

THE ASYMMETRY DILEMMA WITH INTRASTAT – WHICH DATA IS THE BETTER ONE? NATIONAL EXPERIENCES OUT OF THE ESS.VIP “REDESIGN OF INTRASTAT”

Ilda Duarte Fernandes¹

¹ *Federal Statistical Office (Destatis), Wiesbaden, Germany; ilda.duarte-fernandes@destatis.de*

Abstract

Since the implementation of the Intrastat system in 1993 to display movements of goods between the EU Member States, there was a focus on the reduction of administrative burden imposed on the reporting enterprises. In 2006 the Commission suggested a medium-to-long term implementation of the so-called single flow system, implying the compilation of national imports data from partner countries' export data. The subsequent discussion about the feasibility of this system showed that due to existing consistency issues the single flow system is not practicable on a medium-term perspective. Decisive criteria are asymmetries between the EU partner countries' data.

Nevertheless, in November 2011 the European Council decided to envisage further measures on burden reduction, explicitly not excluding the option of a single flow system, while maintaining a sound level of quality. Destatis intensively examined the use of mirror dispatch data of member states on the quality of intra-Community trade statistics. Special focus was laid on the effects caused by asymmetries: As for 2013 German imports from EU Member States were around 1.8 % lower than the respective exports. Broken down by partner country or commodity level those asymmetries might be even higher.

This may be ascribed to errors in reporting. However, to some extent asymmetries are due to the methodology itself. As these asymmetries are partially inherent in the system and may not be totally avoided they should not be considered as quality flaws.

Based on these findings it is analysed how far the degree of asymmetries fits as a tool for quality assessment in intra-Community trade statistics with regard to the use of mirror dispatch data of the member states.

Keywords: Foreign trade, Intrastat, asymmetries, data linking, estimations.

1. From the beginning of Intrastat to the ESS.VIP: “Redesign of Intrastat”

The introduction of the Single Market in January 1993 led to a fundamental change of the collecting system of the foreign trade statistics. With the Single Market, customs declarations were no longer required. As a consequence, a new collecting system for the intra-EU trade in goods was introduced changing from a secondary statistics (by using customs declarations as data source) to a primary survey, the so called Intrastat¹ (where enterprises report directly to the competent national authorities). It is considered to be the largest enterprise survey in the EU, which is why the reduction of the reporting burden on enterprises is one of the key issues ever since then.

In 1996 the European Commission launched the SLIM initiative (Simpler Legislation for the Internal Market) with the aim of identifying ways in which Single Market legislation could be simplified and of reducing the reporting burden on enterprises (Commission of the European Communities, 1996). As a result, the amount of variables to be collected has been reduced and simplification rules were introduced. In addition, the coverage rate for both flows, imports and exports, was reduced from 99 % to 97 % in 2004. Two years later, the Commission suggested a medium-to-long term implementation of the so-called single flow system², implying the compilation of national imports data out of the partner countries’ export data. The subsequent discussion about the feasibility of this system showed that because of existing consistency issues the single flow system is not practicable on a medium-term perspective. Decisive criteria are asymmetries between the EU partner countries’ data.

In November 2011, due to the continuing criticism of insufficient burden reduction in Intrastat, the European Council decided to envisage further measures on burden reduction,

¹ Intrastat is the data collection system for compiling statistics on international trade between the European Union Member States. Its main features consist of monthly statistical declarations sent directly by enterprises to the competent national authorities, a system of thresholds abolishing all statistical formalities for almost two thirds of enterprises and a close link with the tax system (Eurostat, 2014).

² In a single flow reporting system, statistics would be compiled by collecting data in each Member State for only one flow and deriving the other flow from the mirror statistics of the partner country. As there are fewer enterprises exporting than importing, the single flow option that is being discussed is based on collecting only exports as it could imply a significant burden reduction. In this context the reporting on imports would become redundant and a large number of enterprises would be exempted from reporting obligations (Eurostat, n.d.).

explicitly not excluding the option of a single flow system, while maintaining a sound level of quality. Therefore, further reductions of the coverage rate were conducted, although exclusively for imports³. Currently, the coverage rate for intra-EU imports is 93 %. However, the lower the coverage rate, the higher amounts the volume of trade (non-collected) that has to be estimated, which corresponds to a decrease in quality.

In order to address this trade-off between the burden reduction and the level of data quality Eurostat first launched the European Statistical System Vision Implementing Project (ESS.VIP) “SIMSTAT” (SIngle Market STATistics)⁴ in Mai 2012. However, while the work on the ESS.VIP “SIMSTAT” progressed, the lack of alternative options to the ESS.VIP “SIMSTAT” was criticised. Consequently, a second ESS.VIP was launched named “REDESIGN of Intrastat” in December 2014, which ran to April 2016 (Eurostat, 2015). The scope comprises a cost-benefit analysis of all possible options to modernise the intra-EU trade statistics. The focus is to be on the premise of keeping a sound level of quality in the intra-EU trade statistics. In addition, it addresses the administrative burden reduction, as well as key legal and methodological issues. The ESS.VIP “REDESIGN of Intrastat” aims at equipping the ESSC (European Statistical System Committee) in May 2016 with sufficient qualitative and quantitative evidence to allow the ESSC to make a decision on the future orientation concerning the data collection system of Intrastat (Economic Policy Committee, 2015).

2. The options of the ESS.VIP “REDESIGN of Intrastat” in Germany

On the EU-level the ESS.VIP “REDESIGN of Intrastat” comprised several different options not all of them being analysed in each member state. However, the common feature of these options was to focus on the modernisation of the collecting system of the imports in Intrastat.

³ Since introduction of Intrastat the coverage rate on the imports‘ side has been reduced in 2004, from 99 % to 97 %, further in 2010 to 95 % and the last one having been in 2013 to 93 %.

⁴ SIMSTAT is an approach for simplifying Intrastat with the aim of reducing the administrative burden on enterprises while maintaining data quality. For this purpose microdata (on the enterprise level) on intra-EU trade between Member States shall be exchanged and re-used. The SIMSTAT project’s objective is to prove the feasibility of exchanging microdata on intra-EU trade between the Member States, including both technical and statistical aspects (Eurostat, n.d.).

In Germany, two options were under detailed examination: “SIMSTAT-DE” and “Mixed Model”. Both options are based on the approach to keep on collecting national data on the imports’ side, although to a lower share. As a consequence, a higher value has to be estimated for which the distribution across partner countries and commodity codes is not given. In order to mitigate this loss of information, which also corresponds to a lower level of data quality, mirror data of member states is being used.

The mirror dispatch data of member states obtained by the microdata exchange between 20 member states performed within the ESS.VIP “SIMSTAT” provided the database for the simulations within the ESS.VIP “REDESIGN of Intrastat”⁵.

2.1. “SIMSTAT-DE”

Under the “SIMSTAT-DE” option the coverage rate of the national collection of imports is lowered from the actual coverage rate of 93 % to 85 %, e.g. the exemption threshold increases from 800,000 EUR, which is currently being applied, to 5 million EUR. As a consequence, the estimation value increases from 7 % to 15 %. In order to determine the distribution structure across partner countries and commodity codes for these remaining 15 %, mirror dispatch microdata of member states is used as an additional data source. The basic idea is to reduce the information within the mirror dispatch data of the member states by the information already collected on the national level in order to use the remaining information of the mirror dispatch data of the member states for obtaining the distribution structure across partner countries and commodity codes for the estimation value of 15 %.

Practically, in a first step the 85 % of national collected value of imports is aggregated on the enterprise level by partner country and by commodity code. The same procedure is performed for the mirror dispatch data of the member states. The second step consists of merging both datasets on the (micro-) enterprise level in order to extract the information already collected on

⁵ During the third and final phase of SIMSTAT (2nd quarter 2015 – 3rd quarter 2015) the exchange of microdata between the following 20 Member States took place: Austria, Bulgaria, Czech Republic, Germany, Denmark, Estonia, Greece, Finland, France, Croatia, Italy, Lithuania, Luxembourg, Latvia, Malta, Poland, Portugal, Romania, Slovenia, Slovakia (Eurostat, n.d.).

the national level by excluding all matches from the mirror dispatch data of the member states. The remaining information of the mirror dispatch data of the member states is then used to calculate the percentage distribution structure across partner countries and commodity codes, which is applied on the estimation value of 15 %. By joining the national collection of imports and the estimation value distributed according to the information of the mirror dispatch data of the member states, the compilation of national imports under the option of “SIMSTAT-DE” is completed.

2.2. “Mixed Model”

The “Mixed Model” has a very similar design to the “SIMSTAT-DE”: In alternative 1 the coverage rate of the national collection of imports is also lowered to 85 % and in alternative 2 it is lowered to 90 %. The distribution structure across partner countries and commodity codes for the estimation value of 15 % in alternative 1 and the estimation value of 10 % in alternative 2, respectively, is also derived from the mirror dispatches of the member states by reducing the information by the national collected imports.

The difference to the option of “SIMSTAT-DE” is that the merging of both datasets is not performed on the microdata, e.g. the enterprise level, but on a more aggregated level, namely on partner country and commodity code level. Then again, the remaining information of the mirror dispatch data of the member states is used to calculate the percentage distribution structure across partner countries and commodity codes, which is applied on the respective estimation value of the alternatives. The last step is to join the national collection of imports and the estimation value distributed according to the information of the mirror dispatch data of the member states, by which the compilation of national imports under the option of the “Mixed Model” is completed.

3. The challenge of asymmetries in integrating mirror dispatch data of member states

At present, monthly results of foreign trade statistics in Germany are produced without the use of any partner data. By introducing the use of mirror dispatch data of member states, as it is simulated in the two options of “REDESIGN of Intrastat”, there are major challenges, like

data linking, compilation concepts and asymmetries, that need to be addressed in order to mitigate major breaks in time series and to preserve the consistency of the results of the national collection.

The main challenge consists in the treatment and integration of existing differences between mirror figures, e.g. asymmetries⁶. As former analyses showed, the reasons for asymmetries are diverse. It may be distinguished between asymmetries that are not inherent in the system, which are difficult to detect as they are ascribed to errors in reporting. This refers to differing reports of commodity codes or partner countries between mirror data and national data. The latter may be the case for triangular trade or chain transactions as exporters should declare the last country of destination of the goods and importers the country of origin. If the exporter is not aware of the final country of destination when declaring to Intrastat, asymmetries are produced. However, there are also asymmetries inherent in the system as they result from the methodology itself: For example, within Intrastat, different coverage rates for exports (at present 97 %) and imports (currently 93 %) are required which may lead to bilateral asymmetries. In this context, member states apply different estimation methods regarding non-response and below threshold trade, which are closely linked to VAT data. However, mirror VAT data deviates from national VAT data, too. Hence, asymmetries are produced. The treatment of confidentiality is another source for asymmetries as within foreign trade statistics the passive confidentiality⁷ is applied. Furthermore, the treatment of simplification rules is also differing between the member states. As a consequence, these asymmetries inherent in the system may not be avoided, which is why they should not be considered as quality flaws.

As for 2013 German imports from EU member states were around 1.8% lower than the respective exports. Broken down by partner country or commodity level those asymmetries might be even higher. Hence, a special focus of the consistency is laid down on the analysis

⁶ An asymmetry in foreign trade statistics is defined as the difference between the national trade value recorded by countries and the respective trade value in the mirror flow, e.g. the difference between what Germany records as arrivals from France and what France records as the corresponding dispatches to Germany.

⁷ Passive confidentiality in foreign trade statistics is regulated by European law (Regulation (EC) No 471/2009 Article 10, Regulation (EC) No 638/2004 Article 11). In accordance with that law, individual data are kept confidential at the request of the exporters or importers concerned (Destatis, 2016).

how far the degree of asymmetries fits as a tool for quality assessment in intra-Community trade statistics with regard to the use of mirror dispatch data of the member states.

4. Consistency analysis within “REDESIGN of Intrastat”

The analysis of consistency of the simulated results of the option “SIMSTAT-DE” and the “Mixed Model” is one of the key aspects within “REDESIGN of Intrastat”. The approach for measuring consistency consists of the comparison of the results according to the actual national collection of imports to the simulated results derived from “SIMSTAT-DE” and the two described alternatives of the “Mixed Model”. The results presented focus on a consistency analysis by partner country and by the most detailed level of commodity codes, e.g. the 8-digit code level.

Table 1: Comparison by partner country of simulated options to nationally collected Intrastat imports – year 2013

Member states	“SIMSTAT-DE”	“Mixed-Model” coverage rate 85%	“Mixed-Model” coverage rate 90%
	Index: national collection of imports = 100		
AT	97,7	96,7	97,2
BG	97,1	101,2	101,8
DK	99,5	96,7	97,0
EE	95,9	97,5	95,7
FI	100,8	95,5	96,4
FR	100,4	103,2	102,8
GR	98,5	98,8	98,7
IT	101,3	98,0	98,0
LT	103,3	97,1	98,5
LV	98,5	102,9	102,8
MT	91,9	104,1	105,8
SI	99,4	93,1	94,3
SK	101,7	107,9	107,1

4.1. Consistency analysis by partner country

The consistency analysis by partner country was conducted by calculating the shares by partner country of the simulated results according to “SIMSTAT-DE” and both alternatives of the “Mixed Model” on the actual national collection. The results summarised in table 1 do not

enable any prioritisation of the options as no significant difference between the options may be stated. Depending on the partner country, the displayed options show different levels of approximation. However, the percentage difference in value between the results of the simulated options and the national collection are within a range of +10 % and -10 %.

4.2. Consistency analysis by commodity code

The consistency analysis by commodity code on the 8-digit level focuses on the results obtained from the analysis of partner data from France and Italy. This is due to the fact that France and Italy are the only member states who were already collecting the identification number, namely the VAT identification number, of the importing receiver in Germany in their national collection of exports during 2013. Hence, the identification number of the importing receiver in Germany was not simulated as applied for the mirror dispatch data of the other member states that participated in the microdata exchange of the ESS.VIP “SIMSTAT” (see footnote: 6). As the simulation of “SIMSTAT-DE” is based on mirror dispatch microdata of member states, for France and Italy a robust comparison to the “Mixed Model” for these both countries is given.

The percentage differences in value between the simulated results (of the options “SIMSTAT-DE” and the “Mixed Model”) and the national results on imports are calculated for all the bilaterally traded 8-digit commodity codes (some 9,000 different commodity codes) and assigned to 3 different categories:

- Relatively small percentage difference in value of less than 10 %
- A medium percentage difference in value of at least 10 % and up to 50 %
- And a relatively high percentage difference in value of at least 50 %

As stated in table 2, the simulation of “SIMSTAT-DE” for France shows the best approximation to the actual national results: 83.7 % of the imported value is assigned to 8-digit commodity codes that show a small percentage difference in value of less than 10 %. In comparison, the share in value of the same category for the “Mixed Model” with a coverage rate of 85 % for imports (alternative 1) is lower, namely 69.2 %, whereas the category

corresponding to a high percentage difference in value is twice as high as for “SIMSTAT-DE”. The results concerning the consistency for the “Mixed Model” with a coverage rate of 90 % for imports (alternative 2) are situated between the results of “SIMSTAT-DE” and the alternative 1 “Mixed Model”.

Table 2: Percentage difference in value by commodity code between nationally collected Intrastat imports and simulated options for France – year 2013

Percentage difference in value: (Simulated options - National imports)/ National imports	“SIMSTAT-DE”	“Mixed-Model” coverage rate 85%	“Mixed-Model” coverage rate 90%
	Value of national imports	Value of national imports	Value of national imports
	in %	in %	in %
difference \leq 10,0	83,7%	69,2%	75,7%
10,0 < difference \leq 50,0	13,9%	26,2%	21,0%
50 < difference	2,4%	4,6%	3,3%

Table 3 displays the results of the consistency analysis by 8-digit commodity codes for the German imports from Italy. The results show a less satisfactory level than for France, but with regards to the evaluation of the consistency, the same ranking of the options can be stated: “SIMSTAT-DE” gives the best approximation to the actual national results, as 67.5 % of the imported value is assigned to 8-digit commodity codes that show a small percentage difference in value of less than 10 %, whereas for the alternative 1 “Mixed Model” the share of this category is lower, namely 48.5 %. The results for the alternative 2 “Mixed Model” are again between the results for “SIMSTAT-DE” and the alternative 1 of the “Mixed Model” as this category accounts for 58.6 % of the imported value.

Table 3: Percentage difference in value by commodity code between nationally collected Intrastat imports and simulated options for Italy –year 2013

Percentage difference in value (Simulated options - National imports)/ National imports	“SIMSTAT-DE”	“Mixed-Model” coverage rate 85%	“Mixed-Model” coverage rate 90%
	Value of national imports	Value of national imports	Value of national imports
	in %	in %	in %
difference \leq 10,0	67,5%	48,5%	58,6%
10,0 < difference \leq 50,0	27,7%	43,1%	36,1%
50 < difference	4,8%	8,3%	5,2%

5. The asymmetry dilemma with Intrastat: “SIMSTAT-DE” vs. “Mixed Model”

Although asymmetries are measured on (aggregated) commodity code level, most of them result from the enterprise level, like reporting errors, request for confidentiality of exporters and importers and even the usage of simplification rules. Thus, in the case of asymmetries it is hardly possible to identify, whether it is the mirror data or the national data being responsible for the asymmetries. Hence, these asymmetries may not be avoided, which is why they should not be considered as quality flaws. In this respect, the evaluation of the options was not based on the ‘quality’ of both datasets in the sense of measuring the existing asymmetries, but rather on the consistency in order to mitigate breaks in time series as much as possible.

The consistency analysis by the 8-digit commodity code level leads to the conclusion that “SIMSTAT-DE” is the best fit with regards to mitigate breaks in time series and thus to maintain a sound level of quality. This is due to the fact that existing asymmetries are treated differently in “SIMSTAT-DE” than in the “Mixed Model”: For “SIMSTAT-DE” the matching of the mirror dispatch data with reported national imports is performed on the micro level, e.g. the enterprise level, whereas for the “Mixed Model” it is performed on the aggregated level (on partner country and commodity code level). As for “SIMSTAT-DE” the linking of the mirror dispatch data of member states with the nationally collected imports is performed on the (micro-) enterprise level, existing asymmetries to the nationally collected imports are eliminated from the mirror dispatch data of the member states. Hence, the impact of asymmetries is reduced to the distribution structure across partner countries and the commodity code level applied on the remaining estimation value of 15 %. As a consequence, 85 % of the national imports remain coherent to the previous national collections of imports, by which the impact of asymmetries remains moderate. In contrast, the linking within the “Mixed Model” is performed on the aggregated level. Hence, the impact of asymmetries is not only reduced to the estimation value, instead it is extended to the national collection as well, which causes more significant breaks in time series on the 8-digit commodity code level.

With regard to these findings and the main objective of producing results of a sound quality, which is analysed in this context under the premise of consistent results of the national

compilation of imports, asymmetries should be treated carefully and in a distinguished way as a tool for quality assessment in intra-Community trade statistics with regard to the use of mirror dispatch data. To a certain degree asymmetries are inherent in the system as they rely on the methodology itself and in this context it is not possible to decide on the precedence of one data source or the other.

6. References

Commission of the European Communities (1996), COM(96) 520 final: Communication from the Commission to the European Parliament and Council: The impact and effectiveness of the single market. Available: http://www.ec.europa.eu/internal_market/economic-reports/docs/single_en.pdf [2016, April 26].

Economic Policy Committee (2015), Requirements for structural statistics 2015: overview of progress made - Input to the EFC Statistical Package -Available: <http://data.consilium.europa.eu/doc/document/ST-14102-2015-INIT/en/pdf> [2016, April 26].

Statistical Office of the European Communities [Eurostat] (2015), REDESIGN. Re-design of Intrastat. Business project. Available: <http://ec.europa.eu/eurostat/documents/42577/6906243/REDESIGN+fact+sheet/7f8842ef-ca59-443a-bd27-5120ce13e604> [2016, April 26].

Statistical Office of the European Communities [Eurostat] (2014), Statistics explained. Glossary: Intrastat. Available: <http://www.ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Intrastat> [2016, April 26].

Statistical Office of the European Communities [Eurostat] (n.d.), Statistics explained. SIMSTAT - exchange of microdata on intra-EU trade. Available: [http://ec.europa.eu/eurostat/statistics-explained/index.php/SIMSTAT - exchange of microdata on intra-EU trade](http://ec.europa.eu/eurostat/statistics-explained/index.php/SIMSTAT_-_exchange_of_microdata_on_intra-EU_trade) [2016, April 26].

Statistisches Bundesamt [Destatis] (2014), Quality report. Foreign Trade. Available: <https://www.destatis.de/EN/Publications/QualityReports/ForeignTrade/ForeignTrade.html> [2016, April 26].