

Non-Tariff Measures in Agri-Food Trade: What Do the Data Tell Us?

Evidence from a Cluster Analysis on OECD Imports[♦]

Anne-Célia Disdier^a

Frank van Tongeren^b

a: INRA, UMR Economie Publique INRA-AgroParisTech.

b: OECD, Trade and Agriculture Directorate

Forthcoming in *Applied Economic Perspectives and Policy*

Abstract

Non-tariff measures (NTMs) are playing an increasing role in international trade of agri-food products. Although well-recognized, this aspect has not been extensively analyzed at a disaggregated level. This paper focuses on NTMs notified by member countries of the Organisation for Economic Co-operation and Development on 777 products. Using a cluster-analysis, it identifies the correlation between the occurrence of NTMs, their trade coverage and the incidence of trade frictions for these products. The political economy literature on protection provides useful insights, but it is not completely satisfying in explaining cross-product differences in the occurrence of NTMs and trade frictions.

Keywords: non-tariff measures, cluster analysis, political economy of protection

JEL Classification: C83, P16, Q17

[♦] We especially appreciate the suggestions of John Beghin and two anonymous referees. We thank participants at GTAP 2009 and IAAE 2009 for helpful advice and Hsin Huang for providing some of the data used in this paper. The views expressed in this paper are those of the authors and do not reflect the official view of the OECD or of the governments of its member countries.

Corresponding author: Frank van Tongeren, OECD, Trade and Agriculture Directorate, 2, rue André Pascal, 75775 Paris, CEDEX 16, France. Phone: +33 1 45 24 78 81.

Fax: +33 1 44 30 61 01. Email: frank.vantongeren@oecd.org

Introduction

Non-tariff measures (NTMs) are widely believed to assume an increasingly important role in determining international trade, especially between developed and developing economies. Agri-food products are extensively affected, and these products are the ones with the largest number of NTM complaints relative to the sectoral export value (United Nations Conference on Trade and Development, 2005).

NTMs are used for a variety of reasons, including the correction of information asymmetries and other market failures, but also possibly protectionist purposes. The relevance of these motivations will clearly differ across products. In some products only a few NTMs may be found, while other products might have a proliferation of NTMs. Such differences may also imply differences in potential trade frictions between countries. When NTMs address market failures there is more likelihood of a shared understanding between exporters and importers about the necessity of using certain measures, even if they restrict trade, and we would tend to observe less complaints. But if the intent behind a measure is more overtly protectionist, more trade frictions are likely to occur. The purpose of this paper is to study these cross-product differences and to better understand the motivations behind NTMs. We focus on NTMs initiated by governments in member countries of the Organisation for Economic Co-operation and Development countries (OECD) on 777 agri-food products.

A cluster analysis can help to class products into coherent groups, taking a multitude of product characteristics into account. The basic idea of this statistical method is to form groups such that the differences between objects in the same cluster are as small as possible, while the differences between objects in distinct clusters are as large as possible. In this paper, we select three cluster criteria: the occurrence of NTMs, their trade coverage and the occurrence of trade frictions. The latter criterion, although not widely used, provides a rich information and allows to investigate the correlation between the incidence of the notifications of NTMs and the incidence of trade frictions at the product level. We view this approach as an alternative way to test for the

predictions of the literature on protection in a case where econometric estimations are difficult to conduct at a highly disaggregated level due to data unavailability.

There is no shortage of estimates of the trade effect of NTMs, but despite the increasing number of papers investigating these measures, assessments of their trade impact by product or at least by sector tend to be rather scarce and inconclusive. Some sectoral differences are present in Henry de Frahan and Vancauteran (2006). This analysis suggests that the harmonization of standards has a positive and significant impact on intra-European trade (except for condiments). The effect of harmonization is small for meat, dairy, tea and coffee but large for sugar and cacao. Fontagné, Mimouni, and Pasteels (2005) highlight a predominance of negative effects of NTMs on trade of fresh and processed food. Flows of cut flowers, swine meat, vegetables, citrus, sugar, juices, wine, animal feed preparation are significantly reduced by these measures. Finally, using ad-valorem equivalents of NTMs, Disdier, Fontagné, and Mimouni (2008) find a strong negative and significant impact for live plants and cut flowers, gums and resins, beverages, and tobacco and trade-enhancing effects are observed for cereals, and wool and animal hair.¹

Our cluster analysis provides six robust groups of products, significantly different in terms of NTM trade incidence and trade frictions. Our presumption is that such differences cannot be solely explained by differences in health and food safety concerns and where international trade acts as a vector to transmit undesired product attributes. Furthermore, the political economy literature on protection, while providing some useful insights, is not completely satisfying in explaining cross-product differences in the occurrence of NTMs and trade concerns. This literature relies on crude definitions of protection and does not account for the complex forms of NTMs that are presently observed in the agri-food sector of OECD countries.

¹ This work also highlights the sensitivity of results to the measurement used for NTMs and the results are quite different with frequency indexes as opposed to ad-valorem equivalents.

The paper is structured as follows. The next section reviews the political economy literature on protection. We discuss the rationale for cluster analysis and the sources of information used to group agri-food products in the third section. Section 4 implements the cluster analysis and reports the results. Based on the results, possible motivations behind NTMs notifications on agri-food products are discussed. Section 5 concludes.

Political economy literature on protection

The political economy literature on protection has been widely developed both theoretically and empirically to explain the NTM formation process.

The political analysis of Kono (2006) emphasizes that politicians in more democratic societies tend to be more sensitive to public concerns about health, product safety and the environment, which reinforces the tendency to use complex measures and suggests that cross-industry differences could be observed in the degree of NTM coverage. Kono's empirical analysis, cross-country as well as cross-sector, lends support to this hypothesis. NTMs will be more frequent in those agri-food subsectors where consumer interest groups voice concerns relating to food safety, animal welfare and the environment. A coalition of producers and consumers can successfully demand NTMs to address profound health and food safety concerns, where international trade acts as a vector to transmit undesired product attributes.

Many papers focus on endogenous protection to explain cross-product differences in the occurrence of NTMs. Trefler (1993) analyses the trade impact of United States (US) non-tariff protection when the latter is modeled endogenously. The results suggest that the level of import penetration has a positive but not significant impact on protection, while a rise in import penetration leads to greater protection. Furthermore, export-oriented industries receive less protection, since they do not face import competition and/or there is a risk of foreign retaliation. Trefler (1993) also shows that high levels of unemployment increase protection. Finally, when

seller concentration is small, protection is low (due to free-riding in lobbying), while protection is negatively impacted by buyer concentration.

The 'protection for sale' model developed by Grossman and Helpman (1994) postulates rent maximizing lobbying activities in return for political support contributions and takes both import elasticities and industry stakes into account. It predicts that the lower the price elasticity of imports, the higher the level of protection afforded to the industry, because the deadweight loss from import protection increases with the price elasticity of imports. In addition, a low ratio of imports to output favors larger lobbying contributions, and will tend to raise protection in the political economy equilibrium, because low import volumes mean a low social cost of protection. Mitra (1999) endogenizes lobby formation within the Grossman-Helpman framework. The original Grossman-Helpman model was developed for tariffs (although it has also been empirically used to investigate the NTM formation process). However, as highlighted by Facchini, Van Biesebroeck, and Willmann (2006), NTMs - unlike tariffs - do not allow the government to fully capture the rents from protection. Thus using NTM coverage ratios as the dependent variable to empirically test the Grossman-Helpman model could generate biased estimates. To solve this issue, Facchini, Van Biesebroeck, and Willmann (2006) extend the Grossman-Helpman framework to account for NTMs allowing only partial rent capturing.

The Grossman-Helpman model has been tested empirically in several papers. We focus here on papers using information on NTMs. Using US NTMs and lobbying spending in 1983, Goldberg and Maggi (1999) find a negative and significant relationship between protection and import penetration for politically organized sectors, and a non-significant one for nonorganized sectors. On the other hand, using similar data, Gawande and Bandyopadhyay (2000) show that protection significantly decreases with the import penetration ratio if the sector is organized, while protection significantly increases with import penetration if the sector is not organized. Specifically focusing on US food processing industries from 1978 to 1992, Lopez and Matschke

(2006) also find that organized sectors receive more protection than nonorganized ones. Besides, protection decreases with import penetration if the sector is organized. As highlighted by Imai, Katayama, and Krishna (2008), previous mentioned studies face however one main weakness, i.e., the classification of the industries into politically organized and unorganized ones. To solve this issue, they suggest a new approach based on quantile regressions. The basic idea is as follows: if predictions of the Grossman-Helpman model are verified, then the effect of the inverse import penetration ratio on protection at higher quantiles tends to reflect that of politically organized ones. This approach presents the advantage of not requiring any data on political organizations. Using part of the data from Gawande and Bandyopadhyay (2000), the authors show that, unlike previous studies, their results do not provide any evidence favoring the Grossman-Helpman model.

Extending the Grossman-Helpman framework into a heterogeneous firm model, Bombardini (2008) explains why larger firms are more likely to lobby and makes a link between the size distribution of firms and protection. If lobbying involves fixed costs, then the lobbying will be concentrated amongst the larger firms. The more concentrated industry is more effective in its lobbying, as the benefits are kept inside a smaller group. In contrast the marginal benefits of increased protection are declining, and may not outweigh the costs of lobbying, if more firms enter the club. These predictions are not refuted empirically for data on the US.

Another possible cause of differences in NTMs occurrence could be governments' support to ailing sectors. Baldwin and Robert-Nicoud (2007) address the question why declining industries account for most of the protection granted in all industrialized nations. Their framework rests on sunk entry costs, and predicts that in expanding industries rents would attract new entry that would eventually erode the rents generated by protection. In declining industries that is not true, as protection can raise profits sufficiently high, but below the normal return on capital, to make lobbying for protection a rewarding activity.

The analysis by Fischer and Serra (2000) suggests also that some industries would face more pressure from domestic producers to implement a protectionist standard than others. While a non-discriminatory standard raises costs for both domestic and foreign firms, it can be profit enhancing to the domestic firm if the foreign competitor would face high cost of implementing the standard that is specific to one of its export destinations and if the foreign competitor can divert supplies to other markets.

In this paper, our aim is to investigate the motivations governing NTMs notified in the agri-food industry. However, the absence of data at a highly disaggregated level prevents us to test for the assumptions developed in the literature on protection using econometric estimations. The cluster analysis represents an alternative way to test for these predictions. By grouping products into different coherent clusters defined by criteria related to NTMs and trade frictions, it helps to highlight products for which potential protectionist purposes govern NTMs' notifications and products for which sanitary and environmental reasons are the main driver of the notifications.

Cluster analysis: rationale and data

Rationale

First used by Tryon in 1939, cluster analysis is applied in a wide range of fields.² In biology many taxonomies of species or more recently of genes were achieved using such techniques. In the field of medicine, cluster analysis is used to categorize diseases, symptoms and cures. In archeology it is employed to classify objects of past civilizations in order to understand their histories, customs and living habits. Grouping of information is also a key tool in web-based

² For a detailed presentation of cluster analysis, see Everitt, Landau, and Leese (2001); Jain and Dubes (1988); Kaufman and Rousseeuw (2005); and Tan, Steinbach, and Kumar (2005).

information retrieval. The use of cluster analysis in economics is increasing, but in the agricultural economics sub-discipline its application has been limited.³

The two most common clustering techniques are the ‘one-level partitional’ and the ‘hierarchical’ methods. In both techniques, the formation of clusters is based on measures of dissimilarities (similarities) or distances between objects. The choice of measures largely depends on the data used. For continuous data, the most common measure is the Euclidean distance, but other measures can also be computed such as the Manhattan or the Chebyshev distances, etc.⁴ For binary data, a simple matching distance can be used, or alternatively more complex measures based on a pairwise comparison of scores on binary criteria can be employed.

In the so called partitional clustering, objects are assigned into mutually exclusive clusters through an iterative process. The procedure starts with k initial group centers and objects are subsequently assigned to a group based on distance to the nearest center. Cluster centers are typically defined as the mean of the observations, or alternatively the median observation can be used. The k initial group centers are chosen randomly among objects to be clustered, but during the clustering process centers are updated at each pass and objects can move from one cluster to another. The process stops when cluster centers do not shift anymore (or more than a previously defined cut-off value) or when an iteration limit is reached.

By contrast, in the hierarchical clustering, groups are nested and organized as a tree. Hierarchical clustering can be divisive or agglomerative. The divisive clustering starts with only one all-inclusive cluster, which is then split into different sub-clusters. In agglomerative clustering, items are first joined into groups and groups are then joined to each other into larger ones.

³ An Econlit search (made on November 03, 2008) for ‘cluster analysis’ in the paper abstracts provided 399 results, of which 347 were journal articles, 25 working papers and 27 published as a book. Among the 347 journal articles, 36 of them were published between 1985 and 1994, 157 between 1995 and 2004 and 154 since 2005. Restricting the search to papers having “agriculture” as subject yields only 25 journal articles for the whole period 1985-2008.

⁴ Most of these distances are special cases of the more general Minkowski metric. The Euclidean distance is the Minkowski metric with argument 2. The Manhattan distance is the Minkowski metric with argument 1. The Chebyshev distance gives infinite weight.

Hierarchical clustering is very computer-intensive and therefore limited to small samples (typically <250 items). A second weakness is that merges cannot be undone in a further step. Partitional clustering does not suffer from these criticisms, but it is sensitive to outliers. Furthermore, the number of groups has to be specified before the clustering starts. However, the determination of the optimal number of clusters can be guided by test statistics that inform about the likelihood of improving the clustering by increasing or decreasing the number of clusters k .

Cluster analysis on agricultural trade policy issues has recently been used in a number of papers. Diaz-Bonilla, Thomas, Robinson and Cattaneo (2000) use hierarchical and partitional clustering techniques to classify countries in terms of food security. Their sample includes 167 countries and employs five measures of food security (food production per capita, the ratio of total exports to food imports, calories per capita, protein per capita, and nonagricultural population share) to identify 12 country clusters. Their results suggest that the different categories of countries proposed by the World Trade Organization (WTO) (developed, developing, least developed and net food importing developing) do not reflect food security issues very well. Bjørnskov and Lind (2002; 2005) perform cluster analyses to identify developing countries' possible negotiation partners and strategies in the Doha Round. The authors show that African and most Latin America countries pursue quite similar trade policy objectives to the United States and most of the Cairns Group countries. But the positions of most developing countries are in opposition to the European Union (EU). Costantini, Crescenzi, De Filippis and Salvatici (2007) investigate the internal coherence of existing coalitions in the Doha negotiations, with a special focus on agriculture. Instead of relying on countries' self-declared negotiating positions, the authors use 27 indicators concerning economic, social and institutional development level, openness to trade, agricultural productive structure and market access policies as inputs for the cluster analysis. Comparing the country groups obtained from the cluster analysis with the existing ones, they highlight the role of structural features in alliance formation.

Data

Our cluster analysis uses three pieces of publicly available information to group agricultural products: (i) information on the occurrence of NTMs, (ii) information on trade flows, and (iii) information on NTM-related trade frictions amongst countries. Our analysis concentrates on NTMs initiated by OECD governments (with EU countries aggregated to EU). Agricultural products in this study are the ones covered by the WTO Agreement on Agriculture, plus fish and fish products.

Information on the occurrence of NTMs: One of the main sources for research on NTMs is the Trade Analysis and Information System (TRAINS) dataset compiled by the United Nations Conference on Trade and Development (UNCTAD). TRAINS relies heavily on notifications of measures to the WTO (Sanitary and Phytosanitary (SPS) and Technical Barriers to Trade (TBT) Committees). WTO members are required to notify only new or changed measures since 1995. The notification requirement covers measures which differ from international standards, guidelines or recommendations, or situations where no standards exist, and, in addition, may have a significant impact on trade. UNCTAD complements the notifications using national sources. In spite of this effort, TRAINS does not provide a complete coverage of NTMs, and furthermore the country coverage is typically biased towards those countries that diligently notify their measures to the WTO, or those where UNCTAD happens to do complementary data collection efforts. TRAINS data used for this paper have been obtained directly from UNCTAD and contain the most recent notifications for Mexico, the European Union and Japan that were not available before in the TRAINS database and therefore not used in previous empirical works on NTMs.⁵ Our work focuses on NTMs that have not a direct impact on prices or quantities. This excludes all measures such as tariffs, para-tariffs, price control measures, and direct quantity control measures

such as quota (except on sensitive products).⁶ Countries can attribute several motives in their notifications of measures, and the following are singled out for this study: protection of human health, animal health and life, plant health, environment, and wildlife. Appendix 1 describes the filtering and cleaning procedures applied to TRAINS data before their use in the statistical analysis.

Information on trade flows: The trade data used for this study come from the OECD's International Trade by Commodity Statistics database, as maintained by the OECD Statistics Directorate. This dataset is a mirror image of, and has the same contents as, the United Nations Commodity Trade Statistics Database (UN COMTRADE) and provides bilateral import flows of OECD countries from all countries in the world at the Harmonized System (HS) 6-digit level.

Information on WTO specific trade concerns: WTO members can raise specific trade concerns (STCs) in the SPS and TBT Committees. These concerns pertain to issues raised by one (or more) WTO member concerning measures put in place by other members and deemed to restrict trade. However, not all concerns raised relate to perceived trade restrictions, as countries sometimes only seek clarification on a measure put in place by a trading partner, or remind a trading partner of lacking notifications to the SPS or TBT Committee. Raising an issue as a specific trade concern is an important way to start information exchange and bilateral consultations (Organisation for Economic Co-operation and Development, 2002). Although the WTO secretariat keeps a record of concerns raised in both Committees since 1995, only the SPS concerns are accessible in a database format through a web-based portal, and are used in this paper.⁷ The SPS-STC database provides a summary description of cases, as well as pointers to

⁵ The use of recent and unpublished updates largely explains the differences in counts reported in this paper and those obtained by Disdier, Fontagné, and Mimouni (2008).

⁶ Category 6270 'Quotas for sensitive products' primarily relates to products that are deemed harmful for the environment, such as those subject to the Montreal Protocol on Substances that Deplete the Ozone Layer (1987) and its later amendments (CFCs, halons, fully halogenated CFCs, carbon tetrachloride and methyl chloroform).

⁷ <http://spsims.wto.org/web/pages/search/stc/Search.aspx>. Furthermore, the following note by the WTO Secretariat (WTO, 2009) provides the list and some statistics on all specific trade concerns raised at the SPS Committee since 1995:

relevant documents. The data on STC cases include a record of which member(s) raised a concern and when, which country(ies), if any, supported the concern, which country(ies) maintained a measure deemed to restrict exports of the country(ies) raising the concern. The data also give an indication of the products involved using the HS coding system. For example, Ecuador and Israel, supported by Kenya expressed concern at the SPS Committee in 2001 over changes in the European inspection procedures on cut flowers. For the European Union, the new procedures aimed at controlling for the presence of harmful non-European organisms regularly intercepted on cut flowers. Following the discussions at the SPS Committee, bilateral consultations between the EU and complainant countries took place and the EU agreed to postpone entry into force of these new procedures after consideration of the potential difficulties faced by certain exporting countries. For this specific trade concern, the SPS-STC database provides the list of countries involved (Ecuador, Israel, Kenya, and the EU), the official documents related to the concern, the affected product (HS06), two subject keywords (plant health, risk assessment) and a description of the concern and its development over time. One potential caveat of this approach based on available data is that it will not capture cases where trade tensions on NTMs are settled bilaterally without raising the issue at the WTO. Changes and cleanings applied to the trade concerns data are reported in Appendix 1.

Combining the three pieces of information: Data on NTMs, trade flows and trade concerns can be compared. Table 1 reports some comparative statistics from a commodity perspective for the period 1996-2006. EU is defined as EU15 between 1996 and 2003 and as EU25 for 2004-2006. The sample covers 769 products that are subject to at least one NTM. Only eight of the total 777 products for which positive trade flows are observed do not face any NTM in any OECD country (HS 500200 - Raw silk; HS 500310 - Silk waste, not carded or combed; HS 500390 - Silk waste, other; HS 520300 - Cotton, carded or combed; HS 530121 - Flax,

http://docsonline.wto.org/GEN_viewerwindow.asp?http://docsonline.wto.org:80/DDFDocuments/t/G/SPS/GEN204R9.doc

broken or scotched; HS 530129 – Flax hackled or otherwise processed, but not spun; HS 530130 - Flax tow and waste; HS 530290 - True hemp, other).

Column (1) reports the number of notified products by HS 2-digit chapter, while column (2) presents the distribution of notifications. OECD countries often notify several measures on a given product. Not surprisingly, there is a higher concentration of NTMs around fresh products (fish, meat, etc.), with fish and other aquatic products topping the list. However, processed products are also well represented. Column (3) investigates the trade coverage ratio by HS2 chapter. This ratio represents the value of imports subject to notified NTMs relative to total imports. Fish and meat are again at the top of the ranking. The trade coverage ratio is also quite high for products of animal origin (HS05), meat, fish and seafood preparations (HS16) and live animals (HS01). Columns (4) and (5) report the number of SPS trade concerns raised by and against OECD countries for each HS 2-digit chapter. Between 1996 and 2006 a total of 233 specific trade concerns dealing with agri-food products were raised. Out of these 233 cases, 150 were raised by OECD countries. In 139 cases raised, the measure was maintained by at least one OECD country. In column (4), if a concern is raised by several OECD countries, we create a separate record for each country. Furthermore, many concerns involve different HS 2-digit chapters. A separate record is created for each of the chapters. Similarly, in column (5), a separate record is created for each OECD country against which a concern is raised, as well as for each HS2 chapter affected. Most of the SPS concerns are on meat (HS02), fruits (HS08), vegetables (HS08), dairy products (HS04), live animals (HS01), and products of animal origin (HS05).

Insert table 1 here

NTMs' notifications can also be investigated at the country level. Focusing on the US and the EU, which are the main traders of agri-food products, we observe that the EU notifies 708 products, while the US apply NTMs on 521 products. Both concentrate their notifications on

fresh products (e.g. meat and fish products). The US also notify many NTMs on dairy products (HS04), products of animal origin (HS05), and plants, trees and cut flowers (HS06). We also observe a concentration of notifications on meat, fish and seafood preparations (HS16) and beverages, spirits and vinegar (HS22). On the other hand, the EU applies many notifications on vegetables (HS07), fruits and nuts (HS08), coffee, tea, mate and spices (HS09) and vegetable, fruit and nut preparations (HS20).

Even if NTMs are equally applicable to all exporting countries, that is, they do not discriminate between import sources, exporters are differently affected depending on the structure of their exports in terms of products and markets. Our data suggest that developing and emerging countries are more affected by NTMs applied by OECD countries than other OECD countries. 84% of developing and emerging countries' exports to OECD are subject to NTMs, while 76.7% of OECD exports to other OECD countries are subject to notified measures.

Cluster analysis: implementation and results

Combining the data on NTMs with the information on specific trade concerns and the trade statistics allows the correlation between incidence of NTMs and incidence of trade frictions to be identified and the motivations (potentially protectionist, sanitary, environmental, etc.) behind the notifications of NTMs to be highlighted.

The cluster analysis uses three criteria, one from each of the underlying datasets described in the previous section, to provide a statistically sound grouping of agri-food products. The clustering is done for 2006. That is to say, COMTRADE data for 2006 for 12 OECD importing countries (counting the EU15 as 1), 212 exporting countries and 777 HS6 agricultural products are merged with TRAINS data on NTM measures notified by OECD countries since 1995 and that are still in place in 2006 and with concerns brought to the SPS Committee until 2006 and in which OECD countries maintained the measure. Missing and zero trade flows are kept in the database as these observations could be subject to NTMs or specific trade concerns.

The analysis is conducted at the HS6 product level, using 777 observations, and the cluster criteria are as follows:

- Trade coverage ratio for each product: imports of each product in OECD countries subject to at least one NTM relative to total imports of that product in OECD countries;⁸
- NTM notifications for each product: total number of NTMs applied by all OECD countries on each product (if two countries apply the same NTM, it is counted as two NTMs, because the NTM code assigned by UNCTAD could in practice coincide with different import requirements imposed by these two countries);
- Number of SPS concerns by HS6 product (if a concern is raised against two countries, it is counted as two concerns).

Because the three variables are measured in different units they are standardized to zero means and unit variance before clustering. The grouping is obtained by a partitional cluster analysis using the mean of the observations as cluster centers. The distance is measured using the standard Euclidean distance.⁹ Using the Calinski and Harabasz (1974) stopping rule,¹⁰ the optimal number of clusters (based on *k*-mean clustering) is found to be six.

Table 2 presents the number of observations within each cluster and information on their homogeneity. Cluster 1 is much smaller than other clusters. All clusters show high internal cohesion: small within cluster standard deviation, short average and maximum distances from cluster center. Clusters 2 and 3, followed by cluster 4, are the most homogeneous, with low dispersion around their centers. On the other hand, cluster 1 is the least homogeneous. Clusters 5 and 6 are in-between, cluster 6 being slightly more compact than cluster 5.

⁸ This definition is also used by Disdier, Fontagné and Mimouni (2008).

⁹ The Euclidean distance between two observations x and y in the 3-dimensional space (i,j,k) spanned by the 3 clustering variables is defined as the square root of the sum of squared differences, i.e. $d(x, y) = \sqrt{(x_i - y_i)^2 + (x_j - y_j)^2 + (x_k - y_k)^2}$.

¹⁰ The Calinski-Harabasz pseudo- F stopping rule is (with N observations and k clusters): $[\text{trace}(\mathbf{B})/(k-1)]/[\text{trace}(\mathbf{W})/(N-k)]$, where \mathbf{B} is the between-cluster sum of squares and cross-products

Insert table 2 here

Table 3 reports the Euclidean distance between cluster centers. Cluster 1 is relatively distant from all other clusters, while clusters 3 and 4 are closest to each other. The matrix also shows that the distance is increasing when one is moving from cluster 1 to cluster 6, suggesting that clusters are becoming more and more different. For example, the distance is only 3.60 between clusters 1 and 2, while it reaches 6.07 between clusters 1 and 6.

Insert table 3 here

Table 4 provides the mean of each criterion for each cluster and the whole sample. To ease the interpretation of results, the means reported in table 4 are calculated using non-standardized variables. Strong differences exist between clusters. A high share of imports of products included in clusters 1-4 faced NTMs (over 85%), while this share is only 38.6% for products of cluster 5 and 21.6% for those included in cluster 6. Furthermore, the average number of notifications on products of clusters 1-3 is high (on average 23 NTMs). By comparison, this number is only 15.4 for products of cluster 4, 16.7 for products of cluster 5 and 5.3 for those of cluster 6. Finally, the number of SPS trade concerns is very high for products of cluster 1 (57.1 on average), high for products of cluster 2 (16.5) but low for all other clusters.

Insert table 4 here

Figure 1 shows the box plot representation of each of the criteria within each cluster. The lower (upper) limit of a box represents the first (third) quartile of the distribution. The median is shown at the 50th percentile. The vertical lines (or “whiskers”) extend the box to the

matrix and \mathbf{W} is the within-cluster sum of squares and cross-products matrix. A larger value of the statistic suggests a better clustering, i.e. more distinct groups.

lowest and highest adjacent values, excluding extreme values.¹¹ Extreme values are represented with round markers. This figure informs on the distribution of each criterion. Trade coverage ratio is the most dispersed criterion both between and within the clusters. Clusters 5 and 6 have a very high variance for this criterion. The distribution of the two other criteria exhibits much more similarity.

Insert figure 1 here

Three different tests are performed to check the robustness of the clusters. Due to space constraints, results are not reported but are available upon request. First, we search for outliers in our sample. The Hadi (1994) test identifies 25 outliers. However all of them are included in cluster 1 and this cluster does not contain other observations, and hence forms a special group on its own. Second, we test the stability of clusters over time by running the clustering on 2004 data (instead of 2006). The three criteria are the same as before but based on 2004 data. The Calinski and Harabasz (1974) stopping rule indicates that the optimal number of clusters is again six. Results suggest that globally the clusters remain very stable. The Euclidean distance between the centers of a given cluster does not exceed 0.40. Distances between clusters are not significantly different from the ones presented in table 3. Third, we test if the clusters are influenced by the product aggregation by performing the clustering on HS4 data. The optimal number of clusters provided by the Calinski and Harabasz stopping rule is, as before, equal to six. The distances between cluster centers at the HS 6-digit and 4-digit levels are not strongly different from the ones reported in table 3, which confirms the robustness of our cluster analysis.

The product content of each cluster is presented in detail in Appendix 2. We shortly summarize here the main findings:

¹¹ The upper adjacent value is the largest value that is equal or less than the 3rd quartile + 1.5x(3rd quartile – 1st quartile). The lower one is the smallest value that is equal or greater than the 1st quartile - 1.5x(3rd quartile – 1st quartile)

- Cluster 1 (high trade coverage, high number of notifications, very high number of concerns) is the smallest in terms of size and includes two types of products: bovine meat and many dairy products;
- Cluster 2 (high trade coverage, high number of notifications, high number of concerns) contains all meat products (except bovine meat), many vegetables, and edible fruit, nuts, peel of citrus fruit;
- Cluster 3 (high trade coverage, high number of notifications, low number of concerns) includes live trees and plants, cereals, various preparations (such as meat, fish and seafood preparations; cereal, flour, starch, milk preparations; preparations of vegetables, fruit or nuts), and most of beverage, spirit and, vinegar products. This cluster also contains some specific products (e.g. live non-farm animals, fish fillets, crustaceans, birds' eggs, and natural honey);
- Cluster 4 (high trade coverage, low number of notifications, low number of concerns) contains live farm animals, most of fish products, many products of animal origin (such as coral, shell or natural sponges), cut flowers, and soy and cocoa beans;
- Cluster 5 (low trade coverage, low number of notifications, low number of concerns) includes most of milling products, malt, starches, inulin, and wheat gluten, as well as many animal and vegetable fats and oils and some sugar products (such as cane or beet sugar and chemically pure sucrose).
- Cluster 6 (very low trade coverage, very low number of notifications, low number of concerns) mainly includes processed products of chapters 24 and higher (such as tobacco; essential oils, perfumes, and cosmetics; albuminoids and modified starches; silk, cotton and vegetable textile fibers).

The cluster analysis helps to highlight the correlation between the incidence of NTMs and the incidence of trade frictions at the product level. If a high number of NTMs coincides with

a high number of trade concerns, it suggests that the NTMs on products included in this cluster have a potential protectionist effect (typically products of clusters 1 and 2). On the other hand, if a high number of notifications is registered but at the same time a low number of concerns occurs, this would suggest that NTMs are put in place with a shared understanding of sanitary or environmental concerns: exporters do not consider that these NTMs are potentially protectionist (although the share of notified exports could be high) and do not raise a concern at the SPS Committee (typically products of cluster 3).

The analysis of Fischer and Serra (2000) may be helpful in understanding the high incidence of NTMs in terms of both the number of notifications and their trade coverage and at the same time the high level of trade concerns that occur in clusters 1 and 2 (meat products, fresh fruits and vegetables). Fisher and Serra (2000) show that domestic producers may seek protection of their economic interest by limiting foreign competition.

Cluster 3, and to some extent cluster 4, have a relatively high trade coverage but a low incidence of concerns. In these cases one could argue that the NTMs in place are not hindering trade, but that the measures reflect a shared understanding about the importance of regulations to safeguard health, food safety and environmental concerns. Such results tend to be in line with Kono (2006)'s analysis, who suggests that a coalition of producers and consumers can successfully demand NTMs to protect the domestic market from diseases, harmful organisms, etc.

Import penetration and import demand elasticities play a central role in the protection for sale literature (cf. supra the literature review), and these aspects may also contribute to the explanation of cross-product differences. While the proportion of consumption sourced internationally is typically quite small across all agricultural sectors and across countries, there are some differences across the products included in our sample. For example, the international

dairy market is very 'thin' with small trade volumes relative to domestic absorption. The cluster analysis squarely puts dairy products into the group with the highest NTM coverage.

Furthermore, the most dynamic trade growth, and the highest import penetration rates are observed in processed products, but they are mostly found in cluster 6 which registers relatively low levels of NTM notifications and low levels of concerns. Much of the trade in processed products occurs in multinationally operating supply chains, and this suggests that there would be less domestic producer pressure stemming from those chains to erect trade barriers. The works of Bombardini (2008) and of Baldwin and Robert-Nicoud (2007) suggest that we would expect to see relatively more non-tariff measures in sub-sectors that are on the decline in the OECD countries, while we would observe less protection in growing sub-sectors. This pattern is to some extent present in the clustering results, which tend to put the more dynamic processed products (HS24 and higher) mainly in cluster 6. A more serious testing of the inverse relationship between the import penetration and NTM coverage could be done using trade data and domestic consumption figures.

One problem with the existing political economy literature on protection is that it does not tackle sufficiently the question when a measure can be considered protectionist. This literature uses a relatively straightforward definition of protection, either a tariff or quantitative non-tariff border measure (both rising behind the border price of imported goods) or a domestic subsidy (lowering the domestic price of domestically produced goods), and usually focuses on manufacturing industries. NTMs in agriculture can be adopted to protect the safety of food, health or the environment and are therefore considerably more complex and their efficiency costs are much less evident than the standard welfare losses associated with tariffs and quantity barriers. Furthermore, the private sector has its own private standards and associated enforcement mechanisms, which confounds the analysis of public regulations. A fuller theory of NTMs to explain differences across agricultural products should therefore take into account also consumer

benefits, and concomitant incentives to engage in lobbying for regulation, as well as producer incentives to lobby for protection.

Conclusion

This paper classifies 777 agri-food products according to the occurrence of NTMs, their trade coverage and the concerns raised by countries in the WTO SPS Committee. The cluster analysis suggests six robust groups of products. Some part of the cross-product differences can be explained by differences in health, food safety, and environmental concerns. The political economy literature on endogenous protection provides some insights to explain the remaining cross-product differences, but it is not fully satisfactory and invites further research. In particular, investigations at a detailed industry level and across countries represent a promising area of research.

References

- Baldwin, Richard, and Frédéric Robert-Nicoud. 2007. Entry and asymmetric lobbying: why governments pick losers. *Journal of the European Economic Association* 5 (5): 1064-93.
- Bjørnskov, Christian, and Kim M. Lind. 2005. Underlying Policies in the WTO, the Harbinson Proposal and the Modalities Agreement. *Revue Economique* 56 (6): 1385-412.
- _____. 2002. Where Do Developing Countries Go After Doha? An Analysis of WTO Positions and Potential Alliances. *Journal of World Trade* 36 (3): 543-62.
- Bombardini, Matilde. 2008. Firm heterogeneity and lobby participation. *Journal of International Economics* 75 (2): 329-48.
- Calinski, Tadeusz, and Jerzy Harabasz. 1974. A dendrite method for cluster analysis. *Communications in Statistics* 3 (1): 1-27.
- Costantini, Valeria, Riccardo Crescenzi, Fabrizio De Filippis, and Luca Salvatici. 2007. Bargaining Coalitions in the WTO Agricultural Negotiations. *The World Economy* 30 (5): 863-91.
- Diaz-Bonilla, Eugenio, Marcelle Thomas, Sherman Robinson, and Andrea Cattaneo. 2000. Food Security and Trade Negotiations in the World Trade Organization: A Cluster Analysis of Country Groups. Working paper 59, International Food Policy Research Institute.
- Disdier, Anne-Célia, Lionel Fontagné, and Mondher Mimouni. 2008. The Impact of Regulations on Agricultural Trade: Evidence from the SPS and TBT Agreements. *American Journal of Agricultural Economics* 90 (2): 336-50.
- Everitt, Brian, Sabine Landau, and Morven Leese. 2001. *Cluster analysis, 4th ed.* London: Arnold Publishers.
- Facchini, Giovanni, Johannes Van Biesebroeck, and Gerald Willmann. 2006. Protection for sale with imperfect rent capturing. *Canadian Journal of Economics* 39 (3): 845-873.
- Fischer, Ronald, and Pablo Serra. 2000. Standards and protection. *Journal of International Economics* 52 (2): 377-400.
- Fontagné, Lionel, Mondher Mimouni, and Jean-Michel Pasteels. 2005. Estimating the Impact of Environmental SPS and TBT on International Trade. *Integration and Trade Journal* 22: 7-37.
- Gawande, Kishore, and Usree Bandyopadhyay. 2000. Is protection for sale? Evidence on the Grossman-Helpman theory of endogenous protection. *Review of Economics and Statistics* 82 (1): 139-52.
- Grossman, Gene M., and Elhanan Helpman. 1994. Protection for sale. *American Economic Review* 84 (4): 833-50.
- Goldberg, Pinelopi K., and Giovanni Maggi. 1999. Protection for Sale: An Empirical Investigation. *American Economic Review* 89 (5): 1135-55.
- Hadi, Ali S. 1994. A Modification of a Method for the Detection of Outliers in Multivariate Samples. *Journal of the Royal Statistical Society Series (B)* 56: 393-6.
- Henry de Frahan, Bruno, and Mark Vancauteren. 2006. Harmonisation of food regulations and trade in the Single Market: evidence from disaggregated data. *European Review of Agricultural Economics* 33 (3): 337-60.
- Imai, Susumu, Hajime Katayama, and Kala Krishna. 2008. A Quantile-based Test of Protection for Sale Model." Working paper 13900. National Bureau of Economic Research.

- Jain, Anil, and Richard Dubes. 1988. *Algorithms for Clustering Data*. Upper Saddle River: Prentice Hall.
- Kaufman, Leonard, and Peter J. Rousseeuw. 2005. *Finding groups in data: An introduction to cluster analysis*. New York: Wiley.
- Kono, Daniel Y. 2006. Optimal Obfuscation: Democracy and Trade Policy Transparency. *American Political Science Review* 100 (3): 369-84.
- Lopez, Rigoberto A., and Xenia Matschke. 2006. Food Protection for Sale. *Review of International Economics* 14 (3): 380-91.
- Mitra, Devashish. 1999. Endogenous Lobby Formation and Endogenous Protection: A Long-Run Model of Trade Policy Determination. *American Economic Review* 89 (5): 1116-34.
- Organisation for Economic Co-operation and Development. 2002. Sanitary and phytosanitary measures and agricultural trade: a survey of issues and concerns raised in the WTO's SPS Committee. Joint Working Party on Agriculture and Trade, COM/TD/AGR/WP(2002)21/FINAL.
- Tan Pang-Ning, Michael Steinbach, and Vipin Kumar. 2005. *Introduction to Data Mining*. Boston: Addison-Wesley.
- Trefler, Daniel. 1993. Trade Liberalization and the Theory of Endogenous Protection: An Econometric Study of U.S. Import Policy. *Journal of Political Economy* 101 (1): 138-60.
- Tryon, Robert C. 1939. *Cluster Analysis*. Ann Arbor: Edwards Brothers.
- United Nations, Conference on Trade and Development. 2005. Methodologies, Classifications, Quantification and Development Impacts of Non-Tariff Barriers.
- World Trade Organization, Committee on Sanitary and Phytosanitary Measures. 2009. Specific Trade Concerns. G/SPS/GEN/204/Rev.9.

Appendix 1: Filtering and cleaning procedures applied to data

TRAINS database

The raw TRAINS database requires a considerable amount of filtering and cleaning before it can be used for statistical analysis. First, some measures notified by OECD countries address all the above-mentioned motives. In this case we created a separate record for each of the purposes. Second, some countries, such as Mexico, notify the same measure on the same product several times. The only difference is the start year. In these cases only the notification with the oldest start year is kept in order to avoid double-counting. Third, most NTMs are mainly reported on a HS 8-digit level, with some countries (e.g. Japan and the EU), reporting at 2, 4, 9 and 10 digit. To facilitate a linking of the NTM data to detailed trade data, all NTMs have to be assigned a 6-digit HS code, which is the lowest level of aggregation at which internationally comparable trade data are available. This adjustment deflates the number of NTM-counts, because all the branches below the 6 digit level will eventually be collapsed to account for just one measure applied at the HS6 product level. The opposite effect, i.e. an inflation of counts, occurs with notifications filed at a level above the HS6-level. We assume that a NTM notified at the 2 or 4 digit level will affect all HS6 products included in this chapter and a separate record is created for each of the products. Four, three different versions (1992, 1996 and 2002) of the HS classification have been used by OECD countries when providing their notifications and further updates. All HS1992 and HS2002 codes are mapped into the HS1996 system to allow a consistent time series to be constructed. All these changes and cleanings result in a database that includes only one unique observation by notifying country, HS6 product and type of measure.

Specific trade concerns database

To create a usable dataset several manipulations were necessary. First, the product classification reported by the SPS Committee in the specific trade concerns database was carefully checked using the SPS Committee background documents, and where necessary the product coding was adjusted. Second, for each concern, affected products are reported in the HS classification at the 2, 4 or 6 digits. To link these data to the NTM and trade data, all product codes are converted into 6-digit codes of the HS 1996 classification. A concern defined at the 2 or 4 digit level is assumed to affect all HS6 products included in this chapter and a separate record is created for each of the products. Third, EU countries are aggregated into EU and in cases where several EU members raised a concern or maintained a measure, just one observation is kept. After all these cleanings, the database contains one unique observation by country maintaining the measure, country raising the concern and HS6 product. The database is available upon request from the authors.

Appendix 2: Product content of each cluster

Cluster 1 includes bovine meat (HS0201-0202) and many dairy products, such as milk, cream, buttermilk and butter (HS0401-0403, and HS0405), cheese (HS0406) and edible products of animal origin (HS0410).

Cluster 2 contains:

- All meat products (HS02), except bovine meat and HS020732;
- Many vegetables (HS07), i.e. potatoes, (HS0701), tomatoes (HS0702), onions, shallots, garlic, leeks (HS0703), cucumbers and gherkins (HS0707), leguminous vegetables (HS0708), dried vegetables (HS0712), dried and shelled vegetables and leguminous (HS0713), manioc, row root, salep (HS0714 except 071490), and part of vegetables provisionally preserved, not ready to eat (HS0711);
- Products of HS08 “edible fruit, nuts, peel of citrus fruit” (except HS081050 and 081220).

Cluster 3 includes: live trees, plants, bulbs, and roots (HS06 except flowers), cereals (HS10 except HS100640), related to cereals, several products of HS19 “cereal, flour, starch, milk preparations and products” (HS1901, 1903-1905), meat, fish and seafood food preparations (HS16 except HS160210 and 160232), most products of HS20 “preparations of vegetables, fruit, nuts or other parts of plants”, ¾ of HS21 “miscellaneous edible preparations”, most products of HS22 “beverages, spirits, vinegar”. It also contains of some specific products: live non-farm animals (HS0106), fish fillets (HS0304), crustaceans (HS0306 except 030623), birds’ eggs (HS0407-0408 except 040899), and natural honey (HS0409).

Cluster 4 contains:

- Live farm animals (HS01 except 0106);
- Two-third of the fish products (HS03);
- Many products of animal origin (all products of HS0501 “human hair, waste”, HS0502 “bristle, pig, badger’s hair, brush making hair, waste”, HS0503 “horsehair, waste”, HS0504 “guts, bladders and stomachs of animals except fish”, HS0508 “coral, shell, cuttle bone, waste”, HS0509 “natural sponges of animal origin”, and HS0510 “ambergris, civet, musk, for pharmaceutical use”);
- Cut flowers (HS0603);
- Almost all vegetables (HS07) not included in cluster 2;
- All products included in HS1201 “soy beans”, HS1203 “copra”, HS1204 “linseed”, HS1208 “flour, meal of oleaginous seed or fruit except mustard”, HS1213 “cereal straw and husks, unprepared”, and HS1214 “animal fodder and forage products”;
- Half of cocoa products (HS1801 “cocoa beans”, HS1802 “cocoa shells, husks, skins and waste”, and HS180310 “cocoa paste, not defatted”).

Cluster 5 includes:

- Two-third of HS11 “milling products, malt, starches, inulin, wheat gluten” (except HS1105 and 1107);
- Half of the animal and vegetable fats and oils (mostly goods from HS1502 “bovine, sheep, and goat fats”, HS1503 “lard stearin, oleostearin & oils”, HS1509 “olive oil”, HS1510 “other olive oils”, HS1512 “safflower, sunflower, cotton-seed oil”, HS1515 “other fixed vegetable fats and oils”, HS1518 “processed animal and vegetable fats and oils”);
- Half of sugar products (most products included in HS1701 “solid cane or beet sugar and chemically pure sucrose”, HS1702 “other sugars” and HS1703 “molasses”);
- Remaining cocoa products (HS18).

Cluster 6 mainly includes products of chapters 24 and higher: tobacco products (HS24), mannitol and sorbitol (HS2905), essential oils, perfumes, cosmetics (HS33 except 330190), albuminoids and modified starches (HS35 except 350211), amylaceous finishing agents and dye carriers (HS380910), silk (HS50), wool, animal hair products (HS51), cotton (HS52), and vegetable textile fibers (HS53).

Table 1. Product chapters by NTM count, concern count and trade coverage, 1996-2006

HS2 Chapter	Number of products notified by OECD countries	Number of NTMs notified by OECD countries	Share of OECD imports subject to NTMs (%)	Number of SPS concerns raised by OECD countries	Number of concerns raised against OECD countries
01- Live animals	17	286	72.2	27	14
02 – Meat, edible meat offal	53	1,340	84.1	57	42
03 – Fish, crustaceans, molluscs, other aquatic invert.	87	1,573	78.3	15	14
04 - Dairy products, eggs, honey, edible animal pduct.	27	624	71.6	26	24
05 - Products of animal origin	17	317	74.3	23	19
06 - Live trees, plants, bulbs, roots, cut flowers	12	278	69.7	13	16
07 - Edible vegetables, certain roots, tubers	56	1,207	66.2	28	18
08 - Edible fruit, nuts, peel of citrus fruit, melons	55	1,248	62.2	32	35
09 - Coffee, tea, mate, spices	32	630	33.5	8	14
10 – Cereals	16	379	65.5	14	14
11 - Milling products, malt, starches, inulin, wheat gluten	34	609	58.4	9	8
12 - Oil seed, oleagic fruits, grain, seed, fruit	44	804	55.3	8	7
13 - Lac, gums, resins, vegetable saps & extracts	12	118	36.4	8	7
14 - Vegetable plaiting materials, vegetable products	10	69	37.0	8	8
15 - Animal, vegetable fats & oils, cleavage products	46	616	46.2	12	9
16 - Meat, fish, seafood preparations	26	670	72.9	12	11
17 – Sugars, sugar confectionery	16	242	48.7	8	8
18 – Cocoa, cocoa preparations	11	178	44.7	8	7
19 - Cereal, flour, starch, milk preparations & products	17	367	67.9	8	7
20 - Vegetable, fruit, nut, food preparations	44	1,085	68.2	8	9
21 - Miscellaneous edible preparations	16	378	68.3	9	17
22 - Beverages, spirits, vinegar	22	502	67.8	9	7
23 - Residues, wastes of food industry, animal fodder	25	175	51.6	13	12
24 – Tobacco, manufactured tobacco substitutes	9	58	6.3	8	7
29 - Organic chemicals	2	8	35.3	8	7
33 - Essential oils, perfumes, cosmetics, toiletries prep.	14	52	16.0	10	9
35 - Albuminoids, modified starches, glues, enzymes	10	54	24.7	8	7
38 - Miscellaneous chemical products	1	3	3.2	8	7
41 - Raw hides, skins (other than furskins), leather	12	139	38.4	8	7
43 – Furskins, artificial fur, manufactures thereof	9	199	39.4	8	7
50 – Silk	1	5	0.3	8	7
51 - Wool, animal hair, horsehair yarn & fabric thereof	10	75	15.6	8	7
52 – Cotton	4	8	15.4	8	7
53 - Vegetable textile fibers, paper yarn, woven fabric	2	2	0.0	8	7
<i>Total</i>	<i>769</i>	<i>14,298</i>	<i>61.1</i>		

Note: 12 concerns are not reported in column (4). 6 deal with genetically modified organisms (GMOs) and for the 6 other concerns, the WTO SPS-STC database does not provide information on the products. Similarly, 8 concerns are not reported in column (5) (4 deal with GMOs and for the 4 others, information is not provided). In columns (4) and (5), total calculation does not make much sense since some concerns are not reported and since we create separate records to account for all HS2 sectors and OECD countries involved in trade concerns.

Table 2. Clusters characteristics

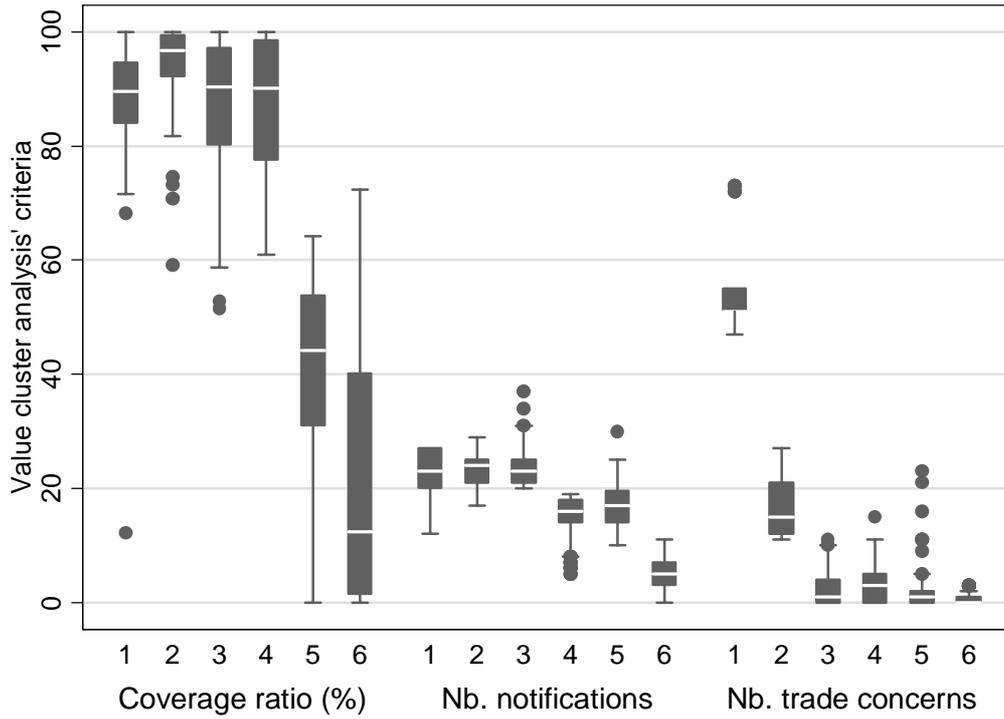
	Number of observations	Within cluster standard deviation	Avg distance from cluster center	Max. distance from cluster center
Cluster 1	25	0.54	1.01	2.84
Cluster 2	131	0.24	0.61	1.23
Cluster 3	195	0.27	0.54	2.13
Cluster 4	216	0.31	0.64	1.65
Cluster 5	116	0.40	0.84	2.12
Cluster 6	94	0.36	0.79	1.71

Table 3. Euclidean distances between cluster centers

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Cluster 1	0					
Cluster 2	3.60	0				
Cluster 3	4.86	1.29	0			
Cluster 4	4.90	1.65	1.19	0		
Cluster 5	5.21	2.46	1.91	1.66	0	
Cluster 6	6.07	3.88	3.49	2.69	1.82	0

Table 4. Mean for each criterion, by cluster and for the whole sample

	Trade coverage ratio (%)	Number of notified NTMs	Number of SPS trade concerns
Cluster 1	85.7	23.2	57.1
Cluster 2	94.6	22.9	16.5
Cluster 3	87.3	23.2	2.2
Cluster 4	87.6	15.4	3.4
Cluster 5	38.6	16.7	2.1
Cluster 6	21.6	5.3	0.7
<i>Whole sample</i>	73.3	17.9	6.5



Note: The ends of the vertical lines (or “whiskers”) represent the minimum and maximum data values, excluding extreme values. Extreme values are represented with round markers.

Figure 1. Box plots for each criterion, by cluster