Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

21-Nov-2001
_____
English - Or. English

STATISTICS DIRECTORATE

National Accounts - International Trade Statistics

RECONCILIATION OF INTRA-EURO TRADE STATISTICS

Paper prepared by Henri Tyrman, Eurostat - Luxembourg

3rd OECD MEETING ON INTERNATIONAL TRADE STATISTICS
Château de la Muette, Paris
5-7 December 2001
Beginning at 9:30 a.m. on the first day

Contact: Henri.Tyrman@cec.eu.int

JT00117057

RECONCILIATION OF INTRA-EU TRADE STATISTICS

1. Before the introduction of Intrastat in January 1993, intra-EU import and export flows were comparable at the aggregated level (all products). Since the implementation of the Intrastat system, a significant asymmetry has been observed between the two flows, with exports higher than imports (see attached graph). This bias is relatively stable and represents, each year, from 4 to 6% of the total amount of exports, i.e. approximately 0.8% of the GDP of the European Union.

    If the external balance of the European Union (15 members) is not affected by this intra-EU asymmetry[1], on the contrary the external balance of the eurozone (12 members) is systematically overestimated.

2. The implementation of the Euro led the European Commission and the European Central Bank to create in 1999 a Working Party ("ad hoc" Committee on asymmetries) in order to determine methods making it possible to build a asymmetry-free extra-eurozone Balance of payments.

    Two types of methods have been proposed: "bottom-up" or "top-down" approach.

    -   The bottom-up approach consists of bilaterally analysing the differences between the Member States so that either the asymmetries can be solved at source, or agreement is reached on general rules which are complied with by each Member State (so that asymmetries are avoided);

    -   With the top-down approach, asymmetries are eliminated on the basis of a model, i.e. starting from a number of mathematical rules and specific assumptions.

3. With regard to trade in goods, the "External and Intra-EU trade" Unit of Eurostat has developed actions following these two approaches.

    The "bottom-up" approach made it possible, starting from bilateral analyses between Member States, to identify the main causes of asymmetries (see annex 1). Since Intrastat has kept essentially the same concepts and definitions used for extra-Community customs declarations, it is clear that the absence of asymmetry before 1993 shows the limited role played by the conceptual aspects. The main causes of asymmetries are rather connected with the introduction of thresholds below which the companies are exempted from declaration, and to a non-response phenomenon from some companies above the thresholds.

    The bilateral analyses however did not make it possible to quantify the effect of each cause of asymmetry and to define rules allowing a reconciliation of mirror flows.

4. On the contrary, the second approach ("top-down") makes it possible to reconcile flows, on the basis of statistical models.

---

[1] Extra-EU trade statistics are still based on Customs declarations, which are not affected by specific Intrastat problems.

A first model, based on time series analysis and detection of outliers, led to a partial reconciliation but was not optimal. A new model, which is more systematic and ensures a full reconciliation, has recently been developed. This "multivariate" model leads to reconciled levels that can be considered as optimal from a statistical point of view (see a brief description of the model in annex 2). The model uses techniques that have already been used in the framework of National Accounts (where inconsistencies between different macro-data is a current situation). The model is flexible and can be combined with exogenous hypotheses, such as the introduction of a preference for one flow in specific cases.

5. The results obtained are satisfactory at the level of the monetary Union as a whole. In particular, dependence to the model specification is limited and results over a given period are robust over time (in relation to the addition of new periods).

Sensitivity to the model is more important at the level of analysis by Member State. This is why it is planned to apply a correction only at the level of the Monetary Union as a whole, and not for each Member State.

6. Experimental results of the model, at the eurozone level, are the following:

Trade balance of eurozone (12 Members) with 3 EU Members "Out" (UK, DK, SW)

Million Euro

	Dispatches		Arrivals		Standard Balance	
	Declared	*Adjusted*	*Declared*	*Adjusted*	*Declared*	*Adjusted*
1997	191 403	188 642	166 039	172 324	25 364	16 318
1998	203 949	200 542	174 608	180 244	29 341	20 298
1999	217 202	213 547	186 671	192 436	30 531	21 111
2000	251 891	245 444	219 319	223 646	32 572	21 798

Results show that, without a correction of the Intrastat asymmetry, the overestimation of the eurozone trade balance is around 10 billion Euro a year.

7. Future work and conclusions

The objective is to produce an experimental set of intra-EU reconciled data and reconciled extra-eurozone trade balance in 2002.

To achieve this goal, Eurostat needs to:
- Work on the refinement of the models used
- Define how the model could be applied on current months.
- Develop an application interfaced with Eurostat environment
- Check the accuracy of SITC results

The work done shows that it is possible to produce reconciled intra-EU data using statistical modelling.

The main advantages of this methodology are the following:
- the estimation of the extra-eurozone trade balance is improved (correction of the bias due to Intrastat)
- the method can be implemented easily, with limited resource.
- the method can be combined with information coming from "bottom-up" approach (for instance in cases where "dispatches are more reliable then "arrivals")

The main drawbacks are the following:

- The method can be applied only at aggregated level of products.
- Results at Member State level are more sensitive to model specification

ANNEX I

What are the main causes of asymmetries in mirror statistics?

This paragraph contains a list of arguments often put forward to explain discrepancies in intra-EU mirror statistics.

1. **Non-response**. Non-response percentages differ widely from one Member State to another. If they are not corrected by adjustments or wrongly corrected, this can cause differences in mirror statistics. Response for dispatches is generally better than response for arrivals.

2. **Thresholds**. The coverage of trade, after application of the statistical thresholds, varies between 94% and 99.8%, depending on the Member State. If not corrected by adjustments - and certain Member States do not make adjustments - these differences in coverage naturally cause differences in mirror statistics. Coverage of dispatches is generally better than coverage of arrivals.

3. Furthermore**, adjustment methods** can cause differences, since they differ, or if not all Member States make adjustments.

4. **Statistical confidentiality**. Confidentiality can affect product or partner-country classification. Asymmetries occur because confidentiality is not applied in the same way in all Member States - one Member State may exclude a transaction from detailed trade statistics whereas the partner country includes it; one Member State may record a transaction under a different commodity or country code compared to the partner country.

5. **Indirect exports**. BE exports to a third country, but the export document is made out in NL. There is probably an Intrastat declaration in BE for the dispatch to NL, but no Intrastat arrival declaration in NL, because there is no one to make this declaration. In theory, NL should use the export document (customs declaration) in order to adjust arrivals from the "actual Member State of export".

6. **Triangular trade** exists where a company in Member State A sells goods to a company in Member State B, which in turn sells it to a company in Member State C, although the goods are physically moved only once - from A to C. In such cases, intra-Community trade statistics should record a dispatch from A to C and an arrival in C of goods from A. There is, however, a risk that A or C will regard Member State B as its trading partner, whereas B will not record this trade because there is no physical movement in B.

7. **Different ways of defining the value of goods**. Differences in value definition (CIF, FOB, transport costs included or not) can cause minor mirror differences in Intra trade. Providers of statistical information (PSIs) have practical difficulties in establishing a FOB or CIF value, because invoice values have to be adjusted on the basis of incomplete data (transport invoices are often not detailed by product and are seldom linked to invoices for commodities).

8. **Exchange rate differences**

9. **Time gap**. The same operation can be recorded under a different reference period because of transport times, holiday months etc.

10. **National and Community concept**. Comparing trade according to the national concept with trade of another Member State according to the Community concept should also be avoided. Only the Community concept should be used, because only this concept is harmonised.

11. **Differences in the classification of goods**. Companies (PSIs) encounter many difficulties in correctly classifying their goods. Errors or different interpretations of CN8 can cause mirror differences at product level or even at the level of major economic sectors such as CN2. Goods classification is generally better for dispatches than for arrivals.

12. **Fraudulent VAT declarations** may influence statistics presumably more often in the country of dispatch than in the country of arrival.

13. It should be borne in mind that:

14. this list of causes is of course not exhaustive

15. to date it has not been possible to quantify the different causes for the various Member States.

ANNEX 2

Process of external trade data reconciliation

The statistics of the European Union trade balance must be made coherent and compatible to satisfy national and European necessities. There are particularly asymmetries inside the zone, especially those relative to the goods exchanges from Intrastat system of declarations. The sources of these asymmetries are multiple and their amplitude is difficult to quantify. Consequently, a modeling of the problem by various statistical methods was carried out. At the conclusion of this analysis, it turned out that the method of data reconciliation developed by the CREST/INSEE constituted the statistical tool allowing the calculation of the best value of mirror reconciled flows.

1. *Methodology of data reconciliation*

This method is based on multivariate models under constraints. The objective of this modelling is to correct existing bias by obtaining the part of the present errors with statistical tools. The basic idea consists in maximising the log-credibility function under constraints. The method considers difference between various series as an error. By modelling the process of error, one considers a series smoothed with regard to disturbances. Several models are tested according to the error type and process. Six tested models are the following ones:

- Stationary white Noise;
- No stationary white Noise,
- Cointegrated white Noise;
- MA (1);
- MA (4);
- MA (12).

All these models are specified and estimated at chapter level (2 first digits of the Harmonised System) for all links between Member States, using sliding monthly time series over 5 years. For each link at chapter level, the best model is selected on a minimisation of the BIC criterion (Bayesian Information Criterion). The method also takes into account various factors relative to the behaviour of the external trade series and particularly the fact that, in theory, the declarations of expeditions are better than the declarations of arrivals in term of quality. Consequently, a weighted BIC criterion as well as a distinction on series were introduced. $\mu1$ and $\mu2$ parameters were so introduced:

- $\mu1$ parameter allows to distinguish series. In case the proportion for which the number of months where expeditions are superior or equal to arrivals, is superior or equals in $\mu1$ then the adjusted BIC criterion is used on these series otherwise, a pure BIC is applied. $\mu1$ parameter was fixed to 0.333 i.e. 12 observations.

- $\mu2$ parameter allows to weight the BIC criterion so as to consider the quality of expeditions. This criterion weighted the model which minimises distance with the BIC criterion (min BIC) and distance with flow of expeditions (Min RMSE). The expression is as follow :

  - *Adjusted BIC = $\mu2$ min RMSE + (1-$\mu2$) min B*IC

  This parameter was fixed to 0.666.

The production of reconciled balances is made by a treatment in several stages. At first, data to be reconciled are extracted at 2 digits and the 00 (confidential European partner) and 99 (other products) chapters of the Combined nomenclature are redistributed in proportion to the importance of declarations by product at 2 digits. In a second step, data are reconciled over five sliding years by relation of a reporter with a partner. Finally, reconciled trade balances are calculated by various aggregations of estimated flows on a temporal and geographic base.

2.   *Reconciliation according to the SITC nomenclature*

At the same time, monthly data were reconciled according to the SITC nomenclature. Data reconciled over five sliding years from the combined nomenclature (2 digits) were used. To obtain correspondence between combined nomenclature at 2 digits and SITC nomenclature at 1 or 2 digits level, an estimation by ratios between declared and adjusted values was calculated. Obtained ratios were applied to data at 8 digits so as to obtain CN8 reconciled data. A correspondence table between CN8 and SITC nomenclatures was applied. The CN8 data aggregation at total level allows to obtain the ventilation of SITC codes at 1 digit and consequently the reconciled SITC data.
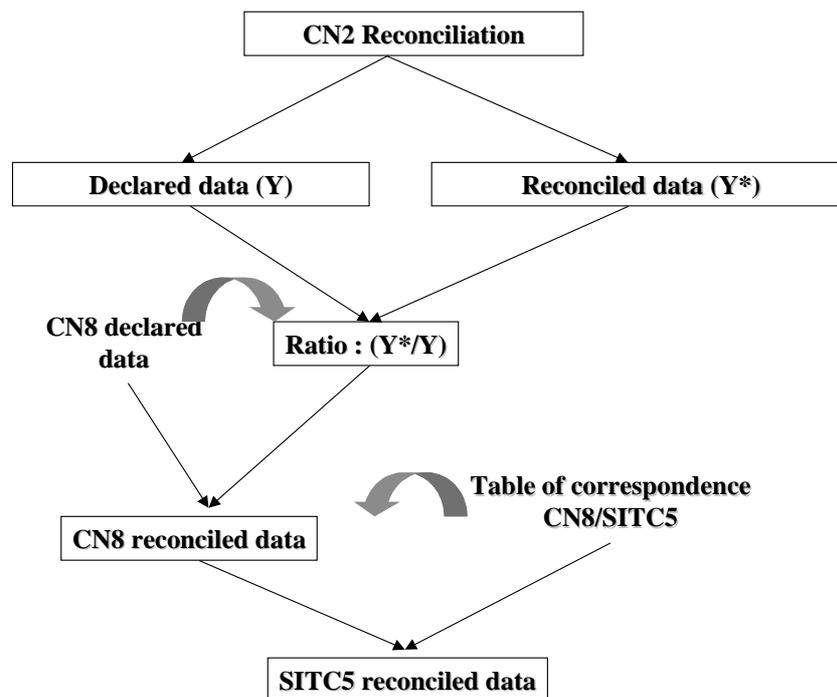
Evolution of intra-EU trade -
Quarterly value of arrivals (imports) and dispatches (exports)