OECONOMIA
COPERNICANA

VOLUME 10 ISSUE 1 MARCH 2019

p-ISSN 2083-1277, e-ISSN 2353-1827
www.oeconomia.pl

ORIGINAL ARTICLE

Citation: Markowicz, I., & Baran, P. (2019). ICA and ICS-based rankings of EU countries according to quality of mirror data on intra-Community trade in goods in the years 2014–2017. *Oeconomia Copernicana, 10*(1), 55–68. doi: 10.24136/oc.2019.003

Contact to corresponding author: iwona.markowicz@usz.edu.pl; University of Szczecin, Faculty of Economics and Management, Mickiewicza 64, 71-101 Szczecin, Poland

Received: 30.11.2018; Revised: 10.01.2019; Accepted: 17.01.2019; Published online: 2.03.2019

Iwona Markowicz
University of Szczecin, Poland
orcid.org/0000-0003-1119-0789

Paweł Baran
University of Szczecin, Poland
orcid.org/0000-0002-7687-4041

ICA and ICS-based rankings of EU countries according to quality of mirror data on intra-Community trade in goods in the years 2014– 2017

JEL Classification: *F14; C10; C82*

Keywords: *official statistics data quality; mirror data; intra-Community trade; EU*

Abstract

Research background: As a system of official EU statistics, Intrastat contains data collected by Member States aggregated by Eurostat on the Union's level in the form of COMEXT database. Country-level data are based on declarations made by businesses dispatching or acquiring goods from other EU Member States. Since the same transaction is declared twice — as an ICS in one country and at the same time as an ICA in another country by the partner — the database contains mirror data. Analysis of mirror data lets us assess the quality of public statistics data on international trade.

Purpose of the article: The aim of the article is to rank EU Member States according to quality of data on intra-Community trade in goods collected by Intrastat. Foreign trade stimulates economic development on one hand and is the development's reflection on the other. Thus it is very important that official statistics in this area be of good quality. Analysis of mirror data from partner states in intra-Community trade in goods allows us to claim that not every Member State pro-vides data of satisfactory quality level.

Methods: We used the authors' methodology of assessing quality of mirror data. These include data asymmetry indices, both proposed by Eurostat and the authors' own proposals. We have also examined the changes in the above mentioned rankings over time.
Findings & Value added: The result of the survey is ordering of EU Member States according to the quality of data on intra-Community trade in goods. The rankings are presented for the period of 2014–2017, during which there were 28 Member States of the EU. Changes in distinct countries' positions were shown as a result of changes in overall quality of statistical data collected in these countries. The research methodology can be used in the process of monitoring data quality of the Intrastat system.

Introduction

Foreign trade stimulates economic development on the one hand, and is the development's reflection on the other. Thus, it is very important that official statistics in this area be of good quality. These data are used to assess the economic situation of a given state and in the process of creating strategies both at the country level and for the entire European Union. This is why Eurostat, together with the national statistical offices, attaches great importance to monitoring and improving the quality of intra-Community trade data.

The Intrastat system was introduced in 1991 by Council Regulation No 3330/91 (7 November 1991) on the statistics relating to the trading of goods between Member States, and has been applicable since 1993. Since then, international trade in goods statistics have been based on two data systems: for intra-EU and for extra-EU trade statistics. Extra-EU trade data, which relate to the trading of goods with non-member countries, continue to be collected by customs administrations, whereas most of the intra-EU trade data are directly collected from traders within the Intrastat system (Eurostat, 2017b).

As a system of official EU statistics, it contains data collected by Member States aggregated by Eurostat on Union's level in the form of COMEXT database. Country-level data are based on statistical declarations made by businesses dispatching or acquiring goods from other EU Member States. Since the same transaction is declared twice: as an intra-Community supply of goods (ICS) in one country, and at the same time as an intra-Community acquisition (ICA) in another country by the partner the database contains mirror data. Analysis of mirror data lets us assess the quality of public statistics data on international trade.

The aim of the article was to prepare an EU Member States ranking according to the quality of data on intra-Community trade in goods collected by Intrastat. An analysis of mirror data from partner states in intra-Community trade in goods allows us to claim that not every Member State

provides data of satisfactory quality level. It should be stressed that the quality of data collected in a given country affects the quality of data in the trading partner countries. Therefore, improving this quality in all EU countries is a priority for Eurostat.

In the article we used the author's methodology of assessing the quality of mirror data. These include data asymmetry indices, both proposed by Eurostat and authors' own proposals. We have also examined the changes in the above mentioned rankings over time.

The result of the survey in ordering of EU Member States according to quality of data on intra-Community trade in goods. The rankings are presented over the period of 2014–2017, during which there were 28 Member States of the EU. Changes in distinct countries' positions were shown as a result of changes in the overall quality of statistical data collected in these countries. The research methodology can be used in the process of monitoring the data quality of the Intrastat system.

Literature review

The topic of mirror data quality is the subject of statistical publications of national statistical offices or Eurostat. It is, however, rarely raised as the subject of scientific research. Foreign trade turnover is an important parameter, and it is used in various analyses, diagnoses and economic forecasts. It is, therefore, crucial to be able to estimate the true value of foreign trade within and outside the EU. The scientific literature more often presents the results of research concerning foreign trade itself, its size and dependence on various factors, and the quality of data is usually neglected.

For example, the goal of an article by Brodzicki *et al*. (2015) was to investigate the determinants of the intensity of Polish exports to its trade partners (country level). The analysis was carried out for 234 trade partners of Poland in the period 1999–2013 with the use of panel gravity modelling. The impact of standard determinants of gravity including partners size and distance on the dependent variable (level of exports) is highly statistically significant and in accordance with the general expectations. The impact of size similarity has not been proven. Adjacency has a robust and positive impact (EU membership).

The problem of discrepancies in mirror data has long been noted by researchers, and literature can be divided into theoretical works, especially concerning the modelling of an unknown, true structure and size of trade between countries, and application works concerning foreign trade research on specific countries or trade within groups of countries. Parniczky (1980)

indicates that such research was carried out at least since the 1920s, and Tsigas *et al*. (1992) date it back to the 1880s. However, all these authors acknowledge that a consensus among economists and statisticians on the need to investigate the mismatch between mirror data has been present since the 1960s. In his work (Parniczky, 1980), the author primarily points out that matrices of data on exports and imports are not as useful as it might seem. Although most practitioners have favoured the use of export information, he argues that the use of a modified import matrix is a better solution. Consequently, according to such a philosophy, in the study of discrepancies between mirror data, it is the importing side that should be given more credit.

The paper Tsigas *et al*. (1992) and other works by creators and users of the GTAP model (Purdue University) and (Ten Cate, 2014) are examples of theoretical articles. The article by Tsigas *et al*. (after Parniczky) points out the causes of discrepancies: the time of registration of transactions, different levels of interest of customs and public statistics, classification errors, transport and insurance costs, inclusion of transit. Another reasons are errors in determining the country of origin or shipment, changes in exchange rates during the reporting period or intentional actions, listed e.g. in the list of reasons for the occurrence of discrepancies in the paper by Hamanaka (2012), which is an example of application work. The authors of the next application study, Ferrantino and Wang (2008) use the measure of asymmetry, which is a slightly modified version of the measure being the basis of the aggregate index presented later in this paper (although they incorrectly describe the formula). These authors then use asymmetry measurement in (Ferrantino *et al*., 2012), among others, to detect evading customs declarations.

Since intra-EU trade statistics are based on statistical declarations of businesses, social and emotional factors in human activity should also be borne in mind (Baran & Markowicz, 2018b). These issues are considered within behavioural economics, initiated by Simon's concept of limited rationality (1972).

Behavioral effects make the data less reliable and more difficult to use, which is an important reason for searching for appropriate methods of data quality assessment.

Research methodology

The study used data on intra-Community supplies of goods from EU Member States and its mirror data on intra-Community acquisitions. Mirror data

(Baran & Markowicz, 2018a) for two countries A and B should be understood as follows: it is the amount of goods declared by a country A based business as dispatched from country A to country B, and acquisition of goods declared in country B as originated from country A (or goods acquired by country A based business from country B, declared in country A and goods declared as supplied from country B to country A, declared in country B).

Data from 2014–2017 were obtained from Eurostat's Comext database. The database is updated on an ongoing basis once the data have been collected by the national statistical offices. The analysed data were downloaded on 18 November 2018.

Examination of the quality of data concerning trade in goods between EU countries is possible thanks to the method of collecting these data. The information is derived from declarations made by entities engaged in intra-Community supplies of goods (ICS) or intra-Community acquisitions of goods (ICA). Data are transmitted from individual Member States to Eurostat. They constitute the Comext database, which then contains mirror data on transactions between all pairs of countries.

Data quality testing is based on an analysis of the differences between mirror data or asymmetries. Several authors (Eurostat, 2017a, 2017b; Javorsek, 2016; Ferrantino & Wang, 2008) recently propose using the following asymmetry index:

$$ {}_{O}W_{E}^{AU} = \frac{E_{AU} - I_{UA}}{K} \tag{1} $$

where:
E_{AU} – declared value of dispatches (supply) from country *A* to all other EU Member States combined,
I_{UA} – declared total value of acquisitions by all other EU Member States delivered from country *A* (mirror data),
$K = \frac{E_{AU} + I_{UA}}{2}$ or $K = I_{UA}$ or $K = E_{AU}$.

We call this index 'general'. What we propose instead is an approach using absolute differences between exports and mirror imports. Such an approach allows for cumulating of all discrepancies and avoids the balancing of the positive and negative differences. All discrepancies are thus taken into account. The authors' indicator is called the 'aggregated' data asymmetry index (mirror data quality index) and is written as:

$$ {}_{Z}W_{E}^{AU} = \frac{\sum_{i=1}^{n}\left|E_{AB_i} - I_{B_iA}\right|}{K} \tag{2}$$

where:
E_{AB_i} – declared value of dispatches (supply) from country *A* to country B_i,
I_{B_iA} – declared value of acquisitions by country B_i delivered from country *A* (mirror data),
$K = \sum_{i=1}^{n} \frac{E_{AB_i} + I_{B_iA}}{2}$.

The aggregate index takes values from the range from 0 (no difference between E_{AB_i} and I_{B_iA} for all countries B_i) to 2 (either E_{AB_i} or I_{B_iA} is equal to zero for all countries B_i, hence the numerator of the fraction is twice larger than the denominator). The higher its value, the lower the quality of the analysed data. A strength of the proposed index is that the positive differences are not balanced with the negative ones. It also enables comparisons between huge and small economies. The fact that a huge, imbalanced transaction can affect the index value, especially in small economies, could be considered a weakness of the proposed measure.

The survey was conducted for 28 EU countries in 2014–2017. The last country to join the EU was Croatia. It has been a member since July 2013, so 2014 is the first year with complete intra-Community trade data for the whole group of 28 Member States. For each country, an aggregated data discrepancy index (2) was calculated for the years 2014 to 2017. Countries were then ranked according to the value of the index in each year. This resulted in joint rankings of EU Member States according to the quality of ICS mirror data in 2014–2017. Such a compilation shows how the quality of each country's data has changed in relation to other countries.

Results

The positions of the EU Member States in the rankings by quality of intra-Community trade data for the period 2014–2017 are shown in Fig. 1. These positions change more or less over the period considered. First of all, attention should be paid to two countries: Cyprus and Malta. They invariably ranked in the last two positions in the ranking, which indicates the lowest quality of data on trade with other EU countries. Germany, Austria, Romania and France are among the top ranked countries. Another group of countries includes: Bulgaria, Belgium, the Netherlands, Spain, Great Britain, Italy, Hungary, Sweden, the Czech Republic, Poland and Slovakia. These countries changed their ranking positions in the years under analysis, and

the above list reflects the ordering from 2017 (positions from 5 to 15). The most significant changes in the rankings from 2014 to 2017 can be observed for Finland, who lost 10 positions (ranked 8th in 2014 and 18th in 2017) and Estonia (who gained 7 positions over this four-year period). This means that the quality of data as compared to other Member States has risen for one and fallen for the other of those two neighbouring countries.

The positions occupied by many individual countries changed in the analysed years. This is due to a similar level of data quality indicators (Fig. 2). The values of indicators for most countries were similar and in 2017 ranged from 0.055 (Germany) to 0.160 (Portugal). Slightly higher values of the indicator (approx. 0.2) were reached by Latvia and Croatia in 2014, and by Luxembourg and Croatia in 2017. On the other hand, outliers are observed in the case of Malta (an increase from 0.380 to 0.520) and Cyprus (a decrease from 1.041 to 0.578). Unfortunately, except for the latter there was no spectacular decrease of the index (or increase of data quality) among EU countries.

It should be added that the value of the difference in mirror data (declared ICS of the analysed country and declared ICA of its trade partners) usually depends on the value of the ICS of the country. Therefore, the discrepancy of data alone cannot be identified with the level of quality. Larger ICS may result in a higher data discrepancy, but the difference in mirror data may still be a small part of the ICS value. These considerations are illustrated in Fig. 3–4. It turns out that in 2017 there was a strong positive correlation between the difference between the mirror data and the ICS value for the EU–28 countries. On the other hand, there is no correlation between the data quality indicators and the value of the ICS. This confirms the usefulness of using indicators in analysing the quality of intra-Community trade data.

Rankings of EU countries by data quality in 2014–2017 were also created for the ICA (Fig. 2). Positions of many countries also changed in this case during the period under consideration. Again, the last two places were occupied by Malta and Cyprus. Ireland, Estonia and Lithuania were also in lower positions compared to most of the EU. Italy, Romania and Spain were on the other end of the spectrum, and occupied three out of four positions with the highest quality of data. They were joined by the Netherlands, which was promoted from 25th to second place (gained a record 23 positions). Latvia (gained 10 positions) and Croatia (gained 9 positions) also showed a significant improvement. On the other hand, Austria (8, 5, 9, 17, respectively) and the Czech Republic (down from 11 to 20) both saw a significant decrease.

Only in the case of the Netherlands, the change was not the result of revisions in a group of countries with good and comparable data quality. This country made a qualitative leap in the period under examination, reducing the value of the index five times (from 0.222 to 0.042). The index for Italy over the whole period did not exceed 0.050 and for Romania it never exceeded 0.060. Lithuania and Estonia were behind the main group of countries with similar data quality index values (about 0.18 in 2017). The last group consists of Ireland, Cyprus (an increase and then decrease to 0.243) and Malta, for which the quality of ICA data declined (an increase of the index value from 0.329 to 0.472).

Discussion

The authors have created a ranking of EU countries according to data quality for both ICS and ICA for 2017 in an earlier work (Markowicz & Baran, 2019). However, these two rankings don't perfectly fit each other which enabled testing the stability of our ranking over a period of constant reconciliation and updates of database. Fig. 7 presents ranking of EU Member States according to ICS in 2017 for provisional data downloaded in April, 2018 (on the left) compared to the ranking for the same period, but created with latest possible data available at the time of writing (database snapshot was from November, 2018). The overall sum of updated or changed figures exceeds EUR 24.3 billion which accounts for over a 0.75% of overall intra-Community trade in goods for the period. And it causes unexpected changes within the ranking positions of several countries (Bulgaria gains seven positions while Sweden loses 5, and Poland or Latvia lose 4 positions in the ranking). Same holds true for ICA reported in Comext database. All this leads to the conclusion that a researcher cannot use provisional data on trade without serious consequences like changing a huge part of the ranking as database improves. On the other hand, it might be interesting to observe how the data converge from a raw and incomplete state to their final and correct form. We could also examine whether the quality of data really grows in the process and asymmetries in mirror data vanish.

Unfortunately, comparing database figures stored at various points in time is not an easy task, since Eurostat does not disseminate versions of their database prior to updates. The only way to follow the whole process of reconciliation and convergence of the intra-Community trade database to its final form would be to download the whole of the Comext data affected after every update (nominally, the updates are performed once a month) in

bulk, store it locally, and compare with previous snapshots on a regular basis.

Conclusions

The ranking positions of many countries have changed in the analysed years mostly due to a similar level of data quality indicators. However, there is a group of Member States that reveal a constantly low level of data quality (Cyprus, Malta), as well as a group of countries with high data quality over the whole period — this group including old Member States like Germany, France, Italy, Spain, and Austria, but also Romania, a country that joined the EU quite recently. Calculated values of data quality index (2) are at similar level except for two countries of the lowest data quality. We tested whether the proposed index was sensitive to a country's overall trade turnover. Our study proves that there is no significant correlation between the trade in goods turnover and the data quality index.

Due to the importance of the data in economic analysis, the quality of the data is constantly monitored by national statistical offices and the Eurostat, the statistical services are working every day in order to improve data on trade. These efforts result in collecting new data on past transactions which fills gaps in the database and instantly improves data quality. In terms of our study the result was a significant change of ranking after seven months of complementing the database. Users of the Comext database should be aware of that limitation and should use recent data with caution.

Research on intra-Community trade raises specific problems. First of all, they concern the quality and reliability of the data. As indicated in the article, statistical data are burdened with various types of errors, causing easy-to-spot discrepancies. Nevertheless, identifying the actual reasons of such discrepancies in mirror data is worth further research.

As we stated in the above chapter, the database constantly improves with time as new data on past transactions are gathered. But this doesn't necessarily mean that the quality of data improves in the process as well. The quality of data will improve when new data emerges on the transactions previously recorded in only one of the countries. Updates in the database may concern new transactions, that have not yet been recorded at all. Consequently, discrepancies in data can also rise with time. Two or more snapshots of the whole database are needed to be compared for a thorough test of mirror data convergence. We plan performing such a task for individual years of 2017–2020 period in order to find a stable, global schema of such convergence, if it only exists.

References

Baran, P., & Markowicz, I. (2018a). Analysis of intra-community supply of goods shipped from Poland. *Socio-Economic Modelling and Forecasting, 1*.

Baran, P., & Markowicz, I. (2018b). Behavioral economics and rationality of certain economic activities. The case of intra-community supplies. In K. Nermend & M. Łatuszyńska (Eds.). *Problems, methods and tools in experimental and behavioral economics*. Cham: Springer.

Brodzicki, T., Śledziewska, K., Ciołek, D., & Umiński, S. (2015). Extended gravity model of Polish trade. Empirical analysis with panel data methods. *Working Paper. Institute for Development, 003*.

Eurostat (2017a). *Compilers guide on European statistics on international trade in goods, 2017 edition*. *Manuals and guidelines*. Luxembourg: Publications Office of the European Union. Retrieved from http://ec.europa.eu/eurostat (15.08.2018).

Eurostat (2017b). *National requirements for the Intrastat system. 2018 edition*. Luxembourg: Publications Office of the European Union. Retrieved from http://ec.europa.eu/eurostat (3.11.2018).

Ferrantino, M. J., Liu, X., & Wang, Z. (2012). Evasion behaviors of exporters and importers: evidence from the U.S.–China trade data discrepancy. *Journal of International Economics, 86*(1). doi: 10.1016/j.jinteco.2011.08.006.

Ferrantino, M. J., & Wang, Z. (2008). Accounting for discrepancies in bilateral trade: the case of China, Hong Kong, and the United States. *China Economic Review, 19*(3). doi: 10.1016/j.chieco.2008.02.002.

Hamanaka, S. (2012). Whose trade statistics are correct? Multiple mirror comparison techniques: a test of Cambodia. *Journal of Economic Policy Reform, 15*(1). doi: 10.1080/17487870.2012.657827.

Javorsek, M. (2016). Asymmetries in international merchandise trade statistics. *UN ESCAP Statistics Division Working Paper, 02*.

Markowicz, I., & Baran, P. (2019). Quality of trade in goods between European Union Member States. *Wiadomości Statystyczne, 1*(692).

Parniczky, G. (1980). On the inconsistency of world trade statistics. *International Statistical Review, 48*(1). doi: 10.2307/1402404.

Simon, H. A. (1972). Theories of bounded rationality. In C. B. McGuire & R. Radner (Eds.). *Decision and organization*. Amsterdam: North Holland Pub. Co.

ten Cate, A. (2014). The identification of reporting accuracies from mirror data. *Jahrbücher für Nationalökonomie und Statistik, 234*(1). doi: 10.1515/jbnst-2014-0106.

Tsigas, M. E., Hertel, T. W., & Binkley, J. K. (1992). Estimates of systematic reporting biases in trade statistics. *Economic Systems Research, 4*(4). doi: 10.1080/09535319200000028.

Annex

Table 1. Correlations between ICS and absolute asymmetries vs. correlations between ICS and index of mirror data quality over the period of 2014–2017

Year	Correlation between ICA and	
	absolute difference between ICA and ICS	aggregated index of data quality
2014	0.9680	-0.2886
2015	0.9662	-0.3193
2016	0.9598	-0.3295
2017	0.9684	-0.3742

Source: own calculations

Figure 1. Changing positions of EU Member States in ranking according to ICS mirror data quality over the period of 2014–2017

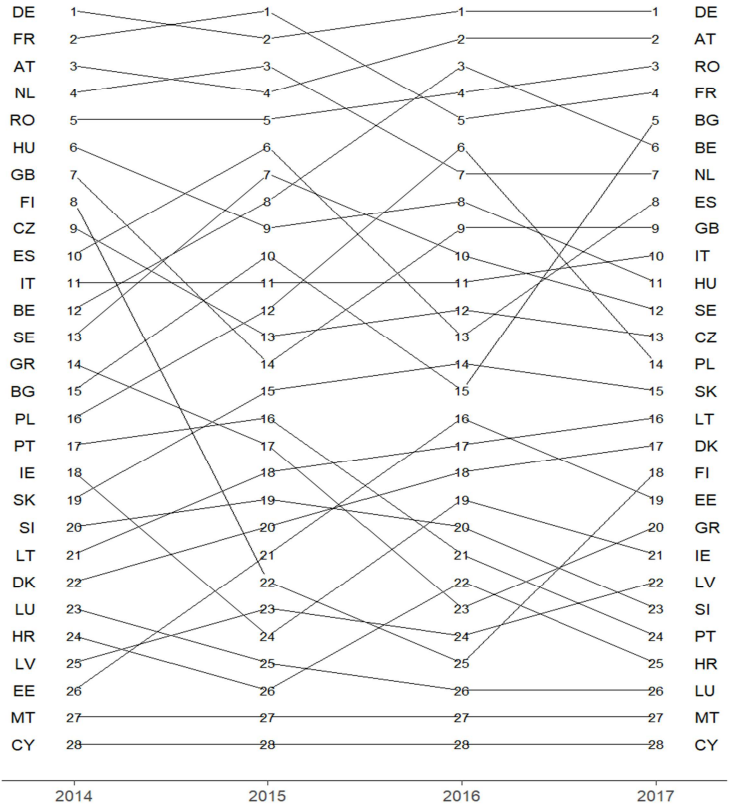

Figure 2. Values of ICS mirror data quality index for EU countries over the period of 2014–2017

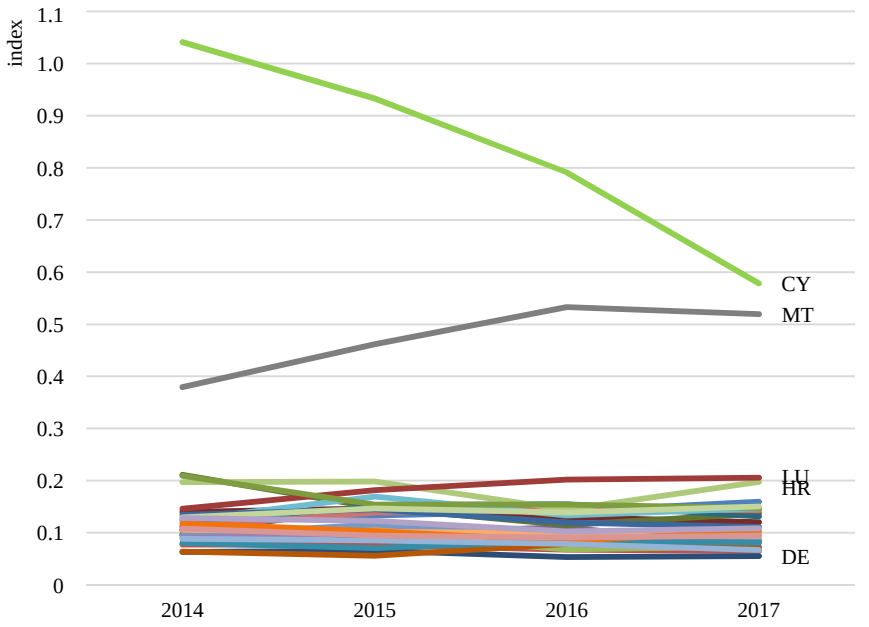

Figure 3. Differences between EU countries' mirror data combined vs. ICS in 2017 (in EUR)

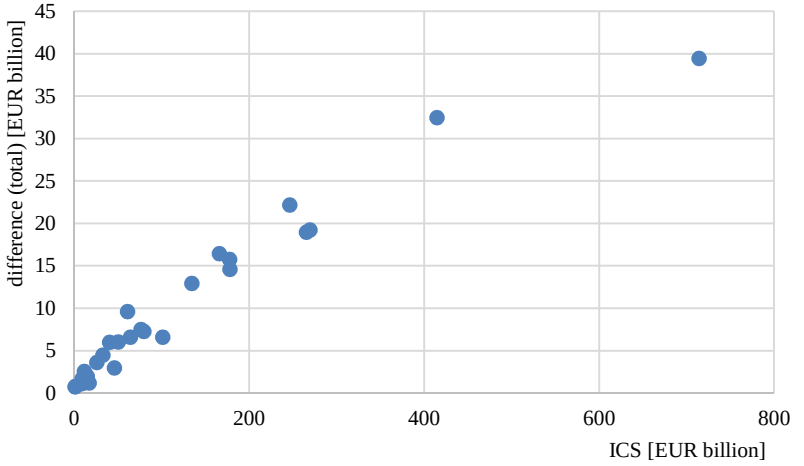

Figure 4. Mirror data quality index vs. ICS in EU countries in 2017 (in EUR)

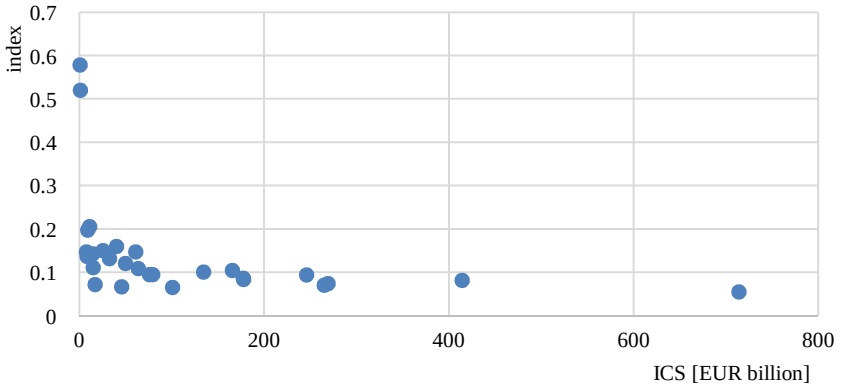

Figure 5. Changing positions of EU Member States in ranking according to ICA mirror data quality over the period of 2014–2017

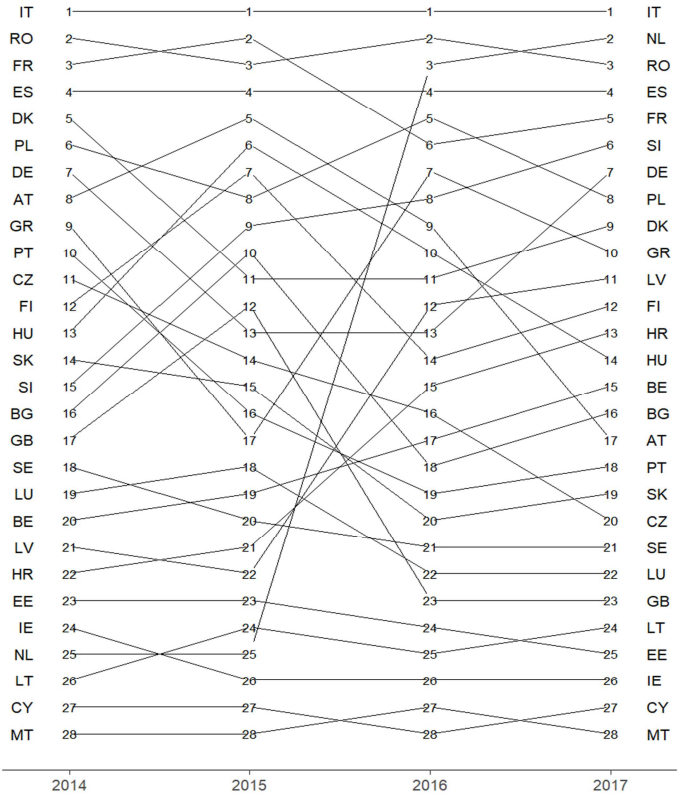

Figure 6. Values of ICA mirror data quality index for EU countries over the period of 2014–2017

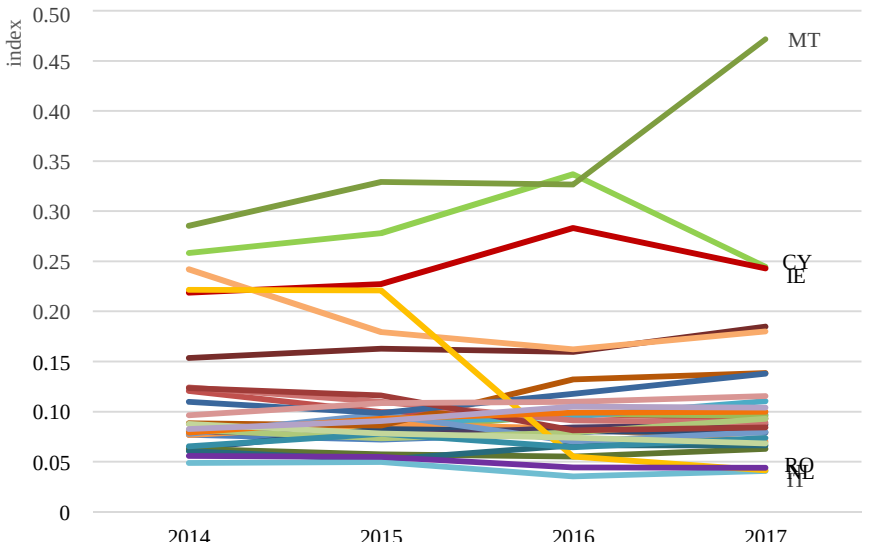

Figure 7. Comparison of rankings of EU countries according to mirror data quality in 2017 calculated with data downloaded in April 2018 and in November 2018