

# „Please take care of our data...!”

Søren Rich, Statistics Denmark, sri@dst.dk

## **Abstract.**

*Exchanging microdata between countries opens up new opportunities. Sharing microdata between national statistical authorities (NSAs) can improve the quality of statistics covering cross-border activities, e.g. external trade in goods and services, balance of payments, foreign investment statistics, etc. The exchanged microdata can be used as a new data source, to produce new statistics or replace collected data and thus reduce the burden on respondents.*

*However, exchanging microdata comes at a price. The microdata must be protected and information about individual companies or people must not fall into the wrong hands. It is usually the data collecting NSA that is responsible for the data protection. However, in a data exchange between countries, the data receiving NSA must also take action to protect the received microdata. This requires some form of cooperation between the NSAs involved in the data exchange.*

*Since 2022, the NSAs of the EU Member States have exchanged microdata on their intra-EU exports through the so-called MDE data exchange (Micro Data Exchange). The NSAs use the received microdata to varying degrees to compile intra-EU import statistics and have achieved significant burden reductions.*

*The EU legislation behind MDE exchange require the receiving Member States to respect the confidentiality of the sending Member State. An Eurostat task force worked for a number of years on how NSAs should respect the confidentiality of other NSAs. The work turned out to cover many aspects of data confidentiality and security.*

*This paper shares some of the experiences and conclusions from this work. In addition to handling classic data suppression when statistics are disseminated, the work showed that when microdata are exchanged and reused by other NSAs, there is also a need to reach agreement on how data are used. Do the use of the microdata put restrictions on how to use the data in contact to reporting units in receiving Member States, how to make data available for researchers, who has ‘ownership’ of the data, etc.*

*The experience from this work can be used in other data exchanges, being the large multi-country exchanges or exchanges at a bilateral level.*

**Keywords:** Keywords: microdata exchange, dissemination, confidentiality, data protection, cooperation between national statistical authorities

## 1. Introduction

Exchanging microdata between countries opens up new opportunities for national statistical authorities (NSAs). Sharing microdata between the NSAs can improve the quality of statistics covering cross-border activities, e.g. international trade in goods and services, balance of payments, foreign investment statistics, etc. The exchanged microdata can be used as a new data source, to produce new statistics or replace collected data and thus reduce the burden on respondents. Quality and burden is parameters NSAs always are measured on and why data exchange and sharing information between NSAs is highly relevant.

The statistics on trade in goods between the EU Member State (In the past called Intrastat) is the most burdensome data collection in all Member States and ever since the EU single market was introduced in 1993, focus has been on how the burden on reporting units could be reduced from this data collection.

Most initiatives have focused on reducing the number of data elements to be reported or reducing the number of small companies being reporting units, thus decreasing the coverage rate and replacing reported data with estimations, typically based on VAT statements of the exempted companies. Gradually, the coverage rate, i.e. the percentage rate of total intra-EU trade being covered by collected data, has been reduced, from comprising trade from all traders in 1993 to covering 97% of total intra-EU exports and 93% of total intra-EU imports in 2021. The result is that almost all small companies and many medium size companies are exempted from reporting to Intrastat.

However, the focus on reducing burden on the reporting units continued. In 2012-2016, a SIMSTAT pilot study on exchange of Intrastat micro data was carried out by the NSAs in the EU Member States. The results pointed toward that it was possible technically to exchange high volume of micro data between the NSAs and that there was a business case in reusing the intra-EU trade micro data from other Member States. The positive outcome lead to the decision to make exchange of intra-EU export micro a new data sources for the NSAs in the EU Member States.

After a preparatory phase, the NSAs of the EU Member States began exchanging microdata on their intra-EU exports through the so-called MDE data exchange (Micro Data Exchange) in 2022. The coverage rate of intra-EU exports has been lowered further to 95% and in intra-EU imports coverage requirement are completely removed.

## **2. The EU MDE microdata exchange**

The mandatory exchange of intra-EU export microdata consists of monthly exchange of micro data from individual exporters covering the export to the other EU Member States. It is data being reported to the survey in the exporting EU Member State and contains information on partner country, detailed 8-digit Combined Nomenclature product codes, value and quantity and VAT ID number of the importer of the goods, together with other data elements supporting the compilation of International Trade in Goods Statistics (ITGS). Various metadata supporting the MDE exchange are also exchanged between the NSAs. The data transmissions are done using the TESTA (Trans European Services for Telematics between Administrations) network that offer a high level of security, because it is available only to public European administrations, and not to the general public. Eurostat host a data hub managing the exchange.

The NSAs use the received microdata to varying degrees to compile intra-EU import statistics and have achieved significant burden reductions since 2022. Two countries, Finland and Estonia, have gone all the way: they have introduced a single flow system where they have stopped collecting data on intra-EU imports at national level and instead use the mirror MDE export microdata from the other EU Member States to compile the intra-EU imports.

However, exchanging microdata comes at a price. The microdata must be protected and information about individual companies or people must not fall into the wrong hands. It is usually the data collecting NSA that is responsible for the data protection. However, in data exchange between countries, the data receiving NSA must also take action to protect the received microdata. This requires some form of cooperation between the NSAs involved in the data

exchange. It also require a trust in that other NSAs have the same high standards with respect to data security, etc.

The EU legislation behind the MDE exchange requires the data receiving Member States to respect the confidentiality of the data sending Member States. The European Statistical Law<sup>1</sup> provide the main principles of good practice applicable to the production of statistics within EU: it regulates the statistical work and cooperation in the European Statistical System (ESS) and provide provisions on how NSAs and Eurostat must comply with data security and confidentiality of data.

The main principle is that individual company or personal data must never be published, distributed or used by NSAs in any way other than for statistical purpose. Individual company or personal data are always *confidential data* and must be highly protected and not transmitted by the NSA to other public administrations or to other users.

However, even when compiling statistics consisting of aggregated data, individual micro data could be revealed indirectly: If only one or two companies are involved in e.g. trade of a certain product with a specific partner country, the published information may reveal individual companies' data. Consequently, the data are considered being *statistical confidential* when they allow a company or person to be directly or indirectly identified.

All business statistics use *active confidentiality* – data cells must be suppressed if only one or two companies are providing data to this cell or if among a number of companies some can be said to be dominating. When using active confidentiality NSAs must take action to protect data, which can directly or indirectly reveal the data of an individual company or person.

Only one business statistics, International Trade in Goods Statistics (ITGS) use the *passive confidentiality* concept: Passive confidentiality means that the NSA must only suppress disseminated data if requested by the trader whose individual data might be indirectly revealed from disseminated results.

---

<sup>1</sup> Regulation (EC) No 223/2009 of the European Parliament and of the Council

Companies typically apply for passive confidentiality and thus suppression of data because the data contain sensitive commercial information.

This concept means that some ITGS data can be disseminated even if the requirements of active confidentiality is not fulfilled, as long as the data does not concern trade in a product code where a company has been granted passive confidentiality. This unique treatment of confidentiality is used in ITGS to ensure that detailed statistics at 8-digit CN level can be disseminated. Without passive confidentiality, detailed ITGS tables would have many empty cells because data had to be suppressed.

How the confidentiality of a sending Member State must be respected in practice proved more difficult to clarify. In which situations must the NSA using MDE data from other EU Member States take action to protect the data? What guidelines should be provided to the NSAs? A Eurostat task force worked for a number of years on how NSAs should respect the confidentiality of other NSAs when using their MDE data in the compilation of ITGS. The work turned out to cover many aspects of data confidentiality and data security.<sup>2</sup>

### 3. Classic data suppression issues

The main problem was how a NSA receiving and using MDE micro data from other NSAs should 'respect' the confidentiality of the micro data. The EU legislation states that *"if the NSAs of the Member State of import make use of microdata records on exports which are subject to confidentiality, they shall ensure that the dissemination of statistical results on intra-Union imports, by the NSAs of the Member State of import, respects the statistical confidentiality granted by the NSAs of the Member State of export."*<sup>3</sup> The term 'respect' in the legal text is rather vague and leave room for interpretation. It also require some

---

<sup>2</sup> This author first had the pleasure of participating in the task force as the Danish representative, and later, from 2021 until the task force concluded its work in early 2025, as a national expert seconded to Eurostat, responsible for the work of the task force.

<sup>3</sup> Article 14(2) in Regulation (EU) 2019/2152 of the European Parliament and of the Council of 27 November 2019 on European business statistics, repealing 10 legal acts in the field of business statistics.

definition of “make use of” because not all use of MDE data may reveal confidential data.

In the MDE exchange the micro data being passive confidential are flagged as statistical confidential in the data files. However, it is up to the NSA receiving the MDE data to take the necessary steps to protect the data being flagged as confidential. Often, this NSA will already have some product codes flagged as confidential in its import statistics based on its own data collection. Trade in these product codes are typically suppressed by hiding the trade under totals or adding the trade together with trade in other product codes. The NSAs thus have the experience in protecting data and the question of ‘respect’ is more about when to take action, not how to take action.

Initially, the task force carried out an analysis simulating confidentiality in a data exchange. The analysis was based on a simulation using ITGS data from CN chapter 46.<sup>4</sup> The work concerned how in practice the statistical confidentiality in intra-EU imports could be kept when using the MDE data.

This rather artificial set-up indicated what could be called a domino effect: To protect import of product X from country A, country B must hide the import of X together with import in product Y, which result in that country C must hide export of Y to B to avoid revealing the confidential import of Y in country B, and so forth.

However, in practice, exchanged MDE data are less likely to become visible in the dissemination of the data receiving Member State. This is because in the compilation of intra-EU trade in goods statistics, Member States typically add estimates for trade from traders which are not reporting units: All small and many medium size companies will typically not report to the intra-EU trade in goods surveys. The added estimates from the exempted companies usually distort the aggregated data so much that MDE data does not become visible.

---

<sup>4</sup> Combined Nomenclature, the product classification applied in EU customs declarations and Intrastat. This CN chapter 46 consist of 22 8-digit CN product codes covering manufactures of straw, basketware and wickerwork and is the only CN chapter where no Member States had any statistical confidentiality.

Furthermore, many methods using the MDE data in the compilation of ITGS result in that the MDE data will not be visible in the compiled statistics. In the compilation, the MDE microdata are often used to allocate estimates for missing trade from exempted companies by partner country and CN product codes. When these estimates are added together with the intra-EU import data collected from national reporting units, it will often create trade data where the (confidential) MDE data cannot be detected, i.e. the MDA data will not be visible in disseminated import statistics. This indirect use of MDE data will thus typically not result in that MDE data flagged as confidential will be disclosed in the compiled statistics.

It is a different matter when MDE data replace part or all import data collected from national reporting units. A single flow is a direct use of MDE data, where the national data collection of intra-EU imports is fully replaced 1:1 by the MDE export data from the other EU Member State. In this case, the MDE data from the exporting Member States will in most cases become directly visible in the import statistics which the NSA of the importing Member State compile and disseminate. The NSA using the received MDE data must take actions to hide the trade being flagged as confidential by other Member States. This also applies if only part of the national data collection is replaced 1:1 by MDE data.

The task force developed a set of main principles for how to act when using MDE data in compilation and produced a number of possible methods to apply to respect the confidentiality of the sending Member States.<sup>5</sup>

#### 4. Using exchanged MDA data in data collection and data validation

During the work of the task forces it also became evident that it was not only in the dissemination of ITGS that MDE data flagged as confidential could be disclosed. It was also a risk that MDE micro data could be disclosed in the processes taking place before the dissemination. Furthermore, the previous

---

<sup>5</sup> The guidelines developed by the task force resulted in a substantial expansion of chapter 9 on confidentiality in the annual publication *"European business statistics compilers' manual for international trade in goods statistics – detailed data"* (<https://ec.europa.eu/eurostat/publications/manuals-and-guidelines>).

section 3 shows that different methods using MDE data in the compilation of ITGS implies different actions to protect confidential MDE data and this points to another problem: how to interpret 'make use of' the MDE data as stated in the article which required the NSAs to respect the confidentiality of the MDE data.

Other use of the MDE data will typically be in the data validation of ITGS or as input in other statistics where ITGS is used as input, e.g. balance of payments or national accounts.

It was necessary to make guidelines on how NSAs can contact importers to validate data if they use MDE microdata in validation of ITGS. If the data receiving NSA e.g. compare import data reported by an importer with the mirroring MDE export data reported by an exporter in the sending Member State, then it is not possible for the NSA in the receiving end to reveal any information from the MDE data to this importer. Even if the export transactions concerns trade with the importer, the importer is considered a third party which is not authorised to gain access to the MDE data. Experts from some Member States claimed that because the MDE micro data contained the VAT ID number of the importer, it was also data 'owned' by this importer and they could be shared with importer. However, the conclusion was clear, the importer cannot be considered a party authorised to access the MDE micro data.

In Intrastat, companies are made reporting units based on their VAT statements or other administrative data. The task force recommended not to use the MDE data to make any importer a reporting unit because not being authorities to access the data, the importer would not be able to be informed about the base for the reporting obligation.

The findings of the task forces also indicated that to ensure the data security and confidentiality of the MDE microdata, good data management internally in the NSA is vital: NSAs must be able to distinct between MDE data and its own data and which data are flagged as confidential and which are not. Otherwise, you risk disclosing confidential data to a third party or may not take the necessary actions to protect data in the dissemination.

## 5. Access to MDE micro data for researchers

Another area where data security considerations must be made in a data exchange setup is when NSAs grant restricted access to micro data for research purposes. If MDE export data are used for compilation of intra-EU import statistics, it is obvious that researchers also want to have access to these data. The majority of EU Member States provide access to ITGS micro data for researchers in some form, where the reporting units are made anonymous.

The EU Regulation states that access to exchanged MDE microdata may be granted to researchers carrying out statistical analyses for scientific purposes subject to the approval of the NSA of the exporting Member State which provided the data.<sup>6</sup> How this approval should be understood in practice also became a topic for the task force. Should a NSA contact NSAs in the 26 partner Member States to get a formal approval?

The guidelines developed by the task force states that if the MDE data are used in an unchanged form in the compilation of intra-EU imports – i.e. the original MDE data are inserted 1:1 in the intra-EU import data with no changes in the data elements of the MDE data – then the NSA using the MDE must get the approval of the NSAs of the other Member States, if access to microdata for research purposes is granted. However, if the MDE data are changed or in some way processed by the NSA using the data in the compilation of intra-EU imports, then it is not necessary to get approval from NSAs sending the MDE data.

In practice, when using the received MDE data in compilation on intra-EU imports, the NSA must assess to what extent the use of the MDE data would require an approval by the sending Member States, if the NSA wishes to provide access for researchers. It would depend on the use of MDE data and if they become directly visible in the data made available to the researchers. However, product codes which are flagged as confidential in the MDE data, must be excluded from the microdata made accessible for researchers.

---

<sup>6</sup> Article 15 in Regulation (EU) 2019/2152 of the European Parliament and of the Council of 27 November 2019 on European business statistics, repealing 10 legal acts in the field of business statistics.

## 6. Conclusions

Data exchange between countries raise many issues. In addition to practical and technical problems of exchanging the data in a secure way, data protection in the country receiving the microdata is an important issue that should not be underestimated in a data exchange project. It not only concerns the protection of statistical confidential data when they are used in the dissemination, the typical issue arising when disseminating statistics based on microdata from individual units. It also concerns a proper and secure handling of data when providing access for researchers or when using the micro data in data validation and contact to statistical units in the data receiving Member State. How data are used in the compilation also has an impact on the necessary actions taken by the NSA in the receiving country to protect the microdata received. Therefore, in a data exchange setup the parties involved should also discuss how data can be used and in which cases the use will require actions to protect the data.

Before a data exchange starts, the involved parties should discuss and agree how data can be used and how they should be protected. The agreed approach can be formalised in a memorandum of understanding, or if necessary, in a more legal binding document.

To some extent the findings in the work of task forces also applies to exchange of data between NSAs in the same country, because the protection of confidential data from individual companies or persons also applies in these cases.