Conjuncture of International Trade (All Countries):

Methodological note

September 2018

Methodological note of ***Conjuncture*** *of International Trade -* ***All Countries***

September 2018

StudiaBo srl
via Santo Stefano, 57
40125 Bologna Italy

Table of Contents

Presentation

This document describes the methodological choices made in the development of the dabase **Conjuncture - All Countries**, concerning the quarterly foreign trade flows reported by a sample of about 70 countries worldwide towards more than 150 partner countries, starting from 2011.

These data, relating to exports and imports, are a highly representative database of foreign trade at an international level: StudiaBo estimates *database Conjuncture - All Countries* to represent about 80% of the overall international trade.

StudiaBo also intends to gradually expand the range of reporting countries considered in the database, as far as their foreign trade data will become available[1].

[1] International institutions are increasingly committed in the dissemination of international trade data. This activity will be continuously monitored to ensure the ability to integrate the database and make it more and more complete.

Data sources

This chapter describes the various sources of the database **Conjuncture - All Countries**.

Eurostat - db Comext

Eurostat Comext(http://bit.ly/2NY1hsk) is the database containing the monthly statements of foreign trade reported by EU and EFTA countries.
For each monthly flow of foreign trade, db Comext reports the measures both in monetary values and in volumes (kilograms and / or other supplementary unit measure). Data are updated monthly at fixed intervals, with a delay of about six weeks for flows reported to partner countries outside Europe and about 10 weeks for flows reported to European countries.
EU member states are using a very detailed level of customs classification, called the Combined Nomenclature, at 8-digit level (CN8). CN8 consists of a further specification with respect to the customs classification 6-digit Harmonized System (HS6). Each HS6 chapter is the sum of one or more sub-NC8s.

United Nations - db Monthly Comtrade

Since 2010 the United Nations Statistics Division has been carrying out an experimental project called Monthly Comtrade, for the construction of an online database, accessible at http://comtrade.un.org/monthly/, on monthly foreign trade flows reported by UN countries.
Data are available in dollars and kilograms, for each product of the Harmonized System customs classification at 6-digits level (HS6).

US Census Bureau - db UsaTrade

US Census Bureau disseminates monthly foreign trade data reported by the United States of America through the portal UsaTrade Online (https://usatrade.census.gov/), according to the Harmonized System customs classification at 10-digit level (HS10). HS10 consists of a further specification with respect to the customs classification 6-digit Harmonized System (HS6). Each HS6 chapter is the sum of one or more sub-HS10s.

Main Data Mining techniques

This chapter describes the main techniques of textit data mining used in the construction of the database **Conjuncture - All Countries**, in order to extract meaningful reliable and up to date information from the original datasets.

Management of Missing Statements

Foreign trade data might suffer, at times, from *missing statements*, especially with regard to developing countries as reporters.
In order to manage this case, as part of the construction of the database, StudiaBo used a technique of *filling missing data*, based on the interpolation of time series, after identification - for each missing reporting country - of its most frequent partner countries.

Management of Missing Measures

Foreign trade data might suffer, at times, from *missing measures* (i.e.: "not available"), especially with regard to quantity.
In order to handle this series of statements, StudiaBo has developed the following technique, which aims to assess the measures "not available" (whether expressed in kilograms or in supplementary unit measures):

- construction, for every foreign trade flow[2], of a characteristic ratio, called Average Unit Value, expressed as the ratio of monetary value and quantity in kilograms;
- construction of a *benchmark* for the characteristic ratio of the flow, given by the geometric mean of two marginal distributions relative to the size of the reporting country (Average Unit Value expressed by the reporting country) and partner country (Average Unit Value inherent in the statements reported towards a given partner country);
- calculation of the time series of the Average Unit Value for single customs code;
- in the cases where the characteristic ratio of the flow is not determined, this is reconstructed taking into account the three components mentioned above (namely: temporal dynamics, average unit value expressed by the reporter, average unit value which is characteristic of the partner).

Note that, secondly, this method also applies to a second characteristic ratio of the flow, which is given by the ratio of the measures of quantity in kilograms and in a supplementary unit measure. It seems appropriate to emphasize that in this second case the distributions at the margin are characterized by greater stability at the level of single fact [export / import by a reporting country to a partner country in a given unit of time], just by virtue of the fact that represent average values, calculated therefore on a greater number of observations.

[2] A flow (or fact) is defined as the combination of the dimension export / import of a reporting country towards a partner country for a given customs code, in a given unit of time.

The stability of the information at the margin is guaranteed also by StudiaBo through specific techniques for the detection and management of *outliers*.

Management of Outliers

Outliers are measurement errors which, if not treated, are affecting the understanding of the economic phenomenon. In the literature on foreign trade flows, such errors are particularly frequent in quantity measures. For this reason StudiaBo, after the construction of the characteristic ratios of each flow, subjects the time series obtained to specific control filters. In particular, through the orderly distribution of these characteristic ratios, StudiaBo sets specific ranges (based on the first and ninth deciles of the distribution) within which the observations are considered *reliable*. The observations excluded from such ranges are consequently recalculated to be in the range of validity, resulting in a *smoothing* distribution of the observations.

Pre-estimates of quarter-end

Foreign trade data can be a very powerful tool for measuring and analyzing in 'real-time' economic trends. One obstacle to that is made by the different timing of the statements made available.
In order to allow a user to monitor even the most recent facts, StudiaBo considered it useful to take a pre-estimate of the current quarter-end. Note that, because of the different characteristics of original databases, the methods applied are different. To describe these methods, it is therefore appropriate to distinguish according to the type of reporting countries.

EU countries

In the case of statements reported by EU countries, taking advantage of the long span of original time-series (from 2000), StudiaBo uses an econometric modelling applied to the time series, called ARMA, which is particularly reliable in measuring short period trends. This econometric modelling is based on the study of the *autocorrelation* (i.e.: the measurement of the relation between an economic variable and its past). It is referred to as *unconditional forecasts*, because they are independent from external assumptions regarding the international scenario and economic policy.
Once available, then official statements are replaced in the databse **Conjuncture - All Countries** instead of those estimated by ARMA models.

Non-EU countries

In the case of statements reported by non-EU countries, given the availability of a shorter span - compared to those by EU countries - of original time-series (from 2010), such statements are less suited to dealing with ARMA modelling. For this reason, the quarter-end statements of non-EU countries are estimated through the use of change rates.
This estimation consists, firstly, in isolating the last two years elapsed. Given the months already covered by official declarations, cumulative change rates are calculated for each flow related to its value and quantity measures. This makes it possible to estimate the information of foreign trade of months still absent in the database and hence derive a first hypothesis of quarter-end. Obviously this pre-estimates are based on the assumption that the months not yet covered by the official declarations have been characterized by the same trend of historic data.

Quantity at constant prices

As mentioned above, the flows of foreign trade concern both monetary values (expressed in the the database **Conjuncture - All Countries** by the measure "V") and physical quantities (expressed in the database by the measures "K" and "U", which indicate, respectively, the weight and the supplementary unit measure of the product).
The analysis of the above mentioned measures might lead to distortions in the interpretation of economic phenomena in the case of non-homogeneous goods. In fact it is known that:

- the monetary value is affected by changes in price over time; therefore, higher values may be potentially traceable only to price increases, for the same amount exchanged;
- physical measurements depend on the product; thus, for baskets of non-homogeneous goods [3], it may have no economic significance to sum the amount of the flows by physical quantities.

In order to make possible the analysis of economic aggregates of non-homogeneous goods, it is appropriate to introduce a measure of *economic quantity*. In order to measure "economic quantities" of non-homogeneous goods, the economic statistics introduced the concept of **constant prices**.
In the database **Conjuncture - All Countries** StudiaBo has introduced the measure *Q: Quantity at constant price*. In essence, the measure Q reflects an operation of *deflation* [4], through which a time-series of monetary values (V) has been transformed into a similar set of values expressed at constant prices, with a reference to a given year (in the case of the Data Warehouse (e.g. 2011).
In other words, the methodology implemented by StudiaBo has deflated all current values, according to the following equation:
$Q_{flow,date} = V_{flow,date}/IP_{date}$
where $Q_{flow,date}$ identifies the value at constant prices in the given economic period for a given combination [Country of origin - Country of destination], $V_{flow,date}$ identifies the monetary value in the given economic period for a given combination [Country of origin - Country of destination], IP_{date} is the price index constructed as follows:
$IP_{date} = P_{date}/P_{2011}$.
The numerator of the index price of P_{date} is the ratio V_{date}/K_{date} which is the average unit value recorded in each unit of time; the denominator is the average unit value calculated in 2011.
By doing so, StudiaBo obtained a series of *economic quantities*, which can reflect changes of the goods considered both in the physical quantities and in their quality.

[3] The inhomogeneity of economic goods may be due to the following factors:
- differences in quality;
- presence of added services;
- availability in different places;
- availability at different times.

[4] In the economic statistics *deflation* assumes great importance in allowing inter-temporal comparisons, in order to analyse the changes in *economic quantities*.