



UNIVERSITÀ DELLA CALABRIA



CEPII



University of Dublin
Trinity College



UNIVERSIDAD
POLITECNICA
DE VALENCIA



A Critical Survey of Databases on Tariffs and Trade Available for the Analysis of EU Agricultural Agreements

Sophie Drogué (INRA, France, coordinator) and *Lubica Bartova* (SAU, Slovakia).

With contributions by *Houssein Boumellassa* (CEPII, France), *Jean-Christophe Bureau* (INAPG, France), *Dusan Drabik* (SAU, Slovakia), *Jozef Galik* (RIAFE, Slovakia), *Jacques Gallezot*, *Monique Harel*, *Fabrice Levert* (INRA, France)

Working Paper 07/6

TRADEAG is a Specific Targeted Research Project financed by the European Commission within its VI Research Framework. Information about the Project, the partners involved and its outputs can be found at <http://www.tradeag.eu>

A Critical Survey of Databases on Tariffs and Trade Available for the Analysis of EU Agricultural Agreements

Sophie Drogué^{a,b} and Lubica Bartova^c

with contributions by

Houssein Boumellassa^b
Jean-Christophe Bureau^{a,b}
Dusan Drabik^c
Jozef Galik^f
Jacques Gallezot^{a,b}
Monique Harel^d
Fabrice Levert^e

a: UMR Economie publique, INRA-INAPG, Grignon, France

b: CEPII, Paris, France

c: Department of Statistics and Operations Research, Faculty of Economics and Management, Slovak Agricultural University, Slovakia

d: LERECO, INRA, Nantes, France

e: INRA, Rennes, France

f: Research Institute of Agricultural and Food Economics, Bratislava, Slovakia

Financial support by the ‘Agricultural Trade Agreements (TRADEAG)’ research project, funded by the European Commission (Specific Targeted Research Project, Contract no. 513666) is acknowledged. The paper is the authors’ sole responsibility and does not reflect the Commission's views.

INTRODUCTION

Applied trade analysis and modeling would not be possible without databases. Over the last few years, a considerable effort has been devoted to put together and harmonize a huge amount of data. This has led to the construction of large-scale datasets on trade (imports and exports in volume and value terms). In parallel, increasing attention has been devoted to measuring tariffs and other trade barriers.

These efforts are still under way. It is only recently that large-scale models, for example, have taken into account the lower tariffs applied under preferential agreements (the former generation of models only considered the "bound" tariffs that act as a ceiling in many countries). The changes resulting from the introduction of better data on tariffs have been very significant, and many institutions have revised considerably their estimates of, say, gains resulting from trade liberalization (Ackerman 2005). However, empirical difficulties faced by modelers suggest that there is still a need for improvement, especially in the area of applied tariffs and tariff rate quotas.

The aim of this study is to review the main data sources dealing with world tariffs and trade. After discussing some issues and methodology, we provide a brief description of the main datasets available in terms of the origin of the data, its accessibility, reliability and shortcomings.

1. SOME TECHNICAL AND METHODOLOGICAL ISSUES

1.1 Data classification

Statistical classifications: focus on the HS. Several classification systems still coexist. A large number of datasets and models still rely on the SITC (Standard International Trade Classification). Some research groups, which play an important role in the modeling of trade policies, use their own classification. This is the case of the GTAP (Global Trade Analysis Project) which groups goods under a set of aggregates, defined as the GTAP classification system, or GSC.

However, more and more data sources use the Harmonized Commodity Description and Coding System, popularly known as the Harmonized System or HS. The HS is a multipurpose system, used as the basis for both customs tariffs and the compilation of trade statistics. Given that most modelers now use international statistics developed under the HS, we will mainly focus on the HS classification in the rest of this paper.¹

The HS was developed by the World Customs Organization (WCO). It was implemented in 1988 by an international convention. The HS is maintained by the WCO, through the Harmonized System Committee. The United Nations (UN) Statistical Commission has adopted the HS as the core classification for goods. The HS now covers 98 percent of world trade, according to the WCO. All the member states of the World Trade Organization (WTO) have (in principle) adopted the HS system. All major goods classification systems are correlated to the HS. There are correspondence tables (although at a relatively aggregated level) between the SITC classification and the HS, for example. The GSC aggregates also rely on groupings, along HS lines.

¹ Note, however, that in most general equilibrium models, the Social accounting matrices rely mainly on the GSC, i.e. the GTAP system of classification.

Caveats of the HS classification. The HS is supposed to be a universal language and code for transportable goods in international trade. Nevertheless, the use of the HS classification does not solve all the problems when working on detailed tariff lines.

The first problem is that the HS classification is "harmonized", only up to the 6-digit level. Beyond that, each country uses its own classification. For example, if one focuses on tariffs, most of the tariff schedules submitted to the WTO in the case of agricultural products are specified at the 8-digit level. In some cases, the countries even have bound tariffs at the 10- or 11- digit level under their international obligations (as is the case of some Asian countries, for example).

The lack of international harmonization beyond the 6-digit level raises several problems. Matching tariff lines with different codes is cumbersome when one attempts to make cross-country comparisons. The exact definition of a good under a particular tariff line is not consistent across countries, and sometimes is not even consistent across datasets in any given country (a frequent problem when matching trade and tariff data). The lack of harmonization of the classification at a detailed level often makes international comparisons of tariffs for a particular commodity almost impossible beyond the 6-digit level. This also makes it particularly difficult to construct *ad valorem* equivalents (i.e. tariffs expressed in percentages) of some specific tariffs (i.e. tariffs expressed in dollars, euros or yen per tonne, bushel, liter, etc.)

The second major empirical problem is that the Harmonized System has experienced several revisions over time. Most countries submitted their tariff schedule to the WTO under the HS-92. All of them were supposed to update the information under the HS-96, right after the Uruguay Round (few of them actually did so on time, and some were still behind schedule in 2002). Recently, many countries have adopted the HS-2002 system. The changes in the successive versions of the HS have created a number of difficulties. As not all countries have shifted to the most recent version, many of the statistical datasets available are still under HS-96. It is possible to establish some correspondence between other sources available, i.e. the HS-92 (for the tariff schedules of several countries), the HS-2002 (for recent information on US and EU tariffs and trade), and the Standard International Trade Classification (for the definition of processed, semi-processed and primary commodities). However, a lot of information is lost by doing so.

The goods covered by the agricultural agreements in the HS classification. The goods covered by the Uruguay Round Agreement on Agriculture include some non-food products of agricultural origin. That is, in the HS classification, they correspond to the items of Chapters 01, 02, 04 to 24 (i.e. all food and agricultural products, with the exception of fish products), and for selected items of Chapters 29, 31, 33, 35, 38, 40, 41, 50, 51, 52, 53 (that is, non-food agricultural products, such as skins for leather, etc.)

Table 1. HS Chapters 2-digit level. Goods covered by the Uruguay Round Agreement on Agriculture

| | |
|----|---|
| 1 | Live animals. |
| 2 | Meat and edible meat offal. |
| 4 | Dairy produce; birds' eggs; natural honey; edible products of animal origin, not elsewhere specified or included. |
| 5 | Products of animal origin, not elsewhere specified or included. |
| 6 | Live trees and other plants; bulbs, roots and the like; cut flowers and ornamental foliage. |
| 7 | Edible vegetables and certain roots and tubers. |
| 8 | Edible fruit and nuts; peel of citrus fruit or melons. |
| 9 | Coffee, tea, mate and spices. |
| 10 | Cereals. |
| 11 | Products of the milling industry; malt; starches; inulin; wheat gluten. |
| 12 | Oil seeds and oleaginous fruits; miscellaneous grains, seeds and fruit; industrial or medicinal plants; straw and fodder. |
| 13 | Lac; gums, resins and other vegetable saps and extracts. |
| 14 | Vegetable plaiting materials; vegetable products not elsewhere specified or included. |
| 15 | Animal or vegetable fats and oils and their cleavage products; prepared edible fats; animal or vegetable waxes. |
| 16 | Preparations of meat, of fish or of crustaceans, mollusks or other aquatic invertebrates. |
| 17 | Sugars and sugar confectionery. |
| 18 | Cocoa and cocoa preparations. |
| 19 | Preparations of cereals, flour, starch or milk; pastrycooks' products. |
| 20 | Preparations of vegetables, fruit, nuts or other parts of plants. |
| 21 | Miscellaneous edible preparations. |
| 22 | Beverages, spirits and vinegar. |
| 23 | Residues and waste from the food industries; prepared animal fodder. |
| 24 | Tobacco and manufactured tobacco substitutes. |
| 29 | Organic chemicals. |
| 31 | Fertilizers. |
| 33 | Essential oils and resinoids; perfumery, cosmetic or toilet preparations. |
| 35 | Albuminoidal substances; modified starches; glues; enzymes. |
| 38 | Miscellaneous chemical products. |
| 40 | Rubber and articles thereof. |
| 41 | Raw hides and skins (other than furskins) and leather. |
| 43 | Furskins and artificial fur; manufactures thereof. |
| 50 | Silk. |
| 51 | Wool, fine or coarse animal hair; horsehair yarn and woven fabric. |
| 52 | Cotton. |
| 53 | Other vegetable textile fibers; paper yarn and woven fabrics of paper yarn. |

Note: beyond Chapter 24, only selected goods of Chapters 29 to 53 are covered by the Agreement.

Source: WTO schedules.

1.2 Tariffs

Bound and applied tariffs. A basic principle of the 1947 General Agreement on Tariffs and Trade (GATT) is that protection of domestic industries, where deemed politically necessary, should be provided through the least-distortionary means, i.e. by customs tariffs administered without discrimination. Maximum tariff levels should be "bound" (i.e. cannot be increased without notification and compensation). This provides third countries with the guarantee that tariffs cannot be raised above negotiated levels without consultation and compensation. In this way, the level of protection is transparent and predictable for traders. It is also more susceptible to negotiation between governments.

Since the Uruguay Round, most countries have published national tariff schedules. However, many apply tariffs at a lower level than the bound tariffs, which act as a ceiling. The official WTO schedules also provide the "bound rates" that prevail at the end of the implementation period of the Agreement, i.e. 2001 for developed countries and 2005 for developing countries. The bound rates are the maximum rates that apply to WTO member countries for trade under the Most Favored Nation (MFN) regime.

Box 1. Some useful definitions

The Most Favored Nation, or MFN. As far as tariffs are concerned, the MFN tariffs are those that apply to all WTO members, even though some particular members can access some markets at preferential (lower) tariffs, under preferential agreements. In practice, the MFN, bound tariff is the ceiling tariff applied on exports of WTO members. Higher than bound tariffs can be imposed on non-WTO members (there are also a few exceptions, such as the imposition of "punitive" tariffs, higher than the bound rates, by the United States to Cuba, which is a WTO Member. This is a violation of the GATT principles). However, not all countries apply the bound tariffs on an MFN basis: some countries may choose to apply lower tariffs than the bound ones, on a non-preferential basis.

Tariff (specific and *ad valorem*). A duty imposed by the importing country on goods entering its territory from another country can be expressed as a percentage of the value of imports. In this case, the rate of duty is called an *ad valorem* tariff. The duty can be expressed in dollars, euros, yens, etc. per tonne of product, per liter, per head, per other physical unit, per tonne of pure alcohol content equivalent. In such cases the rate is called a "specific tariff". Specific and *ad-valorem* tariffs can be combined within the same tariff line.

Water in the tariffs. Several definitions are used (often, water in the tariffs is assimilated to binding overhang, sometimes to cut tariff dilution). In order to avoid confusion, one can consider that there is some water in the tariff when a cut in the tariff will not lead to an increase of trade (more exactly, an increase in effective market access). With this definition, water in the tariffs covers three cases: binding overhang, large preferential margins, and prohibitive mega-tariffs.

Binding overhang. When a country has set a bound tariff at a level higher than the tariff it applies in practice (often, in order to keep a margin for a possible increase in the applied tariff up to the bound tariff), the expression "binding overhang" is used.

Bound tariffs/base tariffs. The bound tariffs are the tariffs that cannot be increased unilaterally, without prior notification and authorization by the members of the World Trade Organization. In the WTO schedules, the term "bound" has a narrower meaning. Indeed, WTO members notified a list (schedule) of "base" tariffs. The base tariffs were the 1995 MFN tariffs, which had to be decreased over an implementation period. The tariffs at the end of the implementation period are called the bound tariffs in the schedules (they should actually be called final bound tariffs). More specifically, in the case of developed countries, the base tariffs are actually tariffs that were bound in 1995, resulting from the tariffication process, and the term "bound tariff" is mostly used to describe the final (i.e. 2001) bound tariffs. For developing countries, the situation is more complicated since a number of them were free to decide the base tariff on which the reduction commitments were applied. The resulting tariff is still called the bound tariff, i.e. a ceiling tariff at the end of the implementation period; 2005.

Applied tariffs are the tariffs imposed by the customs officers at a certain date. The bound rates act as a ceiling for WTO members, but many countries actually apply lower tariffs on an MFN basis. A larger set of countries applies lower than MFN tariffs under preferential agreements (free trade agreements, Generalized System of Preferences, preferential access for certain countries or regions, special agreements with developing countries, etc.). Applied tariffs also include the lower-than-MFN tariffs that are applied within tariff rate quotas (generally called "in-quota tariffs").

Schedules. The official tariff commitments for WTO members are specified in the Schedules, which are legally binding documents defining the bound (MFN) tariffs for a list of commodities.

Current access. The 1994 Marrakech (Uruguay Round) Agreement on agriculture specified that, after tariffication, current market access (i.e. the level of imports that existed during the reference period) had to be maintained or increased. For some countries, this was achieved by the opening of tariff rate quotas. These quotas are called "current access" quotas (as opposed to quotas open under minimum access).

Minimum access. The Marrakech Agreement specified that access to domestic markets had to be open to imports for up to 5 percent of the domestic consumption over 1986-1988 levels, starting in 2001, for developed countries. For countries that still maintained high tariffs, this was achieved by the opening of tariff rate quotas. These quotas are called "minimum access quotas".

Notifications. GATT rules specify that, under the obligations of transparency, member countries must notify the way they fill their obligations and implement their commitments under the market access provisions. A set of documents is submitted to the WTO on a regular basis. It includes modifications in the Schedules, the way tariff rate quotas are filled and administered, etc. Most of these documents are available on-line on the WTO web site.

Tariff cut dilution. The Marrakech Agreement specified that tariffs had to be reduced by a given average (36 percent for developed countries) over the implementation period (1955-2000 for developed countries). The term "tariff cut dilution" is often used to express the idea that many countries have reached this objective by higher percentage cuts on less politically sensitive tariffs, and minor cuts (often 15 percent) on more sensitive products.

Tariff rate quotas, or TRQ. A quantity of imports is set, for a given product that enters the territory with a lower tariff than the tariff applied beyond this quota quantity.

Fill rate. The fill rate of a TRQ describes the proportion of imports, during a given year, relative to the commitment in terms of import quantities, as defined by the tariff rate quota.

Safeguards. The Marrakech agreement on agriculture allows for special temporary safeguard mechanism, for products subject to tariffication. They are imposed if increases in volumes or drops in import prices exceed certain trigger levels.

Tariffication, "dirty tariffication". Tariffication is the conversion of non-tariff barriers that existed before the Uruguay Round, for agricultural products, into tariffs that were supposed to bring an equivalent level of protection. Some countries, however, are said to have set base tariffs at a level that was higher than the one actually provided by the former measures that they replaced. This is described by the term "dirty tariffication". During the tariffication process, developed countries used current bound rates for products that were previously bound. If there were quantitative or other non-tariff barriers, a tariffication formula was used. If there was only a tariff, but that it was not bound, the rate that applied in September 1986 was used for the binding. In the case of developing countries, if the tariff was previously unbound, the country could offer a binding ceiling. The offer was not necessarily derived from a computation of the tariff equivalent.

Specific tariffs and ad valorem equivalents. An important point is the computation of *ad valorem* tariff equivalents. Indeed, in many countries, some tariffs are specified in values per kilo, liter, or head of animal, i.e. as "specific" tariffs. This is particularly the case in the EU, US, Japan, Canada, Switzerland, but the problem exists in many other countries (see Table 2.). There are roughly 5600 tariff lines that include some specific elements, for a set of 127 countries, for which a full Schedule is available for agricultural goods.

Table 2. Importance of non-*ad valorem* tariff lines, in selected countries (WTO schedules).

| | |
|---|--|
| Specific tariffs present in 1% to 20% of tariff lines | |
| | Australia, Brunei, Bulgaria, Egypt, India, Israel, Japan, Korea, Malaysia, Mexico, New Zealand, Papua New Guinea, Singapore, Solomon Islands |
| Specific tariffs present in 20% to 50% of tariff lines | |
| | Canada, Cyprus, European Union, Iceland, Poland, Slovenia, Thailand, United States |
| Specific tariffs present in more than 50% of tariff lines | |
| | Malta, Norway, Switzerland |

Source: WTO

One cannot perform any aggregation, computation and economic analysis on the basis of such tariffs, and they have to be converted into *ad valorem* equivalents. This requires import price data, available at the same level of disaggregation, which is a major obstacle. In some countries, complexity is introduced by the fact that the units in which the specific tariffs are expressed differ. In the case of spirits and wines, the tariff sometimes refers to the percentage of alcohol, for example. In other cases, specific tariffs are specified per head of cattle or per liter, etc. There are also tariffs that vary during the year, especially in the fruit and vegetable sector. There are also some additional tariffs that are proportional to the content of some key agricultural components (such as sugar) within final products. This is a particular problem in

the EU, for example, because a tariff for a given processed product has to be increased by a complex formula taking into account the quantity of sugar, butter, milk-powder and flour used as material.

There is no satisfactory convention for converting specific rates into *ad valorem* tariffs. Local approximations will always have to be used, which are not necessarily robust to changes in tariffs. The calculated value depends on the choice of import prices and the exchange rate, both of which can change over time.

Most databases available use one of the two conventions for converting specific tariffs into *ad valorem* equivalent. The first convention is to use a reference price that is the same for all countries, in order to avoid the bias that results from importing only particular varieties of a good. In other words, the convention is to use the unit value of world trade as a reference price for the conversion of the specific tariff. Given that the statistical systems are not harmonized beyond the 6-digit level, this convention imposes the use of a reference price at the 6-digit level for converting specific tariffs which are often expressed at the 8-digit level. That is, the tariff is converted into an *ad valorem* equivalent using an average price that refers to a more aggregated commodity.

The second convention is to match each tariff with a reference price, i.e. to use a price consistent with the exact definition of the tariff line. The consequence is that the unit value of imports in the country in question has to be used, in most cases. This may introduce a bias when tariffs are high, and there are few imports which refer to a very specific product quality.

If we compare the results of the two conventions, applied to the same vector of specific tariffs in the EU and the US, we find significant discrepancies in the average tariff (see Table 3.).

Table 3. Comparison of average tariffs using two assumptions for converting specific tariffs into *ad valorem* equivalents.

| | European Union | United States |
|--|----------------|---------------|
| Using a world price calculated as a 4-year average of unit values of domestic imports, at the 8-digit level (source COMEXT and USITC). | 17.9% | 7.1% |
| Using a world price calculated as the unit value of the worldwide transactions, at the 6-digit level (source COMTRADE). | 30.1% | 11% |
| "Official" <i>ad valorem</i> equivalents | 23.1% | 11.2% |

Methodology: average tariff for the goods covered by the Agricultural Agreement of the Uruguay Round, non-weighted arithmetic mean on the basis of the WTO schedule, at the 8-digit level. In the first row, the conversion of specific tariff rates into an ad valorem equivalent tariff is made using the 4-year average unit value of imports, at the 8-digit product tariff line for the country in question. In the second row, the conversion of specific tariff rates into an ad valorem tariff uses the reference prices constructed by the Economic Research Service of the US Department of Agriculture in the framework of the AMAD group (unit values of world trade at the 6-digit level, original source COMTRADE, see Gibson et al 2001). The third row corresponds to the "official" calculation of ad valorem equivalents, as computed following the methodology agreed in 2005, under the WTO (note that there is no agreement on an official method for sugar).

Table 3 shows that the assumptions made for converting the specific tariffs into an *ad valorem* equivalent have a major influence on the value of the average tariff. The convention that relies

on a world price reflecting a 4-year average unit value of domestic imports results in a much lower average tariff than the convention that uses a single world price common to all countries (i.e. the unit value of worldwide trade) at the 6-digit level. The problem is that both conventions are equally defensible. Using the value of domestic imports at the 8-digit level may introduce a bias, since sometimes these are only very specific qualities that are imported (e.g. wheat in the EU), and the unit value is very high. The *ad valorem* equivalent of the tariff will therefore be biased downwards. Using the unit value of world trade at the 6-digit level may result in artificial tariff peaks, since the specific tariff at the 8-digit level actually applies to a category whose unit value is much higher than the one we consider, by taking an average unit value at the 6-digit level.

An official method of conversion into ad valorem equivalents. In the WTO negotiation process, it has become essential to express all specific and composite tariffs in *ad valorem* equivalents. Indeed, the August 2004 Framework Agreement states that higher tariffs will experience a larger cut. This requires a tiered approach, and therefore the ability to express all tariffs as a percentage, so as to be allocated to a particular band where a particular reduction rate will be applied. (An alternative to the band system presently under discussion, is the application of the "Swiss formula" which cuts tariffs proportionally to their initial level and introduces a maximum tariff. It also requires expressing all tariffs in an *ad valorem* equivalent).

Even though none of the existing databases has yet included the resulting *ad valorem* equivalent, an official methodology was adopted in June 2005, after 8 months of negotiation. The conversion of specific and composite tariffs relies on the following guidelines:

- Convert the final bound non-*ad valorem* duties into *ad valorem* equivalents using the unit value method based on the import data contained in the WTO Integrated Data Base (IDB), for the 1999-2001 period.
- If the IDB data for the tariff line is either: i/ missing, contains errors, or is lower than US\$2500 on average for 1999-2001 or ii/ the IDB-based *ad valorem* equivalent cannot be considered to reflect the true level of tariff protection afforded by the non-*ad valorem* tariff, an alternative method will be used.
- For the tariff lines that fall into category i/ above, there are four possible alternative methods: a/ extend the base period 1999-2001 by up to two years at either end; b/ use the IDB import value of a closely related tariff line; c/ use the IDB import unit value of the tariff line at issue of a nearby country; d/ use the United Nations database COMTRADE unit value (i.e. a unit value at the HS6 level).
- For the tariff lines that fall into category ii/ above, the "40/20" filter has been developed. In other words, the conversion of non-*ad valorem* duties into their *ad valorem* equivalent will be calculated using the following weightings, based on unit values of COMTRADE and IDB data:
 - (a) For HS Chapters 1 to 16, and the products in Annex 1 of the Agreement on Agriculture in the HS Chapters beyond Chapter 24, a 82.5/17.5 (COMTRADE/IDB) weighting will apply.
 - (b) For HS Chapters 17-24, a 60/40 (COMTRADE/IDB) weighting will apply.
 - (c) For all tariff lines for raw and refined sugar, world prices will apply, "with prices to be agreed."

While this method is used to calculate the tariff equivalents so as to define the cut that would be applied to the particular product (when the Modalities of the Doha Round will eventually

be finalized, assuming an agreement is reached), it is likely that specific and composite tariffs will persist.

1.3. Trade flows

Trade flows are other important variables for assessing trade agreements. They are obviously essential for assessing the quantitative effects of a given trade policy or agreement. In addition, reliable data on imports are necessary for calculating the unit value (i.e. the ratio of the value of imports to the quantity of imports), used for converting specific tariffs into *ad valorem* equivalents and to perform tariff aggregation and inter-country comparisons. Statistics on imports are also necessary for constructing weighted tariff averages and tariff profiles.

International datasets on imports (i.e. the COMTRADE dataset of the United Nations, whose import flows are also reported in the TRAINS dataset by the UN Conference on Trade and Development, hereafter the UNCTAD, and the World Bank's WITS dataset), do not provide information beyond the 6-digit level. More detailed data require using national sources. Nevertheless, these datasets are not harmonized across countries, and the classification also differs sometimes significantly from the one used for tariffs. In addition, the units in which specific tariffs are expressed are not consistent with those used in the trade data (e.g. heads of cattle, dozen of eggs, hectoliters, etc.) National sources are necessary to go beyond the HS6 level.

1.4. The databases considered in this study

In the following sections, we will review the various data sources available for tariffs, trade flows or for both of them. We will successively provide a brief description and an assessment of the datasets listed below. For some of the most-used datasets, we will assess the reliability of all the data by comparing it with a reliable domestic source, the external trade statistics of Slovakia. As Slovakia is representative of countries of particular interest in the TRADEAG project, and as it is a country for which international databases have sometimes provided problematic information, we believe that this is a useful benchmark. The datasets surveyed include:

Trade

- COMTRADE
- BACI
- COMEXT
- FAOSTAT
- NBER (R. Feenstra)

Tariffs

- MacMap
- AMAD
- TARIC
- MAD DG Trade Market Access database
- IDB
- DBTAR
- APEC tariff database

Tariffs and trade combined

- USITC
- TRAINS
- TRADEPREF
- The WITS interface

2. TRADE DATASETS

In this section, we focus on the datasets that provide information on trade statistics. Note that the datasets that provide both access to tariffs and trade data are treated in Section 4. In some cases these latter datasets are among the most useful sources of information on trade, so the reader interested in trade statistics is invited to look both at Section 2 and Section 4.

2.1 The COMTRADE database

2.1.1 Description

COMTRADE is the United Nations trade database (COMTRADE stands for United Nations Commodity Trade Statistics Database). It is available at <http://unstats.un.org/unsd/comtrade/>. COMTRADE is an on-line database of detailed merchandise trade statistics on exports, imports and re-exports, for 275 existing and former countries. Flows are provided by products and country. Data are available on a yearly basis since 1962 for many countries and regularly updated to the most recent, full year. They are supposed to be collected and evaluated homogeneously from the various declaring countries. Values are in US dollars, quantities in kilograms.

Commodities are classified according to different, recognized international classifications. The Standard International Trade Classification (SITC) in its various revisions (Rev.1 since 1962, Rev.2 since 1976 and Rev.3 since 1988) and the Harmonized Commodity Description and Coding System (HS) since 1988 (revised in 1996 and 2002). Most data are now recorded in the 2002 HS nomenclature, but they are also converted and stored under all other classifications.

2.1.2 Origin of the data

The data stem from the declarations of the various countries, generally by national customs authorities. Some 130 countries are supposed to provide the UN Statistics Division with their annual international trade statistics, detailed by commodity and partner country, every year. Their lists and links are available at http://unstats.un.org/unsd/methods/inter-natlinks/sd_natstat.htm. The international statistics are based on the sources listed at http://unstats.un.org/unsd/methods/inter-natlinks/sd_intstat.htm. These data are processed into a standard format, with consistent coding and valuation. All values are converted into US dollars using exchange rates supplied by the countries, or derived from monthly market rates and volume of trade. Quantities are, if possible, converted into metric units. The reporting country (or reporter) is the country declaring to the Statistics Department of the UN.

2.1.3 Accessibility and user friendliness

The database is accessible via on-line query at the following address: <http://unstats.un.org/unsd/comtrade/default.aspx>. Three user statuses are possible: guest user, registered user (with password authentication on the basis of a single license) and registered user with IP authentication (site license). Guest users enjoy restricted access, with a maximum capacity of 1000 records per query and cannot download results. Subscription fees vary between USD100 to USD3500, depending on the country (developed/developing) and the status of the user (academic or not).

The system is rather user-friendly thanks to a graphic interface which allows rapid visualization of simple queries. Several tools are available to handle a query according to the user's needs and experience (*COMTRADE Explorer* is easier to use; *Quick search* dialogue box; *Shortcut query* etc.). *Basic Selection* is recommended for the more complex queries if the

user is not really familiar with the codes. *Express Selection* is useful for experienced users as all queries are carried out by indicating the commodity and country codes. Finally, the registered user can directly download the results in a variety of formats, including those commonly accepted by spreadsheets.

The user can select data: by reporting country, by commodities, by years and modify selection under the classification, commodities, reporting country, partners, years and trade flows. There are additional, advanced options for selecting, sorting order and filtering data, as well as for saving queries. The quantitative units used by exporters or importers, however, are not written explicitly and have to be consulted in a glossary.

Registered users are also provided access to COMTRADE via the WITS interface (World Integrated Trade Solution: <http://wits.worldbank.org/witsweb/>). This provides an integrated shell which is particularly convenient, provided that the users are registered as subscribers to other datasets such as TRAINS.

2.1.4 Reliability and caveats

The UN COMTRADE system is stable and quite efficient. The batch mode allows the user to submit simultaneously several large queries whose results are recoverable later. The wide range of functionalities and nomenclatures is a bit confusing for the inexperienced user. A positive aspect of the database is the possibility of displaying data on bilateral trade flows (if any) between countries.

However, there is some missing information. Import quantities do not always match import values, for example. The pros and cons of COMTRADE stem mainly from the origin of the data. Their collection from various countries guarantees their reliability and theoretically their coherence with the data published at the national level, though the diversity of sources generates many inconsistencies in the database. Even if data collection and recording are harmonized across countries, errors remain and it is very difficult to be certain about the accuracy of the information on commodity flows.

A caveat seems to be that exports to a partner country are generally expressed in FOB price (Free on Board) whereas imports are expressed in CIF price (Cost of Insurance and Freight included). This results in inconsistency in the data. Exports of a country A towards a country B are different in value to imports of B from A. The value of imports expressed in the CIF price must be greater than the value of exports expressed in the FOB price.

In addition to this CIF/FOB problem, other elements may explain the discrepancies between declarations affecting either the values or the volumes, on both sides of a trade flow.

- Typically a shipment leaving one country in December and reaching another one in January of the following year will be recorded as an export in year n and as an import in year $n+1$.
- The recording of re-exportations also poses a problem and it is not clear how the database takes them into account. Data of re-exportation are in theory available in COMTRADE, but the heading is seldom filled, at least for agricultural commodities. It is necessary to bear in mind that many products to or from Europe are forwarded through the large Belgium and Dutch ports of the North Sea.
- In theory, taxes should not be integrated in the value of imports recorded at the CIF price, but in practice there is some doubt on this issue. Generally speaking, exports seem therefore more reliable than imports.

Reliability of COMTRADE data was examined in the case of the Slovak Republic's exports and imports for selected commodities (potatoes, rapeseed and beer), by comparing it with data reported by the Statistical Office of the Slovak Republic (SSO). In the two last columns the ratio of COMTRADE to SSO data is given. There is a close similarity between the data compared (see Table A1 in the appendix). Overall, the two sources match well.

Gehlhar (1996) compared total export and import values for 25 countries, from the database of the UN COMTRADE, the Food and Agriculture Organisation (FAO), the World Bank (WB) and the International Monetary Fund (IMF). He did not find comparability between FAO and COMTRADE, despite their shared primary data sources. Gaulier and Zignago (2005) point out that, at the 6-digit level, the gap between mirror flows exceeds 100% for half of the observations in COMTRADE. Our empirical comparison with Slovak export data indeed reveals significant differences with the amounts of dairy products that other countries declare as being imported from Slovakia: the overall difference being 33%. This exceeds what would normally be expected from differences in statistical treatment, and highlights the problem of mirror flows derived from different sources. They are inconsistent because bilateral trade flows have not been reconciled. Discrepancies make country totals unreliable and lessen the overall integrity of the structure of world trade. Countries are responsible of these discrepancies but they is no information on the reliability of one country compared to another. Gehlhar (1996) pointed out that almost 73% of bilateral flows are unreliable. But they represent only 2% of the total value of trade. For more than 75% of the bilateral flows, the discrepancy was less than 25%, even without accounting for the CIF/FOB difference.

Overall, even though COMTRADE has some serious caveats, it is a database that is often the only source of information on particular countries. Efforts by the CEPII to make import and export declarations more consistent have led to the BACI dataset.

2.2 The BACI database

2.2.1 Description

BACI (*Base pour l'Analyse du Commerce International*, Database for International Trade Analysis in English) aims at providing detailed trade statistics for the largest possible number of countries and years. It was developed by the *Centre d'Etudes Prospectives et d'Informations Internationales* (CEPII), using the same data source as COMTRADE, at the 6-digit level of the HS. Procedures have been developed to ensure consistency and to make the data more convenient for trade modeling and analysis. In particular, the original COMTRADE data are harmonized, in the sense that a unique trade value is computed for each trade flow, initially reported with some discrepancy by the exporter and the importer (see above).

BACI includes various levels of analyses and classifications. Trade flows are characterized in terms of trade types (one-way trade, cross-trade in similar products, cross-trade in vertically differentiated products), product ranges, technological levels and stages of production.

In addition, BACI permits the calculation of price (unit value) indices. Median-Tornqvist price indexes are computed at the country or sector level using unit value ratios.

2.2.2 Origin of the data

The treatment of the current version of BACI covers the period 1989-2003, for all countries declaring their annual international trade statistics to the UN (COMTRADE database). Trade flows are reported in value and quantity by both the exporting and importing country (mirror

flows, when available). Researchers from the CEPII have developed original procedures to harmonize COMTRADE data, including: the evaluation of the quality of country declarations to average mirror flows, the evaluation of CIF rates to reconcile import and export declarations, and the conversion in tonnes for other units of quantities traded. In the current version of BACI, the source data are classified in HS 1988 and 1996 and do not include flows less than USD1000, before 2003.

The evaluation of CIF rates to reconcile import and export declarations intends to solve the fact that, in COMTRADE, import values are reported CIF and the exports are reported FOB. BACI uses mirror flow ratios and a gravity-type equation as estimates of CIF. The equation is estimated by OLS, on pooled data. When there is a strong positive relationship between the ratio of mirror flows in value and those in quantity (discrepancies are likely to be observed simultaneously for values and quantities) observations are weighted in the equation for implicit CIF, by the inverse of the gap between reported mirror quantities ($\text{Min}(QX_{ij}, QM_{ji}) / \text{Max}(QX_{ij}, QM_{ji})$). This gives greater weight to trade flows equally reported by partners, while differences between reported import and export values are then more likely to reflect freight costs. The estimated coefficients for respectively distance and unit value imply that CIF increases with distance and decreases with unit value. Predicted CIF ratios are used to compute FOB import values. This procedure is not applied when it widens the gap between mirror flows.

The assessment of the quality of country declarations is necessary to average mirror flows. Given the great discrepancies observed between reported mirror flows, trade data have to be harmonized (at the 6-digit level, the gap exceeds 100% for half the observations in COMTRADE). Harmonization is carried out in two steps in BACI.

Quality indicators of import and export declarations are calculated for each country. They are used, in the second step, as weights when averaging the mirror flows, to provide the harmonized flow. In order to evaluate the quality of the declarations of countries (as exporters or importers), the absolute value of the ratios of mirror flows is decomposed using a (weighted) variance analysis. The error variable (absolute value of the logarithm of the ratio of mirror flows) is regressed on four sets of fixed effects, for the reporting country, for the partner, for the 6-digit product and for years). An ordinary least squares estimator is used, each observation (trade flow) being weighted by the logarithm of the sum of the two reports. The fixed effects estimated give the marginal impact on discrepancies between flows that can be attributed to all countries or sectors. Therefore, the (relative) quality of declaration of a country i is “cleaned” of the effects of its specialization (the share of poor/good reporters in its trade partners and the share of products with frequent report errors – due to the lack of homogeneity in the 6-digit position, for instance). Regarding the conversion of quantity units, 86% of quantities are declared in tonnes. The other quantities are converted into weights by estimating for each 6-digit product a rate of conversion between each unit (unit, watt, meter, etc.) using mirror flows reported in heterogeneous units.

2.2.3 Accessibility and user-friendliness

During its development phase, BACI is being restricted to internal use at the CEPII. Conditional access may, however, be provided to would-be users, as long as that they have subscribed to COMTRADE. BACI is available using the SAS^(R) software. In other words, some expertise in SAS is required to extract information.

2.2.4 Reliability and caveats

The BACI database is still in a version that has not been fully tested for consistency.

2.3.The COMEXT database

2.3.1 Description

COMEXT is the trade statistics database of the European Union (EU), gathered by the European Communities Statistical Office (Eurostat). COMEXT covers all intra- and extra-EU bilateral trade: i.e. trade between the 25 member States, and between each Member State and their 250 partners in the world. Data are available since 1987, in monthly and annual form (for EU-15), and since 1999 as annual data (for EU-25). Trade flows are reported in values and quantities. The data cover more than 10,000 products. Statistics are available from the more detailed level (8-digit). At 2-digits, there are 99 chapters; the first 24 relating to food and agricultural products. And since 2007, data are also available by import eligibility (only MFN, GSP and/or preferences, unknown) and five import regimes (MFN zero, MFN non-zero, zero preference, non-zero preference, unknown, for the period 2000-2006).

In the EU, goods declared by customs are classified according to the combined nomenclature (NC), in force since 1988. The combined nomenclature is used by the EU for the collection and the processing of trade data. This nomenclature is based on the HS nomenclature, but includes an extra level called the NC sub-position. COMEXT uses a statistical procedure (one digit coding) based on the customs system. SP1= Normal imports and exports; SP3= Imports after/export for outward processing; SP4= Total imports/exports (default for all intra- and extra-EU trade); SP5= Imports for/exports after inward processing, suspension system; SP6= Imports for/exports after inward processing, drawbacks or repayment system; RS7= Textiles, imports after/exports for outward processing (since 1995) (NB: SP4=SP1+SP3+SP5+SP6+SP7).

Data are provided by each Member State. Values are given in local currency and when necessary converted into euros by Eurostat. They are expressed in euros (online version) or in EUR thousands (in the DVD version). Quantities are expressed in 100 kilos (online) or metric tonnes (DVD). For some products quantities are given in supplementary units (liters, number of specimens etc.). The detailed data are broken down by product, according to the combined nomenclature, whereas the aggregate data are broken down according to 2 nomenclatures: the Standard International Trade Classification, Revision 3 (SITC, Rev.3) and the Broad Economic Categories, Revision 3 (BEC, Rev.3).

2.3.2 Origin of the data

EU Member States gather and provide the data to Eurostat. Two different systems are used to build the base: Intrastat for intra-EU trade and Extrastat for extra-EU trade. Regarding intra-EU trade, data relating to trade between the Member States are collected from the firms' monthly statistical declarations, the so-called Intrastat System. Extra-EU trade: data on extra-community trade, the so-called Extrastat system, are gathered from customs declarations, the SAU (Single Administrative Document). Many Member States use simplified procedures (e.g. electronic declarations).

2.3.3 Accessibility and user-friendliness

The COMEXT database is available on the Eurostat website, on CD/DVD or as hardcopies.

Online access: free access to the COMEXT data is provided at the following address: <http://epp.eurostat.cec.eu.int>. It is recommended to use Internet Explorer since, for some strange reason, data selection problems arise with alternative browsers such as Netscape or

Mozilla. Making an on-line query on COMEXT is easy, using a sequential menu. Users can extract a maximum of 30,000 cells, via an interactive interface. They can also access metadata (methodology, nomenclatures, tables of correspondence, etc.) Users can also be given a free login and password, which are required when detailed data on external trade is selected. This makes it possible to download 60,000 cells in batch mode. For example: value, quantity and additional units of meat imports and exports, (79 products, 8-digit level, 25 reporting countries and 5 partners for 2004). On-line extraction thus makes access to COMEXT easier, nevertheless it is limited by the maximum number of cells which can be downloaded. In addition, there is no pre-defined aggregate in the extraction menu. In other words, to obtain the sum of EU imports for a particular tariff line, it is necessary to download the individual imports from each country and sum them, which explains why most users reach the maximum number of cells very rapidly.

Interface/DVD (COMEXT standalone): COMEXT data are available by an interactive interface under Windows, which allows the extraction of monthly or annual data since 1995 according to the combined nomenclature. The extraction can also be done in ASCII format. The file's structure is as follows: reporter (3-digit code), partner (4-digit code), code of the product (CN8 8-digit; HS2 2-digit), flow (1 for imports, 2 for exports), statistical procedure (1-digit code), period, value in EUR thousands, quantities in tonnes, supplementary units. Each file corresponds to a month or a year: http://publications.eu.int/others/sales_agents_en.html .

Hardcopy: Trade statistical yearbook of the EU and monthly bulletins.

2.3.4 Reliability and caveats

Extra-Community export and import statistics are based on the European customs system. Eurostat is considered as a particularly reliable source as far as extra EU-trade is concerned. Several procedures for controlling data are carried out at a very detailed level (10-digit tariff nomenclature). This, along with a well-designed methodology are used to ensure harmonization of the data collected by the Member States.

However, gaps can be observed between COMEXT and the national data, due to differences of methodology in processing the data. The gaps are the following:

- Differences between the national and the Eurostat definitions of the territory (e.g. with or without French overseas territories).
- Goods in transit.
- Country-of-origin or country-of-shipment. EU tariffs depend on the country-of-origin.
- Differences in the treatment of the basic information by national customs.

Obviously, COMEXT only includes EU trade, and cannot be the main source of a world trade model.

Intra-EU trade statistics have now become quite unreliable. Indeed, EU integration has made customs procedures less and less relevant. This has decreased the degree of reliability of the information collected.

A comparison was made between the data of COMEXT and the data reported by the Statistical Office of the Slovak Republic (SSO). This comparison concerns Slovak bilateral trade with the Czech Republic and Germany, for the period 2001-2004. Data reported by COMEXT for Slovak imports from the Czech Republic in 2004 are almost 38% higher than data reported by the SSO (see Tables A3 and A4 in the appendix). Similarly, imports from Germany in 2004 are higher by 12%. Even if exports in COMEXT are given FOB and imports in CIF, such significant differences cannot be justified.

The differences in levels (e.g. around 10%) are quite noticeable for some countries, but for other countries the data are the same or the differences are very small. Moreover, even where differences in levels are noticeable, the differences in growth rates are small, and thus they should not affect analytical uses depending on growth rates (EC, 2005). Data management problems are regarded as major contributory factors to the differences between EU figures and other international sources. These problems usually arise from the following issues (Eurostat, 2005):

- Member States send their trade statistics to the UN, OECD or IMF, as well. The differences that exist between the data published by Eurostat and those published by the Member States will therefore be reflected in differences between Eurostat data and those published by the other international organizations.
- The national practices in revising data to correct past estimates are complex and vary between the Member States, as do their practices in providing revisions to Eurostat and other international organizations.
- Methods to convert national data into a common currency – euros for EU figures, dollars for other sources – may be different (monthly, quarterly, or annual conversions).

Another serious empirical problem for analysts and modelers is that the EU classification at the 8-digit level experiences frequent changes. This creates a considerable problem when using COMEXT, for time series analysis. The NC classification has evolved continuously evolving and the correspondence tables do not make it possible to construct fully consistent series, at the 8-digit level.

COMEXT now distinguishes trade flows between flows under the MFN regime and/or preferential regimes (with a distinction between imports subject to a zero tariff and imports subject to a strictly positive tariff in both cases). This is a valuable improvement. However trade analysts may regret that flows are not available under each specific import regime as some countries are eligible to more than one regime (GSP, ACP etc.) for the same tariff line. In addition, in some cases, a very large percentage of the trade flows remains unaffected between the MFN regime and the preferential regimes (and is included in an "unknown" category), suggesting that the EU statisticians have not been able to track whether these imports take place under a preferential regime or not (olive oil imports by the EU are an illustration of this problem).

2.4. FAOSTAT (Trade section)

2.4.1 Description

The Agriculture and Food Trade database is a part of the FAOSTAT statistical system. The FAOSTAT was developed in the early 1990s. Certain technical and functional limitations of this system have led to its modernization and improvement under the ongoing project FAOSTAT2. In the original version, the Agriculture and Food Trade database is divided into two subsets: "Crops & Livestock Primary & Processed" and "Live Animals". The database provides information on trade flows (exports and imports) of agricultural and processed food commodities. The FAO publishes trade data for aggregate agriculture sectors, as well as individual agriculture commodities. Trade flows are recorded in quantitative units (metric tonnes - Mt) and values (USD thousands). The trade volume is expressed in units, whose abbreviations can be found at: <http://faostat.fao.org/faostat/notes/units-e.htm>.

There are 310 countries and groupings in the database. The groupings can be displayed either as a whole, or disaggregated by particular member states. This is a strong advantage of the

data set. Among other countries, there are data on the CIS and the USA. Time series include the period 1961-2004. Time series are inconsistent for the countries that became independent during observed period (e.g. the Slovak and Czech Republics). A caveat to the database is the impossibility of displaying the territorial structure of a country's trade flows (exports or imports). In other words, the database does not allow the bilateral trade flows between countries to be viewed.

Export values are expressed as FOB and import values are CIF. The exceptions are Australia, Bermuda, Bulgaria, Canada, Czechoslovakia, Dominican Republic, Mexico (from 1992 to 1994), Papua New Guinea, Paraguay, Poland, Solomon Islands, South Africa, the USSR, Venezuela, Zambia and Zimbabwe, for which both imports and exports are FOB.

2.4.2 Origin of the data

The data are obtained from the governmental agencies and national statistical offices listed at <http://faostat.fao.org/abcdq/about.htm>. After a country selection, the user can get information on how the data are collected and treated. The data from EUROSTAT and the United Nations Statistics Division are used as an additional source of information. The dataset is very well documented. Explanatory notes on trade are available at: <http://www.fao.org/waicent/faostat/agricult/trad-e.htm>.

Missing data on quantities or/and values (due to lack of reliable sources, or the unavailability of information) are estimated on the basis of trade returns of trading partners. If there is only information on quantities available, corresponding values are estimated using the unit value data from trading partners.

2.4.3 Accessibility and user-friendliness

The data subsets can be accessed online. Free access is provided to occasional users, but very few cells can be extracted this way, due to strict limits on the data which can be downloaded freely.

<http://faostat.fao.org/faostat/collections?version=ext&hasbulk=0&subset=agriculture>.

We have not experienced difficulties accessing the data this way, free of charge. The data of interest can be displayed in tables with predefined X and Y fields, or saved in the form of a CSV file. In addition, simple statistics, net trade or a three-year average can be computed from the data displayed. The interactive form of the database is very easy to work with. Commodities and countries are displayed in an alphabetical order. It is possible to extract aggregated data and to group them. SITC codes are not indicated and this makes commodity selection more difficult. The database is well documented. Explanation of abbreviations and/or more information on particular items can be obtained in a "Help" section.

2.4.4 Reliability and caveats

In the database there are 638 items included under the "commodity" section. The Agriculture & Food Trade database uses the Standard International Trade Classification – SITC and this should be better indicated. A possible confusion could arise for a user accustomed to the Harmonized System classification (HS). An illustration is given by comparing the database with the Market Access Database (MADB) which uses the HS classification. The MADB states that Chapter 10 refers to "Cereals", while "Cereals" are under Chapter 04 in the FAO database. Another example is Chapter 22 in MADB – Beverages, which are Chapter 15 of the FAO database.

The reliability of FAO data is illustrated by a comparison of selected commodities, FAO data with the data published by the Statistical Office of the Slovak Republic (SSO) (see Table A5 in the Appendix). For the Slovak source data, the official, average exchange rate SKK/USD

according to the National Bank of the Slovak Republic (NBS) was used. There are many discrepancies in data on selected commodities, including rye, barley, maize, malt, rape and sunflowers. Regarding the trade flow of wheat in 2000, it seems that export and import quantities are obviously reversed in the FAO database (Table A6).

The reliability of data reported by the FAO database was assessed by Gehlhar (1996). According to him, FAO merchandise total is generally higher for both imports and exports. The reason for this is that the country totals are not strictly country-reported totals. The FAO substitutes partner's reported trade flows when a country fails to report its own trade flows. This is done at the detailed, commodity level for agriculture. By filling-in gaps, higher levels of trade are obtained in aggregate.

2.5 The National Bureau of Economic Research NBER-UN world trade database

2.5.1 Description

The international Trade Data, NBER-UN world trade database is an effort by R. Feenstra and R. Lipsey of the NBER to put together and distribute data on trade. This database is available at <http://www.nber.org/data> (International trade data, World trade data). It is a set of bilateral trade data by commodity, for 1962-2000, organized at the 4-digit SITC, Rev. 2.

2.5.2 Origin of the data

This database is an updated version of the Statistics Canada World Trade Database. It originates from the UN trade data. Here the information is provided by the importing country assuming that it is more accurate than reports by the exporter. If the importer's report is missing, the corresponding exporter's report and other information are used instead. The set covers two periods: i/ 1962-1983 for all trading partners, data are classified by the SITC Rev 1; ii/ 1984-2000 for 72 reporting countries (98% of global trade), data are classified by the SITC Rev. 2 and include trade quantities. Various corrections have been made in order to provide harmonized information. They are described in the NBER working paper w1140.

2.5.3 Accessibility and user-friendliness

The database is freely available at <http://www.nber.org/data> (International trade data, World trade data), after agreement not to resell or distribute the data. The data is archived in 38 Zip files (one for each year). The NBER-UN trade database is available under the SAS® and STATA® software. Some expertise in this area is thus required to extract the information.

2.5.4 Reliability and caveats

A great deal of work has been done by Feenstra and Lipsey to insure the consistency and the harmonization of the database (see Feenstra *et al.*, 2005). It may however be regretted that they adopt the rule that the importer's report is more accurate than the exporter's, and that they do not try to reconcile the mirror flows. The database also lacks details being aggregated at the 4-digit level. It obliges the user to do extra work, if he/she wants to work with data on tariffs which are classified under the HS nomenclature.

3. SOURCES FOR TARIFF DATA

One major issue faced by modelers is accessing reliable information on tariffs. It is necessary to distinguish data sources that provide information on bound tariffs from sources that provide information on applied tariffs. Indeed, some countries (mainly developing countries) apply only a fraction of the bound tariff as their MFN tariff. Applied tariffs are therefore more relevant for trade analysis than bound tariffs. In addition, developed countries often apply lower tariffs under preferential agreements, or free trade areas.

Focusing on applied tariffs, the issue is made more complicated by the existence of in-quota tariffs, specific eligibility conditions, preferential tariffs under regional agreements, the Generalized System of Preferences, or bilateral agreements. That is, there is not one single applied tariff for a given tariff line in a given country, but the tariff varies according to the country-of-origin, the preferential regime, and the existence of quantitative restrictions. The issue is complex because, for example, under the same regional agreement, the list of commodities covered by a preferential tariff can vary across countries (as with the EU-Africa-Caribbean-Pacific agreement).

All databases do not address these issues in the same way. Some databases provide mainly information on bound tariffs, but more partial information on applied tariffs (AMAD, TRAINS, etc.). Others focus on applied tariffs (EU Market Access Database). Others combine both (USITC, TARIC, MAcMap and CEPII HTS).

In this section, we focus on the datasets that provide information on tariffs. It should be noted that the datasets which provide both access to tariffs and trade data are treated in the next section (Section 4). In some cases they are among the most useful sources of information on tariffs, so the reader interested in statistics on custom duties is invited to look both at Section 3 and Section 4.

3.1 MAcMAP

3.1.1 Description

MAcMap is a recent dataset which has now become the main source of information on tariffs for modelers and analysts of trade negotiations. The MAcMap dataset, developed by the *Centre d'Etudes Prospectives et d'Informations Internationales* (CEPII) and the UNCTAD is a very large set of bilateral, applied tariffs. MAcMap provides information on tariffs applied by 165 countries at the tariff line level, using a bilateral approach. In other words, because of preferential agreements, tariffs applied by a given country often differ according to the partner. The MAcMapHS6 version has addressed this problem by constructing bilateral tariffs presented as a matrix (reporter/partner).

Basically, three products are associated with this MAcMap denomination:

- The MAcMap “Tariff Lines level” database (available from the International Trade Commission, a joint institution involving the WTO and the UNCTAD based in Geneva). This is mainly the basic source, and because it relies on national classification systems, presents little interest for modelers.
- The MAcMapHS6 database. It is a harmonized version of the previous one, for one year (2001). It has been built and updated by the CEPII, but sold by the ITC (see below). It is the main dataset of interest for modelers, since it allows tariffs to be measured on a bilateral basis, within a standardized classification. MAcMapHS6, includes *ad valorem* equivalents for MFN as well as applied protection at HS-6 level for 163 reporting

countries, 208 partners and 5111 products. It is also fully compatible with the CEPII database on bound tariffs and with BACI (trade flows, see Section 2).

- The “MAcMap for GTAP” distribution. This is an aggregated version of MAcMapHS6 using the GTAP nomenclature. MAcMap for GTAP is a highly aggregated version of the MAcMapHS6, ready to use with the GTAP model and software.

Two versions of MacMap exist, with different conditions for access.

MacMap exists as MacMapHS6v1, for the year 2001. This is the version that has led to the tariff dataset of the GTAP6 model. Access to the full version is conditional (see below).

MacMapHS6v2. The base year is 2004. This dataset is still in a beta version form, but will be freely available through the WITS interface at the HS6 level.

3.1.2 Origin of the data

Here, we focus on the MAcMapHS6v1, which provides information within a standardized classification, that is the same for all countries (the Harmonized system at the 6-digit level).

The basic data on protection used to construct MAcMapHS6 originates from country notifications to the WTO, from AMAD and from national customs information reported to the ITC or to the UNCTAD. For each importing country, information on the various instruments of border protection (*ad valorem* tariffs, specific tariffs, TRQs, etc.) is maintained in the base at the most disaggregated level (if possible the tariff line). A first operation aggregates the tariffs and tariff quotas at the HS6 level, using a simple average as trade data are not systematically available at the tariff line level. Then BACI's trade data at HS6 is used to compute *ad valorem* equivalents and to aggregate tariffs at an upper level.

Ad valorem equivalents (at higher levels than HS-6) are computed as trade weighted averages, using the median unit value of a reference group the exporter belongs for weighting. Groups are chosen on the basis of a clustering of countries using the real GDP per capita and the trade openness. This leads to the definition of five reference groups: i/ the richest countries, ii/ high openness, middle income countries, iii/ low openness, middle income countries, iv/ high openness, low income countries, v/ low openness, low income countries. The use of reference group unit values (instead of world unit values or national unit values) allows:

- the noise in unit value data to be limited,
- the economic relevance of *ad valorem* equivalents of specific tariffs to be retained,
- the problem of endogeneity to be addressed. Indeed, if the own country's trade is used as weight for its tariffs, this leads to the well-known problems of inverse correlation between tariffs and trade flows, and to the extreme case that prohibitive tariffs lead to zero imports, and therefore to a zero weight in the aggregate.

In MacMap HS6v2, the EU data are directly drawn from TARIC and the US tariff schedule, with full, in-house treatment conducted by the CEPII.

3.1.3 Accessibility and user-friendliness

Access is restricted to highly aggregated tariffs (at the HS-2 level). The database at the 8-digit level is not accessible outside the UNCTAD, even if country-by-country extractions can be purchased.

Since ITC provides the major part of the raw data, it retains some property rights of the MAcMapHS6 database. In order to access the HS6 version of MAcMapHS6v1, a subscription to MAcMap is needed at the ITC. The Academic community benefits from a special price (USD900), as do institutions from developing countries (USD450). Moreover, Least

Developed Countries currently have free access to MacMap through the USAID program. MacMapHS6 dataset is available on a CD-Rom under the SAS[®] software. Some expertise in this area is required in order to extract the information.

The “MacMap for GTAP” distribution is freely available at:

http://www.cepii.fr/anglaisgraph/bdd/macmap/form_macpmap/access.asp.

Note, however, that this highly-aggregated dataset does not make it possible to simulate the detailed impact of, say, an international agreement. In order to carry out this kind of simulation properly, it is necessary to shock the tariffs at the 6-digit level and re-aggregate them under the GTAP system of classification. Modelers who use the MacMap for GTAP should keep this important point in mind.

Regarding MacMapHSv2, for the base year 2004 (at the time of drafting this section, in July 2006), it was still a beta version that was being subjected to tests by the CEPII.

3.1.4 Reliability and caveats

MacMapHS6 data has gone through a very large number of consistency checks. It clearly provides the best data available to measure tariffs for a large number of countries. The HS6 version of MacMap is therefore particularly adapted to policy simulations and modeling. A limitation is obviously that it only allows changing tariffs at the HS6 level, while tariffs are often applied at the 8-digit level, or more. However, there is no other way to construct a consistent dataset common to several countries, since the 6-digit level is the most detailed one at the international level.

The complexity of the methodology to ensure consistency, as well as the reference group method to compute aggregates and *ad valorem* equivalents are a guarantee of reliability. However, a full understanding of how the final figures are generated requires a significant investment in methodology.

3.1.5. A complement: The CEPII database on Bound tariffs

The CEPII database on bound tariffs was constructed under the TRADEAG project (EU Commission, DG Research contract 513666), so as to match the applied tariffs in MacMap with bound tariffs and to measure binding overhang. It basically contains similar information as AMAD, regarding the bound tariffs of WTO member countries, but the aggregation is consistent with the one of MacMap, and *ad valorem* equivalents are also based on the same methodology. That is, the database is a useful complement of the MacMapHS6 dataset. The database is described in Boüet, Laborde and Jean (2005). So far, it is internal to the TRADEAG project, but access can be granted on a case-by-case basis to would-be users, upon request.

3.2 The AMAD database

3.2.1 Description

The initial source of agricultural and food product bound tariffs are the schedules notified by the WTO member countries in the wake of the Uruguay Round. Official schedules were published in paper version, but the Organisation for Economic Co-operation and Development (OECD) released a non-official electronic version in 1995. Due to the errors contained in the electronic version, a group of institutions decided to put together a more user-friendly source for bound tariffs (and tariff rate quotas). The Agricultural Market Access Database, AMAD, has been a joint effort by Agriculture and AgriFood Canada, the EU

Commission (the Agriculture Directorate-General), the Food and Agriculture Organisation of the United Nations, the Organisation for Economic Co-operation and Development, the World Bank, the United Nations Conference on Trade and Development, the United States Department of Agriculture (the Economic Research Service).

AMAD is a large-scale database constructed to collect information on WTO market access issues in agriculture which includes: bound tariffs, quota volumes, scheduled in-quota tariff rates, scheduled over-quota tariff rates, scheduled MFN tariff rates, indication of special safeguard (SSG) application, notified imports under TRQ, TRQ country allocations, implementation periods, import volumes and values, applied MFN tariffs, Supply/Utilisation data, world reference prices, world import unit values, exchange rates, primary product equivalent factors. Data in AMAD can be classified as:

- Market Access Commitments (WTO Tariff Schedules or bound rate, WTO TRQ Schedules). The bound tariffs are available from HS-6 to HS-8. The AMAD also displays non-*ad valorem* tariffs (e.g. specific tariffs) if a country uses them. If a country applies SSG measures (additional duties), it is also marked in the bindings table.
- Market Access Implementation (TRQ Notifications, TRQ Allocation, applied Tariffs). The level of disaggregation of applied tariffs can differ widely in the database and probably depends on the level used by the reporting country. For example, US applied tariffs are specified at HS-8; while the same information for Poland and Latvia is provided at the HS-9, and HS-10 for Hungary.
- Analytical Tables (Comparison of in- and over- TRQ, comparison of TRQ Schedule and Notification).
- Additional data (FAO Supply and Consumption Data, imports from the UN Trade Data System, world import unit values) The AMAD includes data on imports but does not provide any data on exports. The import data in quantities and values are available from 1995 to 2002. According to the User's guide, the AMAD Release 2.0 covers 53 WTO member countries that have submitted their tariff schedules.

3.2.2 Origin of the data

AMAD is a spontaneous effort by several agencies willing to pool their resources so as to understand better the issue of trade negotiations, at a time when there was little public data. The database uses data which are not necessarily official: WTO tariff schedules, notifications, modifications and rectifications, UNCTAD TRAINS, COMTRADE, FAOSTAT, Eurostat and various national trade data where available (Wainio and Gibson, 2003). The main AMAD data sources are the following:

- For tariffs and TRQ bindings: WTO bindings and WTO member notifications;
- For applied tariffs: UNCTAD Secretariat;
- For trade: UN Trade Data System/National Sources;
- For production and consumption: FAO Database;
- For prices: world import unit values (UN Trade Data System).

3.2.3 Accessibility and user-friendliness

The database is freely accessible at <http://www.amad.org/>. It is not an on-line database. To process the data, a user needs to use Microsoft (MS) Access, even though some tables can be read directly using MS Excel. Some information is stored in Excel (.xls) files. Files have to be

downloaded from the Internet. Data are stored and can be downloaded: i/ for all countries together or ii/ for individual countries separately.

Data in the "Import All" table (located in the zipped file that contains data on all countries) are neither listed in a chronological order, nor according to the HS code. When a country is selected, the commodities are not ordered according to the HS code. An e-mail contact is available to advise on possible problems.

3.2.4 Reliability and caveats

Data selection is convenient, provided that one masters the MS Access. However, import data do not always match the tariff data, since the former often comes from international sources not always available at the 6-digit HS level. The information whether imports are expressed in CIF or FOB is unclear.

The AMAD database is very well documented and provides a useful guide for new users. The information in the guide, however, is sometimes inconsistent with the actual content of the database itself. Regarding reliability of data, users of AMAD have pointed out mistakes in the bound tariffs and in the tariff rate quotas, or at least inconsistencies with the official schedules.

Our tests included a comparison of import data with national Slovak sources, and the comparison of AMAD data and the Harmonized tariff schedules do not show a large number of errors. Certain differences in the import values in USD can be ascribed to the different exchange rates used (See Annex 4 Tables A7 and A8).

In brief, AMAD is a very useful source as far as bound tariffs are concerned. The recent version also provides a lot of useful information on tariff rate quotas. However, the main limitation is perhaps that the AMAD dataset provides little information on the tariffs applied, under preferential agreements. Indeed, it seems that applied tariff data is incomplete, and ignores some important agreements. In addition, it is not available for all countries. In this respect, MAcMap is certainly a more useful source for the economist dealing, for example, with preferential trade agreements. In addition, reliability of the information on bound tariffs is questionable, and there might be discrepancies with the official version of these tariffs in the consolidated schedules.

3.3. The TARIC database

3.3.1. Description

TARIC is the most comprehensive database if one focuses on tariffs of the European Union, which are sometimes very complex. TARIC provides all information, including seasonal tariffs, exclusions, additional duties for agricultural components of processed products, etc.

The European Commission has developed TARIC in conjunction with Member States' Customs and Taxation services. This database is freely accessible on-line, but there is no serious possibility for extractions beyond a very small number of data. TARIC provides information at the 10-digit level of the combined nomenclature (NC10), on the duties levied by the EU at its borders. For each good at the NC10 and partner of the EU, it gives the MFN duty rate and all other duty rates under the various preferential regimes (WTO, ACP, GSP etc.) applying to this country/good as well as the regulation associated.

TARIC is accessible at:

http://europa.eu.int/comm/taxation_customs/dds/cgi-bin/tarchap?Lang=EN (English version).

The origin of the data is national customs, since the database is a part of the information systems of Taxation and Customs Union of the EU.

3.3.2. Accessibility and user-friendliness

On-line consultation of TARIC is rather user-friendly. First, information must be provided on the good. It can be made either by the 10-digit code (a browse button allows the user to choose the right code), or by textual search (filling the box with a character string presenting a good's description). Then, information about the country (partner) or the country group is needed. The result of the query is information on the level of duty. Data are available from 1970 to date, and are regularly updated. This is very useful as it is then possible to draw a historical evolution of the protection.

3.3.3. Reliability and caveats

The information is reliable as it directly emanates from the EC Member States' Customs and Taxation Services but it provides just basic information which must be processed to be useful (e.g. specific tariffs must be transformed in AVE by the user).

The main caveat of TARIC is not due to the database itself, but to the complexity of the EU tariff structure. Tariffs often combine *ad valorem* and specific components include some minimum or maximum tariffs, seasonal tariffs, and a variety of preferential regimes, so that the whole sum of data is very difficult to use. From this point of view, TARIC supplies only the official tariff, but no instrument to compute percentage equivalents or to aggregate tariffs.

3.4. The DG Trade Market Access Database

The EU commission's DG-Trade has had an applied tariff database developed by a private organization. Individual tariff lines can be consulted online, on the web site of the Trade directorate. However the use of the information for other users than would be exporters is prohibited.² In addition, access is limited to browsing a few numbers of tariff lines, individually.

3.4.1 Description

The DG trade market access database has been primarily designed for the EU Trade Directorate. The Market Access Database is an important operational tool of the European Union's Market Access Strategy, supporting a continuous three-way exchange of information between: the EU institutions, Member States and European business.

The database provides:

- information about market access conditions in non-EU countries,
- a way for the European Commission to follow up complaints from businesses about barriers to trade in non-EU countries,

² It is stated that the information in the Market Access Applied Tariffs Database may not be used for any purpose other than reference use by a European Union user personally, located within one of the (CHECK 27) 15 Member States of the European Union, in support of the user's own internal business processes in the European Union. All other use is prohibited unless explicitly approved in writing by the company that developed the data, Eurotariff because such information remains the property of Eurotariff.

- a means of ensuring that the EU trading partners are abiding by their international commitments,
- a better input for defining the EU's trade policy objectives on further trade liberalization in the framework of the WTO, and new free trade agreements between the EU and preferential partners.

The Market Access Database (MADB) allows its users to get information about the accession conditions to the markets of non-member EU countries. MADB contains information on a) applied tariffs; b) trade flows between the EU and non-EU countries; c) sectoral and trade barriers affecting export and investment conditions in non-EU countries; d) an exporters' guide for import formalities; e) sanitary and phytosanitary measures; f) studies and reports concerning market access.

The MADB is an on-line database. Access to the Sectoral, Trade Barriers, Statistics and Studies section is free. However, information in the Exporters' Guide and Applied Tariffs sections is restricted to users located in the EU Member States. The candidate countries Bulgaria, Romania, Turkey and Croatia are also permitted to use it free of charge.

3.4.2. Origin of the primary data

Trade flow data come from the COMEXT database. The owners of the tariff data are Eurotariff, London, UK and Mendel – Verlag, Witten, Germany. The database provides only the most up-to-date information on applied tariffs. The date of the last update is indicated.

The "Applied tariffs database" section covers 91 trading partners of the EU. Countries in the "Statistical database" section are divided into Reporters and Partners. The former numbers 25 EU Member States and EU 25 as a whole. In the latter group there are 209 partner countries of the EU and "All partners" aggregate.

3.4.3. Accessibility

The database is accessible from <http://mkaccdb.eu.int/mkaccdb2/indexPubli.htm>. No problems have been encountered regarding the Internet access. The interactive form of the database is satisfactory. The main page is set up in a user-friendly manner. Documentation is nevertheless minimal ("Frequently asked questions").

However, the main problem is that this data is not available for academic (let alone consulting) purposes. MADB is mainly a database which is designed to provide information to would-be exporters regarding the tariffs their shipments would face, in a particular country. This has two implications.

First, the database is only usable to check tariffs on particular products. The database does not allow the selected data in spreadsheets to be saved or viewed in HTML format. It is not possible either to get aggregations by values or quantities, exports or imports (in kg or EUR) of selected groups of commodities. This is obviously a major limitation for professional use. The data are only provided as small pdf sheets, which make them basically useless except for checking a particular tariff line.

The second point is the intellectual property limitation. Access to the tariff information states clearly that the use for any kind of activity involving building databases or studies such as consulting is forbidden.

3.4.4. Reliability and caveats

The trade flow data (exports and imports) of a selected commodity can be found in the "Statistical database" part. Only limited information is provided and there is no possibility of

aggregation. Mirror flows are not available. The MADB trade flow data are sourced from COMEXT. In spite of this, data are not identical and differences in both values and quantities between them could be found. Non-available and missing data are not unequivocally marked.

Regarding applied tariffs, which are the main originality of the MADB, there is a clear lack of elements for comparison in many areas involving applied tariffs. MADB is the only international database that exists (with the exception of MAcMap), whose information has been aggregated at the 6-digit level. In MADB, data are recorded according to their national classification. This can be a level comprised between the 4-digit and 12-digit one (e.g.: the case of tariff lines in Turkey).

Preferential tariffs and tariffs under TRQs are not included. Eligibility to preferential regimes such as the GSP is mentioned. TRQs are, however, mentioned in the "Sectoral and trade barrier" section. The database gives information only about the existence of TRQs in Canada, Mexico and Norway.³ In the case of Canada and Norway, no information about in- and over-quota tariffs is provided. There is no information on fill rates nor on the management method.

In brief, the MADB is a unique source for applied tariffs. However, it is clearly not designed to be used for research purposes. Access rights and the organization of the database are such that its use is limited to a quick search for EU exporters to the third countries.

3.5 The IDB database

3.5.1 Description

The WTO secretariat has put together a dataset that includes the bound tariffs for agricultural goods, the IDB database. As access has remained restricted, this source has not benefited from an extensive checking of inconsistencies by researchers, as AMAD has. Overall, the IDB dataset indeed seems to be less reliable than AMAD (the bound tariffs, normally fixed since 1994, were changed several times during the year 2003, in the IDB dataset). Nevertheless, within the IDB, the WTO has managed to establish a correspondence between these bound tariffs and trade data.

3.5.2 Origin of the data

The IDB combines data from official WTO schedules, which is the most reliable source of information on bound tariffs, with national sources (applied tariffs, including those in preferential agreements in some countries).

3.5.3 Accessibility and user friendliness

Access to the IDB is limited to official use of WTO member countries' delegations to the WTO. For those who have access, the IDB can be interrogated through the WITS shell developed by the World Bank, which is more user-friendly than the IDB interface itself.

3.5.4 Reliability and caveats

Some of the applied tariffs are not provided, since the information is not always compiled by national countries (e.g.: the EU). In some cases, we found inconsistencies with some national sources (US preferential tariffs for some products under the GSP seem to have been confused with MFN tariffs, at least for year 2002, and they show a significant gap with the USITC data, which is the original source). With this exception, the data is usually highly reliable for major countries.

³ Accessed on November 03, 2005.

3.6 The DBTAR database

3.6.1 Source and structure

DBTAR is a tariff database developed under the TRADEAG project by J. Gallezot (INRA-CEPII). It focuses on the tariffs applied by the EU, at the 10-digit level. Particular attention is paid to the various tariffs under preferential regimes, and to the calculation of representative tariffs for some complex cases, such as the tariff lines for which there is an entry price (fruit and vegetables). In such a case, the tariff actually depends on the price of the products. Attention is also paid to the conversion of specific tariffs into *ad valorem* equivalents, even though the conventions adopted do not correspond to the conventions recently agreed upon under the WTO. The DBTAR provides data from 2001 to 2004.

3.6.2 Accessibility and user friendliness

DBTAR is accessible under the STATA format. This requires knowledge of the software, but is then a powerful way to do extractions as well as to perform statistical analyses and econometric estimates.

Access is so far restricted to the TRADEAG project and to sister projects funded by the DG Research of the EU Commission, under the 6th Framework program.

3.7 The APEC tariff database

This is the Asia Pacific Economic Cooperation (APEC) tariff database. It gathers data on tariff rates of the APEC member economies.⁴ The APEC Tariff Database is updated twice each year. Changes to tariff rates may occur between these update periods. These updates are dependent on the submission of data by the participating economies.

The database is accessible on line (<http://www.apectariff.org/>). For almost each member country, five viewing options are available that allow the search for tariff-related information. Tariffs may be searched by chapter, tariff heading, Harmonized System Code, alphabetical order or description. Important, tariff-related information such as Contacts, Prohibitions, Customs Guides, etc. are provided by member economies.

Tariffs are given in HS-8, they are *ad-valorem* or specific duties. The member economies offer between 91 and 98 chapters and between 932 and 1349 tariff headings. Singapore, which imports a lot of goods duty free, offers only 6 chapters and 16 tariff headings. The most recent update was carried out by Australia in January 2005. Peru has not updated its information since December 1999.

Three countries do not provide their tariff rates. The Hong Kong Special Administrative Region of the People's Republic of China "HKSAR" is a free port and does not levy any customs tariffs on imports. There is also no tariff quota or surcharge. There are no value added taxes nor general service taxes. Excise duties are levied on only four types of goods, irrespective of whether they are imported or manufactured locally. These goods are liquors, tobacco, hydrocarbon oil and methyl alcohol. Furthermore, under the GATT/WTO, the HKSAR has committed itself to bind tariffs at zero, for certain categories of imports which covered about 60% of its imports, in 2003 values. Inquiries on tariff bindings should be made to the Regional Cooperation Division of the Trade and Industry Department.

⁴ Australia, Chile, Hong Kong (China), Korea, New Zealand, Philippines, Thailand, Brunei Darussalam, China, Indonesia, Malaysia, Papua New Guinea, Russia, USA, Canada, Chinese Taipei, Japan, Mexico, Peru, Singapore, Vietnam.

Russia and Vietnam do not provide any tariffs and tariff-related information on the APEC tariff database website.

4. DATASETS COMBINING TARIFFS AND TRADE

4.1 The USITC trade database

4.1.1 Description

The United States International Trade Commission sets very high standards regarding data quality, access and dissemination. The trade dataset of the United States is an example for other countries: it is highly reliable, user-friendly, available from the Internet, and totally free.

International trade data are available from 1989 onwards, on a monthly, quarterly, annual, or year-to-date basis. They can be retrieved in a number of classification systems, including the Harmonized Tariff Schedule (HTS), which is the US version of the HS (again, the classification beyond the 6-digit level is specific to the US). Other classifications are also accessible, in particular the SITC, or the North American Industry Classification System (NAICS). Pre-defined reports on international trade statistics are also available by geographic region and partner country. In addition, the trade data can be matched to data on US tariffs, also available in the same USITC website. Trade flows are available at any degree of aggregation. Trade data under each separate trade regime, including all preferential agreements, is also available. This includes, in particular, the preferential agreements such as the Africa Growth Opportunity Act, the Caribbean Basin Acts, the Generalized System of Preferences, etc.

Several types of information are available, namely US imports for consumption; US domestic exports; US general imports; US total exports; the trade balance. Some useful tools make it possible to define the level of aggregation that is most convenient for the user. Users can extract any trade statistics since 1996. Previous years are available in an archives section, and some particular data can be requested from the USITC by email.

4.1.2 Origin of the data

The USITC maintains the US tariff schedule, but it is unclear which agency actually collects the primary data from the customs.

4.1.3 Accessibility

The USITC Interactive Tariff and Trade DataWeb is an extremely convenient way to access US trade statistics. DataWeb provides international trade statistics and US tariff data to the public, full-time and free of charge, with a simple online connection and an Internet browser. US import statistics, US export statistics, US tariffs, US future tariffs and US tariff preference information are available on a self-service, interactive basis. A user-friendly interface makes it possible to send queries, including complex ones. It allows both expert and non-expert users to create and save customized country and product lists for future re-use, from anywhere in the world.

4.1.4 Reliability and caveats

Obviously, the USITC data provides only trade from and to the US. It should be noted, however, that other US agencies provide worldwide trade data. This is the case of the US Census Bureau (although only under the SITC classification). USITC trade data seems particularly reliable. Some discrepancies can be observed with other sources, but USITC is generally considered as being the reference (for example, we found inconsistencies with the IDB regarding imports under the GSP, but it seems that the problems came from the IDB, not the USITC source).

It may also be noted that apart from the USITC, the US Census Bureau provide access to international trade statistics, on a free, online basis. However, only the SITC classification is available <http://censtats.census.gov/sitc/sitc.shtml>.

4.2. The TRAINS database

4.2.1 Description

TRAINS (Trade Analysis Information System) is a dataset maintained by the UNCTAD. It is based on the HS, and includes data on tariffs, para-tariffs and non-tariff measures as well as import flows by origin, for more than 140 countries. It provides information at the HS6 level. Tariff measures include bound, MFN and some applied tariffs. Para-tariff measures include customs surcharges and additional taxes and charges. Non-tariff measures include price control measures, finance measures, automatic licensing measures, quantity control measures, monopolistic measures and technical measures.

4.2.2 Origin of the data

The tariff information is compiled by the UNCTAD, based on national, inter-governmental and other sources, and is available at the tariff line level.

The trade information is mostly sourced from the United Nations Statistical Division COMTRADE database, and is available at the HS 6-digit level only. It only provides information on imports.

4.2.3 Accessibility and user-friendliness

The database is accessible on-line via the WITS interface developed by the World Bank. A limited version of the database is available at <http://r0.unctad.org/trains/>. Subscription to TRAINS is required in order to obtain active the access through WITS.

TRAINS on Internet provides an on-line access to indicators of Trade Control Measures (Tariff and Non-tariff measures), as well as imports by supplier, at each Harmonized System, 6-digit level for over 130 countries. TRAINS on the Internet does not allow retrieval of the information, unless using fastidious 'copy/paste' techniques. TRAINS on CD-Rom gives more detailed and recent information. TRAINS operated by WITS is far more user-friendly.

4.2.4 Reliability and caveats

The UNCTAD TRAINS database includes applied tariffs, but the data is only available for selected countries (the data for most developing countries actually contains bound tariffs). The data contains only a section on preferential rates, i.e. the ones applied under the GSP, the preferences granted to least developed countries, and on some in-quota tariffs.

4.3 The TRADEPREF database

4.3.1 Description

TRADEPREF is a database put together under the TRADEAG project by J. Gallezot (INRA-CEPII). It focuses on the trade flows under preferential tariff regimes granted by the EU, and information that is missing in COMTRADE and in COMEXT. The database associates the level of imports (in value), the preferential regime and an *ad valorem* equivalent of the duty levied on imports under this peculiar regime. All the data are at the NC8 level (8-digit of the TARIC combined nomenclature).

The 2002 TRADEPREF database allows detailed processing of the utilization of the European preferences, particularly where partners benefit from multiple preferential schemes. The output is a table structured according to the variables summarized in the table below.

| TRADEPREF variable | Description |
|-----------------------------------|--|
| <u>Primary key</u> | |
| Cgeo | Tariff Regime codes |
| 1011 | ERGA OMNES |
| 1031 | ACP African, Caribbean, Pacific |
| 1099 | AL, BA, CS |
| 2005 | GSP (R96/1256 - ANNEX IV) |
| 2020 | GSP (R 01/2501) - General regime |
| 2027 | GSP (R 01/2501) - Annex 1 Column 1 |
| 2080 | OCT Overseas Associates |
| nc8 | Product code (8 digits of the combined nomenclature) |
| num_part | ISO code for country |
| pref_util | Trade regime code |
| <u>Secondary variables</u> | |
| Acp | ACP countries and products eligible |
| Balk | Balkan countries and products eligible |
| Spga | EBA countries and products eligible |
| Spge | GSP countries and products eligible |
| Ptom | OCT countries and products eligible |
| EEA | EEA countries and products eligible |
| MCH | Mashrek countries and products eligible |
| MGB | Maghreb countries and products eligible |
| PHC | Central and Eastern Europe countries and products eligible |
| Spgl | General GSP countries and products eligible |
| Other | Other regimes countries and products eligible |
| PMA | LDC countries (this variable takes a binary value of 0 or 1) |
| Mtot | Total imports by country for all regimes in 1000 Euro |
| Mpref | Total imports by country under preferential regimes in EUR '000s |
| d103 | MFN duty in ad-valorem equivalent (AVE) |
| dutil | Duty used in AVE |
| d142 | Preferential duty in AVE |
| d1243 | Preferential quota duty in AVE |
| d122 | WTO quota duty in AVE |
| exclu | Countries and products exclusion |

4.3.2 Origin of the data

The realization of TRADEPREF relies on the exploitation of the Single Administrative Documents (SAD) which importing companies must fill in. The information is currently of a tax nature (duty collection) and depends on each EU member state (the subsidiarity principle).

The SAD data includes the importer's choice of a preferential regime (and duty), but it represents the required, and not the levied duty. This information must then be verified and assigned to the appropriate regime. To carry out this verification, the TARIC database is used (see above).

4.3.3 Accessibility and user-friendliness

TRADEPREF is accessible under the STATA format. Use requires knowledge of this software, but it is then a powerful way to extract data, as well as to perform statistical analyses and econometric estimates.

Access is so far restricted to the TRADEAG project and to sister projects funded by the DG Research of the EU Commission, under the 6th Framework program.

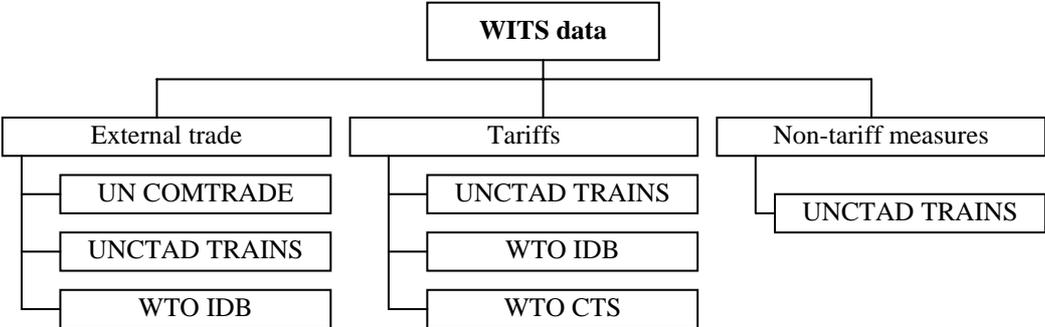
4.3.4 Reliability and caveats

TRADEPREF is so far the only source that details the EU trade under preferential agreements. Part of the trade flows result from estimates (the SAD are not always reliable and need to be made consistent with the overall trade flows in COMEXT). But given the lack of basic information, it is the main source for someone willing to perform detailed sectoral analyses of trade, under various tariffs regimes in the EU. TRADEPREF is available for 2002.

4.4. The WITS interface

WITS is an interface developed by the World Bank that is now used by a large number of agencies. It provides a common access to the four following data bases: COMTRADE and TRAINS of the UNCTAD and IDB/CTS (Integrated Dated Base/Consolidated Tariff Schedule) of the WTO. In the future, WITS should also include MacMap. WITS is particularly user-friendly and even though it contains no original data, it is often easier to access these datasets through WITS than through their specific software. In addition, WITS also proposes some analytical tools to compute tariff cuts under different scenarios: Doha formulas, advanced tariff change simulations and a smart simulation by country, which also gives some results about trade creation effects, welfare effects and revenue impacts.

Figure: WITS architecture



WITS users are required to subscribe to TRAINS and COMTRADE. They must also have access to IDB, which is normally restricted to the delegation of each country, to the WTO and a few international organizations. The use of the WITS interface is free, but the access to each of the three bases is given according to the user’s subscriptions. There is no centralized procedure to access all these data, and the dissemination policy of the various agencies involved is quite puzzling (some research institutions seem to have free access, others are asked very large fees, without any apparent rationale).

REFERENCES

- Ackerman F. (2005). The Shrinking Gains from Trade: A Critical Assessment of Doha Round Projections. Global Development and Environment Institute, Tufts University, Boston, October.
- AMAD. Agricultural Market Access Database, URL: <http://www.amad.org/>
- Bchir M., Jean S. and D. Laborde (2006). Binding overhang and tariff-cutting formulas. Review of world economics. Vol. 142. N° 2. pp.207-232.
- Boonekamp, L. (2004): Agricultural Trade Policy Modeling Workshop. Workshop Summary. Economic Research Service, USDA, November 4-5, 2004.
- Bouët A., Decreux Y., Fontagné L., Jean S. and Laborde D. (2004). A Consistent Ad-Valorem Equivalent Measure of applied protection Across the World: The MacMap-HS6 database. CEPII Working Papers. N°2004-22. December 2004.
- Bureau J.C. and Salvatici, L. (2003). WTO Negotiations on Market Access: What We Know, What We Don't and What We Should. Invited paper presented at the International Conference Agricultural policy reform and the WTO: where are we heading? Capri (Italy), June 23-26, 2003.
- COMEXT database: URL: <http://fd.comext.euostat.cec.eu.int/xtweb/mainxtnet.do>.
- COMTRADE, United Nations Commodity Statistics Database. Downloaded October 28, 2005. ULR: <http://unstats.un.org/unsd/comtrade/>
- Feenstra R.C., Lipsey R.E., Deng H., Ma A.C. and Mo H. (2005). World Trade Flows: 1962-2000. National Bureau of Economic Research Working Paper 11040. January 2005.
- FAOSTAT. <http://faostat.fao.org/faostat/collections?version=ext&hasbulk=0&subset=agriculture>
- Gallezot J. (2005). Data Base of EU preferential trade TRADEPREF. TRADEAG WP. December 2005.
- Gallezot J. (2005). Data Base on Europe, agricultural tariffs DBTAR. TRADEAG WP. July 2005.
- Gaulier G. and Zignago S. (2004). Notes on BACI (Analytical Database of International Trade) 1995 – 2002 Version. CEPII Working Paper. March 2004.
- Gaulier G. and Zignago S. (2005). BACI: An Analytical Database on International Trade. CEPII Working Paper. October 2005.
- Gehlhar M. J. (1996). Reconciling Bilateral Trade Data for Use in GTAP. GTAP Technical Paper No. 10. October 1996.
- Golderman G. and Connolly B. (2005). Numbers Beyond Borders. Library Journal. com. Published April 15, 2005. <http://www.libraryjournal.com/article/CA515802.html>.
- Market Access Database: URL: <http://mkacddb.eu.int/mkacddb2/indexPubli.htm>
- National Bank of the Slovak Republic, URL: <http://www.nbs.sk>
- Peters R. (2004): Agricultural Trade Policy Modeling Workshop. Session 6: Levering Technology. UNCTAD, Washington, November 5, 2004.
- Peters, R. (2005): AMAD Structure and Uses of the Agricultural Market Access Database. A presentation presented at ARTNeT Capacity Building Workshop on Trade Research UN ESCAP. Bangkok, March 2005.
- Stehlo P., Döme G. and Galik, J. (2001). Agrarian External Trade of the Slovak Republic Yearbook 2000. Bratislava: RIAFE, 73 p.
- Stehlo P., Döme G. and Galik, J. (2003). Agrarian External Trade of the Slovak Republic Yearbook 2002. Bratislava: RIAFE, 67p.
- Stehlo P., Döme, G. and Galik, J. (2004). Agrarian External Trade of the Slovak Republic Yearbook 2003. Bratislava: RIAFE, 74p.

A New Deal for Exporters (1999)>. Business Europe; 4/21/99, Vol. 39 Issue 8, p. 6. ISSN 1351-8755.
Statistical Office of the Slovak Republic: URL: <<http://www.statistics.sk/>>. Accessed [2005-11-03].
Stojanovic M. (2005). Harmonisation of International Trade Data Published by the OECD and Eurostat. The Statistics News Letter, Issue 29, OECD.
<http://www.oecd.org/dataoecd/3/3/35723250.pdf>
Wainio J. T., Gibson P. R. (2003). Agricultural Trade Policy Data. USDA, Economic Research Service. September 17, 2003.

Annex 1. Elements of comparison of COMTRADE data and national sources for Slovakia

Table A1. Slovak export and import of Potatoes, Rapeseed and Beer according to COMTRADE and SSO

| Commodity | Year | COMTRADE | | SSO | | COMTRADE/SSO | COMTRADE/SSO |
|------------------------|------|-------------------------|-------------------------|-------------------------|-------------------------|--------------|----------------|
| | | Trade value (USD '000s) | Trade quantity (tonnes) | Trade value (USD '000s) | Trade quantity (tonnes) | Trade value | Trade quantity |
| Potatoes – 0701 | | | | | | | |
| Export | 2004 | 733.232 | 3580.567 | 733 | 3480 | 1.00 | 1.03 |
| Rapeseed – 1205 | | | | | | | |
| Export | 2002 | 17268.028 | 79979.68 | 17 156 | 79 980 | 1.01 | 1.00 |
| | 2003 | 1606.242 | 5344.105 | 1 597 | 5 344 | 1.01 | 1.00 |
| Import | 2002 | 2386.605 | 8473.254 | 2 368 | 8 473 | 1.01 | 1.00 |
| | 2003 | 3296.12 | 11166.101 | 3 278 | 11 166 | 1.01 | 1.00 |
| Beer – 2203 | | | | | | | |
| Export | 2002 | 6594.464 | 19512.088 | 6 552 | 19 512 | 1.01 | 1.00 |
| | 2003 | 8874.361 | 23425.88 | 8 825 | 23 426 | 1.01 | 1.00 |
| | 2004 | 10571.404 | n.a. | 10 563 | 7 036 | 1.00 | - |
| Import | 2002 | 8787.478 | 36280.816 | 8 719 | 36 281 | 1.01 | 1.00 |
| | 2003 | 11272.437 | 41353.865 | 11 211 | 41 354 | 1.01 | 1.00 |
| | 2004 | 15372.691 | n.a. | 15 313 | 30 503 | 1.00 | - |

Source: COMTRADE, RIAFE

Table A2. Bilateral trade: the case of Slovak Milk (0401, HS2002), exports in 2003 (USD)

| | Imports reported by the partner | Exports reported by Slovakia | Imports / Exports | Slovakia's share of partner's total imports | Total imports reported by importer |
|--------------------|---------------------------------|------------------------------|-------------------|---|------------------------------------|
| Czech Rep. | 21998284 | 15964548 | 1.38 | 84.4 | 26073854 |
| Hungary | 818000 | 940902 | 0.87 | 24.2 | 3374000 |
| Croatia | 365055 | 398266 | 0.92 | 4.5 | 8049044 |
| Romania | 282826 | 270946 | 1.04 | 11.8 | 2395991 |
| Russian Federation | 279406 | 247256.00 | 1.13 | 3.2 | 8652057 |
| Bosnia Herzegovina | 195647 | 237616.00 | 0.82 | 1.4 | 14247458 |
| Other | 188706 | 126588 | 1.49 | n.a. | n.a. |
| Total | 24127924 | 18186122 | 1.33 | n.a. | n.a. |

Source: COMTRADE

Annex 2. Elements of comparison of COMEXT data and national sources for Slovakia

Table A3. Slovak total trade flows with the Czech Republic by COMEXT and SSO, in EUR millions (2001-2004)

| Year | Import | | COM/SSO | Export | | COM/SSO | Exchange rate (SKK/€) |
|-------------|---------|---------|---------|---------|---------|---------|-----------------------|
| | COMEXT | SSO | | COMEXT | SSO | | |
| 2001 | 2511.13 | 2484.98 | 1.01 | 2323.80 | 2345.39 | 0.99 | 43.309 |
| 2002 | 2672.85 | 2654.21 | 1.01 | 2301.98 | 2319.77 | 0.99 | 42.699 |
| 2003 | 2868.50 | 2851.64 | 1.01 | 2472.77 | 2498.12 | 0.99 | 41.491 |
| 2004 | 4284.42 | 3112.00 | 1.38 | 2887.72 | 2978.29 | 0.97 | 40.045 |

Source: COMEXT, National Bank of the Slovak Republic, The Statistical Office of the Slovak Republic

Table A4. Slovak total trade flows with Germany by COMEXT and SSO, in EUR millions (2001-2004)

| Year | Import | | COM/SSO | Export | | COM/SSO | Exchange rate (SKK/€) |
|-------------|---------|---------|---------|---------|---------|---------|-----------------------|
| | COMEXT | SSO | | COMEXT | SSO | | |
| 2001 | 4085.24 | 4068.13 | 1.00 | 3818.52 | 3821.96 | 1.00 | 43.309 |
| 2002 | 3976.79 | 3962.51 | 1.00 | 3970.06 | 3970.32 | 1.00 | 42.699 |
| 2003 | 5095.78 | 5076.56 | 1.00 | 5966.21 | 5969.48 | 1.00 | 41.491 |
| 2004 | 6257.21 | 5603.62 | 1.12 | 6431.72 | 6412.56 | 1.00 | 40.045 |

Source: COMEXT, The National Bank of the Slovak Republic, The Statistical Office of the Slovak Republic

Annex 3. Elements of comparison of FAO data for Slovakia and national sources

Table A5. Slovak trade data for Potatoes according to the FAO database and the SSO

| Year | Imports | | | | | | Exports | | | | | |
|------|--------------|--------|---------|-------------------|-------|---------|--------------|-------|---------|-------------------|-----|---------|
| | Quantity (t) | | | Value (USD '000s) | | | Quantity (t) | | | Value (USD '000s) | | |
| | SSO | FAO | FAO/SSO | SSO | FAO | FAO/SSO | SSO | FAO | FAO/SSO | SSO | FAO | FAO/SSO |
| 1996 | 47 550 | 47 550 | 1.00 | 8 795 | 8 414 | 0.96 | 750 | 750 | 1.00 | 167 | 159 | 0.95 |
| 1997 | 3 773 | 3 773 | 1.00 | 1 004 | 1 004 | 1.00 | 225 | 225 | 1.00 | 24 | 24 | 1.00 |
| 1998 | 20 358 | 20 358 | 1.00 | 3 404 | 3 404 | 1.00 | 231 | 231 | 1.00 | 36 | 36 | 1.00 |
| 1999 | 41 675 | 41 450 | 0.99 | 4 649 | 4 638 | 1.00 | 3 | 3 | 1.00 | 1 | 1 | 1.00 |
| 2000 | 50 857 | 5 614 | 0.11 | 3 954 | 1 862 | 0.47 | 11 | 3 094 | 281.27 | 1 | 454 | 454.00 |
| 2001 | 5 614 | 11 228 | 2.00 | 1 778 | 1 778 | 1.00 | 6 187 | 1 229 | 0.20 | 433 | 421 | 0.97 |
| 2002 | 30 021 | 30 591 | 1.02 | 4 368 | 4 431 | 1.01 | 553 | 553 | 1.00 | 104 | 104 | 1.00 |
| 2003 | 31 151 | 31 151 | 1.00 | 4 730 | 4 740 | 1.00 | 2 564 | 2 564 | 1.00 | 468 | 469 | 1.00 |

Source: FAO, RIAFE

Table A6. Slovak trade data for Wheat according to the FAO database and the SSO

| Year | Imports | | | | | | Exports | | | | | |
|------|--------------|---------|---------|-------------------|--------|---------|--------------|---------|---------|-------------------|--------|---------|
| | Quantity (t) | | | Value (USD '000s) | | | Quantity (t) | | | Value (USD '000s) | | |
| | SR | FAO | FAO/SSO | SR | FAO | FAO/SSO | SR | FAO | FAO/SSO | SR | FAO | FAO/SSO |
| 1996 | 15 375 | 15 375 | 1.00 | 3 212 | 3 073 | 0.96 | 11 205 | 11 205 | 1.00 | 1 896 | 1 814 | 0.96 |
| 1997 | 168 267 | 20 231 | 0.12 | 25 612 | 2 561 | 0.10 | 27 669 | 27 669 | 1.00 | 3 054 | 3 067 | 1.00 |
| 1998 | 58 198 | 58 198 | 1.00 | 6 625 | 6 625 | 1.00 | 119 056 | 119 056 | 1.00 | 11 298 | 11 318 | 1.00 |
| 1999 | 13 548 | 13 549 | 1.00 | 1 469 | 1 469 | 1.00 | 105 065 | 105 065 | 1.00 | 9 142 | 9 149 | 1.00 |
| 2000 | 20 237 | 24 142 | 1.19 | 2 147 | 3 066 | 1.43 | 24 143 | 20 235 | 0.84 | 2 509 | 2 276 | 0.91 |
| 2001 | 188 964 | 188 964 | 1.00 | 22 896 | 22 893 | 1.00 | 8 927 | 8 307 | 0.93 | 1 239 | 1 150 | 0.93 |
| 2002 | 1 627 | 1 627 | 1.00 | 203 | 203 | 1.00 | 59 577 | 59 112 | 0.99 | 6 276 | 6 201 | 0.99 |
| 2003 | 23 539 | 23 539 | 1.00 | 2 561 | 2 566 | 1.00 | 26 169 | 26 169 | 1.00 | 3 355 | 3 362 | 1.00 |

Source: FAO, RIAFE

Annex 4. Elements of comparison of AMAD data and national sources for Slovakia

Table A7. Slovak import of Apples–HS 080810

| Year | AMAD | | SSO | | AMAD/SSO | |
|------|-------------|---------------|-------------|---------------|-------------|---------------|
| | Imports (t) | Imports (USD) | Imports (t) | Imports (USD) | Imports (t) | Imports (USD) |
| 1995 | 30808.12 | 10222491.00 | 30808 | 10222489 | 1.00 | 1.00 |
| 1996 | 85845.69 | 11780858.00 | 86457 | 11981302 | 0.99 | 0.98 |
| 1998 | 29503.57 | 8681958.30 | 29504 | 8683866 | 1.00 | 1.00 |
| 1999 | 29138.09 | 9911314.15 | 29138 | 10081113 | 1.00 | 0.98 |
| 2000 | 31752.82 | 8053739.20 | 31753 | 8055236 | 1.00 | 1.00 |
| 2002 | 30010.57 | 6824839.34 | 30011 | 6771632 | 1.00 | 1.01 |

Source: AMAD, RIAFE

Table A8. Slovak import of Beer-HS 220300

| Year | AMAD | | SSO | | AMAD/SSO | |
|------|-------------|---------------|-------------|---------------|-------------|---------------|
| | Imports (t) | Imports (USD) | Imports (t) | Imports (USD) | Imports (t) | Imports (USD) |
| 1996 | 64675.56 | 15836602.01 | 64922 | 15905158 | 1.00 | 1.00 |
| 1998 | 45214.06 | 12349017.40 | 45214 | 12351731 | 1.00 | 1.00 |
| 1999 | 35891.70 | 9519996.83 | 35 892 | 9 683 092 | 1.00 | 0.98 |
| 2002 | 36280.85 | 8787485.38 | 36 281 | 8 718 977 | 1.00 | 1.01 |

Source: AMAD, RIAFE