of the sample and the degree of the relationship: small size clusters or weak relationships result in large differences.

Surprisingly, cluster 1 —the largest one and the most closely associated to the principal dimensions of the socio-economic data— reports 74% of correct assignments compared to 94% in the full sample analysis. The difference of 20 percentage points indicates the wide variance of socio-economic

[19] The general idea behind jacknifing in statistics lies in systematically recalculating the estimates leaving out one observation at a time from the sample set; the method is normally used to estimate the sampling distribution of a statistic. It was used here with the purpose of deriving more robust estimates of the discriminant function.

characteristics that exist in this cluster, including the recently acceded members to the WTO. Cluster 2 appear to be the most robust with 84% of correct assignments.

The remaining smaller size clusters (less than 15 observations), are more vulnerable to the validation procedure. Cluster 5 is also the loosest one, as only half of its members are correctly classified according to the validation procedure. Cluster 5 looks more like a residual grouping of very different data points sharing a few common features (in particular, a low development level). An interesting feature of cluster 3 is that all reassigned countries went to clusters 1 and 2; cluster 3 mainly includes developed countries with relatively protectionist tariff policies.

When looking at individual economies, in the group of developed countries Norway and Iceland are reassigned from Cluster 3 to 1, Japan and Korea from 1 to 2. On the side of developing countries, Kuwait and unexpectedly Bangladesh move from Cluster 5 to 1 and Guyana from 4 to 2. Moves from 3 to 2 relates to relatively advanced developing countries such as Colombia or Tunisia. On the other hand, Gabon and Madagascar are reclassified from 1 to 3 and Botswana, South Africa from 1 to 5 (the re-assignation of SSA to the 5th cluster is nevertheless loose, as the discriminant function identified a 25% probability of pertaining to group 1). China and Guinea move from 1 to 3, and Nicaragua from 2 to 4.

Additional explorations incorporated the regional dimension. Geographical variables improve the direct classification, but deteriorate the cross-validation results and show that regional groupings are poor predictors of tariff patterns. This result (not shown here) was already perceptible earlier in the study, when looking at the large differences across patterns in maps 1 to 4.

V. CONCLUSIONS

In the present study 120 economies are analysed in the basis of their applied duties and binding overhang; the main purpose is to illustrate the margins that economies, under WTO rules, are allowed to use unilaterally at the MFN level, and to identify and understand potential interactions determining tariff policies. The analysis, which proceeds from general to specific, applies exploratory data analysis to identify underlying patterns behind the data, without any pretence at identifying the rational behind the observed structure. Five broad clusters of countries are determined on the basis of their tariff policy, and the relationship between this classification and socio-economic structures is analysed.

A first element of conclusion is that tariff policies seem to protect more the labour intensive sectors, with some variations in the case of agriculture, rather than pursuing traditional protectionist industrial policies through "effective protection". A second element of conclusion is that in tariff policies, the distinction between developing and developed countries is broadly relevant, but not overly determinant. For example, while the extreme clusters 4 and 5 are constituted of developing countries, there is a mixture of developed and developing economies in all remaining clusters. A third conclusion is that agricultural tariff policy becomes a strong discriminating factor across countries as soon as other variables, such as the level of development level, are taken into consideration.

When the product sectors are considered independently, economies seem to follow different tariff strategies to protect their products. Globally and for the majority of countries, the agricultural sector is more protected than the NAMA in terms of applied duties and water, while textiles and clothing is the most protected NAMA sector. For a given country, the tariff structure varies across sectors. Even if developed economies tend to share more common features than developing ones, no clear relationship can be established between the level of development and the overall tariff structure.

From the negotiation margins perspective, developing countries that have joined the WTO recently faced stricter negotiations with their partners resulting in relatively low applied duties and water; as a result, these countries, like the developed economies, have a reduced negotiation margins in almost all sectors.

On the other side of the spectrum, high negotiation margins are not a privilege for LDCs; some developing and developed countries benefit from deep water in some sectors. Unexpectedly, some LDCs show very small negotiation margins as they apply low duties and have only shallow water in some of the product sectors. On the other hand, some developed countries make use of high applied duties in addition to deep water in agriculture. Globally, applied tariffs and water respond to a different logic.

By exploring the relationship between the tariff policy and the socio-economic dimensions, the only clear relation which emerged for individual product groups (excluding agriculture) is the level of development. Specifically, in the NAMA sector, the applied tariffs for raw materials and pharmaceutical/hi-tech products are highly correlated to the level of development. However, the relationship remains limited to the level of the applied tariffs; no strong systematic link has been identified between the water and the socio-economic dimension. Agriculture behaves also differently from other product groups and its level of protection seems not related to the level of development.

The relationship between the overall tariff policy (considering all product groups together) and the socio-economic variables is even more diffuse, and no strong relationship emerges between tariff policy clusters and the socio-economic context. Consequently, we can conclude that trade policy is not over-determined by economic considerations: the decision-making process defining a precise trade policy is the result of more complex interactions.

At last, discriminant analysis is used to validate the robustness of the clustering results and investigates to what extend the belonging of countries to a tariff policy cluster could be predicted based only on their socio-economic profile. The first and most numerous cluster of low protection economies with shallow water, is particularly affected as many countries are reclassified in other clusters. This indicates the wide variance of socio-economic conditions within the cluster, but also the bias created by the countries of recent accession to the WTO, with a tariff structure resulting from strict bilateral negotiations rather than socio-economic factors. Conversely, a few industrialised countries show a wider than expected dispersion of their tariff in relation to their development status.

It is often believed that national trade policies are largely determined by socio-economic concerns, and that structural factors and factor endowments, which orient the productive specialisation of economies should determine the position of each country in the negotiation, as well as its propensity in entering into strategic coalitions. By showing that actual tariff structures are only loosely determined by such structural factors, our results do not support these premises.

* *
*

BIBLIOGRAPHY

Bchir M. H., Jean S. and Laborde D. (2005), " Binding overhang and tariff cutting formulas". CEPII 2005-18.

Costantini V., Crescenzi R., De Filippis F. and Salvatici L. (2007), "Bargaining Coalitions in the WTO Agricultural Negotiations". The World Economy pp. 863891.

Decreux Y. and Fontagné L., (2009), " Economic impact of potential outcome of the DDA". CEPII 2009-01.

Foletti L., Fugazza M., Nicita A. and Olarreaga M. (2009) "Smoke in the Water". World Bank, Policy Research Working Paper 5050.

Fontagné L., Guérin J-L. and Jean S. (2005), " Market Access Liberalisation in the Doha Round: Scenarios and Assessment". Blackwell Publishing Ltd 2005, p. 1073-1094.

François F. and Martin W. (2004),"Commercial policy variability, bindings, and market access". European Economic Review 48 (2004) 665-679.

François F., Martin W. and Manole V. (2005),"Choosing Formulas for Market Access Negotiations: Efficiency and Market Access Considerations" World Bank, Policy Research Working Paper Series number 3474.

Francois J., Van Meijl H., Van Tongeren F (2003), " Trade Liberalization and Developing Countries under the Doha Round". Tinbergen Institute TI 2003-060/2.

Gibson P., Wainio J., Whitley D., And Bohan M. (2001), "Profiles of tariffs in global agricultural markets". Agricultural Economic Report No. (aer796) 44 pp, January 2001.

Goode W. (2003), Dictionary of Trade Policy Terms, Fourth Edition.

Jones K. (2007), "Regionalism and the problem of representation in the WTO". UNU-CRIS Working Papers W-2007/12.

World Bank - World Development Indicators (on-line data base, data extracted in July 2009).

WTO - Integrated Data Base, Consolidated Tariff Schedules database (on-line data bases, extracted in February 2008)

WTO - UNCTAD - ITC (2006) World Tariff Profiles.

VI. ANNEXES

A. ANNEX 1: CORRELATION TABLE BETWEEN SELECTED VARIABLES AND PRINCIPAL COMPONENTS.

	PCA on Tariffs:			PCA on Socio-economic data:				Tariff Policy Clusters:					Regions:							Recent accession	
Variables or components	D1	D2	D3	F1	F2	F3	F4	1	2	3	4	5	ASPA	CIS	EUR	LAC	MENA	NAM	SSA	RAM-0	RAM-1
Inflation (2000)	0.1	-0.1	0.0	0.2	0.5	0.4	0.4	-0.1	0.0	-0.1	0.2	0.0	-0.1	0.0	-0.1	0.0	0.0	0.0	0.2	-0.2	0.2
Inflation (2006)	0.2	0.0	0.0	0.3	0.5	0.5	0.3	-0.2	0.1	-0.1	0.3	0.0	-0.1	0.0	-0.1	0.0	-0.1	-0.1	0.2	0.0	0.0
dInflation(2000-2006)	-0.1	0.1	0.0	-0.2	-0.5	-0.4	-0.4	0.1	0.0	0.0	-0.2	0.1	0.1	0.0	0.0	0.1	0.0	0.0	-0.1	0.2	-0.2
FDI (1995)	0.2	-0.2	-0.1	0.0	0.2	-0.2	0.4	-0.1	-0.1	-0.1	0.3	0.1	0.0	0.0	-0.1	0.2	-0.1	0.0	-0.1	0.0	0.0
dFDI (2000-2006)	0.0	0.0	0.2	0.0	-0.3	-0.5	0.1	-0.1	0.0	0.1	0.1	-0.1	0.0	0.1	0.0	0.1	0.2	-0.1	-0.2	-0.2	0.2
Military 1995	0.0	-0.1	-0.1	-0.3	0.4	0.0	0.2	0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	-0.2	0.6	0.0	-0.1	-0.2	0.2
dMilitary(1995-2000)	-0.1	0.1	0.1	0.3	-0.4	-0.1	-0.4	-0.1	0.2	0.1	-0.1	-0.1	0.0	0.0	-0.2	0.1	-0.2	0.0	0.1	0.2	-0.2
Population 1995	-0.1	0.1	0.1	-0.2	-0.1	0.4	-0.2	0.1	-0.1	0.1	-0.1	0.0	0.3	-0.1	-0.1	-0.1	-0.1	0.1	-0.1	0.0	0.0
dPopulation (1995-2000)	0.3	0.4	-0.2	0.4	0.4	-0.2	-0.4	-0.3	0.0	0.1	-0.1	0.3	-0.1	-0.4	-0.4	-0.2	0.1	-0.1	0.5	0.1	-0.1
dPopulation (2000-2006)	0.3	0.4	-0.2	0.6	0.3	-0.2	-0.4	-0.3	0.1	0.2	0.0	0.2	-0.1	-0.4	-0.3	-0.2	0.1	-0.1	0.6	0.1	-0.1
Rural 1995	0.3	0.4	0.0	0.7	-0.2	0.1	0.0	-0.2	-0.1	0.0	0.2	0.2	0.1	0.0	-0.2	-0.2	-0.3	-0.2	0.5	-0.1	0.1
Trade/GDP 1995	0.1	-0.2	-0.1	-0.2	0.2	-0.4	0.5	0.0	-0.1	-0.1	0.2	0.0	0.1	0.0	-0.1	0.0	0.0	-0.1	-0.1	-0.1	0.1
Current Balance 1995	-0.2	-0.2	0.1	-0.5	0.0	0.1	-0.3	0.1	0.1	0.0	-0.1	-0.1	0.2	-0.1	0.2	-0.1	0.2	0.1	-0.2	0.1	-0.1
dCurrent Bal (1995-2000)	0.0	0.0	0.0	0.0	0.7	0.0	0.1	0.1	-0.1	0.0	-0.1	0.1	0.0	0.0	0.0	-0.2	0.2	-0.1	0.1	-0.1	0.1
Fertilizers 2006	0.1	-0.4	0.0	-0.5	0.2	-0.3	0.1	0.0	-0.1	0.2	-0.1	0.0	0.1	-0.1	0.1	-0.1	0.3	-0.1	-0.3	0.1	-0.1
GNI per cap. 1995	-0.1	-0.5	0.1	-0.8	0.2	-0.2	-0.1	0.2	-0.2	0.1	-0.1	0.0	0.0	-0.1	0.4	-0.1	0.2	0.3	-0.4	0.2	-0.2
dGNI/H (2000-2006)	-0.1	-0.1	0.0	-0.1	0.0	0.6	0.5	0.2	-0.1	-0.1	0.0	0.0	0.2	0.4	0.0	-0.1	0.0	-0.1	-0.2	-0.3	0.3
d2GNI/H (1995-2006)	0.1	0.0	-0.1	0.1	0.1	0.5	0.1	0.0	0.0	-0.1	0.0	0.1	0.1	0.2	-0.2	-0.1	0.1	-0.2	0.0	-0.1	0.1
Gov. Consumption 1995	0.0	-0.3	0.0	-0.4	0.0	-0.2	0.3	0.1	-0.3	0.2	0.1	0.0	-0.2	0.1	0.2	-0.1	0.3	0.0	-0.1	-0.1	0.1
Food imports 1995	0.1	0.3	-0.1	0.6	-0.1	-0.4	0.2	-0.1	-0.1	0.0	0.2	0.1	-0.3	0.1	-0.1	0.0	0.1	-0.2	0.3	-0.2	0.2
Manufacture imports 1995	0.1	-0.2	0.0	-0.5	0.2	0.0	-0.3	-0.1	0.0	0.0	0.0	0.1	0.1	-0.4	0.1	0.2	-0.1	0.2	-0.2	0.2	-0.2
Trade goods/GDP1995	0.1	-0.2	-0.1	-0.2	0.3	-0.3	0.5	0.1	-0.1	-0.1	0.1	0.0	0.1	0.0	-0.1	-0.1	0.1	-0.1	0.0	0.0	0.0
Trade services/GDP1995	0.3	0.0	0.1	0.0	0.1	-0.6	0.5	-0.2	-0.1	-0.1	0.3	0.1	0.0	-0.1	-0.1	0.2	0.0	-0.1	-0.1	-0.1	0.1
Export Food 1995	0.3	0.1	-0.1	0.4	-0.3	-0.3	0.1	-0.3	0.1	-0.1	0.4	0.0	-0.2	0.0	-0.1	0.4	-0.2	-0.1	0.1	0.1	-0.1
Export Fuels 1995	0.0	0.0	0.0	-0.1	0.8	0.0	0.0	0.1	-0.1	-0.1	0.0	0.0	-0.1	0.0	0.0	-0.1	0.3	0.0	0.0	-0.2	0.2
Export manuf. 1995	-0.2	-0.2	0.2	-0.6	-0.2	0.3	0.1	0.2	-0.1	0.0	-0.1	-0.1	0.4	0.0	0.2	-0.2	0.0	0.2	-0.4	0.0	0.0
Agricultural GDP 1995	0.1	0.4	-0.1	0.8	-0.4	0.1	0.1	-0.1	-0.1	0.1	0.0	0.1	0.0	0.3	-0.1	-0.3	-0.2	-0.2	0.4	-0.2	0.2
dAg-GDP (1995-2000)	0.1	0.1	0.1	-0.1	0.2	-0.2	-0.4	-0.2	0.2	0.0	0.0	0.0	0.0	-0.4	-0.1	0.0	0.1	0.1	0.2	0.3	-0.3
Ext. Debt 1995	0.2	0.3	-0.1	0.6	0.4	-0.1	0.1	-0.1	0.0	-0.1	0.1	0.1	-0.2	-0.1	-0.2	0.0	0.0	-0.1	0.4	0.0	0.0
dExt. Debt (1995-2000)	-0.2	-0.1	-0.1	-0.2	-0.6	0.1	-0.1	0.1	0.0	0.1	-0.1	-0.1	0.1	0.3	0.1	0.0	0.0	0.0	-0.2	0.0	0.0
dExt. Debt (2000-2006)	-0.2	-0.3	0.1	-0.6	-0.2	-0.1	0.1	0.0	0.0	0.0	0.0	-0.2	0.2	-0.1	0.2	0.3	0.0	0.1	-0.5	0.0	0.0
Industrial GDP 1995	-0.1	-0.2	0.0	-0.4	0.6	0.3	0.1	0.1	0.0	-0.1	0.0	-0.1	0.1	-0.1	0.0	-0.1	0.2	0.0	-0.1	0.0	0.0
dIndustrial GDP (1995-2000)	-0.1	0.1	-0.1	0.2	0.5	-0.2	0.0	0.2	-0.1	0.0	0.0	0.0	0.0	0.1	-0.1	-0.1	0.0	0.0	0.1	-0.2	0.2
Manuf. GDP 1995	-0.3	-0.3	0.1	-0.5	-0.2	0.3	0.0	0.1	0.2	0.0	-0.3	-0.2	0.3	0.1	0.1	0.0	0.0	0.1	-0.4	0.1	-0.1
Service GDP 1995	-0.1	-0.2	0.1	-0.6	-0.1	-0.4	-0.2	0.0	0.1	0.0	0.0	-0.1	-0.1	-0.3	0.1	0.4	0.1	0.2	-0.4	0.2	-0.2
dServiceGDP (1995-2000)	0.0	-0.2	-0.1	0.0	-0.6	0.3	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.1	0.0	0.0	0.0	-0.2	-0.1	0.1

Note: Correlation coefficients, except for the "D" and "F" PCA components: Factor loading.
Sources: Authors' calculation on the basis of WTO tariff data and World Bank's World Development Indicators.