Selected papers from the 2016 Conference of European Statistics Stakeholders

SPECIAL ISSUE

2017 edition

European Central Bank Eurosystem
ESAC European Statistical Advisory Committee
KSH Hungarian Central Statistical Office
FENStatS
MST

STATISTICAL WORKING PAPERS

eurostat

Selected papers from the 2016 Conference of European Statistics Stakeholders

SPECIAL ISSUE

2017 edition

Printed in Luxembourg.

Manuscript completed in September 2017.

Luxembourg: Publications Office of the European Union, 2017

Print	ISBN 978-92-79-73629-2		doi:10.2785/186842	KS-TC-17-006-EN-C
PDF	ISBN 978-92-79-73628-5	ISSN 2315-0807	doi:10.2785/091435	KS-TC-17-006-EN-N

Table of content

Preface

EMANUELE BALDACCI[1], GYÖRGY BENOIST[2,3], CARSTEN BOLDSEN[4], MONIKA GALAMBOSNE TISZBERGER[2,5], JÁNOS GERENDÁS[6], MARTIN KARLBERG[1], ASTA MANNINEN[7], PER NYMAND-ANDERSEN[8], JEAN-MICHEL POGGI[9], GYÖRGY SÁNDOR[2], ZOLTÁN VERECZKEI[2]

The 2016 Conference of European Statistics Stakeholders

The 2016 Conference of European Statistics Stakeholders (CESS 2016) was held in Budapest at the seat of the Hungarian Academy of Sciences on 20 and 21 October 2016 (http://www.ksh.hu/cess2016). Bringing together around 400 European stakeholders, such as official statisticians, methodologists, private data producers/providers and data users, the two-day event was organised in collaboration between Eurostat, the Hungarian Central Statistical Office, the European Central Bank, the European Statistical Advisory Committee and the Federation of European National Statistical Societies, with the involvement of the Hungarian Statistical Association and the United Nations Economic Commission for Europe.

The conference provided participants with an opportunity to discuss user needs, to exchange good practices in statistical production and outreach , and to illustrate innovative ways of visualising and communicating statistics, and to promote new methodological ideas in data conception, collection and analysis.

In this special Statistical Working Paper (SWP) issue, selected papers from the conference, peer-reviewed by the SWP guest editorial committee (the authors of this preface) are presented, highlighting many of the current and future themes for ensuring good quality statistics fit for use.

The users of official statistics in focus

The conference programme (http://www.ksh.hu/cess2016/programme.html) was organised in seven topical tracks; one of those focused on various aspects related to the study of and interaction with the professional and public users of official statistics. This track was strongly linked to the project for Digital communication, User analytics and Innovative products (DIGICOM; http://ec.europa.eu/eurostat/web/ess/digicom), which is one of the projects of the ESS Vision 2020 portfolio; DIGICOM aims to modernise dissemination and integrate the communication function into European statistics.

One aspect of this track is the exploration of user needs, to see where these are still unmet, as exemplified

(1) Eurostat
(2) Hungarian Central Statistical Office
(3) Hungarian Statistical Association
(4) United Nations Economic Commission for Europe, Statistical Division
(5) University of Pécs
(6) Directorate Statistics, Magyar Nemzeti Bank
(7) European Statistical Advisory Committee
(8) European Central Bank
(9) Federation of European National Statistical Societies

by Manninen *et al*. (2017), who conclude that the demand for small area statistics (to allow for comparative urban and regional studies) is greater than current supply, and issue a number of recommendations regarding geo-referencing and integrating data sets, and involving users early on. Obviously, user involvement is easier to accommodate when data collection is carried out under a specific mandate, as is the case of the recently developed European Banking statistics described by Chionsini *et al*. (2017), but even in such situations, a balance has to be struck between different users (manifested in requirements for national specificities vs. the requirements for EU-wide harmonisation).

An efficient way to communicate official statistics to users is to render it visually with a narrative. As visual communication is largely non-verbal, it lends itself particularly well to digital sharing among practitioners. For this reason, a large team within the aforementioned DIGICOM project is working on shareable visualisation solutions. In this regard, the open source tool presented by Recchini and Capezzuoli (2017) could be a good example, as it appears to be easy to share among National Statistical Institutes (NSIs) and Central Banks (or any institution that performs data dissemination through a Representational State Transfer Application Programme Interface/REST API (as is the case for SDMX Web Services; https://data.europa.eu/euodp/en/data/dataset/estat-web-services).

To allow for a better understanding and use of official statistics, an enhancement of the general level of statistical literacy is required. Actions in this regard could take place at several levels, from educating the public at large (or large subsets thereof, such as primary/secondary school pupils), to more targeted segments, where actual ‘number-crunching’ capabilities are improved for key groups, such as university students with ‘quantitative methods’ modules in their curricula and to facilitate the use of statistics for policy advisers and think tanks. The paper by Dumičić and Žmuk (2017) provides a description of the untapped potential of Croatian enterprises (owing to their limited use of statistical methods). While the scope of their paper is wider than official statistics (‘industrial statistics’ or ‘marketing studies’ might be used to characterise many of the methods that might be of the most direct interest to the enterprises), the general improvement of what Dumičić and Žmuk refer to as *statistical culture* would surely have beneficial effects regarding ‘official statistics’ literacy as well.

Social scientists constitute a specific category of statistics users; they increasingly expect to have access to detailed source microdata for research purposes. As the level of detail increases, data owners worry about ‘spontaneous recognition’, the likelihood that a microdata user believes that he or she has accidentally identified one of the data subjects in the dataset, and may share that information. This concern, particularly in respect of microdata on businesses, leads to excessive restrictions on data use. In his SWP paper, Ritchie (2017) argues that spontaneous recognition presents no meaningful risk to confidentiality. This is becoming an important theme in the future as users have an increasing need for exploring the insights of micro-data, while preserving the confidentiality. This paper is undergoing final review for publishing within the ECB Statistics Paper Series (http://www.ecb.europa.eu/pub/research/statistics-papers/html/index.en.htm).

Statistics beyond borders

In today’s globalised world, the need arises for users to be able to compare the structures and performance of countries and to obtain statistics on cross-border phenomena. . At CESS 2016, one track covered both of these aspects.

The cross-border phenomena in scope of official statistics are rich and varied. This is illustrated by the wide span of papers in this SWP on that general topic; Comune *et al*. (2017) discuss the challenging in accurately estimating cross-border employment, Montvai (2017) tackles (foreign direct) investment, whereas Nilsson (2017) and Obrzut (2017) consider trade and balance of payments.

On the topic of comparability, Baldacci *et al*. (2017) re-emphasise the assertion of (Lynn, 2003) that cross-country statistics harmonisation is important to ensure that data used for evidence related to policy are not biased by important factors associated to unnecessary variations in the way statistics are produced, calling *inter alia* for assessment of the impact of methodological differences on comparability.

Reconciling data from multiple sources

Official statistics is moving away from reliance of sole survey data and administrative data are becoming increasingly used in statistical production. This leads Baldacci et al. (2017) to note that shift to multi-source statistics renders comparability more important than in the past; a change in the data source used could typically be expected to affect the statistics produced unless appropriately controlled for. Reconciliation of various official statistics data sources occur in many applications, for instance the reconciliation of Italian and Swiss data in the efforts to establish cross-border employment statistics discussed in Comune *et al*. (2017), and in the analysis by Obrzut (2017) of the consistency between European balance of payments statistics and the 'rest of the world' account of national accounts.

One track of CESS 2016 dealt with multi-source statistics; in the present SWP, this track is represented by de Bondt and Kosekova (2017), who investigate differences between gross value added and production (for measuring industrial output in the euro area), and by Capezzuoli and Recchini (2017) who present DataSTAT Hub, a tool for the automatic collection of administrative data and metadata to produce official statistics – one possible method for complementing existing data derived from traditional surveys with those from administrative sources.

Modelling

As noted by Baldacci *et al*. (2017), the need to explicitly use models for official statistics production is increasing (not the least due to the abundance of new, large and unstructured, data sources). To capture this trend, CESS 2016 had a track on statistical modelling; including two sessions on small area estimation – one important way in which small area statistics, as called for by Manninen et al. (2017) could be made available.

In this SWP, two very different modelling applications give an impression of the span of modelling that potentially could be applied in the context of official statistics: Koppány (2017) uses structural decomposition analysis to disentangle GDP growth, whereas Majewska (2017) studies life expectancy projection models.

Conclusions

The CESS 2016 conference successfully managed to bring the producers, users and methodologists together, and the papers of this special SWP issue illustrate the wide range of topics discussed.

The focus on users sets the CESS conference series from many other official statistics conferences, and at CESS 2016, the success of this emphasis was evident from the large number of sessions in the 'users' track, some of which are presented in this special SWP issue.

Conferences are normally one-off activities though there are multiple statistical themes to progress and build upon for the next instalment in the CESS conference series due in 2018 at the earliest. It is important that a continuous and sustained effort is maintained to exchange views between users and producers of statistics and to continue to preserve and safeguard the quality statistics as a public good. . In this regard, the DIGICOM project can be of pivotal importance, as it represents a major resource investment into the interaction with official statistics users throughout the rest of this decade.

References

Baldacci, E. Japec, L. and Stoop, I., 'Models of harmonisation: now and in the future' (Published in the present special Statistical Working Paper issue), 2017.

Capezzuoli, A. and Recchini, E., 'DataSTAT Hub: a tool for the automatic collection of administrative data and metadata to produce official statistics' (Published in the present special Statistical Working Paper issue), 2017.

Chionsini, G. Garcia, L. and Quagliariello, M., 'How far have we come on EU banking statistics? Are we there yet?' (Published in the present special Statistical Working Paper issue), 2017.

Comune, M. E., Ronconi, S. and Viviano, L., 'The challenge of permeable borders: estimating cross-border employment between Italy and Switzerland' (Published in the present special Statistical Working Paper issue), 2017.

De Bondt, G. J., and Kosekova, S., 'Measuring industrial output in the euro area: differences between gross value added and production' (Published in the present special Statistical Working Paper issue), 2017.

Dumičić, K. and Žmuk, B., 'Enhancing statistical culture – the unused potential of Croatian enterprises' (Published in the present special Statistical Working Paper issue), 2017.

Koppány, K., 'Estimating Contributions to GDP Growth by Structural Decomposition of Input-Output Tables' (Published in the present special Statistical Working Paper issue), 2017.

Lynn, P., 'Developing Quality Standards for Cross-national Survey Research: Five Approaches', *International Journal of Social Research Methodology*, Vol. 6, No. 4, 2003, pp. 323-336.

Majewska, J., 'An EU cross-country comparison study of life expectancy projection models' (Published in the present special Statistical Working Paper issue), 2017.

Manninen, A., Martin-Guzmán, A. and Bond, D., 'A growing demand for small area statistics' (Published in the present special Statistical Working Paper issue), 2017.

Montvai, B., 'Towards interpretable FDI data in external statistics: Filtering distortions arising from globalisation from data of multinational enterprises' (Published in the present special Statistical Working Paper issue), 2017.

Nilsson, L., 'Trade performance of the EU economies: Inter-country input-output tables as a necessary tool' (Published in the present special Statistical Working Paper issue), 2017.

Obrzut, R., 'Trade asymmetries and Consistency between National Accounts and Balance of Payments' (Published in the present special Statistical Working Paper issue), 2017.

Recchini, E. and Capezzuoli, A., 'An innovative webGIS system for dissemination and visualisation of official statistics and geospatial analysis' (Published in the present special Statistical Working Paper issue), 2017.

Ritchie, F., 'Spontaneous recognition: an unnecessary control on data access?' (Published in the present special Statistical Working Paper issue), 2017.

1 Models of harmonisation: now and in the future

EMANUELE BALDACCI[10], LILLI JAPEC[11], INEKE STOOP[12]

Table of Content

([10]) Eurostat, Luxembourg, Luxembourg; Emanuele.BALDACCI@ec.europa.eu
([11]) Statistics Sweden, Stockholm, Sweden; lilli.japec@scb.se
([12]) The Netherlands Institute for Social Research/SCP; i.stoop@scp.nl

1. Introduction

Ensuring that official statistics are fit for purpose is a critical quality requirement. One important dimension of quality is data comparability. Data should be comparable across time and countries to be used effectively. There are two main methods used in order to improve comparability of data between countries: input and output harmonisation. In this paper we will discuss these methods and how they fit into the changing survey landscape.

Cross-country comparability is becoming more important in the era of globalization. There is an increased demand for statistics that can be used to monitor policy implementation and assess its impacts both at a national and international level. Comparable statistics should ensure that no confounding factors linked to the way data have been collected, processed and disseminated get in the way of the cross-country assessment of policy effects or analysis of cross-country economic, social and environment trends.
The issue of cross-country comparability is particularly important for European statistics, which are increasingly used to monitor progress in society and economy, as it is the case with the macroeconomic and social profile of countries, fiscal governance and Europe 2020 structural indicators. Important decisions are made at European level based on the statistical evidence provided by official statistics across different domains.

While cross-country comparability is an important quality dimension, there are however other equally important dimensions of quality that affect the usability of European statistics. These include relevance, accuracy and reliability, timeliness, clarity and accessibility.

There are trade offs between these dimensions of quality and comparability, which have to be considered having in mind the use of the data and the users. Striking the right balance between the different quality dimensions is therefore a complex issue and statisticians need to aim at fit for purpose quality features in the statistics based on the users' needs. There are also trade offs to be made between national and international interests.

2. Sources of variability across countries

The European Statistics Code of Practice (Eurostat 2013) consists of the following principles regarding statistical outputs:

- *Relevance:* outputs, i.e., European Statistics meet the needs of users.
- *Accuracy and Reliability:* outputs accurately and reliably portray reality.
- *Timeliness and Punctuality:* outputs are released in a timely and punctual manner.
- *Coherence and Comparability:* outputs are consistent internally, over time and comparable between regions and countries; it is possible to combine and make joint use of related data from different sources.
- *Accessibility and Clarity:* outputs are presented in a clear and understandable form, released in a suitable and convenient manner, available and accessible on an impartial basis with supporting metadata and guidance.

The framework highlights different dimensions that are important in statistical production. Good data quality (accuracy and reliability dimension) can be achieved by applying sound methods and quality assurance in the statistical production process. The other dimensions of quality are also important but can be viewed as constraints (e.g., if a user needs data by a certain date and does not get it by then it will be of no consolation to that user that the data is very accurate).

Often there are trade offs between comparability and other quality dimensions, as well as national good practices. Since each step in the statistical production process can affect the final estimates, the design and execution of these steps at national level could affect the degree of comparability across countries in Europe.

Some steps are more crucial than others in order to gain comparability.

Social and cultural environments – concepts that have to be measured may be of different relevance or may manifest themselves differently in different social and cultural environments. What is considered to be sensitive varies between these environments. Harkness et al. (2003) give an example of the effect of cultural environment for the concept of religiosity. In some religious denominations, attendance at a place of worship is an essential element of religiosity, whereas for other religions rites at home are important. In this example, careful considerations must be taken when designing questions if comparing religiosity between countries is of interest. A question that asks about the frequency of church attendance, for example, would not alone provide a basis for comparing religiosity.

Data collection modes and mixes of modes – the mode that is used to collect data can affect statistical outputs. There are many examples in the literature of different mode effects, where statistical results can differ substantially depending on whether data collection is based on face to face interviews, telephone interviews, postal or web questionnaires (De Leeuw 2008).

Questions and questionnaires –the way questions are worded and the order in which they appear in the questionnaire are examples of design decisions that can affect statistics. Ensuring conceptual equivalence is important when designing questions for comparisons between countries. This is problematic when some concepts do not exist in all countries or the meaning varies across countries.

Translation – there are different methods for translating questions in international surveys and they can yield different results. Harkness et al. (2003) provide a good overview of these methods. One translation method is word for word translation. This method will definitely not solve the problem of ensuring conceptual equivalence, as a word for word translation will very often not be meaningful in the target language. Sometimes adaptation of questions, i.e., tailoring questions to better fit the target population, is necessary in order to ensure conceptual equivalence. Recently, more elaborate translation methods have been developed to ensure good translation quality. One such method is team translation.

Methodological and financial resources – differences in available skills and financial resources in different countries and organizations limit what is feasible to achieve in terms of survey design.

Unnecessary variation between countries due to these steps should be minimised. Even when a survey is designed to minimise unnecessary variation between countries, however, necessary variation may remain, e.g., in sampling frames, response rates, experience of interviewers, measurement error structures, prevalence of internet access, and the different languages in which the survey is fielded.

When administrative data is used additional issues arise. Administrative data are collected based on national regulations and the needs of the collecting agency which do not necessarily lead to comparable definitions across countries. Initial quality of administrative data sources may also vary. In the data processing phase, data imputation and integration rely on a variety of assumptions and methods which could lead to additional comparability issues. Finally, data estimation can rely on methods which are not necessarily harmonised across countries. Integration of different data sources to supplement survey data is becoming more common both in the production of statistical indicators and accounting frameworks. With an increased use of administrative and other data sources for the production of European statistics comparability of cross-country statistics is challenging.

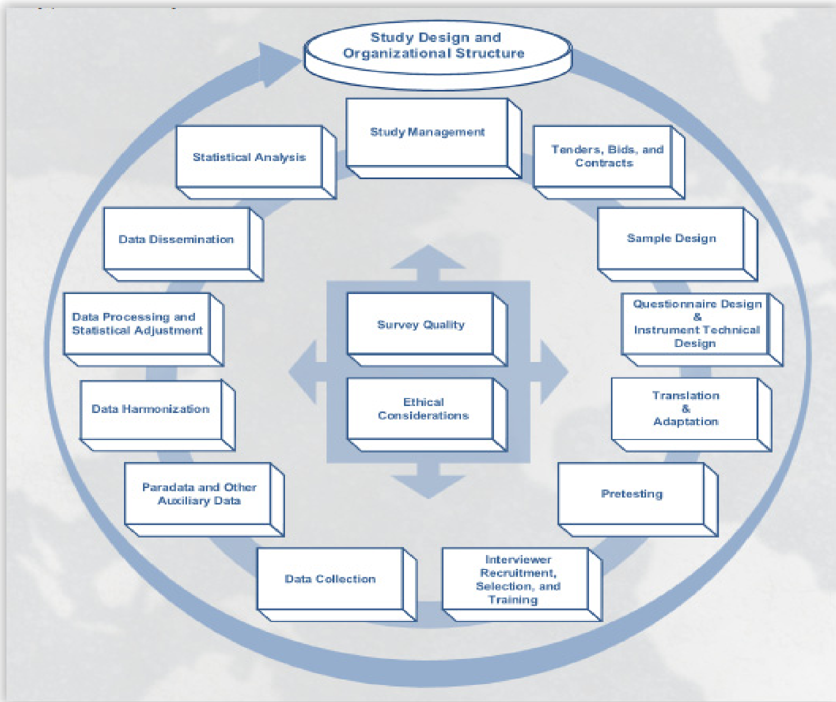
3. Harmonisation methods

Improving harmonisation of concepts and practices used at the different steps of the statistics production process can foster better comparability of European statistics across countries. Reference literature would typically distinguish between two main harmonisation concepts: input and output harmonisation (Körner and Meyer, 2005; Hoffmeyer-Zlotnik & Warner 2013).

Input harmonisation implies the use - at each step of the statistical process - of methods, sources, tools and procedures in order to produce comparable data at European level. This means that not only concepts, definitions and classifications are harmonised, but also that data is collected via the same tool in all countries (i.e., a survey with the same questions) and that the survey design, characteristics and methods are precisely specified. This approach requires a high degree of co-ordination and upfront harmonisation and could result in inflexibility. Some steps of such a harmonised process or even the whole process could be centrally implemented, which in principle could reduce its cost.

Figure 1 gives a conceptual representation of the different phases in comparative surveys. An application of the production process is given by the European Social Survey which is based on strict input harmonisation.

Figure 1: Survey Life Cycle (Survey Research Center (2016). Guidelines for Best Practice in Cross-Cultural Surveys)



The European Social Survey (http://www.europeansocialsurvey.org/) was first conceived in the 1990s, first went into the field in 2002 and seven rounds of data collection have been completed in 2016. Fieldwork has now been conducted in more than 30 countries, and 100 thousand individuals have registered as data users. The Core Scientific Team of the European Social Survey is directly involved in all phases of the survey, from specifying the national design (within the general design framework), testing and translating questions and training interviewers, monitoring fieldwork and coding, to delivering data. The project specifications of the European Social Survey prescribe in detail how the survey should be conducted.

Questionnaire and translation - The core questionnaire was developed based on proposals from experts in social research and survey methodology, and has remained largely identical over the years. In each round, two rotating modules are fielded. The first test looks at the form and structure of questions – its length, the number of complex words it contains, and so on – and then, based on what is known about the quality of items with similar features that have been fielded before, makes a prediction about how well the question will perform (Saris & Gallhofer 2007). A second approach is cognitive interviewing, using human subjects to help predict likely problems with questions (Willis 2005). Finally large-scale pilots are used where whole batches of items are tested on a sample large enough to include meaningful numbers of subgroups. Once the questionnaire has been finalized in English, annotations are added. These annotations do not form part of the final questionnaire but serve as a guide to translators. In the European Social Survey, the Translation, Review, Adjudication, Pretesting and Documentation (TRAPD) approach has been implemented, developed and promoted by Harkness and her colleagues (Harkness 2003; 2007). TRAPD consists of five interrelated procedures (Harkness, 2003), performed by three key agents: translator, reviewer and adjudicator. Reviewers not only need to have good translation skills but also need to be familiar with the principles of questionnaire design and the particular study design and the topics covered. They attend the review sessions and contribute to and guide the revisions. The adjudicator is the person responsible for the final decisions about which translation options to adopt. Pretesting may again result in modifications before the adjudicator signs off on the version for final fielding. Central in TRAPD is the team approach and the detailed documentation required. After translation formal characteristics of the translated questionnaire are compared to the source version using the Survey Quality Predictor (http://sqp.upf.edu/) and independent verifiers check the translations.

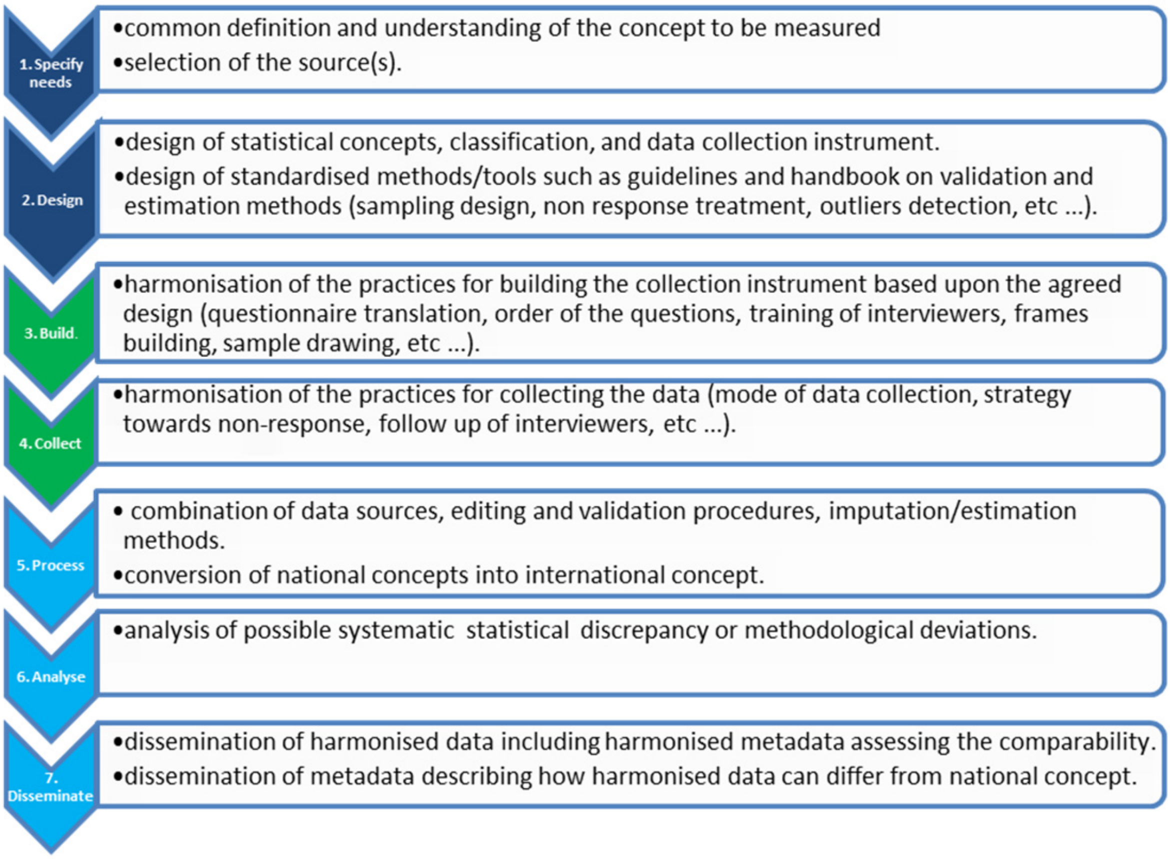
Data collection - The European Social Survey is representative of all persons aged 15 years and over (no upper age limit) resident within private households in each country. The sample is selected using strict random probability methods. Different sampling frames, designs and procedures can be used across countries (see Häder & Lynn 2007). The questionnaire is administered face-to-face. This mode allows the fielding of a rather long questionnaire results in a reasonably high response rate and is feasible in every country. Because of the high costs of face-to-face data collection a range of mixed mode pilots have been conducted. The result of these pilots was that moving to mixed modes would be detrimental to data quality, whereas the savings were very modest because of the relatively small sample size. Detailed contact forms are used to monitor fieldwork and control quality. The data from the contact forms are publicly available for research on nonresponse and data quality. Data from the contact forms can also be used to improve fieldwork (Stoop et al., 2010). After fieldwork, data are coded and sent for further harmonisation to the European Social Survey Data Archive at the Norwegian Centre for Research Data (Kolsrud et al., 2007). The archiving process is aimed at releasing the first integrated dataset around nine months after the first country has finished the fieldwork. After this first release data are accessible.

Such a high degree of input harmonisation is rare to find in the European Statistical System (ESS) reflecting cost considerations and national specificities. The early versions of the European Community Household Panel (ECHP) and the Land Use and Cover Area frame Survey (LUCAS) survey are examples of input harmonisation with the panel being implemented through a common questionnaire, harmonised definitions and sampling requirements and LUCAS being conducted centrally by Eurostat. LUCAS is a survey that covers a range of socioenvironmental aspects such as land take, soil degradation and biodiversity. Another example of input harmonisation in the European Statistical System is the Information and Communication Technologies (ICT) survey, which include several elements of comparability at the design stage.

In contrast to the input harmonisation, *output harmonisation* focuses on the final statistical product. Target variables are defined, however the actual questions may vary between countries. Granda (2010) gives an overview of the issues concerning output harmonisation and especially ex-post output harmonisation. In the ESS, pure output harmonisation is as rare as pure input harmonisation (Clemenceau & Museux 2008). For all surveys, at least the concepts and definitions of variables as well as the classifications to be used are subject to ESS agreements or legislation. Most European statistics, however, use a blend of input and output harmonisation, where some process steps are standardised and some flexibility is allowed for the others. This mainly reflects the need to take into account the different conditions and varying methodological approaches at national level and reflect the subsidiarity approach adopted by the ESS. The European Union Survey on Income and Living Conditions (EU-SILC) for instance, defines a harmonised list of target variables, definitions and concepts, classification, guidelines and recommendations, while the data collection process, including the selection of the appropriate data sources, is left to the NSIs. Structural Business Statistics are another example in the field of economic statistics where the choice of data sources is left to the member states based on commonly agreed methodologies, definitions and classifications.

Actual harmonisation practices in European statistics could be seen as a continuum of interventions along the statistical production process phases ranging from pure input to pure output harmonisation. When looking beyond surveys as the main data source for statistics production and considering alternative data sources such as big data and a mix of data sources, it becomes difficult to classify European statistics harmonisation practices according the input and output dichotomy. In figure 2 the Generic Statistical Business Process Model (GSBPM) has been used to map elements of harmonisation. The model encompasses other data sources as compared to figure 1 that deals with surveys.

Figure 2: Harmonisation through the different steps of the statistical production



Lynn (2003) distinguishes several strategies for cross-national comparability. It is crucial to establish whether cross-country comparability is the main aim, and the trade off with achieving the highest (national) quality. For example, in the "consistent quality" approach for survey data, response rates could be set in advance that are feasible in every country. This could result in response rates that are much lower than usual in several countries. To ensure optimal comparability, paper and pencil interviews (PAPI) could be prescribed and the use of register data be forbidden. This would again result in comparable data, but most likely in poorer national quality than could be achieved had computer-assisted personal interviews (CAPI) and the use of registers been allowed. On the other hand, the "maximum quality" approach aims at achieving national data with the best possible quality. This may harm comparability, e.g., when different survey modes are used. Other strategies usually constrain some quality aspects (e.g., sampling), but leave other design aspects to individual countries (e.g., the survey mode). Still other quality aspects would be formulated as targets rather than constraints (e.g., aim at a response rate of at least 60%). In practice most surveys strike a between these different dimensions.

4. Comparability in a new paradigm

Availability of new data sources reflecting more open access to administrative data and the data revolution, budget constraints and the need to lower the response burden have led producers to move to a new paradigm for statistical production (Citro, 2014). This is based on integrated use of multiple data sources (including new sources), adoption of innovative data analysis tools and methods, and a focus on statistical information services to respond to users (ESS Vision 2020). These changes are leading to the following transformations:

- Increasing use of multiple, integrated data sources.
- New modes of data collection (e.g., administrative registers, smart phones and smart meters).
- Model-based and algorithm-based estimates in the production of statistics.
- Move from pure statistical products to a range of on-demand data analysis and information services.

In the current production of demographic, social and economic statistics in Europe, administrative data play an important part, complementing and in part replacing survey data. Data imputation, estimation of information for sub populations, and missing variables are examples of statistical production processes that already rely on multiple sources and their integration through models to estimate statistical information of interest.

These changes imply that traditional harmonisation methods and practices (both on inputs and outputs) may fail to fully account for significant sources of unnecessary variation in European statistics and need to be complemented by additional efforts to achieve comparability. For example, the combination of data sources may add to variability of statistics across countries, as in the case with uneven quality of original sources when administrative data are used.

Differences in method-based estimations could also contribute to increase unnecessary variations across countries, as harmonised approaches may not be followed for the use of such tools. This calls for an augmented quality conceptual framework (building on existing process models such as GSBPM) to assess the impacts on comparability of the multi-source statistical production environment and more research to ensure adequate harmonisation.

However, not all changes in statistical production systems may harm the comparability of European statistics. For example, more integrated data based on multiple sources can help detect and assess cross-country discrepancies. Administrative data can complement survey data to better measure the tails of the households' income distribution, thereby lowering the risks that comparisons across countries are biased (assuming that administrative data exist for all European countries). And finally, model-based estimations based on larger data sets can lead to more powerful tests of hypothesis.

5. Conclusions

Cross-country statistics harmonisation is important to ensure that data used for evidence related to policy are not biased by important factors associated to unnecessary variations in the way statistics are produced (Lynn, 2003).

A variety of approaches is used in European statistics to ensure that this important quality dimension is fit for purpose, based on the use of the data and the users needs (Hoffmeyer-Zlotnik and Warner, 2013). Cross-country comparability is just one of the different dimensions of quality and as such trade offs could emerge between harmonisation and precision or harmonisation and timeliness, which need to be kept in mind.

Notwithstanding efforts to improve cross-country harmonisation practices, more work is needed to assess the degree to which differences across countries in various phases of the statistical business process have an impact on the usability of results. This includes reviewing harmonisation practices, assessing impacts of methods and tools used in different stages of statistical production and benchmarking surveys.

This focus on comparability is more relevant now than in the past, as official statistics is moving away from reliance of sole survey data and administrative data are becoming increasingly used in statistical production. In the future, new sources of data, including sensor data, geopositioned data, and Internet-based data will increasingly be used to produce statistics (Citro, 2014).

Furthermore, in this new data ecosystem the use of models for statistical production is expanding. The use of algorithms for statistics will be further required by the abundance of new large and unstructured data sources, such as big data. This calls for an assessment of the impacts of different methods and technical approaches to data collection and processing could have on the comparability dimension of quality and the trade offs with other quality dimensions (e.g., timeliness) based on users requirements.

References

Citro, C., 'From multiple modes for surveys to multiple data sources for estimates', *Journal of Official Statistics*, No 30, 3 September, 2014, pp. 381-442.

Clemenceau, A. and Museux, J.-M., 'Harmonisation of household surveys from an Official Statistics Perspective: experience gained from HBS, ECHP and EU-SILC', Paper presented at the 3MC Conference, Berlin, 2008.

De Leeuw, E. D., 'Choosing the Method of Data Collection', *International handbook of Survey Methodology*, NY: Taylor & Francis Group/Lawrence Erlbaum Associates, New York, 2008, pp.113-135.

Granda P, Wolf C, Hadorn R., 'Harmonizing survey data', *Survey methods in multinational, multicultural and multiregional contexts*, Hoboken, NJ: John Wiley & Sons, 2010, pp. 315–32.

Häder, S. & Lynn, P., 'How representative can a multi-nation survey be?', in R. Jowell, C. Roberts, R. Fitzgerald & G. Eva (eds): Measuring Attitudes Cross-Nationally. Lessons from the European Social Survey London, Sage, 2007, pp. 33-52.

Harkness, J. A., Questionnaire Translation, in J. A. Harkness, F. J. R. van de Vijver, & P. Ph Mohler (Eds.), Cross-cultural survey methods, New York: Wiley, 2003, pp. 35-56.

Harkness, J.A., Improving the Comparability of Translations, in: Roger Jowell, Caroline Roberts, Rory Fitzgerald & Gillian Eva (eds) Measuring Attitudes Cross-Nationally. Lessons from the European Social Survey, London, Sage, 2007, pp. 79-93.

Hoffmeyer-Zlotnik, J. & Warner, U., Harmonising Demographic and Socio-Economic Variables for Cross-National Comparative Survey Research, s.l.:Springer Netherlands, 2013

Körner, T., and Meyer, I., Harmonising Socio-Demographic Information in Household Surveys of Official Statistics: Experiences from the Federal Statistical Office Germany. In: Jürgen H.P. Hoffmeyer-Zlotnik and Janet A. Harkness (eds) Methodological Aspects in Cross-National Research, ZUMA-Nachrichten Spezial Band 11, 2005, pp. 149-162.

Kolsrud, K., Kallgraf Skjåk, K. and Henrichsen, B., Free and immediate access to data, In: R. Jowell, C. Roberts: Measuring Attitudes Cross-Nationally. Lessons from the European Social Survey, London, Sage, 2007, pp. 139-156.

Lynn, P., Developing Quality Standards for Cross-national Survey Research: Five Approaches. International Journal of Social Research Methodology, Vol. 6, No. 4, 2003, pp. 323-336.

Stoop, I., Billiet, J., Koch, A. and R. Fitzgerald, Improving Survey Response. Lessons Learned from the European Social Survey, Chichester, John Wiley & Sons. Ltd, 2010.

Survey Research Center, Guidelines for Best Practice in Cross-Cultural Surveys. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan. Retrieved December, 28, 2016, from http://www.ccsg.isr.umich.edu/.

2 Measuring industrial output in the euro area: differences between gross value added and production

GABE J. DE BONDT[13] AND STANIMIRA KOSEKOVA[14]

Table of Content

([13]) Output and Demand Division, European Central Bank, Frankfurt am Main, Germany
([14]) External Statistics and Sector Accounts Division, European Central Bank, Frankfurt am Main, Germany

Abstract

Measuring industrial output in volume terms in the euro area is not trivial, as differences in terms of growth between gross value added and production are not atypical and can last for several quarters in a row. Moreover, persistent level differences between the two measures might exist. The main factors that explain these differences are described: the treatment of prices and seasonality as well as the coverage of economic activity and the country coverage. The treatment of seasonal adjustment and economic activity coverage are more closely analysed. The relative importance of the various factors is, however, difficult to assess, in particular for a specific quarter. It is therefore recommendable that analysts and policy makers monitor both measures to assess the industrial cyclical development, keeping in mind that earlier released industrial production growth is not one-to-one translated into value added growth. Future work for statisticians is a further harmonisation between national accounts and short-term statistics as well as between national practices for seasonal adjustment.

JEL classification: C43, C49, E23.

Keywords: Industrial value added; Industrial production; Statistical measurement; National accounts; Euro area

Note: This paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the authors and do not necessarily reflect those of the ECB.

1. Introduction

Industrial gross value added and industrial production in volume terms are both informative measures of real industrial development. Analysts and policy makers typically focus on monthly industrial production in real time, because of a shorter reporting lag. At the time of the first release of quarterly industrial value added for a specific quarter, industrial production for all three months of that quarter is already published. Industrial production growth is, however, not one-to-one translated into industrial value added growth and even a sign reversal in the growth rates of the two measures is not atypical. In one fourth of all cases of negative euro area industrial production growth since 1999, value added growth has been positive. A recent striking example was the weakness in euro area industrial production excluding construction in the second quarter of 2016 (latest available observation) was not matched by weakness in corresponding value added: quarterly growth rates were -0.4% versus +0.5%. These findings were the main driver to start this comparative study. Although conceptually similar, there are a number of differences between the two measures, deserving a closer analysis.

This paper looks at the differences between these two industrial development measures, both, from a user as well as producer perspective. Industrial production is a short-term statistic that aims to estimate value added on a monthly basis in order to provide a timely measure of industrial output. In practice, however, it is difficult to collect value added data on a monthly basis, which means that the monthly change in industrial production is typically derived from other sources, including deflated turnover, physical production data, labour input and intermediate consumption of raw materials and energy. The gross value added (in general, not only for industry), on the other hand, is a quarterly national accounts indicator and is measured by subtracting intermediate consumption from output. It is one of the main indicators in national accounts and, together with taxes less subsidies on products, it comprises gross domestic product (GDP), which on the income side is equal to the sum of compensation of employees, taxes on production and imports less subsidies, and gross operating surplus. Industrial production therefore only partly describes the development of industrial value added in terms of volume over a longer period, as the link between industrial production and value added may be affected by changes in input ratios and structures of production. Another important factor for differences between the two measures relates to the (lack of) coverage of certain productive activities. For instance, certain adjustments are made in the system of national accounts for including the non-observed economy (Eurostat, 2016a) and not for the short-term production statistics. Gyomai and van de Ven (2014) report some country practices of the applied exhaustiveness adjustments in the national accounts for the non-observed economy, including a breakdown by industrial classifications in terms of GDP as well as in terms of gross value added of the industry. They show that the adjustments for the non-observed economy can be quite considerable, in particular as a percentage of the value added of the relevant activities. Noteworthy is also that the measurement of industrial production index typically, in practice, is applied at the product's level, whereas the value added is a more aggregated concept, for example at the level of the enterprises which can produce more than one product.

Notwithstanding the importance of correctly measuring industrial cyclical development for economic analysis, little literature exists that explicitly deals with this topic. To the best of our knowledge, our study is the first to look in depth into the differences in a consistent way for euro area countries. We are aware of a few academic studies that focus on the differences between industrial gross value added and production. Lucke and Weiß (2002) analyse European countries, Waldmann (1991) the US, Durand (1994) Canada, and Bhatia (1987) India. These studies hint that intermediate input/production and the price of oil or more generally the difference between factor price and input price (in the national accounts intermediate inputs are normally valued at purchasers' prices) might play an important role for the difference between the two measures of industrial development. Durand (1994) argues that the choice of price method is not a question of theoretical nicety: *"it really does change the whole historical record of industrial activity in Canada. It would likely also cause such differences in any industrialized country that went through the same energy and raw material price shocks and, more recently, rapidly declining computer prices."* International organisations also rarely publish on this topic. Exceptions are a United Nation study by Andreoni and Upadhyaya (2014), Bureau of Economic Analysis (2014), ECB (2004 and 2016) and OECD (2005). The Bureau of Economic Analysis is sceptical about gross output: *"Gross output by industry is an essential statistical tool needed to study and understand the interrelationships of the industries that underlie the overall economy. However, because of its duplicative nature, it may not be a good stand-alone indicator of the overall health of an industry or sector. What can one infer about the economic health of an* *industry solely from the fact that gross output increases? Nothing, without understanding what happened to the change in intermediate inputs and to value added. Did the increase in gross output simply reflect a change in the extent of outsourcing or could it reflect a more substantive, fundamental change in the economy? Gross output alone does not provide enough information to answer that question. Moreover, focusing solely on gross output is likely to exaggerate the cyclical-nature of the economy, particularly for sectors that are more sensitive to this cyclicality, like manufacturing."*

Our comparative analysis results in a number of findings. Movements in euro area industrial value added and production (excluding construction) differ in terms of absolute levels and quarterly growth rates. Quarterly growth rates in industrial production are not translated one-to-one in industrial value added growth rates, with the latter having a tendency of being slightly higher at the euro area level. Differences in the movements of the two measures also occur at the euro area country level, but to a varying degree. In addition to conceptual factors, a number of other factors contribute to the differences between the two measures, such as seasonal adjustment on an infra-annual basis, as value added is adjusted at a quarterly frequency and industrial production at a monthly frequency. Prices are treated differently in the two indicators. Gross value added is compiled using basic prices and does not include taxes (less subsidies) on products, whereas industrial production is at factor cost. Different economic activities are included in the two measures. A further source of difference between the two measures is that industrial production typically covers firms above a certain threshold (in terms of turnover or number of employees), with thresholds varying across countries. National accounts attempt to provide a more complete picture by using data from a variety of alternative sources. All in all, despite the close link between value added and industrial production, the differences between the two measures reflect all of the above-mentioned factors to some degree, although their relative importance is difficult to assess, in particular for a specific quarter. From an economic perspective, it is useful to monitor both measures to assess cyclical industrial development, keeping in mind that earlier available industrial production data might provide a misleading signal for industrial gross value added. Further harmonisation between national accounts and short-term statistics, as well as between national practices for seasonal adjustment, would help to reduce these differences.

The remainder of this paper is structured as follows. Section 2 reports the developments in industry value added and production for the aggregated euro area level as well as for euro area countries from a user perspective. Section 3 describes methodological differences between the two measures from a producer perspective. It provides not only a general overview, but also more detailed results about the impact of seasonal adjustment and the coverage of economic activity. Section 4 concludes, including future work on further harmonisation. Note that throughout the paper industry refers to industry excluding construction, unless stated otherwise.

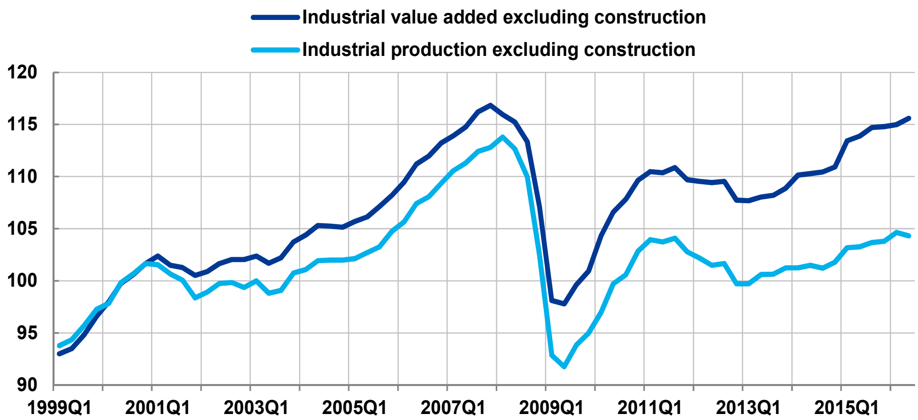
2. Developments in industry value added and production

2.1 Euro area

Movements in euro area industrial value added and production differ in terms of absolute levels and quarterly growth rates. Figure 1 plots both measures in level terms. It shows that for euro area industry excluding construction, the level of value added has, for the most part, been higher than that of industrial production since 1999.

Notwithstanding the observed level difference, both euro area measures tend to show similar cyclical movements in terms of quarter-on-quarter growth (see Figure 2). There have, however, been marked differences of up to 1.5 percentage points, positive or negative, in some quarters since 1999 and the positive or negative growth difference can last for several quarters in a row. The difference between the two growth rates has been 0.2 percentage point on average since 1999. Looking at growth differences in absolute terms, the average as well as standard deviation has been about 0.5 percentage points since 1999. This outcome implies that the differences in growth rates seen in recent quarters were in a range one could expect. Looking at the second quarter of 2016 (most recent available quarter), the weakness in industrial production was not matched by weakness in the corresponding value added (quarterly growth rates were -0.4% versus +0.5%).

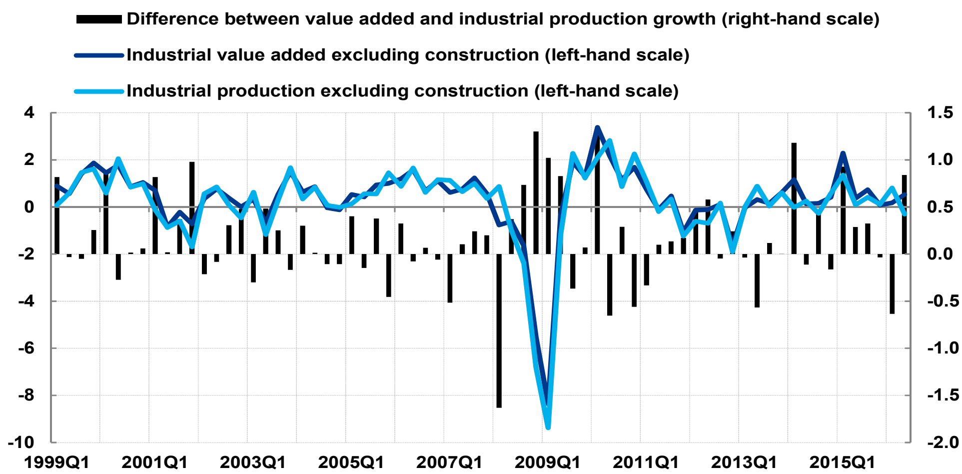
Figure 1: **Euro area industrial value added and production (excluding construction)**
(Index level, 2000=100)



Source: Eurostat and author calculations.

Note: gross value added: calendar and seasonally adjusted chain-linked volumes; industrial production index: quarterly average of working day and seasonally adjusted monthly data.

Figure 2: **Euro area industrial value added and production (excluding construction)**
(Quarter-on-quarter growth rate in %; difference in percentage points)



Source: Eurostat and author calculations.

Note: gross value added: calendar and seasonally adjusted chain-linked volumes; industrial production index: quarterly average of working day and seasonally adjusted monthly data.

2.2 Euro area countries

Differences in the movements of the two industry output measures also occur at the euro area country level, but to a varying degree (see Table 1). Among the four largest euro area countries, the difference between the quarter-on-quarter growth in industrial value added and industrial production since 1999 has been greatest for Spain and France (0.4 percentage point), but small for Germany (0.07). The average absolute growth difference varies between 0.5 for Italy and 1.0 pp. for Spain. For four euro area countries the average difference in growth over the same period has been negative, most markedly for Luxembourg (-0.7 percentage point). The average absolute difference in growth rates since 1999 has been larger than 3 percentage points in Slovakia, Ireland, Lithuania and Luxembourg. It should be borne in mind, however, that these results are also dependent on the period under investigation. For example, for Germany – where more historical data are available – the slight positive bias in growth rates since 1999 turned slightly negative if the observation period starts in 1991.

Table 1: Difference between industrial value added and production growth
(Quarter-on-quarter growth rate differences over 1999Q1-2016Q2)

	EA	BE	DE	EE	IE	GR	ES	FR	IT	CY	MT	LT	LU	LV	NL	AT	PT	SI	SK	FI
Average difference	0.17	-0.51	0.07	-0.19	0.33	0.10	0.41	0.39	0.20	0.10	-	0.12	-0.65	-0.35	0.03	-0.18	0.19	0.11	0.18	0.08
Standard deviation difference	0.50	1.85	1.08	2.80	8.24	2.37	1.44	0.82	0.66	1.32	-	4.70	4.60	1.84	2.05	1.43	1.79	1.20	5.37	1.95
Average absolute difference	0.39	1.49	0.66	0.63	3.91	1.92	0.97	0.63	0.51	1.04	-	3.80	3.40	1.38	1.65	1.13	1.31	0.87	3.90	1.49
Standard deviation absolute difference	0.36	1.20	0.86	0.68	7.24	1.38	1.14	0.65	0.45	0.82	-	2.73	3.13	1.25	1.20	0.89	1.23	0.83	3.67	1.25

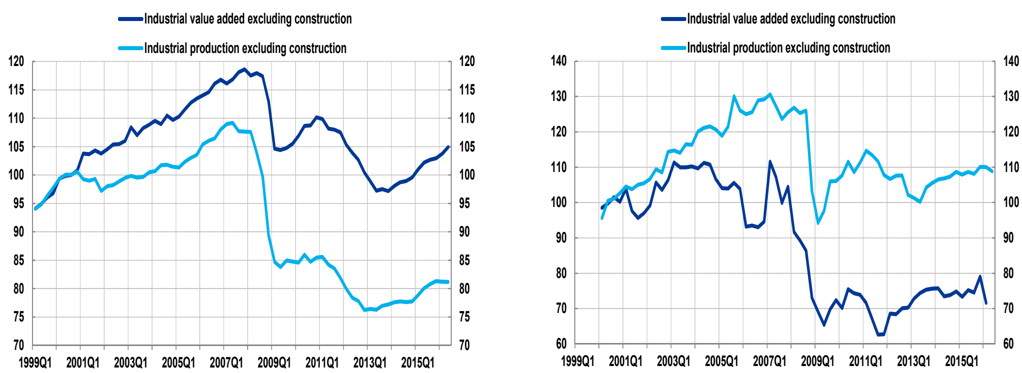
Source: Eurostat and author calculations.

Note: Both growth rates refer to excluding construction, are based on seasonally and calendar / working day adjusted data and are in real terms: chain linked volumes of gross value added and industrial production index. For Slovakia value added data are not working day adjusted. For Malta chain linked value added is not available. Industrial value added for Ireland and Luxembourg up to 2016Q1. Red / green is maximum / minimum value among euro area countries.

Figure 3 plots both industry measures in level terms for Spain and Luxembourg, the two countries where the average difference between the two series in growth rate term is on average most extreme (see bold in the first row of Table 1). In Spain value added has developed far more positively than production, whereas the opposite is the case in Luxembourg. It is quite difficult to explain these developments without having detailed knowledge on the country specificities and compilation practices. When looking at the national accounts data for output, intermediate consumption and the resulting value added for the industry, the Spanish intermediate consumption accounts for about 73% of the output while for Luxembourg it is about 71% which is also the average rate for the euro area; the remainder would be the value added relative to industrial output.[15] One explanation for the relatively lower industrial value added compared to production in Spain could be the automobile industry where intermediate inputs were mainly imported but not accounted for. As a result the volume estimates of the value added would be impacted (import prices may have different trend from the producer prices). In addition, the use of other information such as sales from the VAT registration, industrial turnover index, industrial price index, and production of natural gas and of refined petroleum products also contribute to the estimation of the final value added figure for Spain.[16] In Luxembourg the steel industry is important for the industrial sector (the produced steel is mainly exported) and contributes largely to the industrial value added. The largest steel producer in the world, i.e. ArcelorMittal, is headquartered in Luxembourg.

Figure 3: Country example in level terms: Spain and Luxembourg
(Index level, 2000=100)

Spain **Luxembourg**



Source: Eurostat and author calculations.

Note: gross value added: calendar and seasonally adjusted chain-linked volumes; industrial production index: quarterly average of working day and seasonally adjusted monthly data. Latest observation for value added in Luxembourg is 2016Q1.

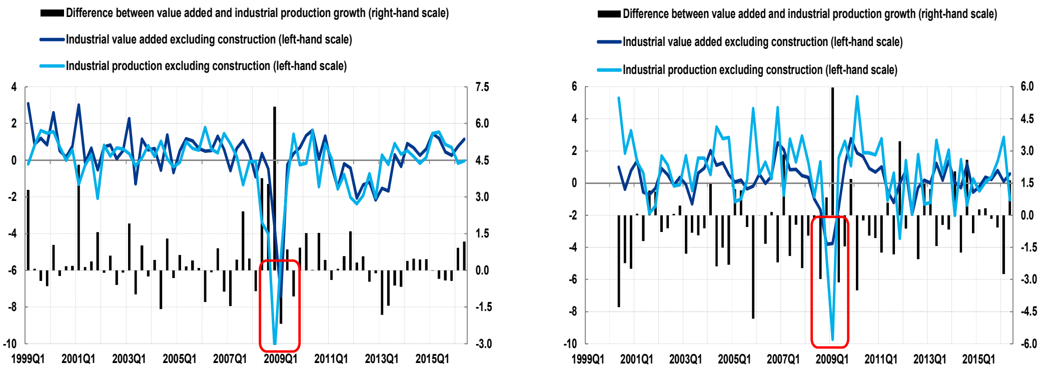
At the country level it is not always the case that both industrial development measures tend to show similar cyclical movements in terms of quarter-on-quarter growth. This finding is illustrated in Figure 4. Most striking is the difference around the Great Recession. In Spain the sharp drop in industrial production growth occurred 1 quarter before the sharp fall in value added (2008Q4 vs. 2009Q1). In Belgium, industrial production growth dropped much stronger than value added growth (-9.7% vs. -3.8%). Even outside the

[15] National accounts output and intermediate consumption are available only in current prices.
[16] Information provided on bilateral basis from Spain (INE).

Great Recession period a difference between the two growth rates of more than 2 percentage point (euro area range), positive or negative, has happened quite often in Spain as well as in Belgium. In Belgium industrial production growth is more volatile than value added growth. This finding relates to a relative Belgian specialisation in intermediate goods, one of the structural features that might explain the leading nature of the Belgian economy for the euro area (Vanhaelen et al., 2000).

Figure 4: Country example in growth rates: Spain and Belgium
(Quarter-on-quarter growth rate in %; difference in percentage points)

Spain **Belgium**



Source: Eurostat and author calculations.

Note: gross value added: calendar and seasonally adjusted chain-linked volumes; industrial production index: quarterly average of working day and seasonally adjusted monthly data.

3. Methodological differences

3.1 Overview

The observed differences between the two measures of industrial development raise the question which factors contribute to these differences and how they can be overcome. In addition to the earlier described different methodological treatment, various factors appear to play a role. Three factors relate to the conceptual treatment of prices, one measurement-related factor to seasonality and three measurement factors to differences in coverage in the euro area (at country or geographical coverage, economic activity or NACE Rev.2[17], as well as firm level). The difference between value added and industrial production probably reflects to some degree all of these factors, but their relative importance is difficult to assess, and it is difficult to determine for the current quarter how they might influence the value added data which are still to be released. Table 2 below provides an overview of these factors which are further explained.

1. Prices at which the two indicators are measured are different and thus are translated into different volume series. Value added is based on basic pricing, which excludes taxes less subsidies on products, whereas industrial production is at factor costs. The link between value added at the basic prices and factor cost are taxes less subsidies on production, which are not available in volume terms and this makes the quantifying of its impact difficult.
2. Another price related factor is the application of a double deflation approach for value added, i.e. deflated output less deflated intermediate consumption, whereas this is not the case for industrial production. The ESA2010 Manual and the Handbook on Prices and Volume[18] recommend the use of double deflation as "the theoretically correct method to calculate value added in volume terms, i.e. deflating separately the two flows of the production account (output and intermediate consumption) and calculating the balance of those two revalued flows". In the case this is not feasible due to data limitations or less reliable data source a “single indicator method” (output or input indicator) can be

[17] The term NACE is the statistical classification of economic activities in the European Union (EU). It is derived from the French “Nomenclature statistique des Activités économiques dans la Communauté Européenne”. Rev. 2 is referring to Revision 2 which was implemented in the official statistics during 2008-2011 (http://ec.europa.eu/eurostat/web/nace-rev2/transition).
[18] See http://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/-/KS-GQ-14-005

applied, as an alternative to the double deflation or the so called "double indicator method", directly to the value added in order to obtain the volume estimate. In such a case however an assumption on the same (input or output) price indices or volume indicators development is made which will impact the further analyses of the data when studying input/output ratios or productivity measures. In practice it is not possible to apply double deflation of quarterly data, unlike the annual national accounts, and most often temporal disaggregation techniques are used by applying synthetic indicators or indicator series (usually a combination of source data, including industrial production as seen below) of high-frequency data (e.g. monthly) that are deemed to be representative for the volume value added development. The resulting quarterly estimates must be aligned with the annual data which is based on the double deflation. For the quarterly industrial value added various data sources are commonly used by the countries:[19]

- NACE "B. Mining and quarrying": *Index of industrial production*; Quantity data; Expenditures on mineral exploration; Metres drilled
- NACE "C. Manufacturing": *Index of industrial production*; Turnover from VAT statistics or business surveys'; Censuses; Production information
- NACE "D. Electricity, gas, steam and air conditioning supply": Quantity data; Sales; *Index of industrial production*; Consumption of inputs
- NACE "E. Water supply; sewerage, waste management and remediation activities": Quantity data; Revenue; *Index of industrial production*; Turnover from VAT statistics or business surveys; Consumption of inputs

The industrial production index on the other hand is constructed as an output volume index which uses the value added of the corresponding industrial activities as weights.[20] Therefore at the lowest aggregation level, i.e. at the individual industrial branches, it is rather a measure of the development of the volume of gross output. Nevertheless, the industrial production index is theoretically designed to approximate the index of value added. Due to this property it is often used as a key input to the indicator series used in the estimation of the quarterly industrial value added.

3. For value added in volume terms, chain-linking is consistently applied by all euro area countries, whereas for industrial production so far only a few euro area countries use chain-linking.[21] Most countries apply a fixed-base approach for industrial production and as a result the euro area industrial production represents a combination of the two approaches. In both cases at present the index base year is 2010. The main difference to the users of the data is that while the chain linking provides more up-to-date weighting structure (i.e. the reference is to the previous year) or more realistic time developments in the most recent quarters, the fixed base indices refer to the weighting structure in the base period and thus tends to provide less accurate picture in the most recent end of the time series. Therefore, the chain linking is preferred over the use of a fixed base year for the index series.
4. Seasonal adjustment may play a role as value added is adjusted at the quarterly frequency and industrial production at the monthly frequency, which can lead to period-to-period differences for both series. Another factor related to the seasonal adjustment is the use of direct vs indirect seasonal adjustment when dealing with both indicators. Seasonal adjustment solely does not explain persistent differences over time.
5. The geographical coverage also differs, as volume value added data is not available for Malta, but the share of Malta – based on value added weights for industrial excl. construction in current prices for the year 2010 – is too small to be considered as significant (0.05%).
6. The NACE Rev.2 coverage differs. Quarterly value added includes NACE sector E (water supply; sewerage, waste management and remediation activities), whereas it is not included in industrial production. The share of this sector in industry value added excluding construction has since 2000 varied between 4.2% in 2007 and 5.0% in 2009 and is relatively small. Chain linked volume value added series of this sector are rather a-cyclical and only published at annual frequency.
7. The coverage at the firm level also differs. Industrial production typically covers firms above a certain

[19] See http://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/-/KS-GQ-13-004
[20] See http://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/-/KS-BG-02-002
[21] To our best knowledge, these are Estonia, Latvia, Netherlands, Slovenia and Finland.

threshold, in terms of turnover or number of employees, with thresholds varying across countries. Value added data aims at an exhaustive coverage, where other sources than those used for industrial production may be used. National accounts data attempt to provide a more complete picture by using data from a variety of alternative sources.

Table 2: Overview conceptual and measurement differences

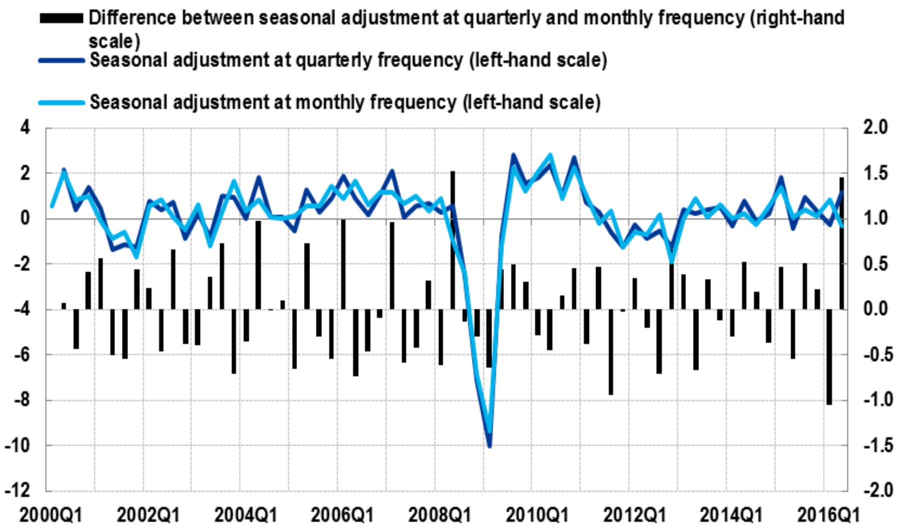
Value added (volume)	Industrial production (volume)
• Basic prices	• Factor costs
• Double deflation approach (deflated output less deflated intermediate consumption)	• Combination of source data is used; output index using value added as weights
• Chain-linking approaches	• Chain-linking and fixed base approaches
• Seasonally adjusted at quarterly frequency	• Seasonally adjusted at monthly frequency
• Euro area based on 18 countries, excl. Malta	• Euro area based on all 19 countries
• NACE Rev. 2: B to E	• NACE Rev. 2: B to D
• Aiming at exhaustive coverage (other sources may be used)	• Covers firms above a certain threshold (turnover or number of employees)

3.2 Seasonal adjustment

Figure 5 plots the estimated impact of the seasonal adjustment when done at monthly and at quarterly frequency for industrial production. In order to quantify this factor industrial production data was seasonally adjusted across euro area countries at a quarterly frequency. The same models were used as much as possible. It appeared that quarterly growth rates can differ substantially depending on whether the seasonal adjustment initially applied to monthly or quarterly time series. Using data that are seasonally adjusted on a quarterly rather than a monthly basis, the average absolute difference between the growth rates of euro area industrial production is shown to have been 0.5 percentage points since 2000. On average, however, the impact of the other factors remains sizeable. It is important to note that the results were dependant on the parameters applied for the seasonal adjustment and more concretely the seasonal filters.

One could also see that rarely there is a quarterly pattern in the differences between the quarterly and monthly seasonal adjustment, while an expectation could be that the same quarters are either under- or overestimated and only the magnitude of the differences vary.

Figure 5: Impact of seasonal adjustment on industrial production
(Quarter-on-quarter growth rate in %; difference in percentage points)



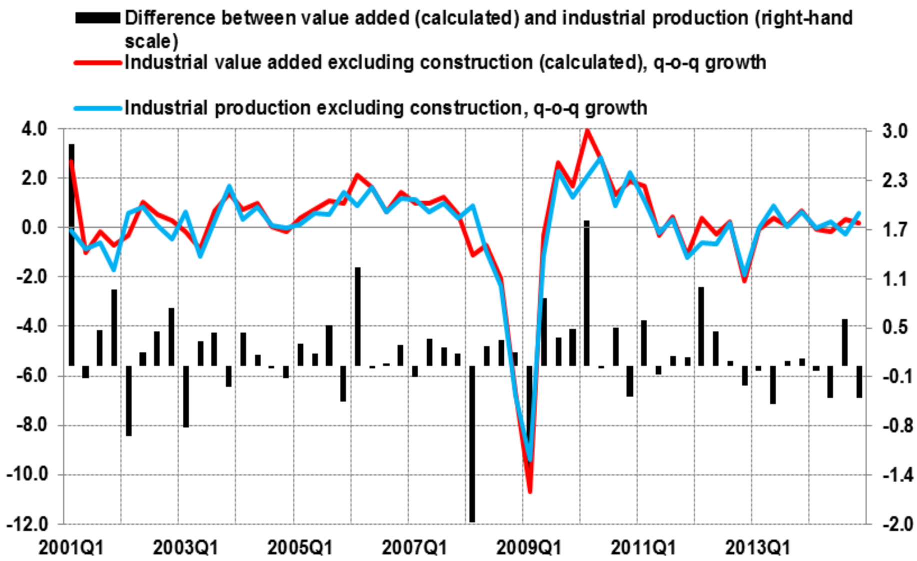
Source: Eurostat and author calculations.

Note: Working day and seasonally adjusted quarterly data were derived for this analysis; non-published data.

3.3 Economic activity coverage

Figure 6 plots the estimated impact of the different coverage in the NACE Rev.2 activity sector between industrial value added and industrial production. The estimated impact was expected to be rather limited given the small share of sector E (about 5% on average for the euro area) in the industrial value added. The precise analysis is, however, hampered by the lack of quarterly national accounts data for NACE Rev.2 E. The growth rate of the calculated industrial value added excluding sector E is generally very close to the officially published data of the industrial sector (NACE "B to E"). Nevertheless, for specific quarters, the impact of Section E is found to be sizeable, i.e. up to 2.2 percentage points during the Great Recession and up to 0.8 percentage points during "normal" times. At the same time, these results confirm the similarly diverging developments of the published value added from industrial production.

Figure 6: Coverage of Sector E
(Quarter-on-quarter growth rate in %; difference in percentage points)



Source: Eurostat and author calculations.

Note: Calculated industrial value added was derived for this analysis by excluding the annual structure of NACE "E" form the quarterly NACE B to E aggregate; non-published data.

4. Conclusions

To the best of our knowledge, this is the first time that differences between industrial gross value added and production are addressed in detail for the euro area and euro area countries. The presented results clearly show that earlier available industrial production data might provide a misleading signal for industrial value added. This is illustrated by quarter-on-quarter growth in 2016Q2: -0.4% for industrial production and +0.5% for value added. The average absolute difference between the two quarterly growth rates has been a sizeable 0.4 percentage points since 1999. This finding implies that the information content of industrial production for conjunctural analysis (read: for industrial value added and thus the industrial contribution to real GDP) is less than usually assumed.

Our work aims to promote future research on this topic and also encourages other users to contribute by addressing similar issues. This paper not only sheds light on in this highly policy-relevant area, but it also lists opportunities for harmonisation. Further improvements in measuring the industrial cyclical development in the euro area are indispensable, because analysts, policy makers and academics monitor and analyse it. Further harmonisation between National Accounts and short-term statistics, i.e. the same activity coverage in NACE available at monthly and quarterly frequency, is required. Improvements in the harmonisation of national practices for seasonal adjustment would also need to be addressed. A common approach to volume data, i.e. the application of chain-linking, is another field for future work. Finally, users will benefit from more information and in principle of more general availability of metadata.

Acknowledgements

The authors would like to acknowledge the contributions on seasonal adjustment of Sylwia Grudkowska. We are grateful for the comments from Heinz Dieden, Michal Doliak, Martin Eiglsperger, Matthias Mohr, Maria Svagrovska, and the European System of Central Banks (ESCB) Working Group on General Economic Statistics and participants of the CESS 2016 conference. The work has also benefited from feedback from Neale Kennedy, Hans-Joachim Klöckers, Henrik Schwartzlose and Caroline Willeke (all ECB). We highly appreciate the helpful comments and suggestions from an anonymous referee. Any errors or omissions are exclusively the responsibility of the authors.

References

Andreoni, A. and Upadhyaya, S., 'Growth and Distribution Pattern of the World Manufacturing Output: A Statistical Profile', *United Nations Industrial Development Organization Working Paper*, 02/2014.

Bhatia, D.P., 'Estimates of Value Added at Constant Prices: An Erroneous Methodology', *Weekly*, Vol. 22(16), April 18 1987, pp. 721-723.

Bureau of Economic Analysis, 'What is Gross Output by Industry and How does it Differ from Gross Domestic Product by Industry?', 22 April 2014, http://www.bea.gov/faq/index.cfm?faq_id=1034

Durand, 'Alternative Estimates of Real Value Added by Industry for Canada', *Economic Systems Research*, Vol. 8(3), 1944, pp. 225-234.

European Central Bank, 'Differences between Industrial Production and Value Added Data in Industry in the First Quarter of 2004', *Monthly Bulletin*, August, 2004, pp. 40-42.

European Central Bank, 'A Closer Look at Differences between Industrial Gross Value Added and Industrial Production', *Economic Bulletin*, Issue 1, 2016, pp. 19-22.

Eurostat, 'Methodology of short-term business statistics', 2003.

Eurostat, 'Handbook on quarterly national accounts', 2013.

Eurostat, 'Building the System of National Accounts - non-observed sector', last modified 20 September 2016, http://ec.europa.eu/eurostat/statistics-explained/index.php/Building_the_System_of_National_Accounts_-_non-observed_sector

Eurostat, 'Handbook on prices and volume measures in national accounts', 2016b.

György Gyomai and Peter van de Ven, "The Non-Observed Economy in the System of National Accounts", *Organisation for Economic Co-operation and Development Statistics Brief*, No.18, June 2014 https://www.oecd.org/std/na/Statistics%20Brief%2018.pdf

Lucke, D. and Weiß, J.-P., 'International Comparison of Industrial Development in the European Context – The Problems', *Economic Bulletin*, Vol. 39(7), German Institute for Economic Research, 2002, pp. 215-220.

Organisation for Economic Co-operation and Development, 'Overview of Industrial Statistics at the OECD', 7 April 2005, http://www.oecd.org/sti/ind/overviewofindustrialstatisticsattheoecd.htm.

Vanhaelen, J.-J., Dresse, L. and de Mulder, J., 'The Belgian Industrial Confidence Indicator: Leading Indicator of Economic Activity in the Euro Area?', *National Bank of Belgium Working Paper*, 12, 2000.

Waldmann, R.J., 'Implausible Results or Implausible Data? Anomalies in the Construction of Value-Added Data and Implications for Estimates of Price-Cost Markups', *Journal of Political Economy*, Vol. 99(6),1991, pp. 1315-1328.

3 DataSTAT Hub: a tool for the automatic collection of administrative data and metadata to produce official statistics

ALESSANDRO CAPEZZUOLI[22], EMANUELA RECCHINI

Table of Content

([22]) Italian National Institute of Statistics (Istat).
The views expressed in this paper are solely those of the authors and do not involve the responsibility of Istat. Addresses for correspondence: alessandro.capezzuoli@istat.it and emanuela.recchini@istat.it.

1. Introduction

The need for relevant, reliable and even more timely statistical data to support decision making process and scientific research has contributed to a growing demand for new statistical information to best analyze and rule, at various levels, the deep social, economic and environmental changes occurred at regional and global scale. Official statistics, characterized by the highest quality possible inasmuch as they are produced in compliance with the United Nations Fundamental Principles of Official Statistics and the European Statistics Code of Practice, are best suited to meet this need.

There is worldwide recognition of the increasing role played by administrative data and metadata in the production of prompt and more disaggregated statistics at higher frequencies than traditional survey data. They offer further information on a wide range of issues, including some which cannot be answered cost-effectively from survey data.

The efficient use of all available information to produce timely, accurate and high quality statistics is a challenge for National Statistical Offices (NSOs), which are even more committed to developing methods and suitable tools for the production, collection, standardization and integration of different types of statistical data. The bringing together of information from different sources makes it possible to fill information gaps or provide insights which cannot be gleaned from unlinked data and to improve the knowledge and understanding about specific phenomena.

In this chapter, a short review on the collection of administrative data and metadata for statistical purposes, with particular reference to Istat experience, is given. Then, some critical issues and possible technological solutions related to data collection processes are considered. including what is offered by the Web and is at the basis of DataSTAT Hub, a tool for collection, and release of data and metadata.In this respect, an overview of technology, model and architecture used to create DataSTAT Hub permits to describe its main features. A specific focus is put on an example of application of this hub to classifications of official statistics.

2. Data collection and official statistics

2.1. Data collection process

Accurate data collection is essential in all fields of study, from social and physical sciences to economy, politics, environment, etc.: it permits to answer research questions, validate hypotheses and evaluate outcomes.

The production of statistics based on administrative data and metadata from different sources is closely related to the methods and techniques of collection, integration and management of archives.

Problems related to automatic data collection are numerous as they involve the production and standardization of outputs, in order to: make them usable by web applications; be stored in a database; be connected (record linkage); be processed by statistical software and/or visualized within *ad hoc* created web platforms.

In this respect, it is worth keeping in mind that in a data collection process both the data collector and data supplier have specific tasks; the main ones are summarized in the box below.

Main tasks of data collector and data supplier

DATA COLLECTOR	DATA SUPPLIER
• manages data requests	• receives data requests
• defines methods and standards	• elaborates data requests
• monitors and handles reminders	• prepares data to be sent
• stores data and metadata	• sends data to data collector
• standardizes and disseminates	

In order to ensure compliance with the United Nations Fundamental Principles of Official Statistics and the European Statistics Code of Practice, particularly as far as the quality and reliability of the data collected by NSOs for statistical purposes is concerned, some aspects related to the different sources of data have to be considered. Data suppliers may provide:

- administrative registers
- synthetic indicators
- metadata
- Big Data

Administrative registers contain data collected by data suppliers with different criteria that often data collectors do not know.

In order to meet quality standards for the data, NSOs change an administrative register into a statistical register, through check, processing, standardization and normalization algorithms. This process ensures the usability of data for statistical purposes and the production of synthetic indicators also by linking records from different sources with each other.

Among the most recent activities carried out by NSOs, Big Data collection plays a key role in producing statistics and monitoring numerous phenomena. Big Data can be collected through two main channels: data supply from institutional and non-institutional agencies (Agency of revenue, telephone companies, Police Headquarters, etc.) and data scraping, namely a technique to extract data published on websites (e.g. online booking sites, advertising sites, etc.) through data capture and storage procedures.

Each data collection process presents critical issues like handling data requests and reminders, complex IT infrastructure, a burden for data supplier, human resources for transactions management. Among the different solutions, there is the possibility to collect data through File Transfer Protocol or upload data through an *ad hoc* website to handle reminders and data supply requests. However, these solutions do not permit to automate the process.

Notably in the case of Big Data, IT architectures used for data collection just mentioned above are not very suitable to store large volumes of data and interface with applications for data mining and/or data analysis. Solutions requiring data standardization (e.g. SDMX, JSON-Stat, etc.) can be very onerous for data suppliers, since they use structured databases based in most cases on *ad hoc* schemes.

The World Wide Web offers a possible solution. In fact, HTTP (Hypertext Transfer Protocol), i.e. the set of rules for transferring files on the Web, can be conveniently used for data collection and data exchange. This will be more fully explained in paragraph 1.3.2.

2.2. Administrative data collected by Istat

At present, the exploitation of administrative data and metadata for statistical purposes is a normal practice for a large number of NSOs. This improves the quality of statistical outputs, eliminates process redundancies, reduces the statistical burden on respondents and minimizes costs.

According to the provisions of the Italian Digital Administration Code([23]), before proceeding to the collection of new data, public administrations are required to verify whether the information they need can be acquired through access to information already in the possession of other public authorities or public bodies.

Technical options for data usability

- web access through the website of the supplier institution or an *ad hoc* thematic website
- **interoperability among public administrations for data collection and data integration**
- the user can process data collected exclusively for the pursuit of its institutional goals; data transfer from one information system to another does not change data ownership
- the transfer of a data from an information system to another does not change the ownership of the given

(Guidelines for the drafting of conventions on the usability Public Administrations data; Legislative Decree n. 82/2005, commonly referred to as the "Digital Administration Code", modified by the Legislative Decree n. 235/2010)

The Italian National Institute of Statistics (Istat) collects and manages a large volume of administrative data and metadata from different sources, among which: Italian Agency of Revenue; Bank of Italy; Ministries; Social Security and government Institutions. From 2009 to 2016, administrative data sets supplied to Istat have trebled. In this respect, the need to engineer processes in order to automate and manage collection and release of administrative data and metadata is increasingly urgent. Data collected by Istat are very different from each other in type, content and structure. Administrative registers may include statistical data, micro data, geographic data, synthetic indicators and many other types of data and metadata.

The main channel used by Istat for the collection of administrative registers from public and private institutions is the ARCAM website([24]). Through this portal it is possible to manage data requests, remainders and upload files containing administrative data and metadata. ARCAM meets the need of managing centralized repositories for data collection in compliance with legislation on the confidentiality of the data.

Another organizational and technological environment to share, integrate and disseminate data and metadata within the Italian National Statistical System (Sistan) is Sistan Hub([25]). This informative system, in line with the objectives of the Digital Agenda, supports the semantic interoperability among different Institutions, facilitates data searching and uses SDMX international standard as main requirement for their representation. Although it adapts to the different needs of NSOs, Sistan Hub requires an investment in human resources for data mapping and creation of the single exit point and this solution is not very suitable for Big Data management and for data suppliers who not always have a thorough understanding of SDMX standard.

The imposition of standards and the creation of IT infrastructures for data-exchange would be very onerous in terms of costs and time of execution for data suppliers.

([23]) For further details, see the Italian Legislative Decree no. 82/2005, subsequently integrated and amended by Supplementary Provisions and the corrective Legislative Decree no. 159/2006.
([24]) For further details on ARCAM, see: https://arcam.istat.it/arcam/
([25]) For further details on Sistan Hub, see: http://sistanhub.istat.it/hub

3. DataSTAT Hub for automatic collection and release of data and metadata

3.1. What is DataSTAT Hub?

DataSTAT Hub is a suitable and easy tool for collection, standardization, integration and release of data and metadata. It automates these processes through HTTP and a model that simplifies the structure of data.

By taking advantage of the potential offered by HTTP, DataSTAT Hub can be used through two different architectures: star or centralized. The former implies that each microservice (hub node) is automatically populated by data supplier through a set of query strings and can be accessed in reading by the central institution that performs data collection. The latter architecture implies the automatic population of the central hub that interfaces with the various institutions through the just mentioned query strings and allows the data collector to store data and metadata for example in a NoSQL database (Cassandra) using the key-value data model for their representation.

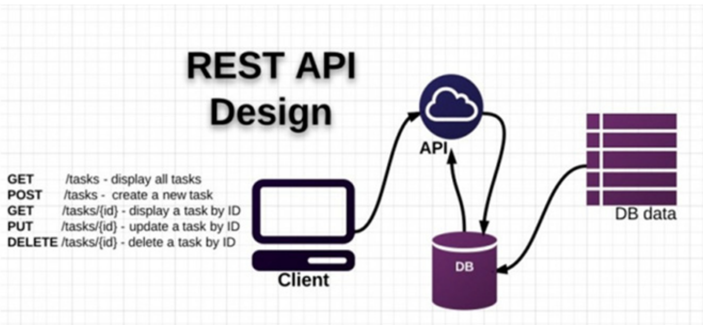
DataSTAT Hub permits to standardize outputs in various formats (XML, JSON, CSV) and models (JSON-stat, SDMX, DDI).

3.2. DataSTAT Hub technology

As mentioned above, HTTP can be conveniently used for data collection and exchange. It is a request/response protocol based on the client-server architecture.

A set of guidelines to use HTTP is defined by REST (Representational State Transfer), an architecture style for designing networked applications. REST architecture permits to separate relational databases from the client through an API (Application Programming Interface), which exploits HTTP to transmit data and exchange information (Figure 1).

Figure 1: REST API architecture



REST allows data supplier to perform CRUD operations(Create, Read, Update and Delete) with a logic similar to that used on any SQL database (Figure 2).

Figure 2: HTTP, CRUD and SQL operations

HTTP	CRUD	SQL	DESCRIPTION
POST	CREATE	INSERT	Create or add a new resource
GET	READ	SELECT	Read, retrieve, search, or view existing resources
PUT	UPDATE	UPDATE	Update or edit existing resources
DELETE	DELETE	DELETE	Delete/deactivate/remove existing resources

3.3. DataSTAT Hub model

The creation of a model to collect different types of data and metadata requires important features to be taken into consideration, in order to simplify, facilitate and generalize the process, among which:

- *unstructured data* – a model collecting data in their essence (key-value) is more convenient and immediate than defining multiple standards for data representation;
- *scalability* – a highly extensible architecture is needed, in case of possible conceptual/architectural future upgrade;
- *intuitive schema* – the model should be easily applied by data suppliers, without resorting to complex studies of any imposed standard;
- *big-data-oriented architecture* – storage is closely linked to the tools used for semantic search, data analysis and data visualization. Elasticsearch, Hadhoop, Solr, Cassandra provide a complete integrated environment for managing them.

When dealing with highly heterogeneous data, it is recommended to use a model to represent them in their simplest form: a key-value pair. The format that is better suited to use HTTP and transmit data objects consisting of key-value pairs is JSON (JavaScript Object Notation) to which different models for data representation can be associated. This storage model, at the basis of DataSTAT Hub, is the same used by some NoSQL databases (for example, Cassandra), well suited for a big data processing architecture.

DataSTAT Hub model is characterized by a five-level recursive structure based on key-value logic.

SKEY Statistical Key Value data model

```
{
"keyspace" :
  {
    "columnfamily" :
      {
        "rowkey" :
          {
          "supercolumn" : { "column name" : "column value"}
          }
      }
  }
}
```

Each data supplier has and autonomously manages one specific keyspace.

There are two column family types in this model: data and metadata.

The number of super columns, namely aggregates of columns, is defined on the basis of data supplier's needs.

Columns contain key-value pairs.

3.4. DataSTAT Hub architecture

DataSTAT Hub uses Elasticsearch, an open source search engine that can be conveniently used for collection and release of data([26]). Through Elasticsearch it is possible to index and map documents/data through query strings to be sent via HTTP in JSON format.

([26])For further details on Elasticsearch, see: https://www.elastic.co/

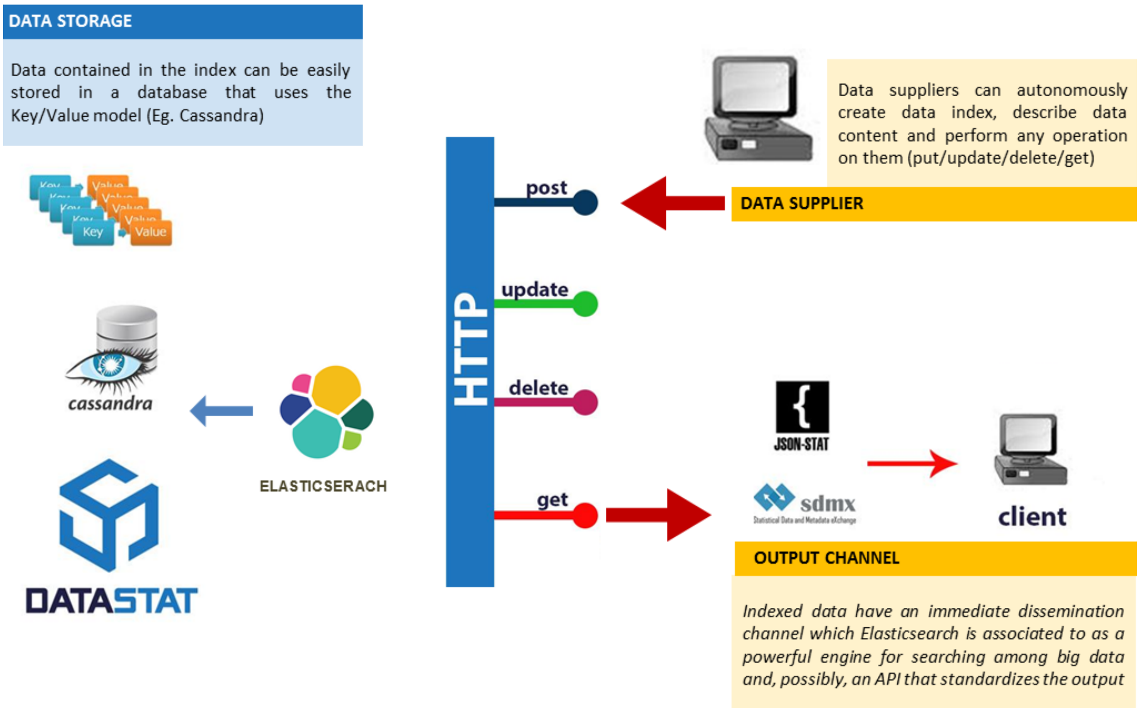
ELASTICSEARCH MAIN FEATURES*

- DOCUMENT: most entities or objects in most applications can be serialized into a JSON object, with keys and values. A key is the name of a field or property, and a value can be a string, a number, a Boolean, another object, an array of values, or some other specialized type such as a string representing a date or an object representing a geolocation.
- INDEX/TYPE: documents are indexed – stored and made searchable – by using the index API, which uniquely identifies the document.
- MAPPING: is the process of defining how a document and the fields it contains are stored and indexed.

* For further details, see: https://www.elastic.co

The phases of data collection and release process through DataSTAT Hub are illustrated in Figure 3.

Figure 3: DataSTAT Hub architecture and data collection and release process



4. Statisticlass: DataSTAT Hub applied to statistical classifications

4.1. Acquisition and management of statistical classifications

DataSTAT Hub is well suited for the solution of some critical issues related to the use of statistical classifications in different fields (surveys, administrative registers, information systems, etc.), such as:

- acquisition, storage, management and updates of classifications;
- multilingual semantic search for coding;
- sharing and dissemination of coding tools.

Each of these issues has been effectively resolved within the project Statisticlass (http://www.statisticlass.eu) thanks to the exploitation of DataSTAT Hub.

Statistical classifications have a generalizable logical structure, described within the Generic Statistical Information Model (GSIM([27])), which provides the organization of content in a complex and structured architecture (see box below on GSIM). Acquisition, semantic search and dissemination of classification items can be managed effectively through DataSTAT Hub.

To this end, it is possible to exploit a very simple JSON object, to which then associate the metadata related to the classification (family, serie, level, etc.).

JSON OBJECT

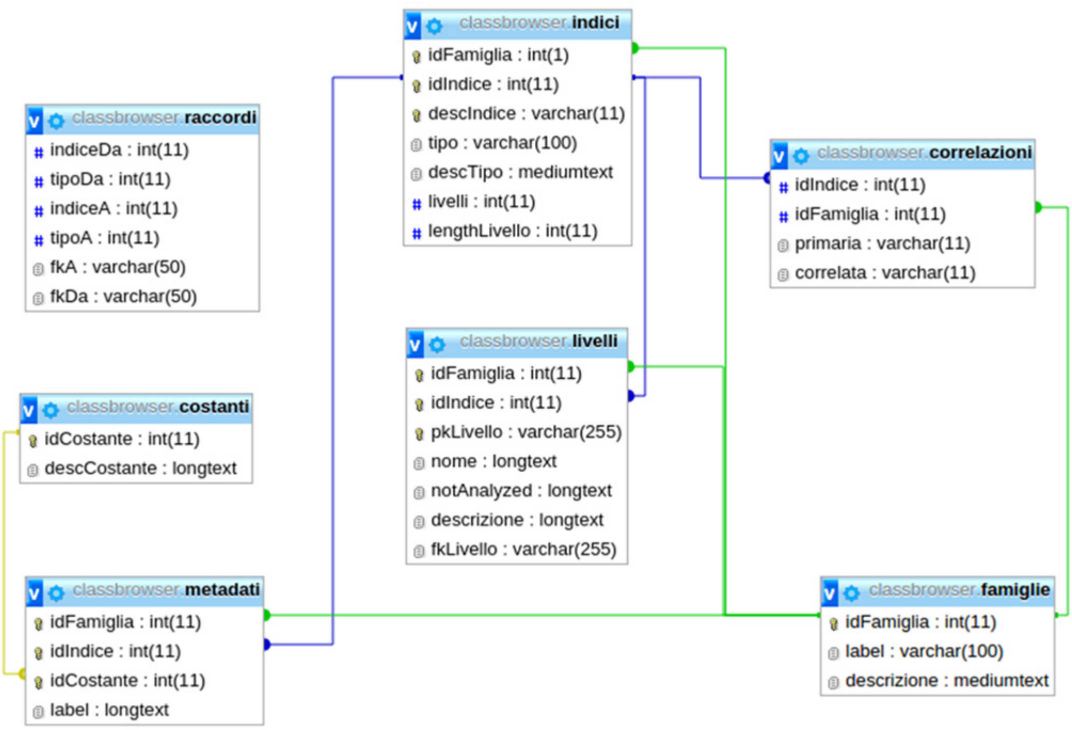
```
{
  "code":"x.x.x",
"name":"NAME",
  "description":"DESCRIPTION"
}
```

GSIM

- **Classification Family**
- **Classification Series**
- **Level**
- **Correspondence Table**
- **Classification Index**
- **Classification Item**
- **Classification index entry**

PUT and GET methods (mentioned in Figure 2) of DataSTAT Hub permit an easy acquisition of classification items which can then be organized through *ad hoc* procedures, on the basis of GSIM model, and stored into a relational database (Figure 4).

Figure 4: Statisticlass entity–relationship model



([27])For further details on GSIM, see: http://www1.unece.org

4.2. Semantic search and coding

Textual search is a very popular technique for users who seek information on the web. It does not require any special skill users have already acquired through surfing the web and it is also suitable to search within statistical classifications and facilitate coding. The most common problem related to semantic searches within taxonomies concerns false-positive and false negative results. The search is usually done through SQL queries allowing users to perform two types of operations: "exact match" and "full text". String parsing algorithms can be associated to the SQL queries.
The main critical issues related to textual searches imply the management of:

- singulars and plurals
- stopwords, namely words irrelevant for research that depend on the dictionary and language
- male and female terms
- differences between a one-word search or a sentence search
- case sensitive
- misspellings

As well as enabling data and metadata collection through HTTP methods, DataSTAT Hub uses Elasticsearch, which is a powerful search engine based on Lucene library and providing numerous RESTful services useful to perform complex semantic searches through JSON query.

A statistical classification can be indexed within Elasticsearch to perform complex and differentiated textual searches through DSL (Domain Specific Language) in JSON format. This solution permits to simplify the formulation of complicated SQL queries and makes the search system from any programming language usable.

Elasticsearch allows users to manipulate large volumes of data thanks to an internal document management, completely independent from relational databases, and the opportunity to create distributed cluster. The creation of index and mapping related to the "occupation" family and the "name" field of the International Standard Classification of Occupations (ISCO-08) are synthesized in the following boxes.

INDEX

Creation of the index related to the "occupation" classification family

```
PUT occupation
{
  "settings" : {
    "number_of_shards" : 3,
    "number_of_replicas" : 2
  }
}'
```

MAPPING

```
PUT occupation
{
 ""mappings": {
  "ISCO08": {
   "properties": {
    "name": {  "type": "string",
    "analyzer": "standard"
    }
   }
  }
```

The analyzer is the tool to build queries through a JSON array. The DSL language is made up of: the tokenizer, namely a system dividing text into individual tokens; a set of filters through which users can fix the criteria for making the search string parsing. The result of a search is an output in JSON format. The items are listed on the basis of a score assigned in function of the discriminating power of the searched string.

The search engine of DataSTAT Hub applied to Statisticlass implies different steps for different levels of precision. The first operation performed by the search system consists of the identification of the exact matching of the string. The results obtained are enriched by a more detailed analyzer, which provides a sophisticated stemming algorithm. This algorithm is combined with a dictionary for stopwords management. Another precision level includes the correction of the syntax of entered strings. In this way, a very precise result is achieved, false positive and negative results are reduced and users can make use of different channels for dissemination and sharing of classifications.

It is worth pointing out that such search systems can be used with different levels of precision to any statistical classification.

4.3. Sharing and dissemination of classifications and search engines

DataSTAT Hub provides a large set of tools for sharing and dissemination of statistical classifications and search engines. In this respect, an immediate channel is the REST web service provided by Elasticsearch. The web service output, in JSON format, is obtained through a query string (see box below). This solution is well suited to the development of web applications that need an efficient, uniform and high level of customization coding system.

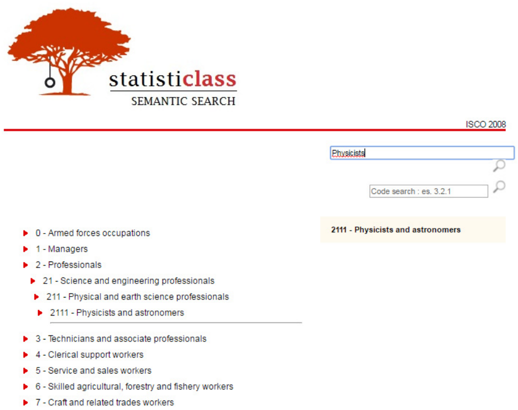
JSON OUTPUT

QUERY STRING


```
{  "took":18,
   "timed_out":false,
   "_shards":{  "total":5, "successful":5,"failed":0 },
   "hits":{ "total":1,  "max_score":3.1464005,
      "hits":[
         {  "_index":"isco08",
            "_type":"isco08",
            "_id":"AVe31A1PIZ7xXoOaTaQB",
            "_score":3.1464005,
            "_source":{
               "pkLivello":"2111",
               "nome":"Physicists and astronomers",
               "fkLivello":"211"
            },
            "highlight":{
               "nome":[   "<em>Physicists<\/em>   and astronomers"  ]
            }
         }
      ]
   }
}
```


Figure 5 summarizes the results of the search within Statisticlass website (http://www.statisticlass.eu).

Figure 5: Statisticlass web interface



The use of web services requires high-level IT skills, therefore it is reserved for a selected group of users (mainly programmers or web developer).

Easy to use widgets have been developed to include coding systems in Statisticlass within web questionnaires or web applications.

A widget is a script that can be included within a web application through a simple "cut and paste" (almost as it happens for multimedia content included in social networks).

Istat experience in using this methodology has been very satisfactory. The coding systems related to the main statistical classifications (ISCO, NACE, ISCED, COFOG, COICOP) were included in several Istat surveys ("Labour Force Survey", multi-purpose survey "Aspects of daily life", "Consumer prices", etc.) and Information system on occupation[28]

[28]For further details on the Information system on occupation, see: http://professionioccupazione.isfol.it

5. Concluding remarks

The importance to complement existing data derived from traditional surveys with those from administrative sources is worldwide recognized. They offer further information on a wide range of issues, including some which cannot be answered cost-effectively from survey data, and make it possible to improve the knowledge and understanding of specific phenomena.

What do problems related to data collection involve? Summarizing: the production and standardization of outputs to be stored in a database, connected, processed by statistical software and/or visualized within *ad hoc* created web platforms.

DataSTAT Hub is a suitable and easy tool for automated collection, standardization, integration and release of administrative data and metadata. It permits to reduce the burden on users, because it does not require the knowledge of the internal data base since the updating is performed through the HTTP query strings and can be used with any programming language; once created, the procedure will be used for each next data supply.

Differently from previously mentioned data collection systems such as File Transfer Protocol and ARCAM, DataSTAT Hub can support all phases of statistical process. In addition to data and metadata collection, this informative system permits to connect to data mining and data analysis applications (e.g. Hadoop) and visualize and disseminate statistical data (Kibana, Spago BI, Micro Strategy, etc.).

Instead, for the purpose of data treatment and validation, the most common techniques, implying the population of traditional data bases linked to the most usual statistical software (SAS, R, SPSS, etc.), can be used.

Not only do users have a reduction of burden but also a reduction of costs in terms of employment of human resources – for organizational, bureaucratic and IT management – and time-saving, by permitting to overcome some critical issues related to the use of administrative data, including those connected with confidentiality of the data. Moreover, DataSTAT Hub can be effectively applied to statistical classifications in order to facilitate semantic search, dissemination and sharing.

It is a user-friendly tool developed by making use of open source technologies (PHP, MySQL, Cassandra) and can be conveniently shared among NSOs, while it is extensible to any other institution interested in the automatic collection and integration of administrative data and metadata.

References

Gormley, C. and Zachary T., *Elasticsearch: The Definitive Guide. A Distributed Real-Time Search and Analytics Engine*, O'Reilly Media, 2015.

Masse, M., *REST API Design Rulebook*, O'Reilly Media, 2011.

Olsen, W., *Data Collection. Key Debates and Methods in Social Research*, SAGE Publishing, 2012.

4 How far have we come on EU banking statistics? Are we there yet?

GAETANO CHIONSINI, LUIS GARCIA, MARIO QUAGLIARIELLO[29]

Table of Content

([29]) European Banking Authority. Views expressed in this text are those of the authors and do not necessarily reflect the position of the EBA.

1. Introduction

In 2011, when the European Banking Authority (EBA) was established as part of the European System of Financial Supervision (ESFS), lack of harmonised and easily available data on EU banks was a major impediment to the supervision of cross-border banking groups as well as an obstacle to market discipline. In fact, limited information on specific institutions and markets, poor transparency and lack of comparable data left investors uncertain on the extent of banks' exposures, thus exacerbating the financial crisis. Since investors were not able to discriminate between safer and riskier intermediaries, contagion and spill over effects were observed across the entire banking sector, despite banks' individual resilience and preparedness to withstand future shocks.

At the global level, one of the first responses of the regulatory community was the start of intensive work on data gaps and the analysis of their impact on supervisors' ability to correctly assess risks. Interestingly, most gaps were not linked to the absolute lack of data but rather to the inadequate use of existing resources and information, hindered by the fragmentation and non-harmonisation of certain macro- and micro-financial data across jurisdictions.

In the EU, the EBA was given the mandate to set common standards for supervisory reporting, as part of the more general task of developing a single set of rules (the Single Rulebook) to be applied consistently by supervisors and banks across Europe.

In just five years, banking statistics have come a long way. EU banks now deal with a single reporting framework, based on the maximum harmonisation principle. This is complemented by common platforms for data collection, validation rules and quality assurance processes. The EBA has established a tradition of disseminating detailed data to improve the understanding of the EU banking sector. The EBA facilitates information sharing among competent authorities and promotes disclosure and transparency of aggregate and bank-level data for external stakeholders, either linked to the EU wide stress tests or as standalone exercises. Targeting a wide range of the end-users – from competent supervisors and policy makers to market analysts, investors and the general public – the EBA is able to provide up-to-date fully comparable information for the largest EU banks.

Before being disclosed, banking statistics go through a rigorous quality assurance process, which is consistent across EU countries. Data is disseminated through different tools in order to reach out to a wider range of users. Interactive and visualisation tools are increasingly provided to facilitate the understanding and concrete use of data. Indeed, transparency requires not only that information is disclosed but also that it is customised according to the specific needs of stakeholders, so as to be properly interpreted.

The objective of this paper is to track the progress made in the EU in developing banking statistics, with a particular focus on the EBA's efforts in setting up a data process which delivers, at the same time, robust and reliable reporting for supervisors and consistent data for external stakeholders.

2. Common definitions as a precondition for comparability and transparency

2.1. Supervisory reporting standards

In 2009, the de Larosière Group([30]) was called to analyse the roots of the crisis and identify a way forward for EU rule-making. As part of the assessment, the Group stressed that there were i) "failures to challenge supervisory practices on a cross-border basis", and ii) "lack of frankness and cooperation between supervisors". With the increasing diversification of banks' activities, especially of cross-border nature, the main obstacle for carrying out peer analyses of supervisory practices and benchmarking exercises was the lack of comparable data across intermediaries.

To be fair, there had been attempts to improve the consistency of supervisory reporting and enhance the comparability of banks' data in the EU. Already in 2005, the Committee of European Banking Supervisors

([30]) Lamfalussy (2000), Initial report of the Committee of wise men on the Regulation of European Securities Markets.

(CEBS) – the EBA's predecessor – published a standardised financial reporting framework (FINREP) for reporting the consolidated financial accounts of EU credit institutions using IFRS. The common framework (COREP) for reporting own funds and own funds requirements and reporting of large exposures was published in early 2006.

However, FINREP and COREP, good examples of the efforts to harmonise and streamline reporting requirements in Europe, were only a partial success, since they were implemented under the Lamfalussy architecture as non-binding guidelines. As a matter of fact, while the framework for own funds and own funds requirements was at least partially implemented by all EU supervisors, FINREP and reporting of large exposures were implemented by fewer supervisors.

The CEBS guidelines followed a data model where supervisors could choose to implement a smaller set of core templates and select just a few non-core templates. Several supervisors also complemented core templates with non-core templates capturing some specific national data needs. Frequency of reporting, as well as remittance periods for institutions, was left to the discretion of national supervisors. As a result, very few supervisory concepts were harmonised and data timeliness was distinct across jurisdictions. FINREP core templates of balance sheet and income statement for example were implemented by 18 EU supervisors but only 3 countries implemented the whole FINREP framework. COREP Guidelines were implemented by all EU supervisors but only 7 of them implemented them in full. From all EU supervisors, only 2 required their banks to report both FINREP and COREP in full (ECB, CEBS 2010([31])). Clearly, status quo then was far from ideal([32]).

With the outbreak of the global financial crisis, and following the report of de Larosière Group, the EBA, as well as its sister authorities for securities markets (European Securities and Markets Authority - ESMA) and for insurance and occupational pensions sector (European Insurance and Occupational Pensions Authority - EIOPA), was entrusted with a number of tasks and instruments for dealing with systemic risk. As already mentioned, one of the central tasks of the EBA is to contribute to the Single Rulebook. The Capital Requirements Regulation (CRR) and the Capital Requirements Directive (CRD) are the backbone of the Single Rulebook, setting rules that are directly applicable to all EU Member States. Among other things, the CRR and the CRD give more formal and binding powers to the EBA, with the mandate to develop harmonised reporting and implementing technical standards (ITS) to specify uniform formats, frequencies, dates of reporting, definitions and the IT solutions.

To fulfil this mandate, in 2014, the EBA finalised and rolled out a single set of supervisory reporting standards[33] , laying down uniform requirements in relation to supervisory reporting for some specific areas of the CRR, such as own funds requirements (in COREP) and financial information (in FINREP). They were followed by additional technical standards, like those on non-performing exposures (NPE) and debt forbearance, on asset encumbrance and on leverage ratios.

The ITS on supervisory reporting are a major step towards the harmonisation of prudential supervision requirements as well as a game-changer to strengthen data comparability. With the ITS on supervisory reporting, all European banks are subject to the same definitions for supervisory reporting, thus ensuring adequate and fully harmonised information and coverage for both micro- and macro-prudential purposes. This Implementing Regulation is the only Union-wide legal framework for supervisory reporting and a "maximum harmonisation" instrument. This means that Competent Authorities cannot add or delete data to be reported, nor can they require the reporting of that data in a different format, neither in a different (less or more granular) breakdown or in a different combination. The opposite would bring back to minimum harmonisation, with banks dealing with different, possibly inconsistent and costly national reporting frameworks.

The level of detail in data covered by the ITS on supervisory reporting is significant. Overall, there are more than 190 tables, with some 150 thousand of a priori admissible occurrences, with the actual data to be reported depending on banks' business model, size and complexity. While comprehensive and addressed to all banks – consistently with the objective to promote a truly Single Market – the reporting requirements aim at being proportionate, with materiality thresholds reducing the burden for smaller and less complex reporters.

([31]) Link to CEBS Guidelines.
([32]) Rimmanen M (2015), "The single EU supervisory reporting", in Quagliariello M. (2015), Europe's New Supervisory Toolkit, RiskBooks.
([33]) Implementing Regulation (EU) No 680/2014 of 16 April 2014 laying down implementing technical standards with regard to supervisory reporting of institutions according to Regulation (EU) No 575/2013 of the European Parliament and the Council.

2.2. IT solutions

When dealing with the complexity of the data collection resulting from the ITS, a robust data model is a prerequisite for achieving a sufficient degree of automation of the processing system. This is also a condition for ensuring short response and lag times, lower cost and high reliability when designing and processing statistical information, adaptability to information requirements evolution and major support towards harmonisation, information documentation and usability. Above all, a robust data model must be scalable and able to cope with a growing complexity of supervisory requirements.

To this end, the EBA has developed the Data Point Model (DPM), a structured formal representation of the data, identifying all the business concepts and their relations. With the DPM, each individual data requirement (basically, each cell in a particular table) is uniquely identified and expressed as a composition of the following characteristics:

- A metric: the characteristic that defines the nature of the measure to be performed. Each metric determines a data type and a period type. The data type (monetary, percentage, Boolean, etc.) corresponds to the nature of the values of the facts to be reported. The period type establishes whether the measure is performed at a specific instant of time (stocks) or during an interval of time (flows).
- A temporal reference: a characteristic that helps determine the specific instant or interval of time in the context a given reporting period. The temporal reference enables the distinction, for instance, of information that refers to end of the reporting period, the beginning of the fiscal year, a quarter, a month, etc.
- A number of dimensional characteristics that qualify and complement the metric and provide the proper context to understand the financial phenomenon represented. Each dimensional characteristic is composed of a dimension and a member. The dimension represents an identifying property and the member is the value given to that property in the context of the data point.

The DPM is, therefore, a metadata([34]) driven model where each data point is defined by the intersection of the metric, the temporal reference and each applicable member. The dimensions identify in detail the information described by data points (e.g. main category, currency of the instruments, sector of the counterparty, residence of the counterparty, location of the activity, etc.).

Within the data production chain in banking statistics, what is not fully harmonised yet is the data exchange format, which is basically the language used for translating each data point in a code that can be transmitted from the reporting entity to the reporting recipient. Within the supervisory reporting context, two main standards are currently used, XBRL and SDMX. While in the US XBRL has been adopted since 2010 and is mandatory, in the EU, Article 17 of the ITS on Supervisory Reporting states that the use of XBRL taxonomies is not mandatory for institutions reporting to National Competent Authorities (primary reporting), and the EBA has decided it shall be mandatory for secondary reporting (from National Competent Authorities to supra-national Authorities). The standard SDMX is currently used by the ECB for other data collections, among which are Balance Sheet Items (BSI) and MFI interest rates (MIR).

3. Data quality as a process and policy objective

3.1. Dealing with trade-offs in data quality assurance

Building a truly European Single Rulebook and reporting framework is, however, not only about common definitions, models and transmission formats. It is also about common practices and similar levels of confidence in the data gathered.

Effectiveness in collecting supervisory data from the reporting institutions requires the following: (i) data timeliness; (ii) completeness and accuracy; and (iii) feedback and follow-up. All in all, timeliness and accuracy of available information for the purpose of performing banking supervision go hand in hand with

([34]) The term "metadata" is used to indicate data describing other data.

good decision-making. A balance has to be found between timeliness and intrinsic quality.

Until very recently, delays to fulfil data requests from supervisors toward banks were usually justified by the effort to compile and submit the best possible and most reliable figures. However, from a supervisory perspective, 100% accurate data gathered with significant delay could barely be used for historical purposes, as it would be useless for taking any proactive and timely decision. On the other hand, information compiled promptly could only be meaningful if it reaches a minimum level of accuracy. This minimum level of accuracy can be assessed by means of validation rules that constitute only the first line of defence against poor data quality. Additionally, this also implies that notwithstanding revisions and/or subsequent amendments, if any needed at all, the broader picture about the information remains unchanged. Resulting decisions and analyses should, therefore, be essentially the same regardless of which data update or version is being considered([35]).

The EBA has been strongly supporting Authorities, and in what concerns the supervisory reporting framework, also data reporters, to boost their systems in order to be more accurate, timely and fully in charge of their processes. Looking back from mid-2014 to the second half of 2016, the state of play of overall timeliness on files submitted to the EBA has improved substantially. Where delays were the norm and less than 70% of the files arrived on time (before the ITS, timeliness was usually below 70%, then improving to around 75% with Q1 2014 and 85% with Q2 2014 data), we now witness almost 100% coverage by the established remittance date for Q1 and Q2 2016, at least in terms of expected files/modules (not necessarily all supervisory reporting templates). Ensuring that financial and prudential information, as well as other banking supervisory statistics, feed supervisors' and researchers' in a timely manner is a must for conducting effective supervision and better understanding financial phenomena.

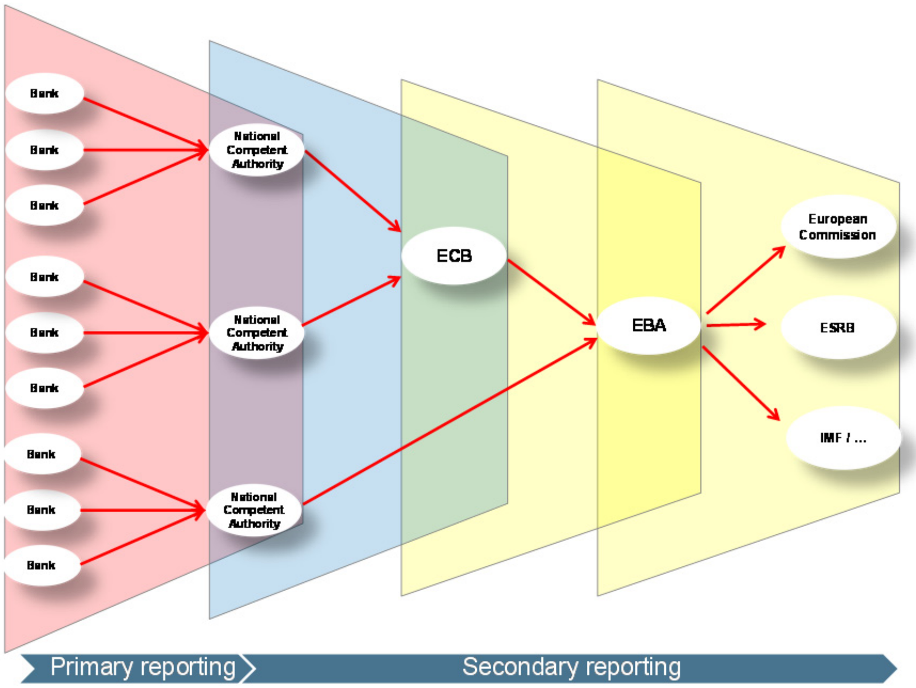
Achieving convergence in production cycles when compiling statistics and agreeing on common timelines were therefore key for the EBA and its objectives. This is particularly important for the data that the EBA is also collecting, via the competent authorities, on the largest EU banks([36]). The technical standards on supervisory reporting require institutions to report to competent authorities six weeks after the reference date (date to which data refers) for quarterly returns, and only with a two weeks lag in case of monthly reports.

It is not only in Europe that increasing demands for higher frequency deliveries, on a timelier basis are requested. In the US, a set of best practices has been recently described([37]) by the Office of Financial Research (OFR), where quality checking procedures against reporting requirements and submitted data are detailed. For the OFR, data needs to be fit for purpose, be available in a timely manner primarily to supervisors but not least to market participants and public in general thereafter. In addition, effective communication channels must be in place to allow for prompt corrections of any data deficiencies or for revisions. Best practices in collecting regulatory data may also play a key role in educating supervisors and reporters, also by facilitating data sharing amongst different authorities.

3.2. Data quality process: validation rules and plausibility checks

Consistently with the spirit of the De Larosière report, a great effort was put in strengthening the cooperation and information sharing among Competent Authorities. Furthermore, with a view of avoiding double submissions and reducing the reporting burden, the principle that banks should not report twice the same data to different authorities has been established. Accordingly, different data flows have been set for primary and secondary reporting.

([35]) Chionsini G. and Garcia L. (2015), "Data Workflow: From Quality Checking to Risk Dashboards", in Quagliariello M. (2015), Europe's New Supervisory Toolkit, RiskBooks.
([36]) See EBA Decision 2015/130.
([37]) "Developing Best Practices for Regulatory Data Collections", May 2016.



Therefore, once data is available at the authorities, an agreed data production cycle begins, simultaneously in all EU jurisdictions, typically with a quarterly frequency([38]), running from 10 to 15 working days. During this period, entities within the list of reporting institutions to the EBA are expected to liaise with their local supervisors in order to enhance data quality by re-submitting files with updated figures, or simply by confirming and explaining the ones reported beforehand. Supervisors at Competent Authorities are expected to enquire, to the extent possible, banks' about outcomes of validation rules, quality checks and supervisory alerts. In fact, the data quality process requires a balance of statistical tools and supervisory expertise: statisticians and supervisors are both necessary for ensuring that the reporting is formally correct and accurate in substance, before being released and used.

The concept of immediate validations performed against the collected data has been developed to a great extent and as far as possible. Over 2000 binding validation rules([39]) have been defined in the first place by the EBA (currently amounting to more than 3000) and are in place to monitor the consistency and plausibility of the submitted data, first to the authorities and thereafter to the EBA. Banks and Competent Authorities are responsible to apply and monitor these validation rules. Both are expected to trigger the needed follow-up when rules are breached. In addition, many validation rules have been given a blocking status due to their importance. Other rules serve more as quality checks and supervisory alerts, in the sense that they aim to capture major developments and do not necessarily include inaccuracies. Some of these may actually be only in the supervisors' remit and are not disclosed to reporting banks.

When validation rules are given a blocking status, they will cause file submission failures, which in turn could mean that reporting obligations are not met if rules are not properly observed. To ease the reporting burden, the majority of these rules are implemented at the technical level as part of XBRL taxonomies. This enables analysts, statisticians and experts at financial institutions to immediately understand which inconsistency in the data is the root cause for submissions' failures.

Ideally, rules should be tested prior to their implementation, to the extent possible and where reliable data is available ex ante. However, in practice, testing data is not always available and the testing phase has not always been possible since some of the definitions and underlying concepts were quite new to the reporting framework, both amongst banks and supervisors. Nonetheless, this to-the-extent-possible testing and the experience gained in the build-up of the ITS on supervisory reporting, as well as the first months of practical implementation, allow the community of supervisors and regulators to be confident about receiving enhanced quality data from banks' regulatory filings.

On top of formal validation rules, which have been given a legally binding status and must be complied with by all institutions and national Competent Authorities, the supervisory community felt the need to go further

([38]) Although monthly reporting exists for specific topics – e.g, liquidity coverage ratio information.
([39]) See EBA validation rules.

and added an additional layer of quality checks and alerts. This layer is being fully implemented and regularly enhanced, complementing checks performed on breakdowns, hierarchies, sign conventions and logical relationships. It goes deeper into the plausibility of figures submitted by banks by comparing year-on-year trends, quarter-on-quarter developments or unlikely, but still plausible, events happening in a given period. When developing informative and effective additional checks, the interaction between statisticians and supervisors becomes crucial.

Banks can explain the reasons for failing these additional quality checks (for instance, a merger can make a comparison over time unreliable). These reasons will prove useful for supervisors and analysts eager to get quickly acquainted with bank's latest activity. Likewise, supervisors have greater freedom in applying these checks when compared to formal validation rules. Country or sector-specific fine-tuning may be performed in order to adjust checks and make them as meaningful as possible for each bank or group of banks. It is up to the Competent Authority to decide whether it wants to provide all these checks and supervisory alerts to the bank. More than announcing what is expected of the data, some alerts may well just serve the purpose of monitoring more closely the bank's activity, which could be in the supervisor's best interest not to be fully disclosed to the bank.

This set of quality checks and supervisory alerts, comprising more than 800 formulae, are split between severe rules and soft alerts. Severe rules are typically rules closer to the concept of validation rules. As a matter of fact, they may well become validation rules in the future, provided they are enforceable and relevant enough to be considered as such[40]. Soft alerts, on the other hand, are more subjective and do not necessarily point to data inaccuracies. These do not fall into the category of one-size fits all validations. They can be fine-tuned to address each bank's activity, business model or even size. They could be seen, from a supervisory perspective, as a sort of early-warning indicator for some of the bank's latest developments as they would help the analyst or direct supervisor to become immediately aware of something that is of interest.

Supervisors are using these quality checks to further assess the quality of the data remitted by the reporting institutions. This layer, which aims especially at enhancing data quality by correcting obvious mistakes or at highlighting unexpected but plausible developments, is particularly important also for promoting harmonisation at the level of quality assurance procedures across different EU jurisdictions. It also ensures that definitions are understood similarly throughout the EU. In addition, comparable implementation techniques and prior testing are put in place before actual implementation. To the extent possible, prior testing will also have occurred before formally using these quality checks, similarly to the process described above regarding validation rules.

3.3. Reporters are key players in the data process

Supervised institutions, particularly the largest banks in Europe, are expected to efficiently manage the data needed to fulfil their regulatory returns. This obviously poses additional challenges for the banking system. Nonetheless, a relevant upside of these improvements is expected to be the enhanced response time and quality of supervisors' own decisions following the full implementation of this new supervisory reporting framework. While the reporting framework will continue to evolve over time, this set of validations and checks will naturally adapt to such developments. In some cases the set will be revised to consider the introduction of new concepts, in others it will be expanded or parts of it considered obsolete.

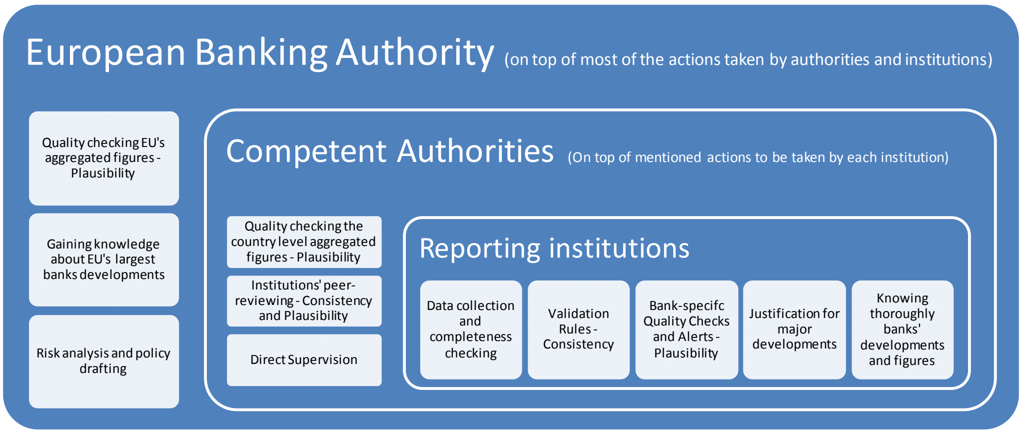
The banking industry itself might prove to be a high-quality setter, by helping supervisors and regulators understand which additional rules or alerts make sense to apply, thus fostering a level playing field for all market players. Such an initiative would have the merits of promoting best practices from within the industry, thus improving the quality and comparability of the data and discouraging less fit-for-purpose practices.

3.4. Feedback and follow-up actions

Another key aspect for data collections is how to provide feedback, communicate it and better monitor and follow-up any resulting actions from it. Bearing in mind the tight submission deadlines, the feedback

[40] Some validation rules may need to be outside the applicable taxonomy due to technical limitations related to the XBRL standard. However, it is important to highlight that from the existing set of validation rules, only around 3% are not implemented in the XBRL taxonomy.

process between the reporting entities and Competent Authorities, and the Authorities and the EBA, needs to happen as quickly and effectively as possible. After compiling and submitting the data to Authorities, reporting entities should be ready for a two-week window where they will confirm the reported figures and provide explanations when prompted, or quickly resubmit the data with improved quality. This will clarify pending questions from direct supervisors, statisticians or analysts looking at institutions' financial and supervisory reporting returns.



Automated processes are crucial to ensure the feedback process can be as quick as possible. Experiences shared among the supervisory community at EBA meetings show that each Authority has its own standards and ways of interacting with supervised institutions. It will be the EBA's role to achieve greater convergence also in this field of action, not by imposing any additional standards but by highlighting and promoting best practices that should be followed across countries and authorities.

4. Disseminating banking statistics

4.1. The role of public authorities

The definition of the supervisory reporting framework was a necessary step for filling the data gaps identified during the crisis, thus improving EU supervisors' ability to assess banks' solidity and identify weaknesses. In parallel, the availability of EU-wide bank statistics allowed the EBA to build its own capacity to collect, quality assure and explore granular bank data. Based on this statistical capacity, the EBA has been able to fulfil its mandate to analyse risks and vulnerabilities in the EU banking sector as well as to evaluate the impact of proposed regulatory measures. Evidence-based policy making has been indeed a core principle of the EBA's activity since its establishment. The EBA has also pushed for making bank data available to external stakeholders, by conducting annual EU-wide transparency exercises.

Since 2011, when the crisis showed that available banking information was scarce, inaccurate and not timely enough, the EBA has been conducting annual transparency exercises, either linked to concurrent stress test exercises (2011, 2014 and 2016), to one-off exercises (2011/2012 recapitalisation exercise) or as standalone exercises like the ones performed in 2013 and 2015. In this way, the EBA has established a tradition of disseminating detailed bank-by-bank data designed to improve the understanding of the EU banking sector and foster market discipline in the Single Market. These data releases complement banks' own Pillar 3 disclosures.

The EBA transparency exercise has evolved over time to a consistent disclosure of a huge amount of banking statistics, produced without extra costs for banks, showing that it could be possible to make use of supervisory data also for purposes other than banking supervision. In that respect, the 2015 EU-wide transparency exercise represented a milestone since it mostly relied on the information reported to the EBA on a regular basis through the supervisory reporting framework[41]. For the first time, data templates were filled directly by the EBA and only sent to banks for verification and sign-off. In this process, the EBA processed and disclosed up to 13,600 data points for each bank involved. That amounted to over 1.3

[41] Only in the case of two topics (sovereign exposures and leverage ratio), for which supervisory reporting data at the required level of detail is not available, data was collected from banks as in previous exercises.

million data points published in aggregate on the EBA website. Not only did the EBA provide raw data for each bank, it also built a dedicated data-model for data-warehousing this information, and provided users with easy-to-use metadata to support interactive tools designed to facilitate data exploration.

However, the EBA's approach had some limitations, mainly related to the sample of banks for which information is disclosed. Indeed, the selection of banks has been traditionally driven by policy actions (stress test, recapitalisation, etc), of which the transparency exercise was an admittedly very important by-product. As a result, EBA's coverage changed over time, thus reducing the possibility for market analysts to carry out time series analyses. Banks themselves would benefit from a more stable sample, so to be able to compare and contrast their key indicators with those of their peers.

Against this background, in order to resolve the problem of differing samples across time, the EBA decided, in 2016, to run the EU-wide transparency exercise on an annual basis and as a stand-alone initiative. This means that as of 2016, the EBA will publish bank-by-bank data for all the banks contributing to the EU averages in the EBA publications (risk assessment reports, risk dashboards). The immediate result will be that the EBA will disclose supervisory data for a sample of about 150 banks selected each year on the basis of clear entry and exit criteria.

4.2. Formats matter

Banking statistics are today serving multiple purposes and are increasingly available to a heterogeneous set of users. For supervisors, the starting point is the knowledge of the reporting standards and the DPM. The seemingly complex framework given by the DPM should not prevent users from exploring the data, but additional tools are needed to make the exploration more user-friendly. With this in mind, the EBA has provided some technical tools (e.g., DPM database, analysis matrix, data dictionary, table layout and data point categorisation) that should guide users in their navigation through the DPM and will continue developing other tools to ease the drill down and metrics/domains/members exploration and providing different alternatives and ways to use, understand and explore such a rich dataset.

However, regardless of the tools, formats and technology used, much more can – and should – be done to ease the analysis of the EU's banking system figures. This is particularly true for external stakeholders, who may prefer simpler – although perhaps less comprehensive and flexible – data exploration tools. Mining and visualisation techniques are indeed playing an increasing role in facilitating the understanding of the banking data currently available and the data published by the EBA can be always exploited relying on interactive tools. The most efficient way of disclosing data, as far as the EBA experience is concerned, is through risk dashboards42 and interactive tools, which leverage on data visualisation tools closer to end-users.

Risk indicators were also in need of more harmonisation and transparency. Having common definitions of the input data indeed allows the computation of indicators in a consistent way across countries. But for this, clarity on how these indicators are built up is important. The EBA methodological guide on the compilation of the risk indicators and detailed risk analysis tools provides all the details for computing and replicate EU figures at different levels of aggregation43. Banks themselves can calculate their own risk indicators for comparing and contrasting their performance with those of their peers. Common banking statistics also allows EU authorities to report to international organisations following the same mapping, as showed in the EBA's guidance on compiling the IMF Financial Stability Indicators starting from ITS data.

A good example of how data visualisation techniques can be useful was the last EU-wide stress-testing exercise, where the EBA provided a broad range of tools and ways in which the disclosed information could be explored. From interactive maps, aggregation tools and comprehensive datasets44, the level of data offered to the market has been unprecedented in the EU context, and matched by very few jurisdictions elsewhere in the world.

From effective data crunching and cleansing to user-friendly ways of exploring the information, the EBA has been developing its internal analytical infrastructure in order to appropriately deal with supervisory reporting data. It is the EBA's objective to put these tools at the service of the supervisory community and the wider public, where possible.

[42] See EBA Risk Dashboard.
[43] See EBA Methodological Guide on Risk Indicators and Detailed Risk Analysis Tools.
[44] See 2016 EU wide stress testing results.

5. Conclusions

Supervisory reporting is – and we expect will continue to be – a sensitive topic for banks and supervisors. Banks tend to emphasise the excessive reporting burden entailed by the EU common standards, while supervisors are sometimes nervous that maximum harmonisation dampens their ability to collect relevant data, thus impairing their ability to perform their tasks. To some extent, both arguments have some ground and this requires that EU rules are drafted – and reviewed as necessary – in order to keep them proportionate to banks' complexity and updated in light of new risks and data needs.

On the other hand, it is fair to acknowledge that a single reporting framework is one of the major achievements of the post-crisis regulation and any step backward would be a defeat for the Single Market. Cross-border banks conduct business subject to one single set of rules, regardless of the number of jurisdictions in which they operate. The exchange of information among authorities in the EU has reached an unprecedented level. Supervisors from different countries in Europe can peer review their institutions against others with similar business models, size or geographical exposures, basing their analyses on harmonised and comparable data, available with the same frequency and following exactly the same production cycles and timeline.

Above all, the disclosure of banking statistics based on harmonised definitions and convergent data processes has contributed to restoring confidence in the banking sector. Remaining concerns on EU banks are very often related to those areas where disclosure is perceived as not adequate, insufficient or unreliable45. Indeed, while the progress is undeniable, there is still a perception that there is a gap in transparency between the EU and other jurisdictions, particularly the US46.

So, are we there yet? We would argue that we are closer but there is still some way to go. In our view, all the relevant actors are to work further in three main areas.

First, authorities need to put some additional effort in ensuring that the maximum harmonisation principle is enforced, reducing ad-hoc data collections and national requirements that are no longer justified. Also, it is important to ensure that further convergence is achieved in the data process, for instance by setting standard procedures that banks must follow for the resubmission of data in case of mistakes or misinterpretation of reporting instructions.

Second, both authorities and banks should work together for embracing a cultural change regarding how regulatory data is seen and treated. The bar is set much higher than in the past. Supervisory reporting will be the main source of information for assessing ongoing risks in the banking system: it serves as a key pillar to conduct supervision such as stress-testing exercises, RWA consistency analyses and supervisory benchmarking. Supervised institutions should be mindful of their reporting requirements: relevant deadlines and the full extent of the required data and details should be submitted smoothly to the supervisors. On the other hand, the reduced number of unforeseen data requests from supervisors would in turn reduce the reporting burden for reporting institutions.

Third, transparency and disclosure in the EU need further improvements, in terms of coverage, scope and frequency. From a technical perspective, there are no obstacles to a more frequent publication of bank-by-bank data. Bank-specific disclosures might even become as frequent as quarterly updates, which would finally put the EU at the same level as the US in terms of data availability. Also, it could be possible to further expand the sample of banks that the EBA covers. However, until today, the EU-wide transparency remains a voluntary exercise where a bank can prevent the EBA from publishing its data. This has not halted the EBA's determination in disclosing bank-level data, but a more explicit legal mandate to do so would facilitate the process.

[45] Andrea Enria (2016), Ensuring transparency in the European financial system. OMFIF City Lecture.
46 Christopher Gandrud, Mark Hallerberg and Nicolas Véron (2016), Bruegel website: The European Union remains a laggard on banking supervisory transparency.

5

The challenge of permeable borders: estimating cross-border employment between Italy and Switzerland

MARIA ELENA COMUNE, SABINA RONCONI, LORENA VIVIANO[47]

Table of Content

([47]) Italian National Institute of Statistics (ISTAT). -Territorial Office for Piedmont and Aosta Valley, Lombardy and Liguria; comune@istat.it, ronconi@istat.it, lovivian@istat.it

1. Introduction

The free movement of workers over borders is enshrined in European Union (EU) treaties and is a pillar upon which Europe was built. The freedom of movement also applies to countries which are part of the European Free Trade Association (EFTA). Notwithstanding a history of more than fifty years, the subject seems underrepresented in socio-economic research. Several studies (Nerb et al. 2009, Fries Tersch et al. 2015, Jorens et al. 2015, InArco Project 2012) point out at the current gap on qualitative and quantitative knowledge of labour mobility patterns, both migration and commuting.

A specific type of labour mobility concerns 'cross-border workers' or 'cross-border commuters', who are citizens that reside in one country but work in another one, and for this purpose move across the border regularly. Even focusing on this specific aspect, the extent and evolution of the phenomenon are difficult to grasp, since harmonised and comparable data are not available at neither European nor national level. According to a recent report prepared for the European Commission (Fries Tersch et al. 2015), however, a first indication of the magnitude of cross-border mobility can be obtained. In 2014, in the EU and EFTA there were about 1.6 million people who worked in a different EU or EFTA country from the one in which they resided. About 1.2 million worked in another EU country (accounting for 0.6 % of the employed EU population) and 379 000 worked in an EFTA country (making up 5.4 % of the employed population in EFTA). In the different countries, the share of cross-border workers relative to the total employed population varies greatly. The small Grand Duchy of Luxembourg, for instance, relies on people residing in other countries for as much as 69 % of its employed population; many cross-border workers are also present in Germany and Austria, although their overall significance is quite low (0.8 % and 3.8 % respectively). In Switzerland, which is one of the main importers in absolute terms (approximately 330 000), they make up 7.6 % of the employed population.

Among European countries, Italy is an important 'supplier' of cross-border labour, though according to a recent study (ISTAT 2014) its significance at national level remains very small relative to the total employed population (approximately 0.3 % or 90 000 in absolute terms). Furthermore, cross-border movements are highly localized on the territory, namely along the longest and traditionally most accessible land border in terms of infrastructures, that is the border between the Italian regions of Lombardy and Piedmont and Switzerland. In the last decade, the Territorial Office for Lombardy of the Italian National Institute of Statistics has been at the forefront of research to study this phenomenon.

In the context of official statistics, two research projects have been carried out with specific reference to the measurement of cross-border workers. The first and less recent one has been limited in geographical scope to the border between Italy and Switzerland. This experience resulted notably in the provision of data on cross-border work and on the economic system of the two neighbouring territories, with a view to comparing and integrating statistical sources and, where possible, to the harmonization of the data produced on either side of the border([48]).

The second and more recent research activity, unlike the first one, has been in principle comprehensive in geographical scope, as it focused on flows across all land borders of Italy, including the Republic of San Marino and neighbouring countries such as the Principality of Monaco([49]) (Figure 1).

([48]) As a result of the project, for three consecutive years starting with 2004, the ISTAT Territorial Office for Lombardy, in collaboration with the Statistical Department of Canton Ticino of the Swiss Federal Statistical Office, published the statistical yearbook on the labour market in the cross-border area of Insubria (ISTAT Territorial Office for Lombardy and USTAT 2004, 2005, 2006).

([49]) The research developed in the context of the activities to ensure compliance with the new national accounts standards, and specifically to allow the provision of updated and revised estimates of labour input consistent with the economic aggregates that contribute to the formation of gross domestic product (GDP). To this end, it was necessary to improve the knowledge of the phenomenon of cross-border work with the objective to measure exhaustively the input of labour employed on the economic territory of our country. It should be noted that the measurement of the input of labour plays a key role in the estimation of Italian GDP. The measure of labour input, and in particular of those employed, includes cross-border workers, i.e. people who cross the border daily or weekly to work in the Italian economic territory, but exclude workers who reside in Italy and cross the border to work in another economic territory.

Figure 1: Map of northern Italian land borders (Alpine arch)

Source: Elaboration on Google maps

The study has been carried out by the ISTAT Working Group on Cross-border Workers (IWGCW), where representatives of the National Accounts Department and of the relevant Territorial offices([50]) collaborated with the aim to analyse the sources available and estimate the in-flows and out-flows of cross-border workers for each border country for the years 2010 and 2011([51]).

The present paper contributes to improving the knowledge base on cross-border movements in a selected border area between Italy and Switzerland, where the largest documented commuting stream flowing out of Italy can be observed. Our work builds on the findings of the IWGCW and specifically updates the results of the previous research by presenting new evidence on the socio-economic profile of cross-border workers. To do this, we use data from the Italian Population Census 2011, which were not yet available for analysis at the time of closure of the IWGCW Report.

Definition adopted of cross-border workers

A cross-border or frontier worker is somebody who commutes regularly (in principle every day or at least on a weekly basis) from the country of residence to a neighbouring country, whether employed or self-employed. Workers employed by productive units located on the economic territory of a neighbouring country and do not commute regularly to work due to teleworking are also included. Seasonal workers are not counted as cross-border workers, as they move albeit temporarily (less than 12 months) from the country of residence to the country where they work (Regulation 883/2004/EC).

The essential elements of the definition are:

- Income: the worker has to perceive an income, whether as an employee or self-employed, regular or not regular, and as long as the economic activity is a productive one according to the national accounts production border.
- Time: the commuting to work from the place of residence has to take place on a regular basis and across a national border. Not to be considered a migrant worker, the person has to go back on a regular basis to the country of residence.
- Space: neighbouring countries not necessarily with a contiguous border in common (for instance, the workers coming to Italy from Croatia, or going from Italy to the Principality of Monaco).

([50]) The territorial offices competent for the following regions were involved: Lombardy, Liguria, Piedmont and Aosta Valley, Veneto and Friuli-Venezia Giulia, Emilia-Romagna and Marche.

([51]) The IWGCW ended its activities in April 2014 and included its main findings and estimations in its final report produced for internal purposes (ISTAT 2014).

2. Sources and methodology

An issue of critical importance is the lack of data for some border areas. The presence or absence of statistical/administrative records plays a key role for data availability. Even when the sources on the labour force exist, the problem of isolating the information regarding the cross-border workers may not be easily overcome. Existing data may allow to observe by approximation the phenomenon in different ways but not to obtain robust estimates.

Information on cross-border workers are collected by tax authorities and social security registers. Their use requires an in-depth knowledge of applicable legislation([52]). To use fiscal data for statistical purposes, it is necessary to establish an active collaboration with the authorities concerned and become familiar with their way of coding and collecting data. It should be noted, however, that the estimates may be affected by systematic underestimation, because labour income is only taxed above minimum thresholds, below which no reporting obligation applies. Therefore cross-border workers with incomes below the threshold would not be counted. Nevertheless, in general, the information gathered from administrative sources which are built by current standards, offers a potential information base regarding social, demographic and economic aspects of the phenomenon.

According to the conclusions drawn by the IWGCW, the most relevant flow of workers crossing daily or weekly the Italian border can be undoubtedly observed in the area between Italy and Switzerland, moving out from the former to the latter.

Good quality data are available on this flow owing to the peculiar nature of the Italian-Swiss border, which differs from all the other Italian land borders as it is the only one which is also an external border of the EU. In fact, while people can move freely within the EU according to the principle of non-discrimination between foreigners and nationals, the Swiss Confederation though in principle allowing the free movement of persons has a peculiar system of residence and work based on the granting of work permits.

The primary reference source on in-flows and out-flows from Switzerland is the Federal Statistical Office (FSO), which is part of the Federal Department of Home Affairs, and is in charge of the production and dissemination of official statistical data on the status and development of the country. Swiss data can be accessed via the FSO web site, fully available in both French and German; time series are available as well as metadata on the different surveys.

Tables can be produced by using individual combinations of the variables, i.e. of the characteristics of the selected data, and they can be exported in various electronic formats. As the FSO is a producer of official statistics, the quality of its data is guaranteed by internationally accepted standard methods and procedures.

Information on the flow of cross-border workers from Italy to Switzerland is provided quarterly by the ‘Cross-border Commuters Statistics’ (CCS), for which data have been updated regularly since 2004 and retrospective overall figures have been calculated from 1996([53]). The CCS are summary statistics prepared using data from the Central Migration Information System (SYMIC), AHV/AVS data (concerning social insurance), the vocational education database, and for time series prior to the 4th quarter 2010, the Job Statistics (JOBSTAT). From 2010, the main sources for the cross-border statistics are exhaustive registers. Hypotheses are made for the groups of cross-border commuters who are not or only partially included in AHV/AVS data (these groups represent only 2-3% of all cross-border commuters).

The data accessible on the official Swiss web site almost always provide information on the phenomenon as a whole. It is however possible for some tables to obtain information on commuters by country of residence, by gender, by canton of work and by age class. Data are disseminated on a quarterly and annual basis.

The unit of analysis is each foreign cross-border worker, that is an employed person of foreign nationality bearing a specific category of work permit (‘G permit’ holder) and gainfully employed in Switzerland. The G permit, i.e. the cross-border commuter permit, is issued by the competent Swiss canton only on the basis of an employment contract secured with a Swiss employer; its type and duration are linked to the

([52]) Namely ‘bilateral agreements’. For Switzerland see: https://www.eda.admin.ch/dam/eda/it/documents/publications/EuropaeischeAngelegenheiten/FS-Bilaterale_it.pdf

([53]) https://www.bfs.admin.ch/bfs/en/home/statistics/work-income/surveys/ggs.html characteristics of the contract. Self-employed workers also need a work permit, which is issued to them according to only slightly different rules. Rules for issuance mainly differ according to the duration of the activity, the following three cases being foreseen:

- if the period of employment is less than three months per calendar year, then no work permit is needed but the notification procedure applies;
- if a fixed-term employment contract is valid for less than a year but for more than three months, then a 'short permit' is issued with period of validity identical to the duration of the contract;
- if the contract has a duration of more than a year or in case of a permanent employment contract, then a 'long permit' called 'G permit' is issued with validity of five years.

It should be highlighted that most cross-border commuters are G-permit holders, since Swiss employers are more likely to hire workers on a permanent basis, much more often than in Italy at least, as they retain a substantial freedom to lay off employees with relative ease.

Given the importance of the G permit for estimating the flows of commuters, it is important to note that there is a possible risk of overestimation of these flows relying on the data on permits, for two reasons. Firstly, workers who lose their job before the end of validity of the permit tend not to declare it, though communication is in principle compulsory. The lack of communication is not sanctioned and the worker is not encouraged to abide by the rule; rather the opposite is true, because in case he/she resumes employment before expiration, the formalities for the issuance of a new permit could be avoided. Secondly, if the workload changes in terms of hours worked, the worker is similarly not encouraged to declare it. Therefore, for instance, if the workload becomes lighter and is carried out only for some days over the week, the estimate of the total number of cross-border workers should not change but there would be a risk of overestimation of the total labour input (or risk of underestimation in the opposite case).

Another source providing official data is the Statistical Department of Canton Ticino (USTAT) of the FSO located in Bellinzona, that is the competent authority for cantonal official statistics. USTAT elaborates its own data for cross-border workers in Canton Ticino by processing the data released by the FSO. With reference to the territory of Lombardy and Piedmont, this information is very valuable because nearly all Italian commuters from these two Italian regions crossing into Switzerland are employed in Ticino. The data released provide highly detailed information such as, for example, the Italian province of residence of cross-border workers (NUTS 3 level)([54]).

As far as Italian sources are concerned, in Italy at present there is no official ad hoc data collection and no ad hoc register, either statistical or administrative, to measure the phenomenon of cross-border workers. Data on cross-border workers are collected every ten years by ISTAT with the Population Census and, more frequently, with the Labour Force Survey (LFS), which is a sample survey; the use of the LFS data is thus subject to the constraints and limitations of the sample results([55]). LFS data at the level of one entire country are not very significant at the local level due to the sampling nature of the statistical survey and the small scale of the phenomenon analysed. The design of the sample involving the rotation of small municipalities can affect its capacity to measure adequately and constantly through time highly concentrated phenomena at territorial level.

It should be pointed out that the IWGCW identified also a number of additional administrative sources on cross-border workers, namely the tax files, the Survey on tourism (Bank of Italy), the Balance of Payments and other sources at local level. The information potential of all these sources to analyse the phenomenon under consideration is however limited, as they have been designed to fulfil different needs.

([54]) Data are available at http://www3.ti.ch/DFE/DR/USTAT/index.php?fuseaction=temi.dati&p1=35&p2=151&p3=160&proId=159

([55]) Official data from the last Population Census and the Labour Force Survey are available in the ISTAT data warehouse: http://dati.istat.it.The data on cross-border workers presented in this paper are the results of the elaborations of the authors.

3. Main findings from official statistical sources

This paper presents the estimates of the flows of cross-border workers between Italy (Lombardy and Piedmont) and Switzerland obtained for different years by integrating Italian and Swiss sources; further, it analyses the distribution of cross-border work across economic activities as well as demographic variables obtained from population census data.

Several factors may contribute to the decision of cross-border commuting. In general, the greater the difference in average earnings or the availability of work, the more likely the border country with more favourable labour market conditions will be to attract cross-border commuters. Cultural affinities and most of all a common language also do play a role in influencing cross-border flows in the Swiss-Italian alpine border area. However, so do other important factors such as good transport infrastructures and the characteristics of the mountain relief, which are very favourable between the Swiss Italian-speaking Ticino and the Italian provinces of Varese and Como.

From the juridical point of view, in recent times flows have been facilitated by bilateral agreements on the free movement of persons between Switzerland, EU and Member states. Nevertheless, Switzerland in practice still maintains a peculiar regime of administrative rules for cross-border workers. Therefore, specific legal rules apply to them and more data are available to monitor the flows in this area as compared to the other border regions in the Alps.

In the following paragraphs we present the estimation of the out-flows of frontier workers from Italy to Switzerland (source: FSO) and their demographic and socio-economic characteristics (elaboration on Italian Population Census data). We also introduce a focus on Italian frontier workers in Canton Ticino (source: USTAT) and on the Swiss frontier workers in Italy (source: FSO, Swiss Population Census).

3.1. Cross-border workers from Italy to Switzerland

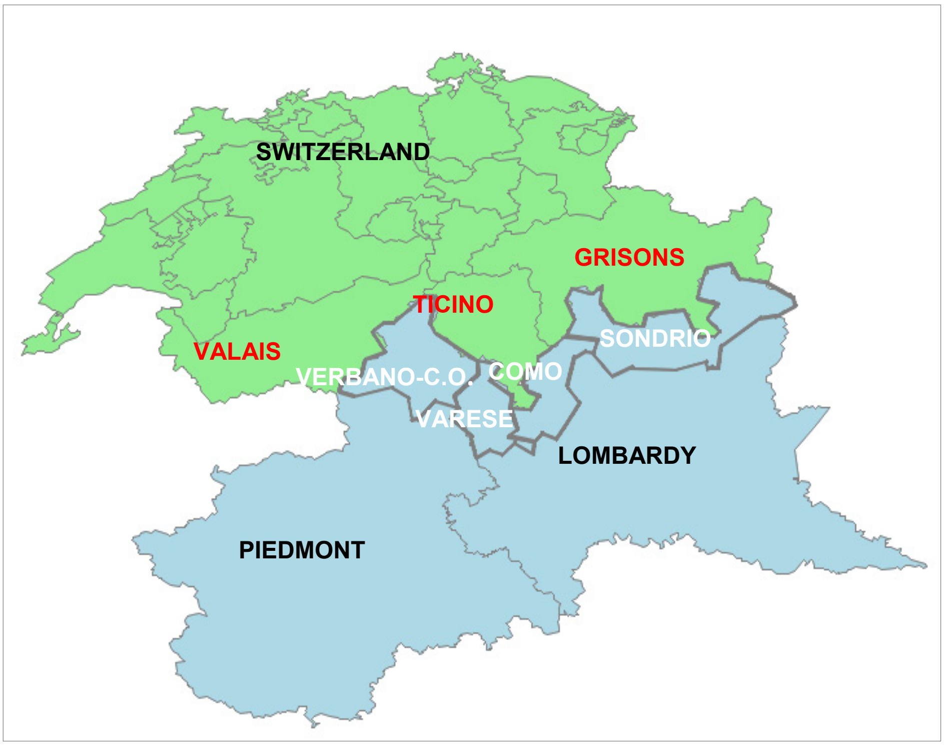
According to Swiss official sources, there were 69 761 cross-border workers from Italy to Switzerland in 2015 (up from 41 072 in 2006). This outbound flow mainly originates from a few provinces in border areas, where it represents a relatively significant share of the employed population in the local labour market, as compared to the national level. In fact evidence from Italian LFS data shows that while the overall figure nationwide only reaches about 0.3 % of the total employed population, in the most affected border provinces percentages are definitely higher: 8.8 % for Verbano-Cusio-Ossola, 8.1 % for Como, 6.0 % for Varese, 5.2 % for Sondrio.

The main destinations are three Swiss cantons: Ticino, Grisons, Valais (Figure 2).

In the same year 2015, Canton Ticino received 62 564 of total cross-border workers from Italy (i.e. 90 % of the total). By far less significant numbers of Italian workers were commuting to the other two main cantons of destination: around 5 000 to the Canton Grisons and around 1 300 to the Canton Valais.

The proportion of cross-border workers in the three cantons has remained almost stable in the last decade (Table 1).

Figure 2: Map of border area between Italy (Lombardy and Piedmont) and Switzerland with provincial borders



Source: Elaboration of the authors using SAS mapping software

Table 1: Out-flows of Italian cross-border workers by canton of destination: trend in the last decade (absolute values)

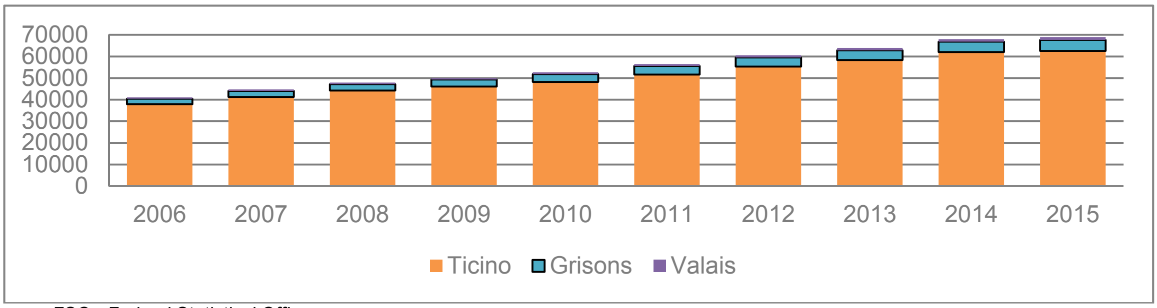
Year	Ticino	Grisons	Valais	Total
2006	37 835	2 588	638	41 072
2007	41 262	2 823	695	44 800
2008	44 195	3 021	677	48 006
2009	46 056	3 255	679	50 224
2010	48 206	3 624	791	52 963
2011	51 618	3 987	928	56 892
2012	55 300	4 256	1 032	60 976
2013	58 358	4 506	1 071	64 439
2014	62 042	4 796	1 212	68 753
2015	62 564	5 072	1 321	69 761

Source: FSO - Federal Statistical Office

The total number of Italian residents commuting to Switzerland for work has increased in the last ten years by around 70 %. Frontier workers to Grisons and Valais in particular have doubled (Figure 3).

To help understand and explain the differences of the flows between the Cantons, it is useful to mention here that transit is not so easy at the Swiss border of Valais and Grisons with the other Italian border provinces. In fact, especially during the winter, travelling by road from Italy into these two cantons can become impossible or extremely difficult owing to the adverse topography and weather conditions, namely heavy snowfall. Besides language barriers, this is therefore an area unlikely to attract high rates of commuting; as such a decision may often involve a relatively time-consuming and potentially risky journey.

Figure 3: Out-flows of Italian cross-border workers by canton of destination: trend in the last decade (absolute values)



Source: FSO - Federal Statistical Office

Using data from the last Italian Population Census, it is possible for the year 2011 to cross check the estimations of the flow and to investigate further the profile of Italian outbound commuters. Since the primary reason for Italians to commute to Switzerland is work (97.7 % of cases), the analysis of census data will only focus on cross-border workers.

The total out-flow of Italian cross-border workers is about 55 000, of which about 21 000 are women. These figures seem quite consistent with the Swiss side (approximately 56 900, of which about 22 000 are women) in terms of measuring the phenomenon as a whole and by gender breakdown, with due allowance being made for the fact that in the two countries the estimations are derived from two different statistical sources. The higher figure provided by the FSO, derived from summary statistics, could be linked among other reasons to the overestimation of the number of valid G permits.

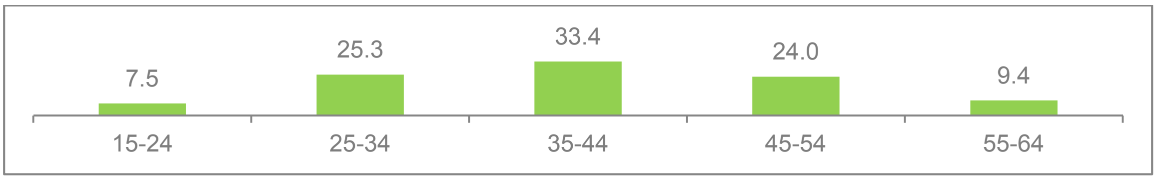
From the data analysed (Figures 4, 5 and 6), we can observe that cross-border workers are mostly men (61.8 %), middle-aged (33.4 % is between 35 and 44 years old), married (50.1 %), with low education (6.5 % completed only primary school, 41.3 % completed only the lower secondary cycle). In Lombardy they mainly reside in the provinces of Varese and Como, in Piedmont in the province of Verbano-Cusio-Ossola (frontier workers from this province are also likely to commute to Ticino). It should be noted that the percentage of people commuting from the province of Sondrio (commuting most probably to Grisons, which is the closest and most accessible canton to them) is not insignificant, and so is the percentage from Milan, which has good rail and road connections. Generally speaking (in 85 % of cases), travel times do not exceed the one hour threshold. A noteworthy exception is the case of workers coming from the city of Milan, which is quite far away from the border but benefits of fast rail connections to Ticino. For instance, travel times by rail from Milan to Bellinzona and Lugano take both approximately one hour and a half.

Analysing further the characteristics of Italian cross-border commuters, we observe that most of them are in employment (97.8 %), with permanent (86.5 %) and full-time contracts (86.5 %). The working week for 80 % of those who are employed is between 40 and 50 hours.

Almost half of the commuters are blue-collar workers, of which 27.1 % are skilled workers and 21.6% unskilled. Lower percentages are employed in service and sales activities (13.8 %) and as technicians and associate professionals (10.5 %).

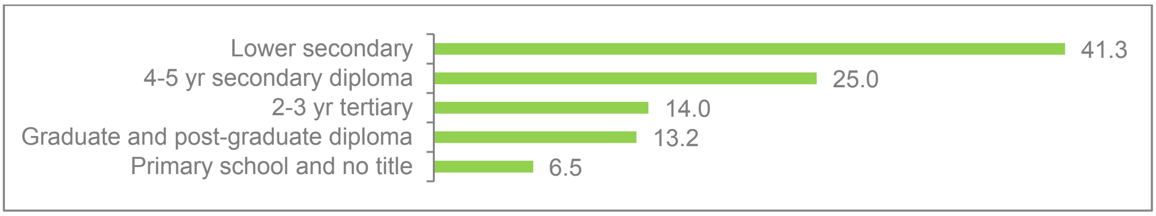
In terms of sectors of activity, Italian cross-border workers are mainly employed in manufacturing, repair and installation of machinery and equipment (28.4 %), followed by the construction of buildings, civil engineering and specialized construction activities (20.1 %).

Figure 4: Italian cross-border workers by age group (%)



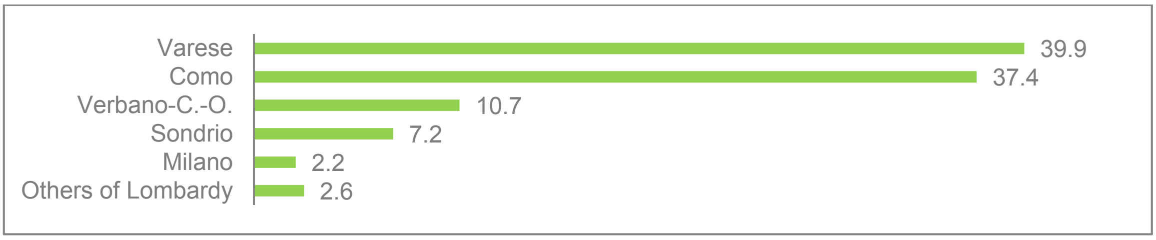
Source: Elaboration on data from ISTAT Population Census 2011

Figure 5 – Italian cross-border workers by educational level (%)



Source: Elaboration on data from ISTAT Population Census 2011

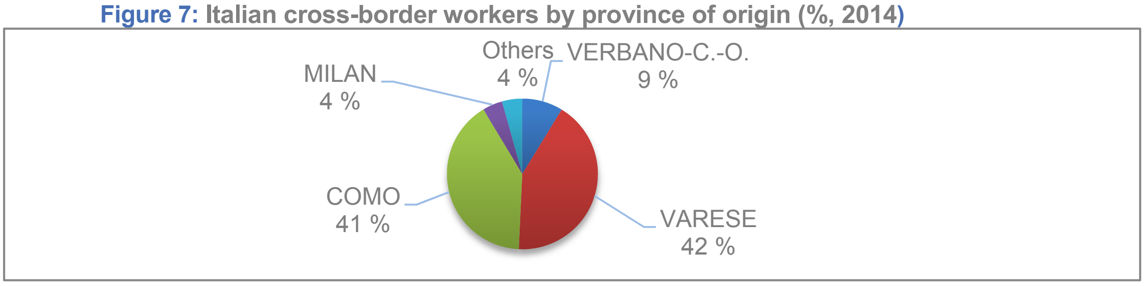
Figure 6: Italian cross-border workers by province of origin (%)



Source: Elaborations on data from ISTAT Population Census 2011

Focus on Canton Ticino

Figure 7: Italian cross-border workers by province of origin (%, 2014)



Source: USTAT - Statistical Department of Canton Ticino of the Swiss Federal Statistical Office

USTAT elaborations provide highly detailed data, such as the breakdown by Italian province of residence of the workers and by economic activity. Figure 7 shows the percentages of commuters with breakdown by main province of origin for 2014. We can see that the provinces with the most important out-flows are Varese (42 %) and Como (41 %), followed by the province of Verbano-Cusio-Ossola (9 %). We have also included the figure for the province of Milan (4 %), though its territory is not strictly located in the border area; nevertheless, good train and road connections to Switzerland facilitate commuting.

3.2. Cross-border workers from Switzerland to Italy

On the other hand, considering the flow of commuters from Switzerland to Italy, we can obtain data from the Structural Survey (SS), which has been implemented annually since 2010 within the framework of the new Swiss Population Census. The SS is a sample survey conducted on at least 200 000 people by means of a written questionnaire or through internet([56]).

([56]) Information on the new Swiss Population Census is available at:
https://www.bfs.admin.ch/bfs/en/home/statistics/population/surveys/census.html

Table 2: Swiss cross-border workers to Italy

Year	*Estimate of number of persons*	*Confidence interval (95 %)*
2010	*1 067*	*25.2 %*
2011	*1 314*	*23.1 %*
2012	*1 170*	*24.9 %*
2013	*1 704*	*21.1 %*
2014	*1 004*	*25.3 %*

Source: FSO - Federal Statistical Office

Swiss Population Census data show a small out-flow to Italy (table 2). In 2014, 1 004 Swiss citizens were commuting to Italy, with a considerable decrease from the previous year. Figures for recent years however do not show a steady trend.

4. Conclusions

Cross-border work is a topic of interest at European and national levels for a number of reasons. For EU integration, frontier workers are a testing ground, because they experience directly the difficulties which need to be addressed to fully realise the free movement of persons. At national level, improving knowledge of this phenomenon including its statistical measurement may be important to understand and portray correctly the dynamics of the labour market. From the national accounts point of view, it may contribute to improving the estimation of labour input underlying gross domestic product.

In large countries like Italy, the relevance of cross-border work on the total workforce is usually low. The focus of our research was deliberately on investigating its quantitative relevance at the local level. In Italy, cross-border workers are mainly found in the Alpine area and their number is significant at provincial level in terms of outbound flows to Switzerland. These flows show an increasing trend and can be monitored on the basis of Swiss official statistical sources, for which quality is guaranteed by standard techniques and procedures. This is not the case for other Italian borders, where no equally reliable sources are available (except for the census year) and estimates can only be attempted on the basis of unofficial information or expert opinions. Previous Istat research suggests that for other borders in Italy the number of commuters is not likely to be significant; nevertheless, this is a phenomenon that affects our ability to provide a correct picture of the local labour market and can help explain differences in wealth between peripheral provinces. Further, it could be subject to changes in accordance with the economic cycle or other drivers of the geographical mobility of the workforce.
Currently neither the Labour Force Survey nor the existing administrative sources allow us to obtain a reliable estimate of the flows of cross-border workers. Data from the population census do allow such estimate and also much more in terms of information on their demographic and socio-economic characteristics, but they are only available every ten years. For 2011, the only recent year for which data are available for the comparison, the integration of data on the Swiss and Italian sides has allowed to cross-check the total estimation of the flow and draw a profile of Italian commuters with fairly detailed demographic and socio-economic variables. The comparisons of Swiss and Italian sources carried out by the authors show that the Italian Population Census is a reliable source, providing consistent estimates in terms of level and gender breakdown for the phenomenon.

However, monitoring regularly and reliably frontier workers at less conspicuous intervals would require the design and implementation of ad hoc statistical methodologies allowing the integration of sample and administrative data, specifically aimed at estimating small-scale, locally concentrated phenomena. At the same time, an important issue for future consideration are the gains and losses in terms of measurement potential as new sources are designed and implemented.

References

Alberton, S., Gonzalez, O. and Guerra, G., *Le trasformazioni nel mercato del lavoro ticinese*, CODE - IRE, Lugano, 2008.

Bednarz, F., 'Lavorare in Svizzera, fra tradizione e nuove opportunità', in *Il capitale umano*, Centro Studi dell'Economia Comasca, Camera di Commercio, Industria, Artigianato e Agricoltura di Como, 2013.

Fries Tersch, E. and Mabilia, V., *Annual Report on Labour Mobility*, European Commission, 2015.

Gonzalez, O., 'La vigorosa progressione dei "nuovi" frontalieri' in *Ticino. Chi sono e dove trovano impiego?,* 'Dati', XIII, n. 01(45-53), USTAT, Giubiasco, 2013.

InArco Project*, Information Actions for Cooperation regarding Cross-border workers in Alpine Arch Countries*, project co-financed by the European Commission, DG Employment, Social Affairs and Inclusion under the Call for proposals VP/2012/004.

ISTAT Working Group on Cross-border Workers, *La stima del lavoro frontaliero: analisi fonti e rilevazione dei flussi in entrata e in uscita*, ISTAT internal report by the Territorial Office for Lombardy for the project *Development of an integrated data base on frontier workers*, April 2014.

ISTAT, Territorial Office for Lombardy and USTAT, *Il mercato del lavoro dell'area transfrontaliera insubrica. Annuario statistico integrato*. ISTAT, Rome 2004.

ISTAT, Territorial Office for Lombardy and USTAT, *Il mercato del lavoro dell'area transfrontaliera insubrica. Annuario statistico 2005*. Comunità di lavoro Regio Insubrica, 2005.

ISTAT, Territorial Office for Lombardy and USTAT, *Il mercato del lavoro dell'area transfrontaliera insubrica. Annuario statistico 2006*. Comunità di lavoro Regio Insubrica, 2006.

Istituto di Ricerche Economiche, Osservatorio delle politiche economiche O-Pol., *Competitività Economica 2011*. *Rapporto sulla struttura economica ticinese*, Università della Svizzera italiana, Lugano 2011.

Jorens, Y., Minderhoud, P. and De Coninck, J., *Comparative Report: Frontier workers in the EU*, FreSco, European Commission, January 2015.

Nerb, G, Hitzelsberger, F., Woidich, A., Pommer, S., Hemmer, S. and Heczko P., *Scientific Report on the Mobility of Cross-Border Workers within the EU-27/EEA/EFTA Countries*, Munich: MKW GmbH. 2009.

Pierini F.*, Les travailleurs frontaliers dans l'Union européenne*, vol. W – 16. Luxembourg: Parlement Europeen - Direction Generale des Etudes, 1997.

Progetto Interreg III, ISTAT and USTAT, *Indicatori e fonti statistiche del mercato del lavoro transfrontaliero*, Documenti di lavoro USTAT n. 6, Bellinzona 2004.

Progetto Interreg III, ISTAT and USTAT, *Rapporto metodologico: procedure di armonizzazione e proposte di integrazione*. Documenti di lavoro USTAT n. 8, Bellinzona, 2004.

Regulation (EC) No 883/2004 of the European Parliament and of the Council of 29 April 2004 on the coordination of social security systems.

Stephani, E. and Mulatero, F., *Disoccupazione: Ticino e Lombardia si allineano. Una lettura transfrontaliera dei dati della rilevazione sulle forze di lavoro*, 'Dati', XIII, n.01 (35-43), USTAT, Giubiasco, 2013.

Touring Club Italiano, *Guida d'Italia: Lombardia*, Milan 2008.

USTAT, *Panorama statistico del mercato del lavoro ticinese*, December 2013.

6 Enhancing statistical culture — the unused potential of Croatian enterprises

KSENIJA DUMIČIĆ, BERISLAV ŽMUK[57]

Table of content

([57]) Department of Statistics, Faculty of Economics and Business, University of Zagreb, Croatia: kdumicic@efzg.hr, bzmuk@efzg.hr

Abstract

Managers in enterprises to get additional information, which can be used to improve business decision-making process, may use statistical methods. On that way, an enterprise can achieve better business results. However, the less often statistical methods are used, the more potential of achieving better business results remained unused. The conducted web survey on the sample of Croatian enterprises has shown that there is a lot of such unused potential. The main problem, as estimated in this survey, is that vast majority of enterprises does not use statistical methods in their businesses. Consequently, the first step of scientists and statisticians should be to convince managers in enterprises that statistical methods could be useful in their businesses.

Keywords: statistical methods use, web survey, Croatia, enterprise.

Acknowledgement: This work has been partially supported by Croatian Science Foundation under the project Statistical Modelling for Response to Crisis and Economic Growth in Western Balkan Countries – STRENGHTS (project no. IP-2013-9402).

1. Introduction

In order to be successful on a modern market, an enterprise must be competitive. The more an enterprise competitive is, the better business results it can expect to have. The level of competitiveness can be achieved on different ways, but it is mainly achieved by higher quality and/or lower price of products and services (O'Donnell et al., 2002). In order to achieve higher quality and lower price an enterprise can use different tools and methods. Consequently, one way to be more competitive is to use statistical methods in everyday business process.

Previous researches have shown that use of statistical methods has positive impact on business results and increases probability of achieving positive net income (Žmuk, 2015a, 2015b). It has been shown that the positive impact of statistical methods use on business results lies in the fact that statistical methods can provide either a new insight, or at least different view, at the certain business process or some problems in conducting the business. On that way, business managers can get additional information, which help them to make better business decisions.

In Ćurlin (2006) and Dumičić, Knežević (2007), forecasting is based on statistical methods application approach, but predicting or foreseeing was defined as an estimation of impact of different factors to the future business with application of qualitative methods based on the subjective estimation. The survey research on Croatian large and middle-sized enterprises' praxis of predicting the changes in business making and business environment is presented. The business sample survey revealed that Croatian large and middle-sized firms mostly used to make predictions of changes by their own staff, based on external information sources, market research and subjective estimation. The difference in approach between enterprises based on their ownership was found. Private enterprises used to make different forecasting approach of business making compared to those in public or mixed private-public ownership.

Large number of changes in short time has become part of the modern business making for most of the firms. The paper by Pejić Bach, Dumičić and Gogala (1999) presents the results of research whose aim was to determine how the biggest Croatian companies are preparing sales forecasts, which methods are used and what is the accuracy of the methods used. In particular, the relationship to forecast sales and information technology and the factors that influence the use of quantitative statistical methods to forecast sales are discussed. Research has shown that Croatian enterprises are commonly used subjective methods of forecasting sales and the use of information technology in forecasting sales of common practice especially in the application of quantitative methods. Use of quantitative methods increases with an increase in the total number of used forecasting methods. There is a correlation between the degree of conversion of enterprises and the use of quantitative methods, which are usually used by companies whose conversion is fully completed, as well as public companies, and rarely companies in which the conversion process is still ongoing.

In the paper, the level of acceptance of statistical methods is observed on the case of Croatian enterprises only. According to the Croatian economic situation and its development (Moore, Vamvakidis, 2007, Sever, Drezgić, Blažić, 2011), it is assumed that enterprises in Croatia did not have recognized usefulness of statistical methods use.

The research question set in this paper is whether enterprises have recognized possible advantages of statistical methods use or they still have high aversion to statistical methods applications as something, which is too complicated to be used. The more the enterprises use statistical methods, the higher are the benefits of statistical methods use on a country level. However, if vast majority of enterprises do not use statistical methods, a certain potential of benefits, which can be achieved by using statistical methods, remains unexploited. On that way, a country cannot achieve its maximum competitive level what has strong impact on the development of that country.

This paper is organized as follows. After the brief introduction into the research problem, in the second chapter description of data and methods that will be used in the analysis is provided. In the third chapter, results of conducted web survey among Croatian enterprises are presented and discussed. In the final chapter, conclusions are brought.

2. Data and methods

In order to inspect characteristics and development level of statistical culture in Croatian enterprises a web survey (Couper, 2008) by using a simple random sampling (Kish, 1995) was conducted.

Overall 26,186 Croatian enterprises, that are registered in the Court Register of the Republic of Croatia as limited liability enterprises (Official Gazette, 2011) and that are subject to the submission of annual financial statements in accordance with the Accounting Act (Official Gazette, 2007), were contacted by e-mail starting in October 2012, and finishing in February 2013. In the invitation e-mail there was a unique hyperlink for each enterprise, which led them to the questionnaire. The respondents were determined as managers at any of the management levels or persons who are competent to speak about statistical methods use in their enterprise.

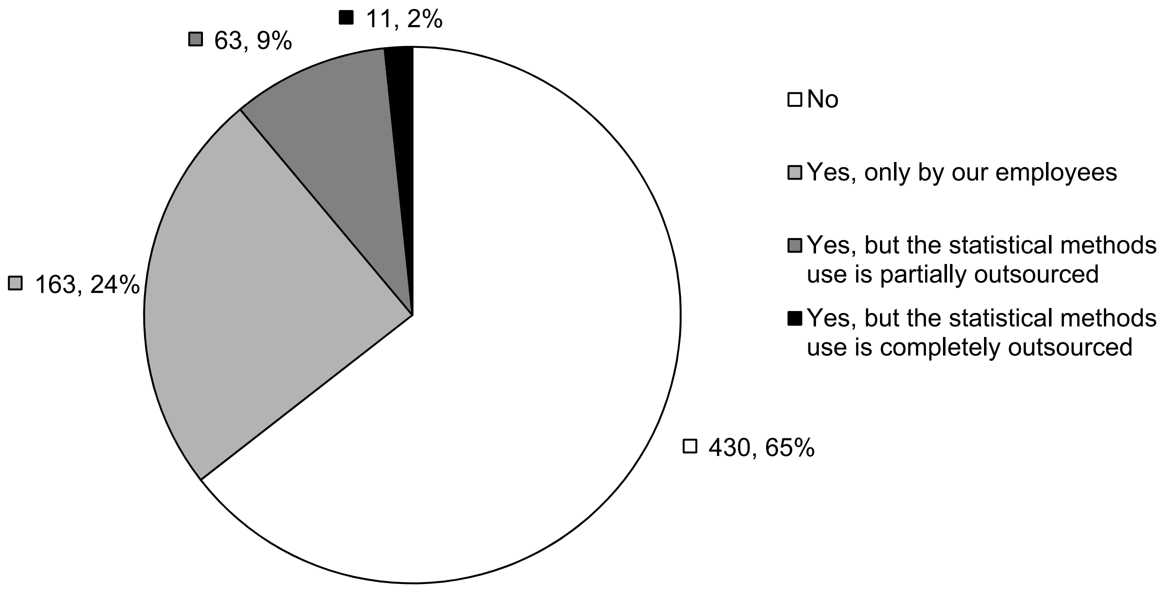
There were two questionnaire versions. The first one was designated for enterprises, which use statistical methods whereas the second one was designated for enterprises, which do not use statistical methods. In order to distinguish between enterprises which use and which do not use statistical methods the first question in the survey was a filter type of question. Depending on the answer on that filter question, an enterprise was redirected to the first or to the second version of the questionnaire. The questionnaire version for enterprises that use statistical methods is consisted of overall 28 questions, whereas the questionnaire version for enterprises that do not use these methods is shorter and it is consisted of 11 questions, only. It has been estimated that enterprises, which use statistical methods would need about 15-20 minutes to complete the questionnaire whereas 5 minutes is the threshold for enterprises, which do not use statistical methods.

Overall 667 Croatian enterprises participated in the survey. Consequently, the response rate of the survey is 2.55 %, being the percentage of eligible sample units that cooperate in a survey. In this paper, the analysis of the collected data will be conducted by using selected inferential statistics methods (Field, 2011).

3. Analysis and discussion

As mentioned before, the first question was a filter question. This filter question enabled appropriate different analysis approach for enterprises that use and that do not use statistical methods. There were four possible answers on the filter question whereas three answers were oriented to those enterprises that use statistical methods, in one of the three presented ways, and one answer was applicable for enterprises, which do not use such methods. It has to be emphasized that a brief explanation, what under statistical methods use has to be meant in the survey, was provided in the questionnaire. Figure 1 shows the distribution of answers on the filter question.

Figure 1: Distribution of Croatian enterprises whether they use or do not use statistical methods in their business processes, nTotal = 667



Source: survey.

According to Figure 1, the survey results have shown that the vast majority of 64.47 % enterprises in the sample do not use statistical methods (SE (proportion) = 0.018, CV (proportion) = 2.84 %). On the other side, in 24.44 % of enterprises, the statistical methods are used only by employees (SE (proportion) = 0.016, CV (proportion) = 6.72 %). In 9.45 % of enterprises statistical methods applications are partially outsourced (SE (proportion) = 0.011, CV (proportion) = 11.84 %), whereas in 1.65 % of enterprises statistical methods use is completely outsourced (SE (proportion) = 0.005, CV (proportion) = 29.52 %). These results are speaking in great favor of a lack of statistical method use by Croatian enterprises. In Table 1 proportion point estimates and 95% confidence intervals, based on the normal approximation for the sampling distribution, with the coefficient of confidence of z = 1.96 (Cochran, 1977) for the filter question is given.

Table 1: Proportion estimates and 95 % confidence limits for the filter question whether Croatian enterprises use or do not use statistical methods in their business processes, $n_{Total} = 667$

Are statistical methods used in your enterprise?	Number of enterprises in the sample	Estimated proportion of enterprises	Standard error	Coeff. of variation (%)	Margin of error	95 % confidence limit for the proportion	
						Lower limit	Upper limit
No	430	0.6447	0.0183	2.84	0.0366	0.6080	0.6813
Yes, only by our employees	163	0.2444	0.0164	6.72	0.0330	0.2114	0.2773
Yes, but the statistical methods use is partially outsourced	63	0.0945	0.0112	11.84	0.0227	0.0718	0.1171
Yes, but the statistical methods use is completely outsourced	11	0.0165	0.0049	29.52	0.0103	0.0062	0.0268

Source: survey.

Overall, 237 enterprises in the sample are using statistical methods, in one or another way, in their business processes. However, 121 enterprises think that statistical methods are not used enough in their enterprise and that there is space for more intensive use of such methods. On the other side, 87 enterprises support the opinion that statistical methods in their enterprise are used at satisfactory level. However, 29 enterprises were not able to put themselves on one or another side or they have provided “do not know” answer. If those unsure enterprises are left from the further analysis, by using statistical one-tail proportion test, at 5 % significance level, it may be concluded that in more than 50 % of enterprises, in which statistical methods are used, they are not used at satisfactory level ($H_0 \ldots p \leq 0.5$, $H_1 \ldots p > 0.5$, standard error = 0.0355, z-value = 3.1269, p-value = 0.0009).

Table 2: Frequency of statistical methods use in Croatian enterprises, nUsers = 237

Statistical methods	Frequency of statistical methods use									
	Monthly		Yearly		Once in 3 years		Rarely than once in 3 years		The statistical method is not used	
	No.	%	No.	%	No.	%	No.	%	No.	%
Descriptive statistics	127	54	52	22	8	3	9	4	41	17
Outlier detection methods	23	10	23	10	5	2	7	3	179	76
Inferential methods	12	5	20	8	7	3	8	3	190	80
Sampling methods	36	15	43	18	10	4	9	4	139	59
Clustering methods	12	5	22	9	5	2	10	4	188	79
Design of experiments	9	4	14	6	6	3	6	3	202	85
Statistical process control	22	9	21	9	5	2	9	4	180	76
Acceptance sampling	20	8	19	8	9	4	11	5	178	75
Indicators of dynamics	51	22	41	17	7	3	8	3	130	55
Naive forecasting models	35	15	28	12	9	4	6	3	159	67
Smoothing forecasting models	13	5	18	8	5	2	7	3	194	82
Trend forecasting models	43	18	39	16	9	4	13	5	133	56
Regression models	16	7	16	7	6	3	8	3	191	81

Source: survey.

In Table 2, frequency of the observed statistical methods use in Croatian enterprises is analyzed. According to the results presented in Table 2, it may concluded that statistical methods, except those descriptive ones, are very rarely, if they even are, used. Descriptive statistical methods are used by more than 50 % of enterprises in the sample, among those that use at least one statistical method in their business processes on monthly base or even often. It has been expected that the majority of enterprises would use descriptive statistical methods, because these methods are basic and unavoidable part of each statistical analysis, but it has not been expected that there is so great share of those enterprises, which even do not use other considered statistical methods.

Table 3: **Motivation for increased and more intensive use of statistical methods in Croatian enterprises, n_{Users} = 237**

Motivation reason	Agree		Do not agree		Do not know	
	No.	%	No.	%	No.	%
Stricter legislation	83	35	103	43	51	22
Increased market demands	205	86	13	5	19	8
Better education level of employees	178	75	30	13	29	12

Source: survey.

In Table 3, the results for motivation reasons for increased and more intensive use of statistical methods in enterprises are presented. The results suggest that motivation for increased and more intensive use of statistical methods by enterprises is seen in an increased market demands and in better education level of employees, whereas legislation is not a key motivation for more intensive use of statistical methods. These results are going in favor of the conclusion that Croatian enterprises are starting to recognize why they should use statistical methods in their everyday business processes at the greater extent.

Table 4: **Attitudes toward reasons for initial statistical methods use, nUsers = 237**

Reason	Agree		Do not agree		Do not know	
	No.	%	No.	%	No.	%
Statistical methods facilitate the business decision-making process	213	90	11	5	13	5
Statistical methods usage improves business results	194	82	22	9	21	9
Statistical methods application leads to cost efficiency	177	75	26	11	34	14
The management demands statistical methods usage	99	42	116	49	22	9
Statistical methods use has been a prerequisite for business certification	59	25	145	61	33	14
Statistical methods have been used by competitors	52	22	115	49	70	30

Source: survey.

As regards the reasons for initial statistical methods use, the majority of enterprises have recognized real reasons for statistical methods. So, according to the results from Table 4, the most enterprises that are users of statistical methods use them because the applications of these methods facilitate the business decision-making process, improve business results, and lead to cost efficiency. On the other side, the most of corresponding enterprises that are users of statistical methods do not consider or neglect the following reasons as key ones for statistical methods use: the management demands statistical methods usage, statistical methods have been used by competitors, and statistical methods use has been a prerequisite for business certification.

In the selected enterprises, which do not use statistical methods, the respondents have been asked for reasons why they do not use statistical methods. According to the results from Table 5, they emphasized the current overload of employees with other jobs and assignments as a main reason why they do not use statistical methods.

Table 5: **Attitudes toward reasons for not using statistical methods, $n_{NonUsers}$ = 430**

Reason	Agree		Do not agree		Do not know	
	No.	%	No.	%	No.	%
The existing overloaded of employees with other jobs and assignments	275	64	116	27	39	9
The nature of products/services	247	57	90	21	93	22
Lack of financial resources	216	50	154	36	60	14
Lack of statistical methods use knowledge	210	49	153	36	67	16
Statistical methods usage would not have significant impact on business results	192	45	116	27	122	28
Lack of educated employees	164	38	204	47	62	14

Source: survey.

Statistical methods could be very helpful in making decisions in key business situations. Regarding this issue, among statistical methods users, 54 enterprises always use statistical methods whereas 145 enterprises use them only occasionally. On the other hand, 31 enterprises do not use statistical methods as additional tool for making decisions in key business situations. In the sample, seven enterprises were indecisive and gave “do not know” answer.

Table 6: **Attitudes toward reasons for a lack of statistical methods use, n_{Users} = 237**

Reason	Agree		Do not agree		Do not know	
	No.	%	No.	%	No.	%
The existing overloaded of employees with other jobs and assignments	164	69	57	24	16	7
Statistical methods are not well known in the enterprise	162	68	58	24	17	7
The lack of qualified employees	122	51	91	38	24	10
Lack of financial resources	112	47	101	43	24	10
The nature of the product / service	106	45	104	44	27	11
The additional statistical methods use would not have an impact on business results	51	22	150	63	36	15

Source: survey.

Similar to the question “why they do not use statistical method”, which was offered to the respondents in enterprises, which declared as the non-users of statistical methods, the question “why they do not use them more often”, or, to be more precise, “what are the reasons for a lack statistical methods use in their business processes” was asked the respondents in the enterprises which declared as the users of these methods. According to the results from Table 6, as the main reason for a lack of statistical methods use, the enterprises have emphasized the two following reasons: the existing overloaded of employees with other jobs and assignments, and statistical methods are not well known in the enterprise. It can be noticed that enterprises, which do not use statistical methods, have also marked of the employees work overload level as a main obstacle to the beginning of statistical methods use.

The confirmation that statistical methods are not well known in the enterprises is given by asking the respondents in the sampled enterprises about the average grade of employees' statistical methods use knowledge, which would they gave to their employees. According to the results presented in Table 7, the most of sampled enterprises among statistical methods users rated employees' knowledge about statistical methods use with really low grades, with the mode being assessed as "C", that might be recognized as the grade "good" on common rank scales, respectively.

Table 7: Average grade of employees' statistical methods use knowledge, n_{Users} = 237

Grade	No.	%
A	9	4
B	38	16
C	78	33
D	58	24
E	20	8
I cannot evaluate	34	14
Total	237	100

Source: survey.

Regarding enterprises' investments into employees' education in the field of statistical methods use, among 237 enterprises in the sample that comprised to be statistical methods users, 119 of them do not invest any amount for that purpose. In the sample, 66 enterprises in average per year invest into employees' education in the field of statistical methods use less than HRK 10,000 whereas 21 enterprises invest between HRK 10,000 and 50,000. Only one enterprise in the sample in average per year invests into employees' education in the field of statistical methods use more than HRK 50,000. However, 30 enterprises selected "do not know" answer.

From 237 enterprises from the sample, which declared as being the users of statistical methods, only 21 of them, or 9 %, are satisfied with statistical education possibilities available on the Croatian educational market. On the other side, 41 enterprises, or 17 % of enterprises that are statistical methods users, are not satisfied with statistical courses offered, which are available on the Croatian educational market. However, what is the most important is the fact that the rest of enterprises in the sample, 175 enterprises, or 74 % of enterprises that are statistical methods users, do not know anything about available statistical courses on the Croatian market.

4. Conclusions

Statistical methods can improve business results of enterprises by providing additional information, which is used in the decision-making process. However, the research results have pointed out that Croatian enterprises in vast majority do not use statistical methods. What's more, even enterprises, which use statistical methods, do not use their full potential.

The web survey on the sample of Croatian enterprises has shown that 2/3 of Croatian enterprises do not use statistical methods at all. Among enterprises that use these methods, 10% of them outsource statistical methods applications partially and 2% outsource them completely. Regarding the simplest statistical methods, the descriptive methods are used in 1/2 of enterprises that are declared as the statistical methods users, and dynamics indicators are applied in only 1/5 of them. The main reasons for statistical methods use appeared to be: they facilitate decision-making process and improve business results. In addition, enterprises, which are statistical methods users, do not use their full potential, with the main reason of the current overload of employees with other jobs and assignments.

Statistical knowledge of the employees was assessed to be at quite low level, with an average grade being C, i.e. “good”. Low investment level in employees' additional statistical education is recognized as a common attitude.

Based on the survey research results, the general recommendations for Croatian enterprises arose: the awareness of statistical methods usefulness should be improved, and, Croatian enterprises' intrinsic motivation to incorporate the tradition of investing into employees' continuous statistical education must be increased. Educational programs offered at the educational market should be fitted to the enterprises' needs. In other words, it might be concluded that there is a lot of space for improvements by using statistical methods in enterprises in Croatia. In order to improve existing situation, the emphasize must be given to the rise of enterprises' awareness of statistical methods application usefulness.

In addition, possibilities of attending statistical educational programs, in both methods and software applications, should be better advertised. There is a lot of work in front of educators, scientists and expert statisticians, official statistics stakeholders and statistical associations to convince managers in enterprises that statistical methods are just a tool, as part of overall business literacy, which could help them to make better business decisions. They should start planning more financial resources for employees to attend statistical educational programs in a common way. More statistical programs suitable for those already employed should be developed, as well as for those who have just entered the enterprises or took part in some kind of business activities. Furthermore, statistical associations should invest more efforts to promote usefulness of statistics, not only among enterprises, but also among general population. On that way different conferences, round tables, press releases and similar, will additionally promote and popularize statistics use in everyday life of individuals and of enterprises.

References

Cochran, W. G., Sampling Technique, New York: John Wiley and Sons, 1977.

Couper, M. P., Designing Effective Web Surveys. New York: Cambridge University Press, 2008.

Ćurlin, S., Primjena metoda predviđanja u strateškom planiranju hrvatskih poduzeća. Scientific master thesis. Zagreb: Faculty of Economics and Business, University of Zagreb, 2006.

Dumičić, K., Knežević, S., Survey research on Croatian enterprise's praxis of predicting the changes in business making and business environment. Ekonomski pregled 58(3-4), 2007, pp. 158-180.

Field, A., Discovering Statistics Using SPSS, London: SAGE Publications, 2011.

Kish, L., Survey Sampling, New York: John Wiley & Sons, 1995.

Moore, D., Vamvakidis, A., 'Economic Growth in Croatia: Potential and Constraints', IMF Working Paper WP/07/198, 2007, pp. 1-38

O'Donnell, A., Gilmore, A., Carson, D and Cummins, D., 'Competitive advantage in small to medium-sized enterprises', Journal of Strategic Marketing 10(3, 2002, pp. 205-223.

Official Gazette, Accounting act. 16(109), 2007.

Official Gazette, Company act. 20(152), 2011.

Pejić Bach, M., Dumičić, K. and Gogala, Z., Korištenje metoda prognoziranja prodaje u hrvatskim velikim poduzećima, Ekonomski pregled 50(7-8), 1999, pp. 780-803.

Sever, I., Drezgić, S., Blažić, H., 'Budget spending and economic growth in Croatia: Dynamics and relationships over the past two decades', Proceedings *of Rijeka Faculty of Economics: Journal of* Economics and Business 29(2), 2011, pp. 291-331.

Žmuk, B., 'Adoption and benefits of statistical methods in enterprises: Differences between Croatian regions', South East European Journal of Economics and Business 10(1), 2015a, pp. 55-65.

Žmuk, B., 'Business Sample Survey Measurement on Statistical Thinking and Methods Adoption: the Case of Croatian Small Enterprises', Interdisciplinary description of complex systems 13(1), 2016b, pp. 154-166.

7 Estimating Contributions to GDP Growth by Structural Decomposition of Input-Output Tables

KRISZTIÁN KOPPÁNY[58]

Table of Contents

([58]) Széchenyi István University, koppanyk@sze.hu

Introduction

This paper presents a case study[59] to demonstrate the calculation methods of GDP growth contributions using structural decomposition analysis (SDA) of input-output tables and their Hungarian applications. Although the required data are available with a considerable time-lag, results show that taking supplier relations and value chain multipliers into account can significantly alter the picture on growth effects of industries and final demand categories by the conventional approach. This can be instructive for analysts, policy and decision makers, not only in Hungary, but also in other countries. The study was performed by using public macroeconomic and sectoral data obtained from the Hungarian Central Statistical Office's (HCSO) dissemination database and STADAT tables.[60]

Input-output tables

Input-output tables are published with a much longer time-lag than flash estimates of GDP. Therefore the case study is not on the last quarter but, according to the annual horizon and the publication schedule of input-output tables,[61] on an earlier year.

For analysing GDP volume change, two successive years' input-output tables are needed, of which, the latter is expressed at previous year prices. In Eurostat Database only three countries, Denmark, Netherlands, and Romania have industry by industry tables of this type. Estimating a constant price table is a challenge for other countries.

At the time of writing this paper, the latest input-output table published by the HCSO is valid for the year 2010.[62] This is produced from the supply and use tables[63] by the "fixed product sales structure" transformation (*Eurostat* (2008), p. 351., Model D), calculated with 88 industries and published in a 65 by 65 aggregation depth.

Supply and use tables at current prices are available also for the subsequent two years, so using these and the method referred to above I could generate current price input-output tables for 2011 and 2012, as well. Data available for next steps required unification of industries 68A: Imputed rents of owner-occupied dwellings and 68B: Real estate activities (excluding imputed rents), so from this point I worked with 64 industries.

The former detailed dataset at previous year's prices was not available, consequently a constant price table for 2012 was developed from the 2012 year current price table with the RAS method (*Miller–Blair* (2009) sections 7.4.1–3) using the previous year price margins available in the dissemination database and STADAT (see Figure 1).

[59] This research was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and the PADS Foundation.

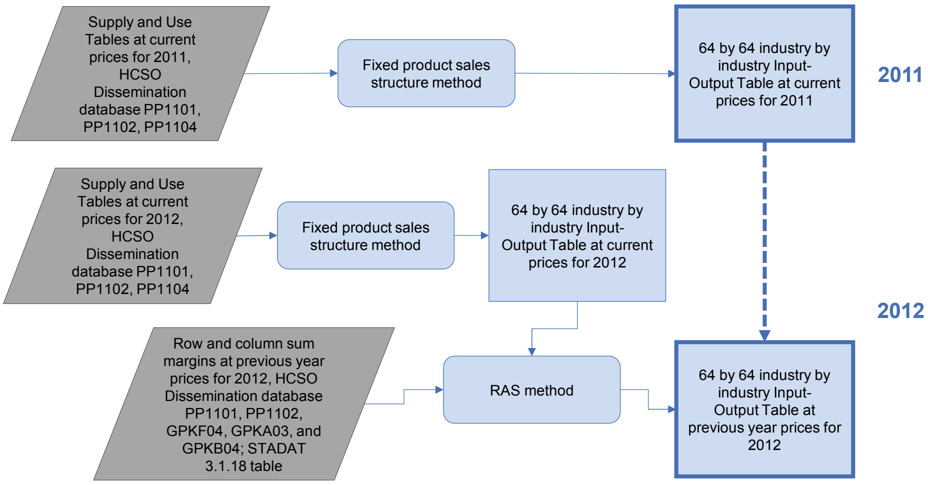
[60] This paper is a short version of *Koppány* (2016) published in the Hungarian Statistical Review, http://www.ksh.hu/statszemle_archive/2016/2016_08-09/2016_08-09_881.pdf. Full-length analysis in English is forthcoming in Acta Oeconomica. See these studies for technical background and results in detail.

[61] According to the European guidelines, input-output tables are published by HCSO every five years, with a three-year time lag.

[62] Dissemination database / National accounts, GDP / Input-output tables, supply and use tables / Symmetric input-output table (industry by industry), at current basic prices NACE Rev. 2 (ESA2010) (technical code PP1109)

[63] I.b. PP1101, PP1102 and PP1104.

Figure 1: Generating input-output tables



Although margins are available for 2013–2014 too, updating for these years, even for current price tables can only be done by RAS or other approximation techniques. This would made the results precarious. For this reason, I use 2011 current price and 2012 previous year price input-output tables (Table group 1) for the demonstration of the application of SDA, by which we can analyse GDP growth and growth contributions of the year 2012.

Due to the size of the tables, Table group 1 shows the simplified, four industry, three final demand component and only one value added row version of these, which will be of assistance to us in the demonstration and comprehension of the decomposition methods, and comparisons between the numbers by conventional and SDA techniques. In spite of the short form presentation of the data and the results as well, calculations are made on 64 industry levels.

Table group 1: Simplified input-output tables for Hungary
(billion HUF)

Simplified input-output table for base year 2011 at current prices

Industries	Intermediate consumption					Final use				Total use / output
	A Agriculture, forestry and fishing	B-E Mining; manufactu-ring etc.	F Construction	G-T Services	Total intermediate consumption	Household final consumption expenditures	Other domestic final demand	Export	Total final use	
A Agriculture, forestry and fishing	564	719	4	101	1 389	317	242	652	1 211	2 600
B-E Mining; manufacturing etc.	352	3 558	422	1 746	6 077	2 260	737	16 747	19 744	25 822
F Construction	3	67	63	217	349	26	1 947	98	2 071	2 420
G-T Services	256	2 382	447	6 077	9 163	7 583	6 819	3 979	18 382	27 545
Import	285	12 683	456	3 061	16 485	2 239	1 942	1 961	6 142	22 627
Taxes less subsidies on products	33	206	40	756	1 036	2 649	378	183	3 211	4 246
Total intermediate / final use	1 493	19 616	1 432	11 958	34 499	15 076	12 065	23 620	50 761	85 260
Gross value added	1 106	6 206	988	15 586	23 887					
Output	2 600	25 822	2 420	27 545	58 386					

Simplified input-output table for current year 2012 at previous year prices

Industries	Intermediate consumption					Final use				Total use / output
	A Agriculture, forestry and fishing	B-E Mining; manufactu-ring etc.	F Construction	G-T Services	Total intermediate consumption	Household final consumption expenditures	Other domestic final demand	Export	Total final use	
A Agriculture, forestry and fishing	499	721	4	90	1 314	296	82	642	1 020	2 334
B-E Mining; manufacturing etc.	353	3 153	363	1 640	5 509	2 250	677	16 158	19 085	24 593
F Construction	3	58	70	218	349	23	1 779	104	1 906	2 255
G-T Services	257	2 272	419	5 787	8 735	7 514	6 761	3 963	18 238	26 973
Import	305	12 123	436	2 905	15 769	2 015	1 899	2 218	6 132	21 901
Taxes less subsidies on products	44	198	37	741	1 021	2 622	368	188	3 178	4 199
Total intermediate / final use	1 460	18 525	1 329	11 381	32 695	14 721	11 566	23 273	49 560	82 255
Gross value added	874	6 068	926	15 592	23 460					
Output	2 334	24 593	2 255	26 973	56 155					

Conventional growth contributions

The data required for calculating growth contributions by the widely-used conventional method can be acquired from the input-output tables, as well.

Arranging industries' values added and taxes less subsidies on products (grey cells in Table 1) to Table 2, branches' value added and the whole economy's GDP changes can be obtained as the differences of constant price current and base year numbers (in the case of industry A Agriculture, forestry and fishing, for example, 874 – 1 106 = –232). Expressing these in proportion to base year gross domestic product, we have growth contributions of industries in a percentage form (–232 / 28 134 = –0,82 %), which are exactly the same as the statistics in STADAT 3.1.6 table.[64]

To quantify demand side effects, we need to assemble the components of the well-known expenditure approach GDP identity (dark blue cells in Table 1). Totals of household consumption, other domestic final use, and export can be found in the sums of the same columns. Last cells of the fifth rows is subtracted from them, which are the sums of all intermediate and final use of imports. Using these, similarly to the production approach in the upper table of group 2, we can calculate growth contributions of demand components as well. Results differ slightly from STADAT 3.1.6 only because of the variance of national account and input-output table valuation standards.[65] The method is the same.

For the compatibility of the result from the conventional method reviewed above and the SDA, some changes were made in Tables 2 that do not affect the main point. First, seeing that growth effects of industries are of great importance, we omit taxes less subsidies, and express contributions not for the GDP, but the fully industry-divisible gross value added (GVA) (Table group 3). Although percentage GVA contributions somewhat differ from those based on GDP, relative weights of branches remain the same. Furthermore, these numbers are directly comparable to the results gained from the input-output model in the next section.

Table group 2: Conventional GDP growth contributions
(billion HUF and percentage)

Production approach contributions to GDP growth based on industries own value added

Industries	Base year, 2011 (at current prices)	Current year, 2012 (at previous year prices)	Change	
			in value	in proportion to base total
A Agriculture, forestry and fishing	1 106	874	-232	-0,82%
B-E Mining and quarrying; manufacturing; electricity, gas, steam and air conditioning supply etc.	6 206	6 068	-138	-0,49%
F Construction	988	926	-63	-0,22%
G-T Services	15 586	15 592	5	0,02%
Taxes less subsidies on products	4 246	4 199	-48	-0,17%
Gross domestic product (at purchaser's prices)	**28 134**	**27 659**	**-475**	**-1,69%**

Expenditure approach contributions to GDP growth

Components of final use		Base year, 2011 (at current prices)	Current year, 2012 (at previous year prices)	Change	
				in value	in proportion to base total
Household final consumption expenditures		15 076	14 721	-355	-1,26%
Other domestic final demand		12 065	11 566	-500	-1,78%
of which	Final consumption expenditures by non-profit organisations serving households (NPISH)	444	448	4	0,01%
	Final consumption expenditures by government	5 847	5 761	-85	-0,30%
	Gross fixed capital formation	5 569	5 324	-245	-0,87%
	Changes in inventories	206	33	-173	-0,62%
Export		23 620	23 273	-347	-1,23%
Import (-)		-22 627	-21 902	725	2,58%
Gross domestic product (at purchaser's prices)		**28 134**	**27 657**	**-476**	**-1,69%**

[64] http://www.ksh.hu/docs/eng/xstadat/xstadat_annual/i_qpt017a.html, for the conventional methodology see http://www.ksh.hu/docs/eng/modsz/modsz31.html

[65] Import is valued at fob (free on board) parity in national accounts, and at cif (cost, insurance and freight) in the input-output tables. Cif/fob adjustments, direct purchases abroad by residents and purchases on domestic territory by non-residents cause differences in trade and household consumption.

Table group 3: Contributions to GVA growth
(billion HUF and percentage)

Production approach contributions to GVA growth based on industries own value added

Industries	Gross value added: Base year, 2011 (current prices) (v^0)	Gross value added: Current year, 2012 (at prev year prices) (v^1)	Change: in value (Δv)	Change: in proportion to base total
A Agriculture, forestry and fishing	1 106	874	-232	-0,97%
B-E Mining and quarrying; manufacturing; electricity, gas, steam and air conditioning supply etc.	6 206	6 068	-138	-0,58%
F Construction	988	926	-63	-0,26%
G-T Services	15 586	15 592	5	0,02%
Gross value added total	**23 887**	**23 460**	**-427**	**-1,79%**

Expendtirus approach contributions of final demand for domestic products to GVA growth

Components of final use		Base year, 2011 (at current prices)	Current year, 2012 (at previous year prices)	Change: in value	Change: in proportion to base total
Household final consumption expenditures for domestic products		10 187	10 084	-103	-0,43%
Other domestic final demand for domestic products		9 745	9 299	-447	-1,87%
of which	Final consumption expenditures by non-profit organisations serving households (NPISH)	443	447	4	0,02%
	Final consumption expenditures by government	5 511	5 475	-36	-0,15%
	Gross fixed capital formation	3 784	3 488	-296	-1,24%
	Changes in inventories	7	-111	-118	-0,50%
Export from domestic products		21 476	20 867	-609	-2,55%
Intermediate use from imports and taxes less subsidies on intermediate products (-)		-17 521	-16 789	732	3,06%
Gross value added total		**23 887**	**23 460**	**-427**	**-1,79%**

A second modification is that direct import content of final demand components (dark green cells in Table 1) is ignored in the expenditure table, so only final use from domestic sources is taken into account. The import row includes only intermediate consumption henceforth.[66] Changes in the final demand for domestic products can, of course, alter the intermediate use from imports, which has an adverse effect on GVA. Thus, growth contributions of domestic product demand components indicated in Table 2 can be imprecise. Assessment of their value added effect depends on the industry mix of final demand change, domestic and foreign supply chains of the concerned industries, and companies' value added ratios. Multiplicative processes taking place can be kept track of by the input-output model, and factoring the changes can be made by a structural decomposition analysis. These techniques will be covered in the following section.

When comparing SDA and the conventional method, values of Table 3 will serve as reference points. These are the growth contributions calculated separately from the supply and demand side surface of the economy, from the margins of the input-output tables. Only such calculations can be accomplished using current GDP statistics, which ignore the interconnections between industries captured by the numbers in the light blue highlighted cells of Table 1. A more profound investigation based on these can penetrate deeper into the growth relationships and discover details that cannot be revealed from above. For this, however, we need to recall the basics of the input-output model.

Structural decomposition analysis

Using the input-output model we can draw on the well-known mechanism that

- changes in the final demand (A) affects output of industries, which in turn,
- modifies the purchases from other industries (B), i.e. intermediate consumption too, generating a circular, multiplicative process in the economy, and as a result of this,
- value added of the industries will vary (C) (see Figure 2).

[66] For the sake of switching from GDP to GVA we correct with product taxes of intermediate consumption also in this row.

Figure 2: GVA multiplication process in the demand-pull input-output model
(billion HUF)

Simplified input-output table for base year 2011 at current prices

Industries	Intermediate consumption					Final use				
	A Agriculture, forestry and fishing	B-E Mining; manufacturing etc.	F Construction	G-T Services	Total intermediate consumption	Household final consumption expenditures	Other domestic final demand	Export	Total final use	Total use / output
A Agriculture, forestry and fishing	564	719	4	101	1 389	317	242	652	1 211	2 600
B-E Mining; manufacturing etc.	352	3 558	422	1 746	6 077	2 260	737	16 747	19 744	25 822
F Construction	3	67	63	217	349	26	1 947	98	2 071	2 420
G-T Services	256	2 382	447	6 077	9 163	7 583	6 819	3 979	18 382	27 545
Import	285	12 683	456	3 061	16 485	2 239	1 942	1 961	6 142	22 627
Taxes less subsidies on products	33	206	40	756	1 036	2 649	378	183	3 211	4 246
Total intermediate / final use	1 493	19 616	1 432	11 958	34 499	15 076	12 065	23 620	50 761	85 260
Gross value added	1 106	6 206	988	15 586	23 887					
Output	2 600	25 822	2 420	27 545	58 386					

Value added between two years can change because of three main reasons: changes in the

- final demand (D),
- domestic requirements of the supply chains (E), and
- industries' value added (F) ratios (relative to output) (see Figure 3).

These are the three terms on the right hand side of the following basic value added SDA equation

$$\mathbf{v} = \langle \mathbf{c} \rangle \mathbf{L}\mathbf{f} \, ,$$

where $\mathbf{v}$ is the vector of values added, $\langle \mathbf{c} \rangle$ is the diagonalized vector of industry value added ratios, $\mathbf{L}$ is the so-called Leontief inverse of direct requirements of intermediate inputs, and $\mathbf{f}$ is the vector of the final demand.

Figure 3: Reasons of GVA changes
(billion HUF)

Simplified input-output table for base year 2011 at current prices

Industries	Intermediate consumption					Final use				
	A Agriculture, forestry and fishing	B-E Mining; manufacturing etc.	F Construction	G-T Services	Total intermediate consumption	Household final consumption expenditures	Other domestic final demand	Export	Total final use	Total use / output
A Agriculture, forestry and fishing	564	719	4	101	1 389	317	242	652	1 211	2 600
B-E Mining; manufacturing etc.	352	3 558	422	1 746	6 077	2 260	737	16 747	19 744	25 822
F Construction	3	67	63	217	349	26	1 947	98	2 071	2 420
G-T Services	256	2 382	447	6 077	9 163	7 583	6 819	3 979	18 382	27 545
Import	285	12 683	456	3 061	16 485	2 239	1 942	1 961	6 142	22 627
Taxes less subsidies on products	33	206	40	756	1 036	2 649	378	183	3 211	4 246
Total intermediate / final use	1 493	19 616	1 432	11 958	34 499	15 076	12 065	23 620	50 761	85 260
Gross value added	1 106	6 206	988	15 586	23 887					
Output	2 600	25 822	2 420	27 545	58 386					

Simplified input-output table for current year 2012 at previous year prices

Industries	Intermediate consumption					Final use				
	A Agriculture, forestry and fishing	B-E Mining; manufacturing etc.	F Construction	G-T Services	Total intermediate consumption	Household final consumption expenditures	Other domestic final demand	Export	Total final use	Total use / output
A Agriculture, forestry and fishing	499	721	4	90	1 314	296	82	642	1 020	2 334
B-E Mining; manufacturing etc.	353	3 153	363	1 640	5 509	2 250	677	16 158	19 085	24 593
F Construction	3	58	70	218	349	23	1 779	104	1 906	2 255
G-T Services	257	2 272	419	5 787	8 735	7 514	6 761	3 963	18 238	26 973
Import	305	12 123	436	2 905	15 769	2 015	1 899	2 218	6 132	21 901
Taxes less subsidies on products	44	198	37	741	1 021	2 622	368	188	3 178	4 199
Total intermediate / final use	1 460	18 525	1 329	11 381	32 695	14 721	11 566	23 273	49 560	82 255
Gross value added	874	6 068	926	15 592	23 460					
Output	2 334	24 593	2 255	26 973	56 155					

Factoring the differences between two years' value added vector, $\Delta\mathbf{v} = \mathbf{v}^1 - \mathbf{v}^0$ (superscripts indicate the relating time periods: 0 is for the base, and 1 for the current year) can be done in several ways. The two so-called polar decompositions, stepping from current to base year with weights,

$$\Delta\mathbf{v} = \langle \Delta\mathbf{c} \rangle \mathbf{L}^1\mathbf{f}^1 + \langle \mathbf{c}^0 \rangle (\Delta\mathbf{L})\mathbf{f}^1 + \langle \mathbf{c}^0 \rangle \mathbf{L}^0 \Delta\mathbf{f} \, ,$$

and vice versa

$$\Delta \mathbf{v} = \langle \Delta \mathbf{c} \rangle \mathbf{L}^0 \mathbf{f}^0 + \langle \mathbf{c}^1 \rangle (\Delta \mathbf{L}) \mathbf{f}^0 + \langle \mathbf{c}^1 \rangle \mathbf{L}^1 \Delta \mathbf{f} ,$$

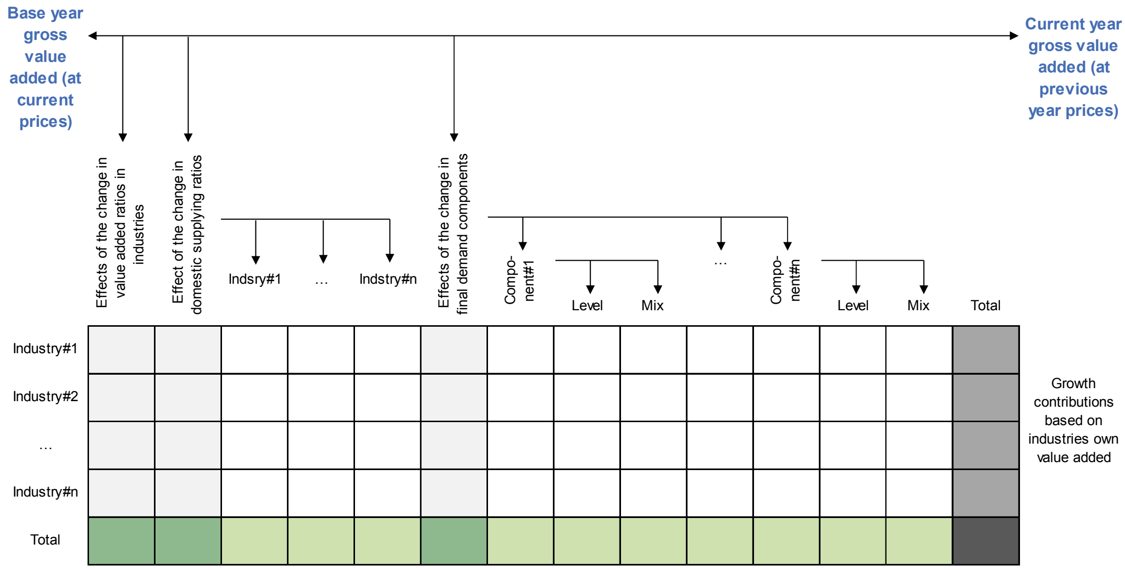
however, in the case of three or more components are not unique, don't cover all possible formulations. Empirical evidence suggests that the average of them still gives a good approximation. The other way for coping with the non-uniqueness problem is bracketing two adjacent terms, and make hierarchical or nested decompositions. Both of them was used in the case study.

Another problem was the interdependence between value added and domestic supplying ratios, which called for a special solution.[67]

Figure 4 shows the design of the SDA. Variations between two years' value added are broken to the three main part effects mentioned above, with the second and third one divided further to sublevels. Results can be obtained in a matrix or cross-tab format, where

- column sums show the SDA version of the expenditure approach, i.e. the demand side originated effects, and
- row sums give the production approach industry break-up with exactly the same conventional growth numbers on the whole.

Figure 4: SDA design



In order to reveal new aspects on growth from the supply side, as well, I performed the variance analysis with the following, modified version of the basic equation:

$$\bar{\mathbf{v}} = \left\langle (\mathbf{c})' \mathbf{L} \right\rangle \mathbf{f} .$$

Vectors $\bar{\mathbf{v}}$ in SDA#2, in contrast to former vectors $\mathbf{v}$ in SDA#1, allocate domestic value added to industries not on the basis of where they appear, but according to all the direct and indirect nationwide effects that an industry's final demand can have. So this alternative gives not the own value added of industries, but the value added of all industries contributing to a given industry's final demand through the supply chains, of course, to the extent of their contribution. Using this model, we have a different production approach, which also yields significant deviations from conventional growth contributions in certain industries. Results in the short form Table 4 serve for an easy comparison to Table 3, and will be evaluated in the next section.

[67] On these problems and on the SDA itself in genereal see *Rose–Casler* (1996), *Dietzenbacher–Los* (1998, 2000), *Pei et al* (2012), and *Koppány* (2016). Detailed mathematical apparatus used in this study is also available in *Koppány* (2016).

Discussion of the results

Figures 5 and 6 help give an overall assessment. The waterfall chart shows that, according to the most important column sums of the SDA, the change of value added ratios have the only significant positive effect on 2012 growth. Shifts in domestic direct requirements, particularly those of manufacturing, and the fall in final demand decreased total value added.

Benchmark Tables 2 and 3 indicated export as a considerable negative factor, which was overcompensated by the more declining import. Thus from the demand side, international trade was the only positive force. SDA results indicate these differently. Taking the industry mix of export and the multiplication processes through the value chains into account and fixing the supplier structure and value added ratios at an average of two years, we can say changes in export hardly affected the growth on the whole. Cutdown of domestic final use of domestic products, mainly the decrease in investments, was the greatest retractive force. The growth effect order of the components of domestic final demand in SDA, however, is the same as in Table 3.

An in-depth discussion of the various industry part effects behind the column sums, and unfolding the complexities of the levels and mixes is beyond the limits of this paper; however, highlighting variances between industries own value added and those of their supply chains definitely deserves mention. These can be followed by a row-by-row comparison of Tables 3 and 4.

Value added production of an industry, according to the “accounting” used in SDA#2, depends, on the one hand, on its final output, and on the other hand, on its value added multiplier. Agriculture, for example, sells more for intermediate, than final use, so, despite its relatively high multiplier, it has a lower value added from final demand supply chains than its own realized measure (a part of the latter, in supply chain approach, will be accounted to other industries, for which agriculture is a supplier). Supply chain values added of manufacturing and construction, however, far exceed their own one. These are due to the prodigious production and export volumes of the key growth manufacturing sub-branches, and the high multiplier value of construction. Hence, decline of the final demand for construction, in Table 4, decreased economic growth more than the fall in its own value added in Table 3.

Table 4: Production approach growth contributions by SDA#2
(billion HUF and percent)

Production approach contributions to GVA growth based on final demand industry supply chains' value added

Industries	Gross value added		Change	
	Base year, 2011 (at current prices) (v^{-0})	Current year, 2012 (at previous year prices) (v^{-1})	in value (Δv^{-})	in proportion to base total
A Agriculture, forestry and fishing	877	690	-187	-0,78%
B-E Mining and quarrying; manufacturing etc.	7 336	7 192	-144	-0,60%
F Construction	1 341	1 238	-104	-0,43%
G-T Services	14 333	14 340	7	0,03%
Gross value added total	23 887	23 460	-427	-1,79%

Figure 5: Column sum SDA results
(billion HUF and percentage)

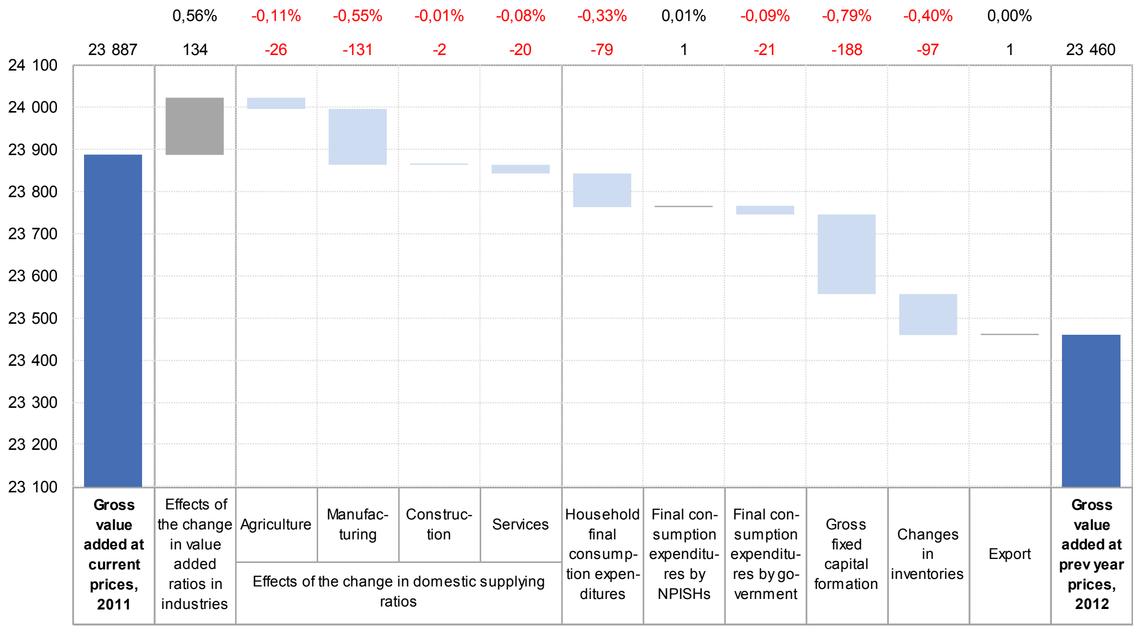
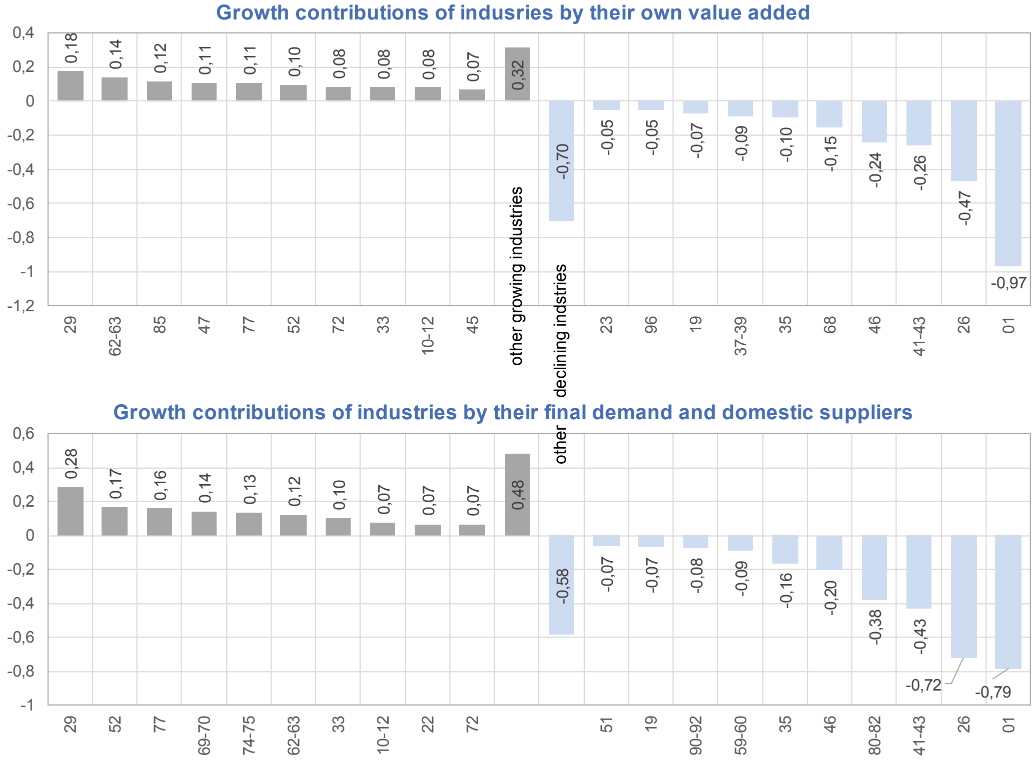


Figure 6 shows the effect of the most and least growth-contributing industries in 2012 estimated by both methods. The ranking is headed by the manufacture of motor vehicle in both cases, although value according to the second approach was more than a one and a half times higher. Growth contribution of the automotive industry by its own value added was 0.176 %; however, it bore a 0.284 % effect through its suppliers, in spite of its almost minimum and somewhat decreasing multiplier value, caused by its high import, and low domestic supply and value added rates. Nevertheless, low and declining multipliers, coupled with a high and increasing export volume, resulted in an ascent from second to first position in the ranking of final use effect, the direct and indirect consequences of which overcompensated the negative growth effects of declining domestic supplying rates.

When making a comparison of the lists of the first and last ten indusries of the upper and lower diagram of Figure 6, a significant overlap can be seen. The most and least own value added growth-contributing industries generally have the greatest effects through their supply chains, too. The ranking between them, however, is somewhat different. Warehousing and support activites for transportation, for example, is second by its supply chains, and only sixth with its own value added.

Figure 6: **Effects of the most and least growth-contributing industries in 2012**
(percent)



Legend: 01: Crop and animal production, hunting and related service activities; 10-12: Manufacture of food products, beverages and tobacco products; 19: Manufacture of coke and refined petroleum products; 22: Manufacture of rubber and plastic products; 23: Manufacture of other non-metallic mineral products; 26: Manufacture of computer, electronic and optical products; 29: Manufacture of motor vehicles, trailers and semi-trailers; 33: Repair and installation of machinery and equipment; 35: Electricity, gas, steam and air conditioning supply; 37-39: Sewerage; waste collection, treatment and disposal activities; materials recovery; remediation activities and other waste management services; 41-43: Construction; 45: Wholesale and retail trade and repair of motor vehicles and motorcycles; 46: Wholesale trade, except of motor vehicles and motorcycles; 47: Retail trade, except of motor vehicles and motorcycles; 51: Air transport; 52: Warehousing and support activities for transportation; 59-60: Motion picture, video and television programme production, sound recording and music publishing activities; programming and broadcasting activities; 62-63: Computer programming, consultancy and related activities; information service activities; 68: Real estate activities and imputed rents of owner-occupied dweellings; 69-70: Legal and accounting activities; activities of head offices; management consultancy activities; 72: Scientific research and development; 74-75: Other professional, scientific and technical activities; veterinary activities; 77: Rental and leasing activities; 80-82: Security and investigation activities; services to buildings and landscape activities; office administrative, office support and other business support activities; 85: Education; 90-92: Creative, arts and entertainment activities; libraries, archives, museums and other cultural activities; gambling and betting activities; 96: Other personal service activities.

Summary and comparison, pros and cons

The calculation of growth contributions by SDA, like any method, has both advantages and disadvantages. As a conclusion we present a brief overview of these. Theoretical and methodological limitations are not explained here, instead, difficulties evident from the choice of investigated periods are emphasized. The time-lag of several years in producing and publishing supply, use and input-output tables, the assumptions, limitations, and imprecisions of the models, updating and approximation techniques impede an up to date and accurate operation of the analysis.[68] Undoubtedly, flash estimates of quarterly GDP by statistical offices also need re-examinations and sometimes corrections; however, conventional methods of calculating growth contributions can be applied immediately, even by the most current and simple structure

[68] The general reason of official statistics for constructing and publishing input-output tables only every five years is that the structures of the economies modify relatively slow. It might have been true for the past, but not for the future. Being round the corner of the large scale robotization, virtualization, IoT, big data and hopefully green revolution, the world, including technological and economic structures and so the driving forces of growth, will probably change faster than ever before. Statistical offices definitely perceive these phenomena and the pressure from analysts and policymakers for the most current and high quality data on economic structures, at the same time. Timely estimates of several statistical indicators, especially those of GDP, improved significantly in the last decades (see *Kokkinen–Wouters* (2016)). There must be some possibilities also in reducing the production time of input-output tables. A decrease of the time-lags will boost the applicability and the relevance of the growth decomposition analysis presented here.

data, providing very quick indicators for analysis and policy.

Structural decomposition of the factors of economic growth offers extra information to the standard production and expenditure approach contributions calculated independently from the changes of own value added of industries and the levels of final demand components. Conventional methods show only the surface from two separate sides. Both methods presented here, however, consider multiplicative effects of final use from domestic output through the supply chains, and decompose them to part effects of changes in value added ratios, supplying structure and final demand, and further subcomponents. The effects are allocated between industries, as well, so the demand side and the value added generation of the producers (in SDA#1) and supply chains (in SDA#2) are connected as two dimensions of growth and shown together in a crosstab format.

Different approaches yield different insights and significant variance in the results. Consequently, SDA, in spite of the time-lag of data and the imprecision of updating techniques, can be a useful complement to standard techniques. Structural decomposition and variance analysis of input-output tables show a deeper structure of the economy, thus offering a different approach to assessing GDP generation and growth contributions of industries, supply chains and final demand components for a better understanding of the driving forces of growth. As a complementary tool for analysis, it can support economic, development and policy decisions of the private sector and the government.

References

Dietzenbacher, E.–Los, B., 'Structural Decomposition Techniques: Sense and Sensitivity', *Economic Systems Research*, 10:4, 1998, pp. 307-324.

Dietzenbacher, E.–Los, B., 'Structural Decomposition Analyses with Dependent Determinants', *Economic Systems Research*, 12:4, 2000, pp. 497-514.

Eurostat, Eurostat Manual of Supply, Use and Input-Output Tables, 2008.

Kokkinen, A. and Wouters, H., 'GDP growth estimates for Europe at 30 days – is that feasible?' Conference of European Statistics Stakeholders. Budapest, 20–21 October 2016.

Koppány, K., Növekedési hozzájárulások számítása input-output táblák strukturális felbontása alapján. (Calculation of Growth Contributions Using Structural Decomposition of Input-Output Tables.) Hungarian Statistical Review, 94: 8–9881-914, 2016.

Miller, R. E. and lair, P. D., 'Input-Output Analysis', *Foundations and Extensions*, Second Edition. Cambridge University Press, Cambridge, 2009.

Pei, J., Oosterhaven J. and Dietzenbacher, E., 'How Much Do Exports Contribute to China's Income Growth?', *Economic Systems Research*, 24:3 275-297, 2012.

Rose, A. and Casler, S., 'Input–Output Structural Decomposition Analysis: A Critical Appraisal', *Economic Systems Research*, 8:1, 33-62, DOI: 10.1080/09535319600000003, 1996.

8

An EU cross-country comparison study of life expectancy projection models

JUSTYNA MAJEWSKA[69]

Table of content

([69]) University of Economics in Katowice: justyna.majewska@ue.katowice.pl

1. Introduction

Life expectancy (LE) at birth is defined as the mean number of years still to be lived by a person at birth, if subjected throughout the rest of his or her life to the current mortality conditions (Eurostat, 2016). Over the last 170-year period European countries have experienced a continuation of the pattern of falling mortality rates that began in the 19^{th} century. Life expectancy consistently increased by 0.24 years of life per calendar year of time, or at the rate of 24 years per century. Over long historical period we have observed linear trend in life expectancy. However, today there is no consensus on trends over the very long term, e.g. whether there is a natural biological limit to longevity, the impact of future medical breakthroughs, long-term impact of public health programmes and societal behaviour such as reduction of smoking rates or increased prevalence of obesity European Commission (2014). Extrapolation methods, mostly used so far by producers of official statistics, would lead us to expect a female life expectancy of around 108 years at the end of the 21st century (Lee, 2003).

Accuracy is crucial in a life expectancy projection. It allows governments and other institutions to plan wisely and helps individuals comprehend the likely futures for their countries and the world. Projection of mortality and life expectancy for the next 50-100 years is a great challenge, because official projections generally assume that gains in future life expectancy at birth will slow down compared with historical trends. Stochastic mortality models provide the good base for creation of different scenarios for future level of mortality and life expectancy. Stochastic models are well suited for longevity risk monitoring and decision-making.

As Li and Lee (2005) indicated the convergence in mortality levels for closely related populations can lead to unsuitable mortality projections, if the projections for individual populations are obtained in isolation from one another. Similar historical trends in long-run life expectancy patterns can be useful for countries. Besides, analyses of the main determinants of life expectancy (the socio-economic, environmental or behavioural factors) of associated populations are crucial. Knowledge of existence of some common stochastic trends in mortality rate in cluster of European countries (see Lazar et al. 2016) can be used for projections mortality rates and life expectancy.

Selected stochastic mortality models are considered to achieve good life expectancy projections in homogeneous clusters of EU countries. The empirical analysis is preceded by theoretical remarks on modelling and projection of mortality and life expectancy.

2. Remarks on mortality model

A large number of mortality models have been proposed to analyse historic mortality rates and project them into the future. As Cairns (2013) indicates it is easy to create new mortality model. However, many of them are over-parameterized or include some additional age/period terms, added in an ad hoc manner that cannot be justified in terms of demographic significance. Consequence of poor specification of a model can lead to period effects in the data being wrongly attributed to cohort effects, which results in the model making implausible projections (Hunt and Blake, 2015).

2.1. Stochastic mortality modelling

Nowadays research and development efforts are focused on improving mortality models. First generation of stochastic models includes Lee-Carter model (1992), original CBD model (2006) and P-spline model (Currie et. al 2004). In response to their disadvantages new, more complex, models appeared. Some of these new models can be thought of as direct descendants of the Lee-Carter model, other models have significantly different roots. In most cases, increased complexity improves the model fit to historical data, but providing a conclusive assessment of their ability to project future developments in mortality is more difficult. In recent years multi-population models are of great interest. The genealogy of mortality models was described by Cairns (2013). All models are well documented in the literature.

The majority of existing stochastic models can be presented in the age-period-cohort framework. The model is written as (Hunt and Blake, 2015):

$$\eta\left(E\left(\frac{D_{xt}}{E_{xt}}\right)\right)=\alpha_x+\sum_{i=1}^{N} f^{(i)}(x,\theta^{(i)})\kappa_t^{(i)}+\gamma_{t-x} \quad (1)$$

This equation has the following components:

- a link function η transforms the observed data into a form suitable for modelling (the raw data usually consists of death counts D_{xt} and exposures to risk E_{xt} at ages x and for years t),
- a static age function α_x captures the general shape of the mortality curve,
- N age/period terms $f^{(i)}(x,\theta^{(i)})\kappa_t^{(i)}$, consisting of companion pairs of t period terms $\kappa_t^{(i)}$, which give the evolution of mortality t rates through time, and age functions $f^{(i)}(x,\theta^{(i)})$ that determine which segments of the age range these trends affect,
- cohort parameters γ_{t-x} which determine the lifelong effects that are specific to different generations denoted by their year of birth.

The procedure for constructing mortality model, developed by Hunt and Blake (2015), helps in capturing all the significant information present in the age, period and cohort dimensions of the data. The procedure provides evidence for the addition of each term to an existing model (Hunt and Blake, 2015). Assessment whether the model captures all demographically significant age/period terms is desired at each stage of the procedure.

2.2. Criteria for a good model

Consistency with historical data and 'optimal' level of complexity are the minimum to consider the particular model as a good one for projection purpose. Moreover, there should be a trade off between model fit and accuracy projection. Table 1 shows some existing models and criteria that the models satisfied([70]). Most of existing models meet most of above criteria. It is important to note that beyond qualitative criteria the quantitative criteria have to be considered (log-likelihood measure, information criteria AIC and BIC). The goodness-of-fit of mortality models is also typically analysed by inspecting the residuals of the fitted model (Villegas et.al 2015).

In the literature we find lots of empirical studies that compare different models, goodness-of fit analysis, and scenarios for future mortality rates. Booth and Tickle (2008) discussed in details the most important methods in mortality modelling and forecasting. Cairns et al. (2008) compared quantitatively eight stochastic models explaining improvements in mortality rates in England and Wales and in the United States. Multi-population models were compared by Enchev at al. (2015). Hyndman at al. (2011) used the age- and sex-specific data of 14 developed countries, and compared the point and interval forecast accuracy and bias of ten principal component methods for forecasting mortality rates and life expectancy. There is no possibility to determine the best model for modelling and projection of mortality. The model should be country-specific and adjusted to the trends of mortality rates, economic and demographic environment in specific country.

Table 1: Quality of mortality models - criteria comparison for selected mortality models (+

([70]) Criteria are described in details in Cairns et al. 2008, Plat 2009, Hunt and Blake 2015

meets criterion, +/- partly meets criterion, - does not meet criterion)

Criterion	LC (1992)	RH (2006)	APC (2006)	Currie at al. (2004)	CBD (2006)	CBD M6 (2007)	CBD M7 (2007)	CBD M8 (2007)	Hyndman, Ullah (2007)	Plat (2009)
Positive mortality rates	+	+	+	+	+	+	+	+	+	+
Consistency historical data	+/-	+	+	+	+	+	+	+	+	+
Long-term biological reasonableness	+	+	+	+	+	+	+	+	+	+
Robustness	+	-	+	+/-	+	+	+/-	+/-	+	+
Forecasts biological reasonable	+/-	+	+	+	+	+	+	+	+	+
Ease of implementation	+	+	+	+/-	+	+	+	+/-	+/-	+
Parsimony	+	+/-	+/-	+	+	+/-	+/-	+/-	+/-	+/-
Possibility generating sample paths	+	+	+	-	+	+	+	+	+	+
Allowance for parameter uncertainty	+	+	+	+	+	+	+	+	+	+
Incorporation cohort effects	-	+	+	+	-	+	+	+	+	+
Non-trivial correlation structure	-	+/-	+/-	+/-	+	+	+	+	+	+
Applicable for full age range	+/-	+/-	+/-	+/-	-	+/-	+/-	+/-	+	+

Note: M1 – Lee Carter (LC), M2 – Renshawn -Haberman (RH), M3 – APC, M4 – Currie et al. (2006) with B-splines and P-spline, M5 – original CBD, M6–M8 – generalizations of CBD (each of models M6 to M8 is an extension of model M5 with some allowance for the cohort effect)

Source: own compilation based on Plat (2011) and Cairns et al. (2007)

3. Life expectancy projection

Between mortality and longevity modelling is close relationship particularly clear when considering the survival probability. As a consequence, the mortality models are used for both mortality and longevity.

Observing historical trends is required when predicting future mortality rates. The mortality decline that we can observe today has its roots in improvements achieved long ago in living standards and diet, public health institutions, medicine, and other areas relevant to the physical well-being of the population. Speaking in general terms, living conditions improve from one period to the next. Secondly, the health and remaining life span for people living today is determined not only by contemporary period factors but also by living conditions earlier in life. Since even conditions in the foetal stage have an influence on longevity, improvements during early childhood could have an effect on mortality trends today. Predicting future mortality trends also requires a multivariate point of departure. No single factor, but a variety of factors, determines health and remaining life span, and we do not know a priori which one is the most important.

Three general approaches of mortality projection are distinguished: explanation, extrapolation and expectation (Booth and Tickle, 2008). The most popular methods of mortality projection are extrapolative. It makes use of the regularity typically found in both age patterns and trends over time. This approach includes the traditional and relatively simple extrapolation of aggregate measures such as life expectancy, as well as more complex methods such as the Lee-Carter method. The explanation approach makes use of structural or epidemiological models of mortality from certain causes of death, where the key exogenous variables are known and can be measured. The expectation approach projections are based on the subjective opinions of experts involving varying degrees of formality.

Methods of mortality projection used in the EU countries vary significantly. Such comparison has been prepared by Stoeldraijer et al. (2013). Information on the methods were given for Austria, Belgium, Denmark, France, Italy, Ireland, Luxemburg, the Netherlands, Norway, Poland, Portugal, Spain, Sweden, and the UK. Authors found out that extrapolation methods were used most frequently, mainly some variants and extensions of the original Lee-Carter model. Large group of European countries produce scenarios (usually optimistic, pessimistic and neutral) based on experts' opinions and knowledge, informed judgement, and assumed target life expectancy. Moreover, the historical period used by national statistical offices differs considerably by country, and variation in the length of the forecasted period. It also appears that the observed past trends determine which method and historical period is used.

Different institutions and companies produce the mortality and life expectancy projections in Europe (e.g. Eurostat, WHO, OECD, research groups mainly of actuarial profession, SwissRe). It happens that projections of different institutions differ from those published by the national statistical offices. It is due to methodological differences, like calculation of mortality tables, and various assumptions on mortality and life expectancy.

4. Methodology

SKATER (Spatial 'K'luster Analysis by Tree Edge Removal) is an alternative to other regionalization methods, based on minimum spanning trees (Assunção et al. 2006). SKATER first constructs a spatially contiguous graph by removing edges that do not connect geographic neighbours and then builds a minimum spanning tree from the graph (any tree built from this graph is spatially contiguous). The tree is then recursively partitioned to generate a given number of regions. Two important limitations of this algorithm are: (1) it uses contiguity constraints in a static way, where the contiguity matrix is not dynamically updated during the clustering process and (2) minimum spanning tree has a well-known 'chaining' problem.

The spatial clustering algorithm involves some variables, representing the main indicators for similarities in the historic mortality rates of countries, or predictors for these similarities. Life expectancy is affected by many factors such as: socioeconomic status, including employment, income, education and economic wellbeing; the quality of the health system and the ability of people to access it; health behaviours such as tobacco and excessive alcohol consumption, poor nutrition and lack of exercise; social factors; genetic factors; and environmental factors including overcrowded housing, lack of clean drinking water and adequate sanitation. In the empirical study, we use gross domestic product per capita, educational attainment, fertility and access to health care. The SKATER algorithm divides the EU countries into homogeneous contiguous clusters according to these variables.

We have chosen seven models for a comparative analysis of life expectancy projection models: the *Poisson* log-bilinear *Lee Carter* (LC), APC, CBD, M6, M7, M8 models, and PLAT model. All the models belong to the Lee-Carter family. Formulas are described in details in Villegas at al. (2016) and Cairns at al. (2007). LC uses principal component analysis to decompose the age-time matrix of mortality rates into a bilinear combination of age and period parameters, with the latter being treated as time series to produce mortality projections. The CBD model relies on the linearity of the logit of one-year death probabilities at older ages. It assumes that for a given year the logit of the one-year death probability is a linear function of age, and treats the intercept and slope parameters across years as stochastic processes. Three extensions (M6, M7, M8) to the original CBD model incorporate combinations of a quadratic age term and a cohort effect term. Plat (2009) combines features of the CBD and the Lee-Carter models to produce a model that is suitable for full age ranges and captures the cohort effect.

Life tables were obtained from Human Mortality Datasets (HDM, 2016)[71]. Demographic and economic data (recorded in 2012) used for spatial clustering – from Eurostat database([72]), and geographical information (administrative boundaries) from the Geographic Information System of the Commission (GISCO)([73]). All calculations were made in R with spdep, demography and StMoMo packages([74]).

5. Empirical results

The empirical analysis starts with obtaining homogenous spatial clusters of EU countries according to the following demographic and economic variables: gross domestic product per capita, educational attainment, fertility and health status. The highest GDP per capita was recorded by Luxembourg, Denmark and Sweden, the lowest – Bulgaria, Romania and Hungary. Czech Republic, Lithuania, Slovakia, Poland and Estonia recorded the highest level of people with upper secondary, post-secondary non-tertiary and tertiary education, meanwhile Portugal, Malta, Spain — the lowest level (in aging group 15–64). The highest fertility

[71] http://mortality.org

([72]) http://ec.europa.eu/eurostat/data/database

([73]) http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units

([74]) Package spdep is a collection of functions to create spatial weights matrix objects from polygon contiguities, package demography - functions for demographic analysis, StMoMo - implementation of the family of generalised age-period-cohort stochastic mortality models.

rate was in France, followed by Ireland, Sweden, Norway, Finland, and the lowest — in Portugal, Spain, Poland, Slovakia, Hungary. Health status is measured here by self-reported an unmet needs (for medical examination for reasons of barriers of access)([75]). The highest percentage of population aged 16 and over in Latvia, Greece, Estonia, Poland and Romania reported that they had unmet needs for medical examinations or treatment. Netherland, Austria and Slovenia reported relatively low shares for this indicator.

As a result of skater algorithm([76]) we got five clusters:

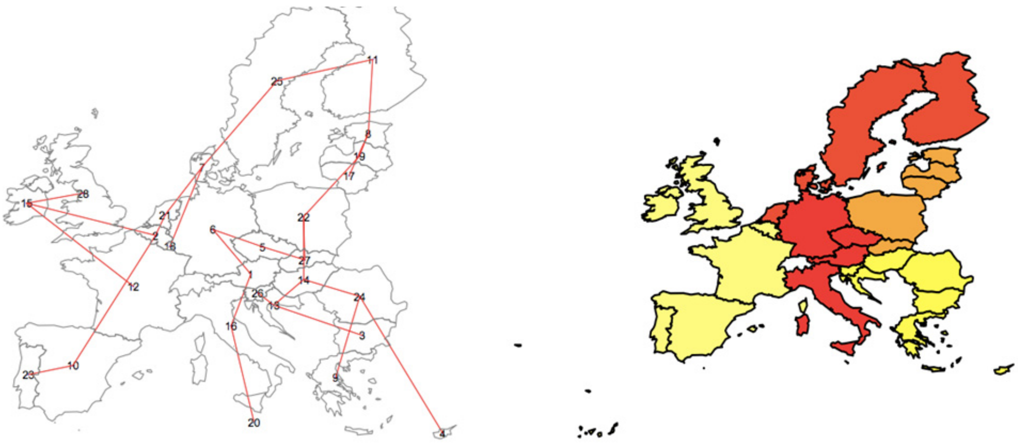
- 1st cluster: Netherlands, Luxembourg, Denmark, Sweden, Finland
- 2nd cluster: Ireland, UK, Belgium, France, Portugal, Spain
- 3rd cluster: Czech Republic, Germany, Austria, Italy, Malta
- 4th cluster: Estonia, Latvia, Lithuania, Poland, Slovakia
- 5th cluster: Slovenia, Croatia, Hungary, Romania, Bulgaria, Greece, Cyprus

The spatial diversification is shown in figure 1. Countries were grouped according to geographical proximity and levels of the selected variables. Malta, Cyprus, Romania and Croatia had to be excluded from the analysis due to lack of data in HDM.

Created clusters are quite natural. Based on means of each cluster for each variable the following distinctiveness of the clusters is possible: (a) first cluster groups developed countries with the highest GDP per capita and very high fertility rate, (b) to the second cluster belong countries with the highest fertility rate and the lowest education level, (c) the third cluster groups countries with the lowest percentage of population aged 16 and over that they had an unmet needs for medical examinations or treatment, (d) countries from the last two clusters have the lowest level of GDP per capita, however, the fourth cluster groups countries with the highest education level and the highest percentage of population aged 16 and over that they had an unmet needs for medical examinations or treatment, (e) the fifth cluster groups countries with very low level of education and high percentage of population reported that they had an unmet needs for medical examinations or treatment.

Considering this diversification in terms of life expectancy trends the gap between the former communist countries and the Baltic States, and the rest of Europe is evident. Countries grouped in the clusters have recorded in general comparable levels of LE and experienced a similar evolution of LE in log run. The exceptional case is Portugal in the second cluster and Czech Republic in the third cluster.

Figure 1: Spatial diversification of EU countries in 2012 in terms of GDP per capita, educational attainment, fertility and access to health care



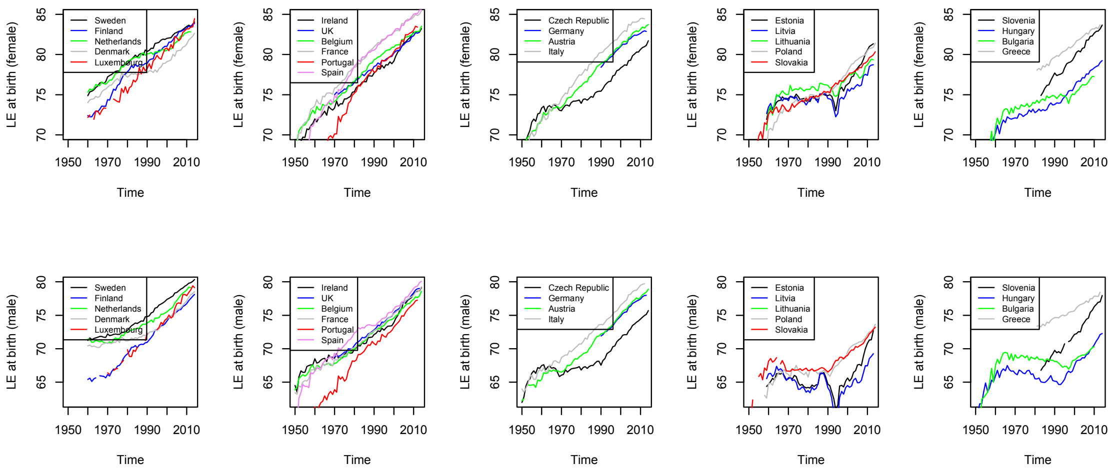
Source: own calculation

([75]) According to Eurostat: **self-reported unmet** needs means: person's own assessment of whether he or she needed examination or treatment for a specific type of health care, but didn't have it or didn't seek for it. EU-SILC collects data on two types of health care services: medical care and dental care. The indicator is treated as destimulant of life expectancy.

([76]) A restriction to the procedure was added by establishing a minimum number of units in each cluster, what results in avoiding clusters with one country, like Luxembourg.

Since 1960, life expectancy in Western Europe has risen by around 10 years (from 70 to 80 years). The equivalent figure in Central and Eastern Europe is 7 years (from 67 to 74 years). This difference in life expectancy between the two parts of Europe is greater today than it was in 1960. Trends in life expectancy for the countries grouped in a spatial cluster are presented in figure 2. A visual inspection of figure 2 suggests the possibility of common stochastic trends in LE.

Figure 2: Trends in life expectancy in EU countries



Source: own calculation

Seven stochastic methods were applied to each country separately, for both male and female. Availability of data in HDM database determined the length of historical period (50, 40, 30, or 20 years). The same periods were considered for all the countries from each cluster. Projections of LE at birth were made for 2012 with the best model derived for each country separately, according to the AIC and BIC criteria. None of the following models APC, M6, M7 and M8 appeared in the top 3 highest ranked models. For the best-fitted model sample errors (defined as projection value less actual value) were calculated. Results, for countries in each cluster, are given in tables 2-6.

Table 2: The best stochastic mortality models and error rates in LE at birth projections for countries in cluster 1

	The best-fitted model in historical period (LE at birth projection error)			
	Male		Female	
Historical period:	1960–2000	1970–2000	1960–2000	1970–2000
Sweden	PLAT (–1.12)	PLAT (–0.67)	LC (0.78)	LC (0.64)
Finland	PLAT (–1.57)	PLAT (–0.79)	LC (0.07)	LC (0.01)
Netherlands	CBD (–2.10)	CBD (–1.63)	LC (–0.60)	LC (–0.60)
Denmark	CBD (–1.90)	CBD (–1.70)	CBD (–0.57)	CBD (–0.83)
Luxembourg	CBD (–2.08)	CBD (–1.45)	LC (–0.78)	LC (–0.47)

Source: own calculation

For Netherlands, Denmark and Luxembourg CBD model provides for male the best fits the historical data. In case of Sweden, Finland, Netherlands and Luxembourg projections of female LE at birth are very close to the actual LE, due to the LC model. The best projections we have noted for Finland for male.

Table 3: **The best stochastic mortality models and error rates in LE at birth projections for countries in cluster 2**

	The best-fitted model in historical period (LE at birth projection error)					
	Male			Female		
Historical period:	1950–2000	1960–2000	1970–2000	1950–2000	1960–2000	1970–2000
Ireland	PLAT (–3,45)	PLAT (–3,43)	PLAT (–2,69)	LC (–2,18)	LC (–2,21)	LC (–1,99)
UK	PLAT (–2.26)	PLAT (–2.05)	PLAT (–1.52)	LC (–0.99)	PLAT (–1.24)	PLAT (–0.95)
Belgium	PLAT (–2.02)	PLAT (–1.70)	CBD (–0.22)	LC (0.50)	LC (0.41)	PLAT (0.30)
France	PLAT (–2.30)	PLAT (–2.06)	PLAT (–1.64)	LC (–0.19)	LC (–0.33)	PLAT (–0.40)
Portugal	LC (–7.49)	LC (–7.49)	LC (–7.49)	PLAT (-6.55)	PLAT (–6.56)	PLAT (–6.56)
Spain	PLAT (–2.44)	PLAT (–2.57)	PLAT (–2.27)	LC (–0.25)	LC (–0.34)	PLAT (–0.08)

Source: own calculation

Better projections we got for female in all countries in cluster, regardless of the length of historic period. The worst LE projections are recorded for Portugal. Portugal is an outlier on the European landscape. As the Western European country had the lowest male and female LE in 1970. However, by 2009, it had reached 76 and 82 years, respectively — a huge increase.

Table 4: **The best stochastic mortality models and error rates in LE at birth projections for countries in cluster 3**

	The best-fitted model in historical period (LE at birth projection error)			
	Male		Female	
Historical period:	1972–2000	1982–2000	1972–2000	1982–2000
Czech Republic	CBD (–1.44)	PLAT (–1.53)	PLAT (–1.69)	PLAT (–0.97)
Austria	PLAT (–1.07)	PLAT (–0.37)	PLAT (–0.45)	PLAT (0.02)
Italy	PLAT (–1.13)	PLAT (–0.50)	PLAT (–0.11)	PLAT (0.27)

Note: Germany excluded due to insufficient length of historical data. Source: own calculation

For countries from cluster 3 projected LE at birth for 2012 is close to the actual. The best results were obtained for Austria and Italy for female.

Table 5: **The best stochastic mortality models and error rates in LE at birth projections for countries in cluster 4**

	The best-fitted model in historical period (LE at birth projection error)			
	Male		Female	
Historical period:	1960–2000	1970–2000	1960–2000	1970–2000
Estonia	PLAT (–5.82)	PLAT (–5.98)	LC (–4.80)	PLAT (–5.30)
Latvia	PLAT (–5.3)	PLAT (–5.05)	PLAT (–3.1)	PLAT (–3.00)
Lithuania	PLAT (–2.38)	PLAT (–2.04)	PLAT (–1.64)	PLAT (–1.51)
Poland	PLAT (–3.15)	PLAT (–2.93)	PLAT (–2.28)	PLAT (–2.17)
Slovakia	PLAT (–3.65)	PLAT (–3.07)	PLAT (–1.83)	PLAT (–1.49)

Source: own calculation

Plat model (except Estonia, female, 1960–2000) provides the best — among considered methods, but not good in general — fits to the historical data, regardless to the sex and length of historical period. The errors in female life expectancy at birth are considerably smaller than for male, in case of Lithuania, Poland and Slovakia the errors span from 1.49 to 2.28 years.

Table 6: The best stochastic mortality models and error rates in LE at birth projections for countries in cluster 5

	The best-fitted model in historical period (LE at birth projection error)					
	Male			Female		
Historical period:	1950–2000	1960–2000	1970–2000	1950–2000	1960—2000	1970–2000
Hungary	PLAT (–4.50)	PLAT (–5.18)	PLAT (–5.1)	PLAT (0.00)	PLAT (–0.38)	PLAT (–0.34)
Bulgaria	PLAT (–2.68)	PLAT (–3.30)	PLAT (–3.50)	PLAT (0.03)	PLAT (–0.33)	PLAT (–0.51)

Note: Greece and Slovenia excluded due to insufficient length of historical data

Source: own calculation

For countries from cluster 5, regardless of sex and historical period, PLAT model provides the best fit to the historical data. However, for male errors were considerably larger than for female. Point projections of LE for 2012, based on 40–years length historical period, are very close (for Hungary even the same) to the actual life expectancy.

6. Conclusions

In last years insurance industry favours a stochastic approach to modelling and projection mortality and life expectancy.

All of models used in this study share the same underlying assumption that the age, period and cohort effects are qualitatively different in nature. Specifically, there is randomness from one year to the next, perhaps caused by local environmental factors, which we do not observe between ages. Original CBD model and generalizations of CBD differ from LC and APC model in that they assume a functional relationship (and hence smoothness) between ages. The choice of the model M6 to M8 depends on one's personal beliefs about the underlying randomness in the age, period and cohort effects, e.g. if one believes that there should be an underlying smoothness between ages, that there is randomness between cohorts, and randomness from one year to the next, then greater weight might be placed on models M6 to M8 (Cairns at al 2007). PLAT model combines the CBD model with some features of the Lee-Carter model to produce a model that is suitable for full age ranges and captures the cohort effect. Our results are not fully consistent with the existing studies, where LC and CBD models are not treated as good candidates for modelling mortality and life expectancy. However it is worth to notice that producers of national statistics do not use sophisticated models.

Life expectancy at birth projections, resulting from seven mortality models, differs between countries. In general PLAT model provides the best fits the historical data, followed by LC and original CBD. Generalization of CBD models was not found as the best for modelling mortality rates. In general, better fit in sample and life expectancy projections were obtained for females.

We used different historical periods to fit the models. In point of fact institutions use different historical period – long period (e.g. 80-100 years) and relatively short (15-20 years). The choice of the period may produce different outcomes. However, we do not notice any correlation between historical period and error rates in life expectancy projections.

For each cluster projection resulted in a wide range of life expectancy at birth for 2012. The worst projections were obtained for countries from cluster 4 containing Estonia, Latvia, Lithuania, Poland and Slovakia, especially for male. Life expectancy at birth in these countries has increased by 6 to 10 years since 1970, after years of stagnation or decline. Such a trend in life expectancy results in best fits of PLAT model, but worst projections (e.g. male in Hungary and Bulgaria). For these countries other models or different historical period should be considered. The best projections were obtained for countries from cluster 3 (Czech Republic, Austria and Italy) and cluster 1 (Sweden, Finland, Luxembourg, Denmark, Netherlands). Increase in life expectancy continued at a steady rate in historical period results in good fit of LC, and also good projections. CBD provides a good fit when non-linear trend is observed (e.g. Denmark, male Netherland).

It is important to note that conducting the spatial cluster analysis by sex would improve identification of common trends in life expectancy in clusters. Besides, we have settled on the four economic and demographic variables, according to their relevance and variation across space, but the research on finding the best set of predictor variables of historic and future mortality level should be continued.

References

Assunção, R.M., Neves, M.C., Camara, G. and Freitas, C.D.C., 'Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees', International Journal of Geographical Information Science, 20, 2006, pp. 797–811.

Cairns, A.J.G., Modelling and Management of Longevity Risk, 2013, https://www.actuaries.org/lyon2013/papers/AFIR_Cairns.pdf.

Cairns, A.J.G., Blake, D. and Dowd, K., 'A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration', Journal of Risk and Insurance, 73, 2006, pp. 687–718.

Cairns, A.J.G., Blake, D. and Dowd, K., Modelling and Management of Mortality Risk: A Review, Working paper, Heriot-Watt University, and Pensions Institute Discussion Paper PI-0814, 2008.

Cairns, A.J.G, Blake, D., Dowd, K., Coughlan, G., Epstein, D., Ong, A., and Balevich, I., A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. North American Actuarial Journal, 13(1), 2009, pp. 1–35.

Currie, I.D., Durban, M. and Eilers P.H.C., Smoothing and forecasting mortality rates, Statistical Modelling 4, 2004, pp. 279-298.

European Commission, 'The 2015 Ageing Report, Underlying Assumptions and Projection Methodologies', European economy, 8, 2014.

Eurostat, http://ec.europa.eu/eurostat/cache/metadata/en/demo_mor_esms.htm, 2016.

Haberman, S. and Renshaw, A.., A comparative study of parametric mortality models. Insurance: Mathematics and Economics, 48, 2011, pp. 35–55

Hunt, A., Blake, D.P., A General Procedure for Constructing Mortality Models. The Pensions Institute, Pensions Institute, Discussion Paper PI-1301, 2015.

Hyndman, R.J., Ullah, M.S., 'Robust forecasting of mortality and fertility rates: A functional data approach', *Computational Statistics & Data Analysis*, 51(10), 2007, pp. 4942–4956.

Lazar, D., Buiga, A. and Deaconu A., 'Common stochastic trends in European mortality levels: testing and consequences for modeling longevity risk in insurance', *Romanian Journal for Economic Forecasting*, 2, 2016, pp. 152–168.

Lee, R., 'Mortality Forecasts and Linear Life Expectancy Trends' in *Bengtsson, T. (ed.) Perspectives on Mortality Forecasting, III. The Linear Rise in Life Expectancy: History and Prospects, Social Insurance Studies* 3, 2003, pp. 19–40.

Lee, R. and Carter, L., 'Modeling and forecasting the time series of US mortality', Journal of the American Statistical Association, 87, 1992, pp. 659–671.

Li, N. and Lee, R., Coherent mortality forecasts for a group of populations: an extension of the Lee-Carter method. Demography, 42(3), 2005, pp. 575-594.

Plat, R., 'On stochastic mortality modelling', *Insurance: Mathematics and Economics*, 45(3), 2009, pp. 393–404.

Shang, H.L., Booth H. and Hyndman R.J., *Point and interval forecasts of mortality rates and life expectancy: A comparison of ten principal component methods*, Demographic Research, 25, 2011, pp. 173–214.

Stoeldraijer, L., van Duin C., van Wissen L. and Janssen, F., *Impact of different mortality forecasting methods and explicit assumptions on projected future life expectancy: The case of the Netherlands*, Demographic Research, 29, 2013, pp. 323–354.

Villegas, A.M., Kaishev, V. and Millossovich P., *StMoMo: An R Package for Stochastic Mortality Modelling*, 2015, https://cran.r- project.org/web/packages/StMoMo/vignettes/StMoMoVignette.pdf.

9 A growing demand for small area statistics

ASTA MANNINEN[77], PILAR MARTIN-GUZMÁN[78], DEREK BOND[79]

Table of Content

([77]) Member of the ESAC, European Statistical Advisory Committee
(78) Member of the ESGAB, European Statistical Governance Board
(79) Independent Consultant in Statistics and Development

1. Introduction

The demand for small area statistics is coming from all sectors: government, commercial and business, research, media, non-profit and voluntary organisations, and citizens and local communities. The users of small area statistics are organisations, agencies, groups and networks, research and innovation, education, information services, and individuals. Their interests and responsibilities range from local to international level. Thus, the information requests also embrace international comparative small area statistics.

Timely and spatially fine-scaled statistics of high quality are crucial for mapping, monitoring and understanding essential dimensions and changes of our society and people´s everyday life. In the case of the European Statistics we share the opinion of the European Statistical Advisory Committee (ESAC) which states that there is a continuing need for adequate geographical breakdowns of the published data and that the aim should be to provide spatial statistics fine-tuned enough to support national and regional policy-making. ESAC has also paid attention to the increasing role of indicators in policy- and decision-making at all government and geographical levels. From the users´ point of view this ´indicator movement´ should be thoroughly addressed in terms of how the indicators are developed, made available and presented on various geographical levels. In other words, the indicators should meet the standard of high quality small area statistics.

The request of comprehensive small area statistics is a great challenge to official statistics, especially if the provision of small area statistics mainly relies upon surveys. Official statistical surveys are high-quality data sources, though expensive, time-consuming and burdensome. Therefore, small are estimation techniques, methodologies and practices are developed (80). Administrative data and new geospatial data are also being explored as part of a new integrated multisource approach for producing small area statistics.

The paper will deal with various demands for small area statistics from *the users´ point of view*. How official statistics currently meet the demands will also be discussed in the presentation of selected examples and case studies from Spain, United Kingdom and Finland.

2. Why are small area statistics gaining in importance?

Small area statistics are gaining in importance. There are many reasons for the increasing demand of the broad scope of small area statistics. Firstly, we would like to mention the major trends shaping our world, such as urbanisation, globalisation, the international migration, the increasing inequality in many aspects (well-being, income, wealth and so forth), digitalisation and climate change. Monitoring these trends requests special attention on cities and regions, which are the prime scene of these trends. Consequently, sustainability is to be monitored. This in turn raises the question how the current UN Sustainability Development Goals might be properly translated into statistical and other information actions. Territorialisation of relevant indicators will certainly be one of the key actions (ESAC 2016).

Secondly, advancing open government and democracy is a major trend in Europe and also overseas. For official statistics this means supply of coherent statistics not only on national, city- or regional level but also on district- and sub-city-level, even on residential areas and neighbourhoods. From the users´ point of view the supply of statistics should preferably be as open data allowing for customised use free of charge.

(80) The European Commission has financed research projects such as SAMPLE (Small Area Methods for Poverty and Living Conditions Estimation) and AMELI (Advanced Methodology for European Laeken Indicators).

Thirdly, small area statistics are an invaluable tool in running a city, advancing economic development in a region, providing the means for management and decision making on various administrative levels, and empowering citizens and local communities for active participation. Everyday life is local and thus small area statistics come close to citizens lives and their living areas. We may add that small area statistics through empowering citizens also improve citizens´ statistical literacy at large.

Targeted policies at the EU, national and local levels serve as a fourth example. These require high quality, timely and relevant small area statistics. The Europe 2020 strategy aiming at creating a smart, sustainable and inclusive economy is closely linked to the developments happening in cities and regions. The EU cohesion policy being the principal investment tool for Europe 2020 targets is crucial to the successful development in regions and cities. Adequate small area statistics should be available for monitoring and impact assessment, and for development measures.

Territorial cooperation (and sometimes competition) requests comprehensive small area statistics that are comparable. Examples on territorial cooperation are: cooperation on national and transnational level, empowering cities and regions through traditional and new forms of cooperation, and re-enforcing neighbourhoods and residential areas. Comparative statistics contributes significantly to an understanding of the dynamics and challenges, opportunities and diversity that explain or inform development in cities and regions across time, countries and cultures. For example in the case of cities, comparable statistics allow city leaders to know how their city is doing relative to their peers. Small area statistics contribute to capacity building of a territory.

Finally, many cities and regions draw annual population projections not only on the city and the region but also on sub-regional and sub-city level down to city-districts. This is a request because of the rapidly changing operating environment and because of strict financial resources. In the Nordic countries, land use planning, planning for housing and services and so forth are challenging responsibilities and tasks for a municipality. The information needs accordingly embrace a good amount of small area statistics rich in content and of high quality, including also projections and research. In terms of quality and value added, it is important that the small area statistics are consistent over time and thus allowing the users to picture the progress and developments per region, city, city district or neighbourhood over time.

As we may see from the above said, the request for comprehensive statistics of high granularity both in terms of geography and subject content is obvious and urgent.

3. What is the value added of small area statistics?

Small area statistics add new content and characteristics to the information stock. Small area statistics also encourage development of new tools for access and analyses of geospatial statistical data. Providing small area statistics in practice usually involves a process of merging various data sources, which gives further advantages, such as improved data quality through increased use and new uses of already existing data.

Integration of statistical and geospatial data makes it possible to produce statistics based on new spatial entities of high relevance, for example new types of traffic and commuting statistics of better fit for purpose. Statisticians and planners may use real routes, distances and travel time by vehicle.

In the case of small area statistics there is one major characteristic which is critical to the value added of these statistics, namely the delimitation of territorial units. The units should be relevant in terms of at least function, size and comparability. This is challenging especially in the case of international comparisons.

From the users´ point of view small area statistics support and facilitate better informed decisions, evidence based policies, better allocation of resources, better performance, improved location based business intelligence, detection of assets and shortcomings on various geographical levels, and monitoring and measuring impact of territorial and place-based policies. All in all, small area statistics contribute to better quality of life from many aspects.

An appreciated value added is the geography allowing for various graphical presentations and visualisations. Geographical presentations are also a great opportunity to make statistics attractive and more widely used.

The most recent uses of small area statistics and geo-referenced data are to be found in application areas such as smart cities, and smart transport and energy solutions.

4. Examples and case studies

In this section some examples of the type of small area statistics that are available to a power user will be discussed. Examples of internationally compatible datasets are given, this is followed by a brief discussion of nationally compatible small area datasets developed mainly by three national statistics offices (NSOs), but also by city statistical offices in cooperation with the NSOs. The section finishes by giving two examples, both based in Finland, of how small area statistics can be integrated to help enable the power user to arrive at sound policy decisions. The examples chosen are regularly produced and maintained statistical data or derived from a series of repeated surveys. In the selection of examples special attention is paid also to open access.

Geography is at the heart of small area statistics and provides many challenges for official statistics in providing small area statistics Three user focused challenges are: the need for different geographies for different topics; small enough areas are required to allow for local focus; and a high number of boundary changes makes it difficult to merge or combine data sets, and difficult or impossible to monitor change over time. A major challenge facing the producers of the data is that the identified and delineated geographical units should be sufficiently populous to provide statistically significant data.

To respond to these various challenges official statistics has to recognize and include several types of spatial entities in their programmes. Today it is common practices for official statistics to provide data according to legal and standard administrative delineations of geographical units (country, county, municipality, etc.), to use stable building blocks such as statistical units, and to provide georeferenced or geocoded statistical data (grid references, addresses, and other geographical identifiers). Geocoded data considerably improve the availability and usefulness of small area statistics. The spatial data infrastructure of today, supported e.g. by EU´s Inspire directive, facilitates interface services to the users.

As geography is at the heart of small area statistics the issue of how to access and visualise them has attracted much attention. With the growth in geographic information systems there has been many attempts by official statisticians to provide a map based visual interface to access the data as well as display the results. Whilst those traditionally involved in analysing data, such as most current power users, are happy with traditional text based menus and searches to discover the relevant small area data they need, casual users and new power users used to commercial products such as google maps and ArcGIS are more likely to use visual interfaces. The complex cognitive issues combined with the problem of choosing spatial units means that, as presented below, success of such approaches is variable.

4.1. International comparative statistics

International comparative statistics on cities and regions are requested on an ongoing basis by many power users. The real interest of these users can be very varied. For example, at an international government level they can be trying to decide where or how to allocate fund or assess the impact of policy. At a local government level they might be trying to measure how effective local initiatives are compared to those adopted by other areas. In the commercial sectors they are required for market assessment to help decide where to launch new products and services or locate a new manufacturing plant.

For Europe, the Eurostat´s website offers four dedicated, web accessible sections on comparative sub-national official statistics: City statistics, Cities (Urban Audit) http://ec.europa.eu/eurostat/web/cities/data/database, Degree of urbanisation http://ec.europa.eu/eurostat/web/degree-of-urbanisation/overview, Metropolitan regions http://ec.europa.eu/eurostat/web/metropolitan-regions/overview, and Regional Statistics http://ec.europa.eu/eurostat/web/regions/data/database. Of these sections the Urban Audit was developed to allow for transnational comparisons of city level data. However, the task has not been easy and because of different approaches, three of which are discussed below, few cities have been able to provide a complete set of comparable statistics.

As well as the online data sets many publications are available online. Examples are: 'Statistics on cities, towns and suburbs' a new flagship online publication issued 2016, 'Quality of Life in European Cities 2015´ a perception survey conducted every three years and 'The Regional Yearbook' the newest online publication being issued 2016.

GISCO, the 'Geographical information system of the Commission', is a permanent service of Eurostat providing geographical information and maps. Selected datasets are available also to the public.

Examples of results of the cooperation between Eurostat and ESAC concerning availability of regional and territorial breakdowns of official statistics are that Eurostat has taken ESAC Opinions into account in specific statistical priority areas such as the regional and territorial dimension of the Europe 2020 strategy, People´s Europe, and the renewed strategy for population and housing census focusing on geocoding census 2021. ESAC has actively put forward that the work on indicators within the ESS (European Statistical System) need to reach tangible progress towards more geocoding of European data in order to provide a deeper understanding and insight of the developments at regional and local levels. Currently, ESAC is contributing to the DIGICOM project on digital communication, user analytics and innovative products.

At the international level two of the main power users and also providers of internationally comparable small area statistics are the OECD and the UN. Examples of OECD and UN datasets are 'Regional Statistics and Indicators' and the 'UN-Habitat Urban Data Portal', respectively. Examples of how these datasets are used include: 'How's Life in your region? Better Life Index' (http://www.oecdbetterlifeindex.org/), 'Regions at a Glance', the newest one being issued 2016 (http://www.oecd.org/regional/oecd-regions-at-a-glance-19990057.htm), 'OECD Regional Outlook' (http://www.oecd.org/regional/oecd-regional-outlook-2016-9789264260245-en.htm from the OECD) and 'The World Urbanization Prospects series' (https://esa.un.org/unpd/wpp/) and the UN-Habitat's *City Prosperity Initiative (CPI)' and the corresponding online publication 'State of the World´s Cities 2012/2013. Prosperity of Cities'.

4.2. Spain. Selected examples and development measures

Establishing infrastructure and other capacities for producing small area statistics is challenging for official statistics in any country. *The Population Register (Padrón Continuo)* is the result of a cooperation among the 1,825 municipalities existing now in the country and the National Statistical Office. It provides rich municipality level statistics and for big cities also sub-city level statistics (e.g. for the 21 districts of which the municipality of Madrid consists). At municipality level information on demographic statistics, nationality, country of birth and maximum level of education attained are provided. The labour market statistics are a particularly sensitive issue in Spain, due to the high level of unemployment in the country. The Labour Force Survey provides information only at provincial levels. There is a strong demand for information of higher granularity on the labour market.

This demand is particularly significant in the two insular regions, as public policies are there partly implemented at insular level, and many strategic plans are specifically designed for each of the islands. The project CANAREA, run by ISTAC (the regional Statistical Office of the Canary Islands), provides rich information on the labour market for each of the seven islands. The islands differ greatly from each other in terms of population and also climate conditions. Consequently, a considerable number of policy decisions has to be taken at insular level. Data are estimated, using small area techniques, on the results of the Labour Force Survey in combination with the records from the Social Security and the Public Employment Service files. These two administrative sources do not follow international definitions and, in addition, the definitions of variables can be - and often are - changed through a ministerial decree. Therefore, they cannot be used directly for the provision of comparable series on labour force statistics for fine-tuned breakdowns.

Some other regional statistical offices are also adopting similar practices. GDP at municipal level is being calculated for some of the big cities (e.g. Barcelona).

The Central Directory of Enterprises can provide information at municipality level on the number of enterprises, their legal status, main economic activity and size (in terms of employees' stratum). It is a source of indicators for enterprise demographic analysis. Unfortunately, there are some problems in assigning activities to the different branches in the multi-locational enterprises.

4.3. United Kingdom. Neighbourhood statistics in Northern Ireland

In 2000 the United Kingdom's Office of National Statistics (ONS) introduced the *Neighbourhood Statistics Service (NESS)*([81]) to address significant gaps in the information required for evidence-based sub regional policy making, and to contribute to achieving the Government's National Strategy for Neighbourhood Renewal. To compliment NESS which covered England and Wales, the Northern Ireland Statistics and Research Agency (NISRA), in 2004, introduced the corresponding *Northern Ireland Neighbourhood Information System (NINIS)*([82]) to provide a similar web based method of accessing sub-regional information. Obtaining the information needed for these systems has provided the Government Statistical Service with many challenges as well as opportunities.

In the United Kingdom the approach to collecting official statistics is based on the use of censuses and sample surveys. This raises many interesting issues surrounding the development of sub-regional statistics that are not encountered in countries that have adopted a register based approach to their official statistics. These issues have resulted in the need for both NESS and NINIS to develop a flexible approach to proving small area statistics. The issues can be loosely addressed under two main headings, the appropriate geographies to use and the frequency and granularity of the statistics that are available at the sub-regional level. The geography issue is complex as the basic spatial identifier in most surveys and censuses is the post code. Post code areas are not directly related to administrative areas so various assumptions have to be made. As for administrative areas these change over time and the sub-regions (normally wards) also change to reflect population movements. Thus deciding on the appropriate geographic units is difficult. So both NESS and NINIS provide information for various possible spatial units. For example, NINIS provides information for 13 possible spatial units, though few series are available for more than three or four of these units. Great care has to be taken that sensitive individual data cannot be obtained by manipulation of the spatial units.

The availability of sub-regional data is limited as very few surveys are large enough to facilitate the provision of small area data. The lack of timely and relevant census data makes the adoption of standard small area estimation techniques difficult as there is a lack of variables that can be used to add weight. Therefore, there are very few examples of where small area statistics have been successfully used in the United Kingdom for policy formation. Those that do exist are mainly in the area of health where there are large amounts of epidemiological data available with high granularity.

The situation is however changing, in recent years the Government has been trying to move to an outcome based approach thus requiring the availability of high quality spatially disaggregated economic and social indicators([83]). In addition, the `Digital Economy' bill([84]) currently being discussed will enable more use to be made of administrative data. The need for spatially disaggregated economic and social data was

(81) http://www.neighbourhood.statistics.gov.uk/dissemination/
(82) http://www.ninis2.nisra.gov.uk/public/Home.aspx
(83) https://www.northernireland.gov.uk/publications/draft-programme-government-framework-2016-21-0
(84) https://services.parliament.uk/bills/2016-17/digitaleconomy.html

highlighted also in the recent official ´Bean Report' on economic statistics. The next few years should see major improvements in coverage and hence relevance and usefulness of both NESS and NINIS.

4.4. Finland and the case of Helsinki. Integrated geospatially enabled statistics

The National Statistical Offices (NSOs) in Finland, Sweden, Norway and Denmark have had access to large amounts of administrative data for statistical purposes for years. Thus these NSOs have been forerunners in developing and adopting register based censuses, which are a prime example of integrated multisource small area statistics. In the Nordic countries we talk about the 'register tradition' in statistics on national, regional and local levels. This tradition is beneficial to all users. In addition, it´s cost-effective and enables timely dissemination and publication.

In the Helsinki region, small area statistics based on national and municipal administrative registers, break downs of official statistics and other regular data sources represent a best practice dating back to the 70's. One example is *Helsinki small area statistics database* (http://www.aluesarjat.fi/) containing compatible statistics on small areas ranging from residential areas and neighbourhoods to districts and major districts of the cities in the region. The database offers continuously updated statistical data on phenomena such as demography, housing, construction, education, employment and income. The data is available as open data in the *Helsinki Region Infoshare*([85]) in various formats, also as application programming interfaces (API). A spin-off product of the *Helsinki small area statistics database* is the annual web-publication *Helsinki by District*, which is widely used and well received among the citizens. Enterprises, service providers and media are also active users of these small area statistics. Not to forget that the small area statistics database with its long time series is invaluable in assisting researchers, statisticians and information officials engaged in urban and regional issues.

In the case of Helsinki, the Urban Facts department of the city([86]) provides comprehensive small area statistics and applies the standards of official statistics. The data sources used are multisource originating from the small area statistics database and other statistical data sources of the city, administrative registers of the city, Statistics Finland, the National Institute for Health and Welfare, and many other organisations or institutes. The cooperation with Statistics Finland is essential and invaluable.

A joint Nordic urban database entitled NORDSTAT was set up in 1990 on the initiative of the capital cities. NORDSTAT consists of the major Nordic cities and their functional urban regions and uses the framework and design of *Helsinki small area statistics database.* The content of the database follows the standards of official statistics. The population projections of the cities and regions are also part of the database. The responsibilities of the management of the database rotate among the capital cities.

Advancing urban mobility and accessibility with the aim of reaching smart and sustainable traffic and transport solutions is a challenge task requesting a lot of integrated fine-grained geocoded data. New research([87]) in this field has successfully utilised small area statistics accessible as open data and innovative collaboration.

The next two examples are about geospatially enabled statistics which utilise the approach of integrating and customising multisource information. The success factor behind these examples is data interoperability, open data and cooperation. In the case of the City of Helsinki it means that the city can enrich its own data through cooperation with Statistics Finland and other information-intensive public organisations with an open data policy.

(85) http://www.hri.fi/en/http://www.hri.fi/en/

(86) City of Helsinki Urban Facts (http://www.hel.fi/www/tieke/en) is the expert organisation responsible for urban statistics, research, information service and open data as well as records management at the City of Helsinki.

(11) Henrikki Tenkanen,Vuokko Heikinheimo, Olle Järv, Maria Salonen &Tuuli Toivonen: Open data for accessibility and travel time analyses: Helsinki Region Travel Time and CO2 Matrix. University of Helsinki. Department of Geosciences and Geography. Helsinki 2016.

The first example in this category is a recently published information service called the *Liiteri service* (http://liiteri.ymparisto.fi/) which includes over 1,000 different statistics and hundreds of map layers. The service is maintained by the *Finnish Environment Institute*. Originally, the data derive from numerous different sources and are maintained by different authorities: Statistics Finland, Finnish Environment Institute, Population Register Centre, National Land Survey of Finland, Finnish Transport Agency, National Board of Antiquities, Centre for Economic Development, Transport and the Environment, cities and municipalities, Ministry of the Environment, The Housing Finance and Development Centre of Finland and the Geological Survey of Finland. The *Liiteri service* collects the data from these sources in a coherent and interoperable form and offers one source and interface to all data. Thus it is easy to use the data for the end-users, even for those who are not so familiar with geographical information systems. Moreover, the users may combine statistics, maps and their own data and download it all for own uses. The *Liiteri service* is well received among users on local, regional and national level.

The second example of integrated multisource geospatially enabled statistics is about commuting and comes from *Statistics Finland.* Statistics Finland produces annual population and employment statistics by regions. The data are mainly based on around 40 administrative and statistical data files. The unit-specific data can be linked to map coordinates and then compiled by all regional divisions including small areas such as postal code areas or grids. Statistics Finland has produced commuting statistics since 2005. Data sources including map coordinates of buildings where people live and work have allowed estimates of distances to work and directions of commuting. These calculations have been based on the shortest distance between two points ‘as the crow flies´.

Today, Statistics Finland is able of considerably improving the commuting estimates and extending the use of commuting statistics by integrating new multisource geospatial data to its register data. These new data are the Digiroad - a national road and street database (by the Finnish Transport Agency), data of automatic traffic measuring devices (by the Finnish Transport Agency) and timetable data of public transportation (by Helsinki Region Transport HSL/Journey Planner and the Finnish Transport Agency/Matka.fi). The new commuting statistics are for example distances and time to work by private car, bicycle or public transportation, all of them based on using real commuting routes (Pasila 2016 and Piela 2015).

5 Conclusions and discussion

According to experiences from work on urban and regional statistics([88]) the demand for comparative urban and regional statistics is still today greater than the supply. The increasing pressure on timely data is critical. There are also information gaps. New phenomena usually enter cities first and thus statisticians, researchers and planners engaged in analysing and interpreting urban phenomena are in the forefront asking for official statistics on new and emerging issues. A question remains as to what extent research at universities has been invited to contribute with new insight and suggestions to fill information gaps.

The users´ request of timely international comparative city statistics and indicators is to some extent met by commercially produced city indicators. There are many private agencies on the international information market and many of them have a long tradition of producing city indicators on annual bases (Mercer, The Economist Intelligence Unit, Monocle and UBS). Moreover, there is the World Council on City Data Open Data Portal ([89]) offering standardized city data.

Advancing small area statistics suggests a number of recommendations based on experiences and achievements so far. An attempt to summarize useful recommendations would appear as follows. The users would welcome geo-referenced data sets accessible as open data to as great extent as possible. Another next option would be geo-referenced data made available as public use files. Thirdly, it would be desirable to get geographical breakdowns of already published and new data, statistics and indictors. Access to integrated and interoperable small area statistics and data would enable analyses and research of complex urban phenomena and interdependences. User involvement from an early stage on in

([88]) The Standing Committee for Urban and Regional Statistics and Research (SCORUS) of the International Association for Official Statistics (IAOS).
([89]) http://open.dataforcities.org/

developing official statistics, also small area statistics, would be a good tool for official statistics to address and understand user needs. It could significantly add to the insight of capacity building for providing small area statistics on an ongoing basis.

The importance of comparative small area statistics has been underlined in earlier chapters. International cooperation in general and especially focused on international frameworks and standards would support comparability and integration of statistics and geospatial data[90]. Fortunately, these competence areas are strengths of national and international official statistics.

References

AMELI. Advanced Methodology for European Laeken Indicators. https://ec.europa.eu/eurostat/cros/content/ameli_en

Bean, C., 'Independent review of UK Economic Statistics: Final Report', London: HM Treasury/Cabinet Office, online https://www.gov.uk/government/publications/independent-review-of-uk-economic-statistics-final-report, 2016.

City of Helsinki Urban Facts. www.hel.fi/tietokeskus.

ESS Vision 2020. http://ec.europa.eu/eurostat/web/ess/about-us/ess-vision-2020.

Europe 2020 in a nutshell. http://ec.europa.eu/europe2020/europe-2020-in-a-nutshell/index_en.htm. Last update: 18.08.2016.

Finnish Environment Centre. www.ymparisto.fi and http://liiteri.ymparisto.fi/.

Helsinki by District. http://www.hel.fi/hel2/tietokeskus/julkaisut/pdf/16_05_27_Helsinki_alueittain_2015_Tikkanen.pdf

Helsinki Region Infoshare - open data service. www.hri.fi

INSPIRE Directive. Infrastructure for Spatial Information in the European Community. http://inspire.ec.europa.eu/

Jaakola A. (25.9.2016): Memo on small area statistics in Helsinki. City of Helsinki Urban Facts.

Mercer 2017. 2017 Quality of Living Rankings. https://www.imercer.com/content/mobility/quality-of-living-city-rankings.html.

Monocle 2016. https://monocle.com/film/affairs/top-25-cities-2016/.

Opinion on a pre-final draft EU SDG indicator set. ESAC Doc. 2017/22. 29.3.2017.

Office for National Statistics, UK. https://www.ons.gov.uk/ .

Pasila A., *Enriching census data with the Journey Planner to produce statistics on commuting*. Presentation at Nordiskt Statistikermöte, Stockholm 22-24 August 2016.

Piela P., 'Entä jos kaikki polkisivat töihin. - If everybody would bike to work', An Article in Tieto & Trendit – magazine, Statistics Finland. 10 March 2015.

Quality of life in European Cities 2015. Flash Eurobarometer 419. http://ec.europa.eu/regional_policy/sources/docgener/studies/pdf/urban/survey2015_en.pdf .

Regions and Cities. Statistics Explained, Eurostat. http://ec.europa.eu/eurostat/statistics-explained/index.php/Regions_and_cities.

Regional and geographic information services at Statistics Finland. www.stat.fi and https://tilastokeskus.fi/tup/alue/index_en.html.

[90] The UN Statistical Commission is currently developing a global statistical and geospatial framework which will allow for harmonised small area statistics (see United Nations Expert Group on the Integration of Statistical and Geospatial Information).

Report on the ESAC workshop on indicators: user requirements, methodological issues and communication challenges, 10 May 2016, The Hague. ESAC 2016. http://ec.europa.eu/eurostat/documents/7330775/7339476/Report+on+the+Hague+workshop+on+indicators_10062016.pdf/60b456f1-d50c-435d-bf50-df46f9912a1d.

SAMPLE. Small Area Methods for Poverty and Living Conditions Estimates. http://www.sample-project.eu/.

Tammilehto-Luode M., Nieminen J. and Ruotsalainen K. (20.9.2016): Memo on small area statistics in Finland. Statistics Finland.

Tammisto R., Tammilehto-Luode M., and Pihlajamaa T. (15.11.2016): Geostat 2. Some experiences and conclusions – applying the GSPBM to geospatial statistical data production. European Forum for Geography and Statistics Conference 2016, Paris 15 November 2016.

http://www.efgs2016.eu/access-by-session/session-3-integration-of-geography-in-the-statistical-production-process/.

Tenkanen H., Heikinheimo V., Järv O., Salonen M. and Toivonen T. (2016): Open data for accessibility and travel time analyses: Helsinki Region Travel time and CO2 Matrix. University of Helsinki. https://agile-online.org/Conference_Paper/cds/agile_2016/shortpapers/184_Paper_in_PDF.pdf.

The DIGICOM project, Eurostat. https://ec.europa.eu/eurostat/cros/content/essvip-digicom_en.

The Economist Intelligence Unit 2016. Global Liveability Ranking 2016. https://www.eiu.com/public/topical_report.aspx?campaignid=liveability2016.

The European Statistical Advisory Committee (ESAC) is the voice of the user needs in the European Statistical System (ESS). http://ec.europa.eu/eurostat/web/european-statistical-advisory-committee-esac/about-esac/assignment.

The National Statistics Institute, Spain. http://www.ine.es/en/welcome.shtml.

The OECD Regional Database. http://www.oecd.org/gov/regional-policy/regionalstatisticsandindicators.htm.

The OECD Metropolitan database. http://www.oecd.org/gov/regional-policy/regionalstatisticsandindicators.htm.

The World Urbanization Prospects. United Nations. https://esa.un.org/unpd/wup/.

The UN Statistical Commission. United Nations Expert Group on the Integration of Statistical and Geospatial Information.http://ggim.un.org/docs/meetings/3rd%20UN-EG-ISGI/Proposal-for-a-global-statistical-geospatial-framework.pdf.

UN Sustainable Development Goal indicators website. http://unstats.un.org/sdgs.

UN-Habitat Urban Data. Free access to data about countries and cities around the globe. http://urbandata.unhabitat.org/.

UN Sustainable Development Goals. 17 goals to transform our world. http://www.un.org/sustainabledevelopment/.

United Nations Expert Group on the Integration of Statistical and Geospatial Information. http://ggim.un.org/docs/meetings/3rd%20UN-EG-ISGI/Proposal-for-a-global-statistical-geospatial-framework.pdf.

Urban Europe – statistics on cities, towns and suburbs. 7.9.2016. http://ec.europa.eu/eurostat/statistics-explained/index.php/Urban_Europe_%E2%80%94_statistics_on_cities,_towns_and_suburbs.

UBS Prices and Earnings 2016. https://www.ubs.com/microsites/prices-earnings/prices-earnings.html.

10 Towards interpretable FDI data in external statistics Filtering distortions arising from globalisation from data of multinational enterprises

BEATA MONTVAI[91]

Table of Content

1. Summary

FDI is a key element in the evolving international economic integration, however, statistical data collection, compilation of statistics and data dissemination are primarily designed and run by individual countries.

Due to globalisation, multinational enterprises become key players of cross border investments. In order to optimise profitability, they establish complex production and financial structures of enterprises, including production facilities and special purpose entities (SPEs). These multinational group structures are designed

([91]) Magyar Nemzti Bank; montvaib@mnb.hu

to achieve best results on group level profiting from different national facilities (including differences in national Law and legislation) in different countries. **Although SPEs are residents of a host country, basically they perform economic activities abroad, they are marginally connected to the domestic economy.** They are primarily involved in intra-group intermediation of financial resources, and may result in substantial external financial flows and positions of the host country.

Furthermore, there are some other resident subsidiaries where these **special activities and intra-group functions are combined with real economic operations**. Because of their real economic activities, they cannot be classified as SPEs, despite they also take part in financial intermediation (i.e. their foreign parent companies pass through large amounts from one foreign subsidiary to another one within the multinational enterprise group).

These special entities and activities are different in many countries but may have significant distortive effects on national statistical data. To avoid these distortions, even data collections, processing, publications- and interpretation - of national FDI-statistical publications need to be supplemented.

This paper presents the history, development and recent findings and practices applied by the Magyar Nemzeti Bank in order to improve the quality of BoP and FDI statistics and to reduce these distortive elements in FDI statistics. The MNB (the Central Bank of Hungary) is responsible for the compilation of BoP and IIP statistics, identifies and publishes FDI statistics with the **following supplements**:

- separate publication of BoP and IIP for Special purpose entities (SPE);
- separate publication of capital in transit transactions and asset portfolio restructuring transactions
- separating mergers and acquisitions within FDI equity transactions,
- publication of inward FDI stock by the country of the ultimate investor.

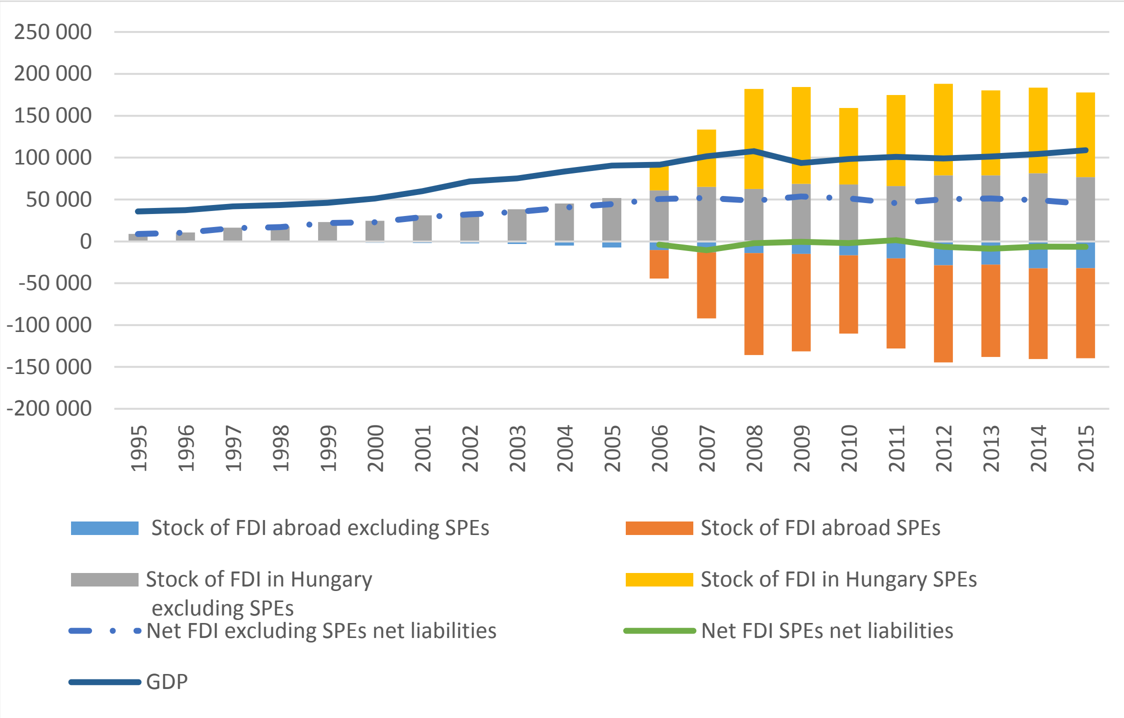
The new editions of international statistical guidelines, such as the Balance of Payments Manual 6th edition (BPM6), and especially the OECD Benchmark Definition on FDI 4th edition (BMD4) gives new measures of foreign direct investment. However, SPEs, capital in transit and asset portfolio restructuring transactions had appeared and tackled in Hungary's compilation practice well before the completion and implementation of the updated international statistical guidelines. Separating mergers and acquisitions within equity transactions and compiling inward FDI positions by the country of ultimate investor are considered by the OECD as important supplemental items needed to support users better understand and interpret FDI statistics.

2. Distortive effects of globalisation are visible in FDI statistics

Hungary is an open economy in Central Europe, member of the EU as of 2004. Foreign direct investment (FDI) is a key driver of international economic integration. The significance of FDI activity among OECD countries is shown in Figure A in Annex 1. Inward and Outward FDI stocks (excluding SPEs) as % of GDP at the end 2015.

The source of FDI data is BOP statistics. In Hungary BOP statistics data with the same methodology (BPM6 which is consistent with BD4([92])) is available since 1995. Separating distortive elements of globalisation in BOP and FDI statistics in Hungary was essential well before the introduction of the updated international guidelines implemented in the member states of the European Union in 2014.

Figure 1: Hungary: GDP, Inward and outward FDI stocks of SPEs and excluding SPEs (million EUR)



Source: MNB

Figure 1 shows that there are different periods of FDI activities in Hungary:

- Between 1995 and 2006 inward FDI had dynamically expanded (reaching 50 billion euros by 2005) and significantly exceeded (marginal) outward FDI. From the end of the nineties path-through activities were present via off-shore companies, at that time their net transactions were recorded in the BOP statistics.
- The off-shore status has been terminated, but path through activities had remained. Statistical category had to be implemented for separating path-through entities - Special Purpose Entities (SPEs). As from 2006 **BOP and IIP statistics** has been published **including SPEs** fulfilling the international guidelines and **excluding SPEs** for economic analytical purposes. Total (Including SPEs) FDI stocks significantly differ from FDI excluding SPEs. The level of inward and outward stock excluding SPEs remained stable between 2006-2011.
- After 2011, besides SPEs new distortive elements had appeared, i.e. capital in transit and asset

([92]) IMF Balance of Payments and International Investment Position Manual, Sixth edition, 2009
OECD Benchmark Definition of Foreign Direct Investment Fourth edition, 2008

portfolio restructuring transactions which can be separated in flows but until now cannot be filtered from stock data. Thus, concerning FDI stocks some slight further expansion has been recorded, mostly due to new distortive effects.

This presentation describes the **present practice of the MNB**, including recent developments in **separating new distortive effects and further supplemental breakdowns in the FDI publications** of MNB supporting users in the interpretation of FDI data.

2.1. 1999-2005: Off-shore companies

At the nineties, it was possible to establish off-shore companies in Hungary. Huge transactions – in and outflows in a short period of time had appeared in the magnitude of billions of Euros without affecting the Hungarian economy. Off-shore companies were defined by an administrative concept, we had an information registry based on administrative data source. Their statistical treatment was based on their type of activities. There were two types of off-shore companies: one with pass-through activity: these off-shore enterprises were passively rechannelling huge funds. For that type of entities, we applied the net recording of their transactions. At that time positions were calculated as cumulated net flows. The other type of off-shore entities was trading with property rights. We applied no difference in recording their transactions compared to „normal” enterprises.

There were no separate set of data published on off-shore companies.

As of 1 January 2003, the Corporate Taxes Law was amended - enterprises with off-shore status no longer could be established in Hungary. The existing off-shore firms had to be transformed into normal businesses by 1 January 2006 at the latest. Some off-shore companies had left Hungary. However, there were some newly established companies as well with pass-through activity and not registered as off-shore companies. **We had to find definition for identifying pass-through entities**.

In the meantime, the revision of the international standards (BPM6 and BD4) had started and showed that Hungary is not alone: other countries face similar problems. These pass-through enterprises are called in the international guidance special purpose entities.

2.2. From 2006: Special Purpose Entities

Due to the widening globalisation, multinational enterprises (MNEs) become the key players of cross border investments. To optimise their profitability, they establish complex structures of enterprises including special purpose entities. In 2006 off-shore companies were outlawed in several countries (including Hungary). Therefore, MNEs had to find new ways and means to realise group level profitability requirements. Different types of SPEs have been created worldwide (including: financing subsidiaries, conduits, holding companies, shell-companies).

There is no internationally agreed definition of SPEs. We define SPEs by the criteria set elaborated in the OECD. (OECD BMD4, p 188):

An enterprise is considered as an SPE if it meets the following criteria:

1. The enterprise is a legal entity,
   - a) Formally registered within a national authority; and
   - b) subject to fiscal and other legal obligations of the economy in which it is resident.
2. The enterprise is ultimately controlled by a non-resident parent, directly or indirectly.
3. The enterprise has no or few employees, little or no production in the host economy and little or no physical presence.
4. Almost all the assets and liabilities of the enterprise represent investments in or from other countries.
5. The core business of the enterprise consists of group financing or holding activities, that is –

viewed from the perspective of the compiler in each country – the channelling of funds from non-residents to other non-residents. However, in its daily activities, managing and directing local operations plays only a minor role.

SPEs in Hungary are resident enterprises basically performing their activities abroad, and their connection with the domestic economy is minimal. SPEs are primarily involved in intra-group intermediation of financial resources, but in fact, their parent companies decide upon the direction and the amount of the funds flowing through them. They are not targets of the direct investment, they just participate in the reallocation of funds within the enterprise group in different countries. Their net investment registered through various financial instruments is close to zero over longer periods of time in the host country. Large amounts are moved through them, and their transactions inflate particularly the assets and liabilities of the financial account. This distorts the statistics describing the real economic and financial processes of the host country.

The major issue is, how to reconcile the two competing principles:

1. SPEs are resident entities: thus, they should be covered by BoP and IIP statistics regardless of their specificities, and

2. Recording their transactions and positions in the core accounts highly distorts and limits the analytical power of BoP and IIP for economic analysis

For proper decision making and analysis we had been segregating transactions and positions of SPEs (as capital with no real effect on domestic economy).

Since 2006 there is no difference in statistical treatment of SPEs. We consider that it is a publication and interpretation issue, not a compilation one.

In line with the international requirements, from January 2006 the MNB has also been preparing BOP and IIP statistics **including SPEs (to enable the analysis of international data by mirror statistics**). Beside these internationally required statistics, the BOP and IIP of Hungary **can also be analysed from an economic aspect based on data excluding SPEs**. In its statistical publications and reports, the MNB analyses data excluding SPEs.

Activities and characteristics of SPEs are changing in time. We have experienced that the characteristics of SPEs can change not only by countries but in time in the same country (i.e. Hungary).

- The early pattern was „incoming FDI equity flow/ and direct loan extension”.
- After 2005, this pattern changed to establishment of non-resident branch financed by loan of a related company.
- Indirectly (through a resident affiliate) having equity stake in a non-resident enterprise
- An enterprise previously providing services sets up a non-resident branch and continues its business through the branch
- Nowadays they remove funds from the branches and include them into another subsidiary

- Occasionally, the changing activity and characteristics of enterprises might result in a need for statistical reclassification from/into SPEs.

SPEs are defined jointly by the Hungarian Statistical Office (HCSO) and the MNB. (The maintenance of SPE registry is a joint effort, too). The source of information is the BOP data collection survey system, the corporate tax declarations.

Before the implementation of BPM6 Hungary has continuously experienced the distortive effects of data including SPEs published in the international institutions. Therefore, we have made initiatives to achieve, that

- all affected countries report their relevant statistics excluding and including SPEs, and
- all international organizations publish both national data

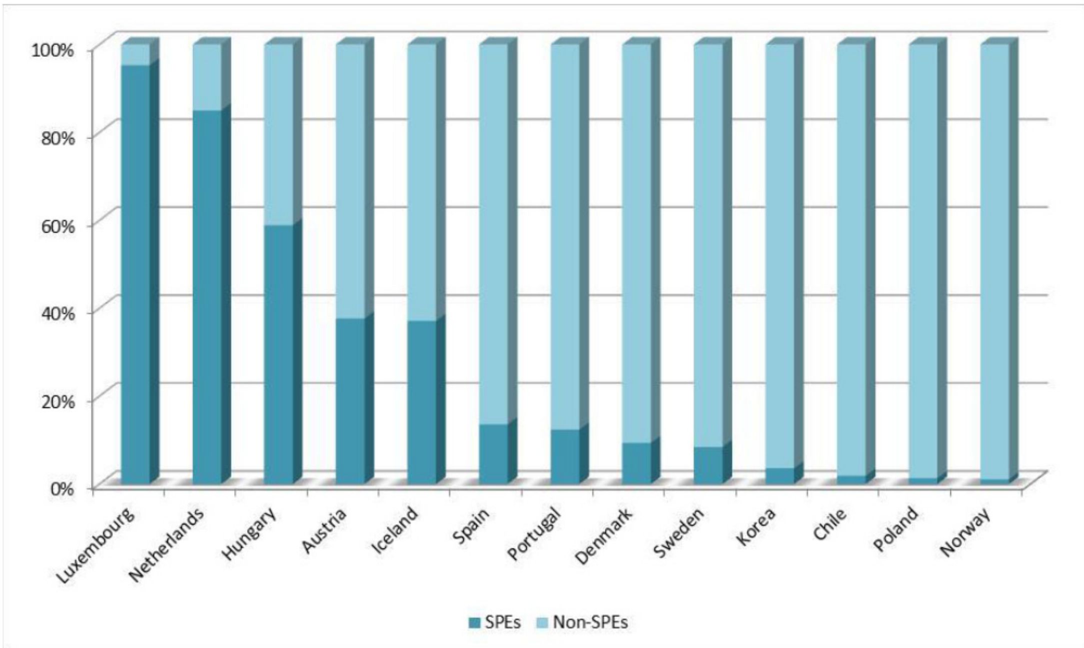
- draw attention to the confusing distortion of data including SPEs, that they are not appropriate for analysing country data, interpretation of economic trends, debt indicators, etc.
- support the initiatives for a common register of SPEs
- separate SPEs' data within the Financial Account Statistics in a specific sector breakdown.

Despite our efforts after the implementation of BPM6/BMD4/ESA2010 in some international organizations' database BOP data are available only including SPEs.

However, we are aware that only a few countries are seriously affected with these issues (See Figure 2). The major affected countries are Luxemburg, the Netherlands and Hungary, and partly Austria and Iceland (within the OECD). They are all engaged in the importance of publishing FDI data excluding SPEs.

Figure 2. shows the percentage of inward positions accounted for by resident SPEs in 2014. For Hungary, the percentage of outward position is even more significant reaching 80%.

Figure 2: Impact of resident SPEs on inward stocks for selected countries, 2014



Source: http://www.oecd.org/corporate/FDI-BMD4-brochure.pdf

2.3. In 2010's: New distortive effects - capital in transit and restructuring of asset portfolio transactions

Within a multinational enterprise, mixed groups can be formed from SPE and non-SPE affiliates in a country. These companies take part in intermediary activities and they also perform real economic operations in the host country, (thus they cannot be classified as SPEs). Their foreign parents pass through them large amounts from one foreign subsidiary to a third country within the group.) There are no internationally available guidelines for the treatment of capital transiting through operational affiliates of multinational enterprises. It is in the Research Agenda of the OECD Working Group of International Investment Statistics yet.

In Hungary, we face the difficulty: the outstanding magnitude of these transactions made it necessary to identify and separate them, we call these transactions **capital in transit**, passing through resident subsidiaries also engaged in real economic operation. The difficulty is, that beside the enterprise approach, individual transactions must be investigated. We have identified two types of activities that result huge transactions without any effect on the domestic economy: one is capital in transit, the other is restructuring of asset portfolios, when transactions are related to a **financial restructuring of some affiliates of the MNEs**. Besides capital in transit, restructuring of asset portfolios has similar effects on FDI flows (sharp increase and decrease in flows), these impacts on statistical data are like that of SPEs.

3. Supplemental breakdowns for better understand foreign direct investments

To improve the interpretability of FDI statistics and to reduce the impact of distortive elements of globalization, the MNB publishes supplemental breakdowns for FDI statistics.

3.1. Separating Capital in transit transactions

At the end of 2011 huge capital in transit transactions have appeared within the scope of non-SPE companies in which the inflow and outflow took place within the same quarter. As mentioned above, there is no general international methodological guidance on how to treat these special transactions. However, the magnitude of these transactions made it necessary to be identified and separated because these transactions significantly distorted our statistics already excluded SPEs. In the second half of 2011, the MNB launched a project exploring the phenomenon of capital in transit, with the purpose of identifying and separately presenting them. Capital in transit transactions are identified during the monthly compilation of balance of payments. These large transactions are usually completed within one month.

This micro-level approach, checking every company/transaction individually, is possible in Hungary because the relevant scope of companies includes 10- 20 enterprises and there are less than 10 such transactions quarterly. At the same time, the size of the transactions is sometimes one magnitude larger than the value of the regular transactions, which needs to be explained to users. Thus, since September 2012 capital in transit transactions has been presented in a separate table on the website of the MNB, retrospectively to 2008. Furthermore, to make the direct investment data more interpretable, FDI flows excluding capital in transit are published separately.

We refer to capital in transit, when a resident, non-SPE company belonging to a multinational company group is also active in passing through capital within the enterprise group, in addition to its normal activities (production, service). This activity increases the total value of both capital inflows and outflows in the statistics. Similarly, to the activities of SPEs, this flow of capital has no effect on the economy of the host country. As the activities of SPEs, capital in transit transactions are usually FDI, but they can take the form of portfolio investment or other investment.

Capital in transit is usually related to special function(s) of Hungarian affiliates of multinational enterprises. It is important that the foreign owner decides if the special activity is performed through the affiliate in Hungary or third countries. To recognise a capital in transit or asset portfolio restructuring transaction (details see below) it helps that the volume of inflow and outflow is typically same or similar within a specific (short) period.

In Hungary capital in transit transactions are identified based on company reporting during the compilation of the monthly or quarterly balance of payments data.

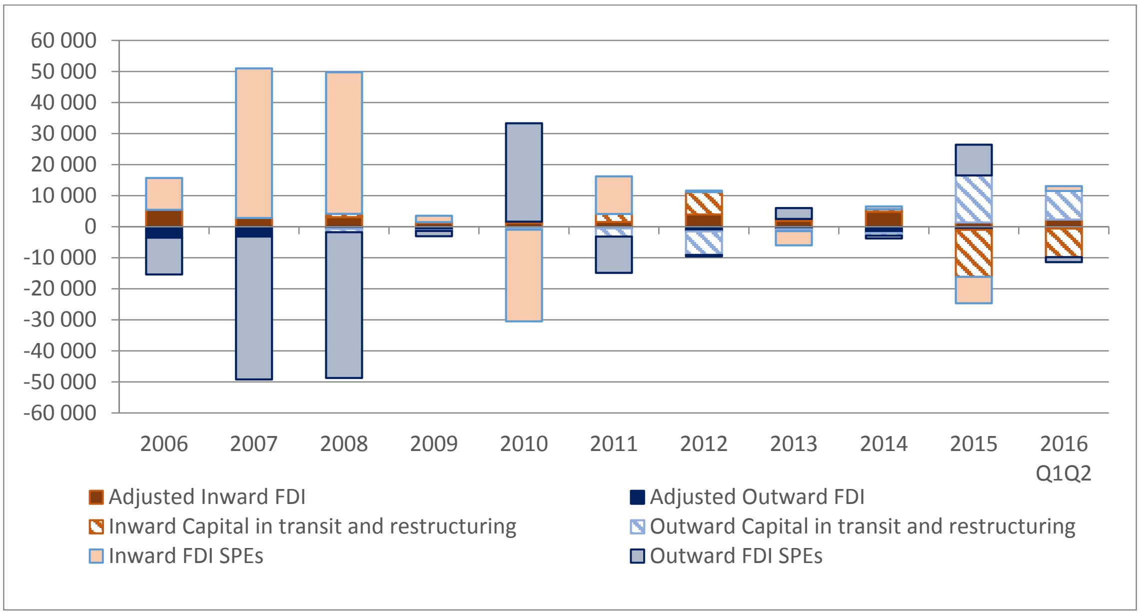
3.2. Separating restructuring of the asset portfolios transactions

If **a multinational corporation realigns its asset portfolio in a cross-border fashion**: liquidating one subsidiary, establishing a new subsidiary, contributing the assets of one subsidiary into another, etc., then extremely high capital withdrawal and equity investment transactions must be recorded in the balance of payments without any real capital withdrawal or equity investment taking place into the country. Therefore, together with capital in transit, we also classify these transactions as transactions to be separated. To easier interpret the direct investment transactions data, we publish the data on foreign direct investment transactions excluding capital in transit and asset portfolio restructuring available on the website of the MNB.

In Hungary, the effect of capital in transit and asset portfolio restructuring was especially significant in 2012, 2015 and 2016. However, the magnitude of these flows varies annually. For users interpreting FDI statistics data it is especially important to adjust the data for capital in transit and asset portfolio restructuring in our publications. The different magnitudes of FDI flows of SPEs, capital in transit and asset portfolio restructuring transactions and adjusted([93]) flows are shown in Figure 3.

Data including SPEs, excluding SPEs, adjusted FDI data are available in the MNB website. **Country and activity break down are also available** for the adjusted data.

Figure 3: FDI flows data (million EUR)



Source: MNB

To present the difficulties of FDI data interpretation, note that the total column size data represents the internationally available FDI flow data in the IMF and Eurostat database (including SPEs). Data excluding SPEs are available in the OECD and UNCTAD database. Adjusted FDI data are available only on the MNB website. They represent different possible conclusions about the significance of FDI activity in Hungary, and different ranking of countries.

3.3. Separating mergers and acquisitions from equity transactions

Mergers and acquisitions (M&A) provides **supplementary information on FDI equity flows**. The separation of M&A from other type of equity relates to the specific feature that M&A contains **purchase/sale of existing shares** by a direct investor/direct investment enterprise. While greenfield investments or extension of capital increase the total assets of the resident enterprises and therefore directly impact on production, etc, M&A usually doesn't have the same impact, there is a change only in the structure of foreign direct investors/direct investment enterprises. A merger means a combination of two companies to form a new company, while an acquisition is the purchase of one company by another where

([93]) Capital in transit and asset portfolio restructuring transactions are filtered together due to confidentiality constraints in the publications of MNB.

no new company is established. An M&A transaction usually does not result in any new investments, only the ownership structure of the already existing equity investments changes. Separate M&A transactions data are available on the MNB website: Data according to BPM6 methodology

3.4. Inward FDI position by country of the ultimate investor

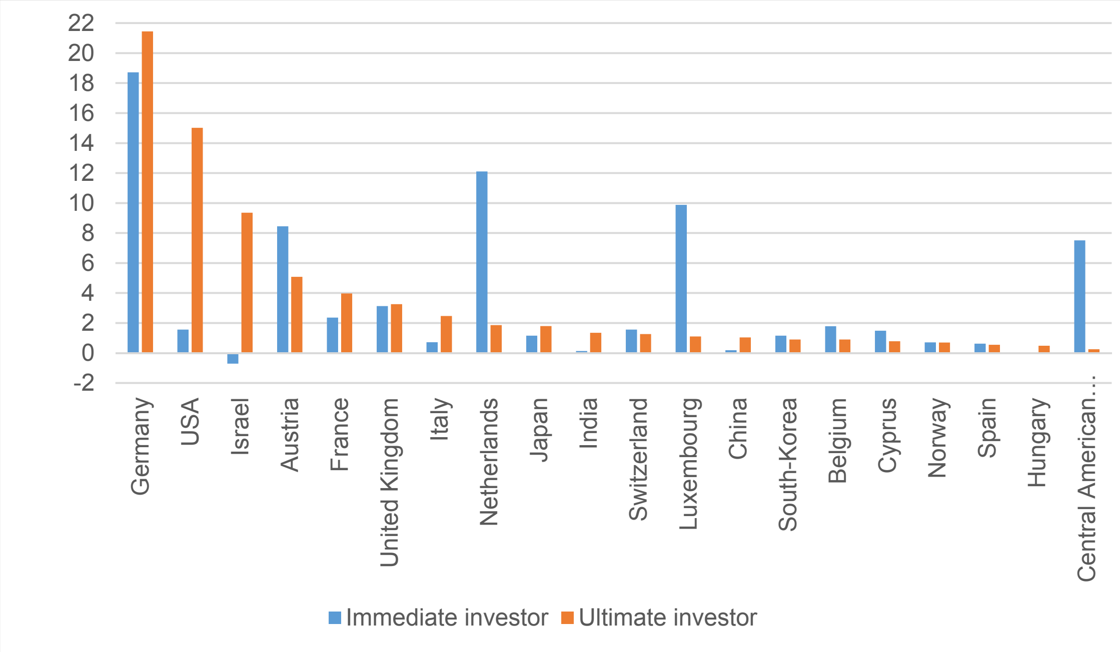
In line with international statistical recommendations (BPM6, BMD4) the MNB publishes stock, flow and income data of direct investments in a breakdown by immediate investor and investment. However, among users there is **an increasing need for Inward FDI position by country of the ultimate investor.** As from September 2016, this supplemental breakdown of FDI stock data in Hungary by the ultimate investing country is published, first for 2014 annual data. It will be published regularly with the annual data release in September, with a time lag of 21 months (T+21 months). The breakdown is compiled based on company reports. Already about 10 countries has published this supplementary statistical breakdown by the OECD website as well.

Ownership structures have become more complex and several investors invest indirectly through intermediate entities such as holding companies and regional management centres. Contrary to the country breakdown that takes into account the immediate partner, the country of the ultimate investor can be Hungary as well. This case is called **round tripping**. It means that in fact a Hungarian majority investor invests in Hungary through a foreign company (companies). Accordingly, by separating the round tripping within stock data, the breakdown by ultimate investor allows the identification of foreign investment sources that are actually not new for the country. In economic terms, **round tripping is a specific type of capital in transit**, which may be related to tax optimisation, just like in the case of other capital in transit.

Irrespective of the number of direct investors, the entire value of foreign direct investments of the resident enterprise is allocated t**o the only majority investor at the top of the ownership chain** of the multinational enterprise group. The country breakdown by the ultimate investor shows which countries exercises the ultimate control over the foreign direct investments in Hungary([94]).

([94]) There is no definition, neither any practice for breakdown by the ultimate host country yet, it is in the Research Agenda of the OECD WGIIS. Thus, mirror data for the ultimate investing country breakdown doesn't exist yet.

Figure 4: Inward FDI position by the country of the immediate and ultimate investor (billion EUR)



Source: MNB

Figure 4. shows that the order of the countries investing in Hungary by size of the investment becomes significantly rearranged in the breakdown by ultimate investor: the largest investors are Germany, the United States, Israel and Austria, while in the breakdown by immediate investor Germany is followed by the Netherlands and Luxemburg as the second and third largest investors, preceding Austria, and the weight of Central American countries is also significant as a result of capital-in-transit.

In case of Hungary **more than 40 per cent of the inward FDI stock has been invested via transit countries** including The Netherlands, Luxembourg and some Central American countries.

4. Conclusive remarks

Globalisation may have significant distortive effects thus needs be internationally agreed supplements on the national statistical data publication - thus interpretation - of national FDI-statistics.

Weight of FDI compared to GDP significantly varies among countries and in time**. In those countries where the role and effects of FDI is high, and the distortive effects of globalisations are significant, more detailed FDI statistics are needed inward and outward FDI alike.**

To improve the interpretability of these statistics and to reduce the impact of distortive elements of globalization, we have developed a special practice in Hungary. Besides separating Special Purpose Enterprises as required in BMD4/ BPM6, we found informative/necessary for our users to publish detailed data excluding capital in transit and asset portfolio restructuring transactions Furthermore, MNB publishes supplemental breakdowns as mergers and acquisitions in equity flows and inward FDI stock by the country of ultimate investing country.

Annex I.

The weight of FDI compared to GDP varies by countries. The country breakdown of inward and outward FDI to GDP is shown in Figure A for OECD countries. The data of Hungary and OECD-total are highlighted to show the significance of FDI activity in Hungary. (These ranks are based on FDI data are excluding SPEs. Note that for Hungary the **inward FDI stock including SPEs compared to the GDP is 180 %, in the rank Hungary would be just after Ireland**, while the **outward FDI stock including SPEs compared to the GDP is 140 %, in the rank Hungary would be just after the Netherlands**.)

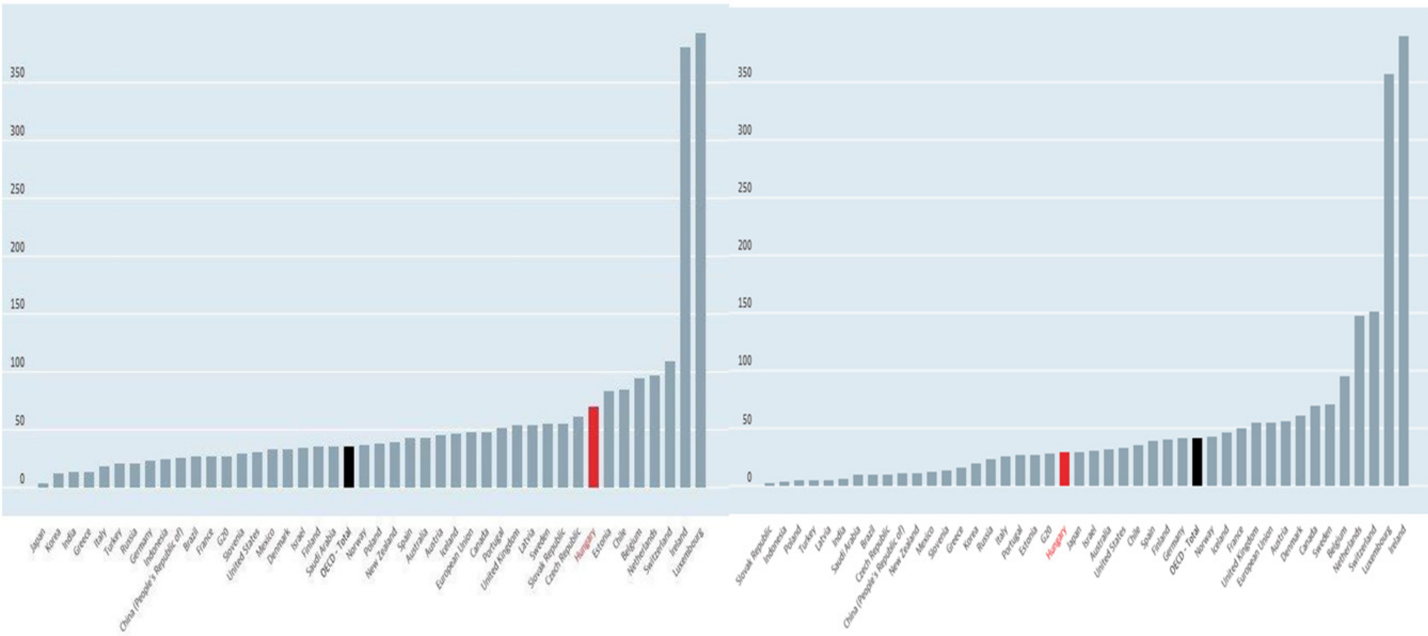


Figure A Inward and Outward FDI stocks (excluding SPEs) as % of GDP at the end 2015

Source: https://data.oecd.org/fdi/fdi-stocks.htm

11 Trade performance of the EU economies: Inter-country input-output tables as a necessary tool

LARS NILSSON[95]

Table of Content

1. Introduction

The EU's trade and investment policy – or the common commercial policy – is an exclusive power of the EU, which is carried out by the European Commission (the Commission). It relates to trade in goods as well as to trade in services and to areas such as foreign direct investment (FDI), trade related aspects of

([95])European Commission, DG TRADE; lars.nilsson@ec.europa.eu

intellectual property rights, public procurement and technical barriers to trade, etc. The EU's trade and investment policy is often carried out through negotiations, which are conducted by the Commission on behalf of all EU countries.
Trade policy affects EU citizens and the Commission aims to listen to all input provided so that trade policy most accurately reflects EU society's values and interests as a whole. This takes place through a variety of channels before, during and after the negotiating process, when the Commission evaluates the impact of all significant new trade policy proposals.

The basic motivation for opening up to trade is that it leads to increased specialisation and improved resource allocation, allowing firms to fully exploit economies of scale and to lower production costs. At the same time the increased presence of foreign competitors puts a downward pressure on prices and offers greater product variety for consumers. In addition over time, trade openness allows ideas and technologies to spread and spurs innovation and productivity growth. All these reinforcing channels amount to profound changes to how the economy works. However the many inter-linkages at play make these effects difficult to quantify.

Section 2 of this paper briefly reviews the type of analyses that are carried out to assess the impact of EU trade policy initiatives. The main methodology used by the Commission in assessing the impact of major trade policy initiatives – so called computable general equilibrium (CGE) models – is looked at in Section 3 together with the associated data requirements, in particular input-output (IO) tables. Section 4 briefly highlights complementary methodologies, while Section 5 and discusses analytical and modelling limitations. Section 6 concludes.

2. Assessment cycle in case of free trade agreements (FTAs)

EU FTAs are thoroughly scrutinized. Their impact is assessed before negotiations begin, during the negotiations and at several points in time once negotiations have been concluded.

Firstly, impact assessments (IAs) examine whether major new free trade negotiations should be launched at all. IAs include consultation of interested parties and they identify and describe the problem to be tackled, establish objectives, formulate policy options and assess the impacts of these options. IAs assess the environmental, social and human rights and economic impacts of various policy options and thus prepares evidence for the College of Commissioners of the advantages and disadvantages of possible alternatives. All IAs are treated by the Regulatory Scrutiny Board (RSB, which provides a central quality control and support function for Commission impact assessments and evaluation work. The Board issues opinions based on the requirements of the Better Regulation Guidelines.([96])

Secondly, trade sustainability impact assessments (SIAs) are carried out during the negotiations to help the Commission as a negotiator to shape the negotiating process in a direction coherent with overall EU policy. They are independent studies conducted by external consultants, on the basis of which the Commission sets out its' own views on the identified impacts and on the policy measures proposed to address them. Trade SIAs consist of two complementary elements: an analysis of the potential economic,([97]) environmental and social impacts that the trade agreement might have, both in the EU and in the partner countries; and a transparent and wide consultation process which inter alia aims to engage stakeholders both in the EU and in partner countries and allow the studies to take into account knowledge and concerns of relevant interest groups. They are distinct from Commission IAs.

The economic assessment of the negotiated outcome (EANO) focusses on the economic value of trade barrier reductions following the final, precise outcome of the negotiations. The analysis follows the actual text of the agreement (including the tariff dismantling schedules) which makes it possible also to assess the reduction in non-tariff barriers (NTBs) in the form of trade cost reduction of separate provisions of the agreement. Hence, a thorough analysis of all trade cost reducing elements is made and quantified and there is no need to assume the extent to which e.g. NTBs will be liberalised as is done in the ex-ante IAs. One should note, though, that the EANO does not focus on the potential social and environmental impact of the concluded agreement.

(96) European Commission (2015), Better Regulation Guidelines, *COM (2015) 215, final* and *SWD (2015) 110,* final.
(97) In case there is a recent IA containing economic analysis, an SIA does not undertake a new one.

Finally, ex-post evaluations, or interim ex-post evaluations aim to provide a reliable and objective assessment of how efficient and effective an initiative has been and how performance can be improved in the future. Ex-post evaluations take place during and after implementation of an FTA. Civil society organisations participate in the monitoring of trade agreements that have been concluded between the EU and partner countries and provide input specifically on social and environmental issues.

3. Computable general equilibrium (CGE) modelling and data requirements

3.1. CGE models

Most studies the Commission undertakes to assess the general impact of FTAs at overall EU level are carried out using CGE models, which are state of the art tools for overall assessments of trade agreements at region, country and broad sector level.

These models are computer-based simulations which calculate the future state of the global economy (including any country or region specifically analysed) as a consequence of a specified set of (trade) policy changes, based on the most advanced and well established economic theories. Over the past decades(s) CGE models have undergone important changes to keep up with the economic theory on which they are grounded.

The main advantage of CGE models is that they analyse the effects of trade policy taking into account the main links between the domestic and international production of goods and services. The also include consumption and investment decisions of firms across sectors as well as of consumers and the government. Output comes in the form of results on a wide range of sector socioeconomic indicators (tariff revenues, imports, exports, production prices, wages, CO2 emissions, land use, etc.). The models also account for the fact that different sectors compete for capital, labour and land.

CGE models help answer 'what if…' questions by simulating the price, income and substitution effects of different policy changes and comparing them against predictions about what would happen without a policy change, a so called baseline. The baseline is key since it is the counterfactual against which the economic outcome of the initiative is assessed. Hence, the models allow economists to simulate at the same time how all sectors and actors adjust to the changes to costs, prices and/or incentives that a trade policy change would cause. This allows for an assessment of all the direct and indirect effects of changes to trade policy.

For example, assume that policymakers decide to raise import barriers on steel to relieve the competition pressure on the domestic industry. A CGE model would then also show how detrimental protecting this one sector from competition would be to downstream industries that use steel as inputs (due to higher steel prices). Furthermore, the inter-linkages in the CGE model would also pick up the impact on upstream industries, since the steel producers and downstream industries would make less use of business services like logistics. CGE models are therefore important for evaluating economy-wide effects of specific policy decisions.

One of the most commonly used CGE models has been founded by the Global Trade Analysis Project (GTAP). It comes in a dynamic version and in a static version. The dynamic GTAP model allows international capital mobility and capital accumulation, while it preserves all the features of the standard GTAP model, such as constant returns to production technology, perfectly competitive markets, and product differentiation by countries of origin – the so-called Armington assumption. At the same time, it also incorporates international capital mobility and ownership.

In this way it captures important FTA effects on investment and wealth that are missed by a static model.([98]) In the dynamic GTAP model, each of the regions is endowed with fixed physical capital stock owned by domestic firms. The dynamics are driven by net investment, which is sourced from regional households' savings. The savings in one region are invested directly in domestic firms and indirectly in

([98]) Ianchovichina, E. I. and R. A. McDougall (2001), Structure of Dynamic GTAP, *GTAP Technical Paper 17,* Center for Global Trade Analysis.

foreign firms, which are in turn reinvested in all regions. The dynamics arising from positive savings in one region is related to the dynamics from the net investment in other regions. Overall, at the global level, the savings across regions are fully invested in home and overseas markets.

Attempts have also been made to incorporate FDI in CGE models explicitly, using underlying data on FDI stocks and flows. For example, as a result of a project initiated by the European Commission (DG Trade), Gouel et al (2012) provided GTAP with a global bilateral multi-region multi-sector FDI stocks and flows database, fully consistent, balanced and suitable for use in CGE work.([99]) However, the GTAP center does not provide any publicly available CGE model including an FDI module and the FDI database produced by Gouel et al (2012) is gradually becoming outdated.([100])

CGE models can also be used for ex-post assessments of trade policy changes.([101]) In an on-going evaluation of the EU-Korea FTA which came provisionally into force in July 2011, Civic (2017) first ran an extensive econometric exercise using gravity type of regressions to isolate all factors that may influence trade between the EU and Korea, except for the FTA.([102]) Secondly, tariff cuts as observed in the data and reductions in other trade costs as implied by the results of the econometric exercise were fed into a CGE model and "undone" to construct a counterfactual in which the EU Korea FTA is assumed never to have materialised. This counterfactual is then compared to the observed situation in 2015. The advantage of this approach is that no direct measures of observed reductions in NTBs are needed and the simulation exercise is tied to the gravity estimation.

Attempts have also been made to account for global value chains (GVCs) in CGE models. The OECD launched its own CGE model known as METRO (ModElling TRade at the OECD) in 2015. The underlying database for the model is based on the GTAP sectoral structure but covers less countries/regions. The main novelty of the model is that by making use of the OECD-WTO Trade in Value Added (TiVA) database, it allows for trade policy analyses of trade flows by use category, i.e., by intermediate, household, government and capital consumption. This implies improved analyses of GVCs.

Similarly, Walmsely and Minor (2016) base themselves on the GTAP model but then go on to classify the importing agent and exporting country of each imported products, along with differential tariff rates by end-user and origin.[103] As above, this is of analytical importance, since, depending on product category and end-use, both the bundle of goods in the product category and the tariff rates they face may vary significantly.

3.2. Data requirements

Data for CGE models are usually drawn from the GTAP database, which is a global database characterizing economic linkages between sectors, countries and regions, combining detailed bilateral trade, transport and protection data as well as data on energy, emissions and power technologies. It is built on the most reliable international data sources (including Eurostat data for EU countries). It undergoes constant scrutiny by the different stakeholders and users such as the Commission, the World Bank, OECD, IMF, WTO, United Nations, FAO, etc.

The underlying input-output tables are heterogeneous in sources, base years, and sector details, thus for achieving consistency, substantial efforts are made to make the disparate sources comparable. The objective of the database is to facilitate the operation of economic simulation models ensuring users a consistent set of economic facts, not to provide a repository of IO tables.

The latest release represents 140 countries/regions and 57 sectors and features three reference years. With its wide country and sector coverage, the GTAP database, which is fully documented, is the only global database available for this type of analyses which can guarantee long-term continuity and regular

([99]) Gouel, C., Guimbard, H. and D. Laborde (2012), A Foreign Direct Investment database for global CGE models, CEPII, *Working Paper No. 2012-08*.
([100]) However, such models do exist. See for example, Lakatos, C. and T. Fukui (2014), The Liberalization of Retail Services in India, *World Development, Vol. 59*, pp. 327–340 and Tsigas, M. and Yuan, W. J., (2017), Addressing Excess Capacity--The Effect of China's FDI in the Iron and Steel Industry in Five Central Asian States: a GTAP-FDI Model Perspective, GTAP Resource 5376.
([101]) See e.g. ITAQA (2012), Evaluation of the Economic Impact of the Trade Pillar of The EU-Chile Association Agreement, *Final Report*, Contract SI2.575484, For the European Commission, Directorate General for Trade.
([102]) Civic (2017), Evaluation of the Implementation of the Free Trade Agreement between the EU and its Member States and the Republic of Korea, *forthcoming*.
([103]) Walmsley T. L., and P. Minor, (2016), ImpactECON Global Supply Chain Model: Documentation of Model Changes, *ImpactECON Working Paper No. 06*, ImpactECON: Boulder, CO, USA.

updates.

Following the European Court of Auditors audit of EU preferential trade arrangements,([104]) the Court recommended updating the EU IO tables available in the GTAP database. The Commission has subsequently launched a project together with DG JRC with the objective of complying with Court's recommendation so as to base its trade modelling analysis on the most reliable and recent such data. DG JRC has produced new tables for all 28 Member States and based on the latest available year (i.e. 2010) under the new European System of Accounts (ESA10) methodology and submitted those tables to GTAP. The project has greatly benefited from the support of Eurostat in respect to the quality of the European Statistics used, their good practices guidelines and through consultations during the project.

Regulation (EU) No 549/2013 of the European Parliament and of the Council of 21 May 2013 on the European System of National and Regional Accounts in the European Union determines the methodology to be used for the compilation of national accounts data to be submitted to Eurostat. The regulation further defines that, while Supply and Use Tables should be submitted annually, IO tables should be submitted every five years.

From an analytical point of view, it would be welcome with more frequent updates of EU (and other countries') IO tables. On the other hand, one should recall that most economic data in the underlying GTAP database used to simulate the impact of EU FTAs, such as trade flows, protection patterns, GDP, final consumption, etc., is updated to the latest year available before the analysis begins. Partly as result, the simulated impact of EU FTAs appears to be only marginally affected. For example, a report by the Commission's Joint Research Centre shows that updating EU technical coefficients in EU supply and use tables to the most recent year available only marginally changed the simulated impact of the EU-Korea FTA.([105])

4. Complementary methodologies

The advantage of the CGE methodology comes at a cost, notably the high level of aggregation (57 sectors) required to be able to use comparable and consistent data across countries to run these models. This contrasts with the fact that trade liberalisation takes place at the much finer tariff line level.

The Harmonised System (HS) comprises about 5000 products at 6-digit level. In the EU, the Combined Nomenclature (CN) contains two subheadings of the HS and thus breaks it down to the 8-digit level, while, additionally, TARIC([106]) contains two subheadings of the CN which brings us to the 10-digit level at which EU customs duties normally are defined. If products at this fine level of aggregation are considered sensitive, the assessment of trade policy changes would have to rely on complementary analyses based on other methodologies.

4.1. Partial equilibrium (PE) models

Partial equilibrium models can handle impacts at a detailed level of the product classification, for example on passenger cars or a specific set of agricultural products. In addition to this ability, the main advantage of PE models is its minimal data requirement. In fact, the data required is in principle only for trade flows and duties. Analysing the impact of a trade policy change at a fine level of aggregation also reduces aggregation biases which may be present when using CGE models. However, due to their simplicity PE models do not cover the input/output (or upstream/downstream) linkages across and between sectors and countries. The PE approach also misses constraints that apply to factors of production (e.g., labour, capital, land) and their movement across sectors.

In principle, PE models can also be used to assess the impact on trade in goods according to their end-use category. The United Nations (UN) trade statistics has recently carried out a 5th revision of the

([104]) European Court of Auditors (2014), Are Preferential Trade Arrangements appropriately managed? *Special Report No 2/2014, 2014/C 154/05*.
([105]) European Commission (2014), Replies of the Commission to the special report of the European Court of Auditors "Are Preferential Trade Arrangements Appropriately Managed?", *194 final*.
([106]) The integrated Tariff of the European Union, http://ec.europa.eu/taxation_customs/business/calculation-customs-duties/what-is-common-customs-tariff/taric_en.

Classification by Broad Economic Categories (BEC),([107]) which covers seven broad such categories and 19 unique ones, eight out of which are classified as intermediates. This data exist with conversion tables into the Standard International Trade Classification (SITC), the Harmonised System (HS) and the Central Product Classification (CPC). Nevertheless, this type of analysis lends itself best in a general equilibrium setting since the impact on intermediates is interesting not least because of the subsequent impact they have on production of the final good.

4.2. Econometrics

Econometrics applies statistical methods to economic theory and economic data and thus providing empirical content to economic relations. The development of bilateral trade between two trading partners is in itself not a good indicator of the success of an FTA. Many other factors affect the evolution of trade. Therefore, econometrics and counterfactual analysis are required to isolate the impact of the FTA on the volume of bilateral trade.

In international trade, one of the most successful models used is the so called gravity model. The gravity model explains trade between two countries by their economic size and the distance between them. Additional (binary) variable are added to the model to account for various trade facilitating and/or trade hindering factors (common border and/or language, being landlocked, etc.). This equation has proven to be an excellent estimator of bilateral trade flows and is therefore used to control for non-FTA factors that matter for trade. It is usually applied ex-post and its result can be used for further ex-post CGE modelling, c.f. Section 3.

5. Analytical and modelling limitations

5.1. The impact of non-tariff barriers (NTBs)

Irrespective of the type of model used to assess the impact of trade (and other policies); model output will never be of higher quality than model input. As far as trade policy analysis is concerned, one area of limitation concerns NTBs.

The main challenge regards quantifying the costs NTBs impose, ideally in ad-valorem equivalents. For example, if there is a restriction on imports of eggs in the form of additional sanitary controls, how much, in percentage terms, does it add to the price of the foreign good? In services, the cost of restrictions is even harder to quantify. How economically significant is a cap on the number of foreign engineers allowed to deliver a particular service? Can it be quantified in ad-valorem equivalents?

The costs of such limits are difficult to analyse and they can easily spill into goods trade if for example foreign engineering services are needed to install imported technically advanced goods such as solar panels or wind turbines. These are increasingly important questions to address from a policy standpoint as NTBs are fast becoming the main friction to trade as tariffs have come down worldwide following multilateral (and unilateral) trade liberalisation in many countries.

Another challenge is to determine the extent to what is to be negotiated, or has been negotiated, actually reduces NTBs. In the EU-Korea FTA, NTBs in the motor vehicles sector were perceived by the EU industry as the most significant obstacles to exporting to Korea. As a result, the FTA contains provisions under which Korea accepts the equivalence of international standards or EU standards for its major technical regulations. For example, now, Korea accepts EU on-board diagnostic devices conforming to the so called Euro 6 standards as compliant with Korean standards. From an analytical point of view, it is however challenging to assess what the value of such a provision is. In many cases as specific as this one, industry input is essential.

[107] https://unstats.un.org/unsd/trade/classifications/bec.asp. One should note that a breakdown of end-use categories of services does not exist.

5.2. The choice of model closure(s)

Another area affecting the interpretation of CGE results concerns the so called model closure(s). CGE models contain more variables than equations; hence some variables have to be determined exogenously. The choice of variables to be exogenous is called model closure and the type of closures chosen may influence the results of CGE simulations significantly. The most common closure rules relate to the labour and capital markets, the current account and the government balance.

The default closure in the GTAP model fixes the capital and labour supply and requires the model to restore equilibrium by adjusting the rate of return to capital and the wage rate. This is sometimes described as reflecting a medium-term time horizon in which labour supply is relatively “sticky”. Under an alternative closure rule, the return to capital and/or wages can be fixed. The supply of capital and/or labour then adjusts to restore equilibrium. This is sometimes described as reflecting longer-run “steady-state” growth conditions. Each of these closure rules are extreme; capital and labour supply is neither perfectly elastic nor perfectly inelastic. The reality is likely to be somewhere in between.

The 'fixed employment closure' is commonly used for analyses of EU FTAs since there is no established theoretical framework linking the functioning of labour markets to CGE models. In addition, in an EU context it would be highly complex to model the reaction of 28 labour markets to a trade shock, when the reservation wage differs across EU Member States and the incentives for people already in employment to change jobs are different across sectors and countries as well. However, model output using this closure provides information on reallocation of labour between sectors, thus indicating in which sectors employment is likely to increase or decrease as a result of the new agreement. As such the closure also affects the simulation results output since sectors react to positive demand shocks by drawing labour from sectors which are less, negatively or not affected, instead of adding labour from the pool of unemployed in the economy.

The current account closure relates to whether or not the current account balance should be fixed. A fixed current account implies that when a trade policy shock results in unbalanced changes in imports and exports, the original trade balance is restored by (implicit) exchange rate adjustments. Alternatively, the current account can be allowed to adjust to the trade shock. The change in the current account then must be offset by equivalent changes in capital flows. In reality, unbalanced trade impacts are likely to have both effects: inducing subsequent exchange rate adjustments and also offsetting capital flows.

No one closure is more correct than another one and the specific closure adopted should be suited to the circumstances of the economies affected by the model. For example, the 'fixed wage closure', as opposed to the 'fixed employment closure', could be used to model trade impacts on developing countries that have a large reserve pool of labour in subsistence rural agriculture and for which a perfectly elastic supply of unskilled labour would be an appropriate assumption. In other words, an analysis of a policy implemented in a period of high capacity utilization should adopt a different closure than an analysis of a policy implemented in a period of high excess capacity.

5.3. How to capture productivity changes

Trade affects productivity through a variety of different links. Traditionally, it does so through the reallocation of resources according to each country's comparative advantage. More recently, accompanied by an increase in access to firm-level data to verify theoretical advances (Melitz, 2003),[108] research has shown that there are significant productivity differences among firms within the same industry and that trade supports productivity growth also by forcing the least productive firms to exit and reallocating their resources to more productive firms.

The reason being that firms in liberalised economies face greater competition than in closed economies; at the same time improved market access abroad provides the most productive firms opportunities to increase exports while drawing resources from the less productive firms. The latter eventually exit thereby raising the overall level of productivity in the economy.

CGE models capture the first type of productivity gains through the reallocation of production and factors of production across sectors of the economy analysed. The increased firm-level productivity effect as a result

(108) Melitz, M.J. (2003), The impact of trade on intra-industry reallocations and aggregate industry productivity, Econometrica, vol. 71(6), pp. 1695–725.

of trade liberalisation is harder to capture and quantification attempts is still at an early stage. Balistreri et al (2011)([109]) find welfare gains from trade liberalization that are four times larger than in a standard trade policy simulation while Corcos et al (2012),([110]) inter alia looking into efficiency gains via firms selection, note that gains from trade are much larger in the presence of selection effects, but with considerable variability across countries and sectors. Looking into a range of trade models, Breinlich and Cuñat (2016)([111]) conclude that models substantially underestimate productivity growth as observed in the data.

5.4. Outdated elasticities

The extent to which the macroeconomic impact of changes in trade policy changes is estimated depends on how demand and supply react to prices changes. Higher so called elasticities lead to larger substitution effects between imports and domestic products and increased welfare gains. Traditionally, the CGE modellers have made use of elasticities which have been based on econometric time series estimations of price variations between domestic goods and imports. Hertel et al (2004) identify problems related to this approach (e.g. insufficient observed variation in relative prices) which they address to produce a new set of elasticities of substitution (EoSs) between imported goods.([112])

This new set of EoSs is currently incorporated into the most recent version of the GTAP database (v. 9). However, the elasticities obtained by Hertel et al (2004) are based on a dataset used by Hummels (1999), who in turn used data from 1992 on the USA, New Zealand, Argentina, Brazil, Chile, and Paraguay.([113]) That is, the Armington elasticities used for the lion's share of CGE analyses using GTAP data date back to the early 1990s and are based on empirical work on only six countries out of which none is European. Furthermore, the EoSs for a given sector are the same across all regions.

In addition, the EoSs between imported commodities follows the "rule of two", i.e., it equals the EoSs between domestic and imported goods multiplied by two. This approach which was first proposed by Jomini et al (1991) and has been retained in the GTAP database, but does not seem to have a particularly strong or recent empirical foundation.([114])

6. Conclusions and way forward

Arguably, never before has so much attention been paid to the EU's trade and investment policy. Modern trade agreements include provisions not only on tariffs but on a range of issues, some of which fall in the category of NTBs. As tariffs have come down worldwide over the past decades, the relative importance of NTBs has increased. Similarly, the role played by services trade both directly, or indirectly as input to goods production, has perhaps never been greater.

Macro-level assessments of the economic impact of FTAs are usually provided through the use of CGE models. IO tables are indispensable for assessing the economy wide impact of trade agreements using CGE models and the Commission has recently undertaken and finalised work to update the IO tables for EU28 with the most recent data available. Clearly, this type of analyses would benefit from regularly updated and recent IO tables for all countries. Ex-post analyses have traditionally relied upon econometrics, but recently, CGE models are also increasingly being employed. Should there be a particular interest in the impact at detailed product level, PE models may be more suitable to use.

Nevertheless, irrespective of methodology used, assessing the impact of an FTA either ex-ante or ex-post

(109) Balistreri, E., Hillberry, R. and T. Rutherford (2011), Structural estimation and solution of international trade models with heterogeneous firms, *Journal of International Economics, vol. 83(2)*, pp. 95–108.
(110) Corcos, G., Del Gatto, M., Mion, G. and G. I.P. Ottaviano (2012), Productivity and Firm Selection: Quantifying the 'New' Gains from Trade, *The Economic Journal, 122: 754–798*, doi:10.1111/j.1468-0297.2011.02487.x.
(111) Breinlich, H. and A. Cuñat (2016), Tariffs, Trade and Productivity: A Quantitative Evaluation of Heterogeneous Firm Models, *Economic Journal, 126*, pp. 1660–1702. doi:10.1111/ecoj.12218.
(112) Hertel, T.H., Hummels, D. Ivanic, M. and R. Keeney (2004), How Confident Can We Be in CGE-Based Assessments of Free Trade Agreements? *GTAP Working Paper No. 26*, Center for Global Trade Analysis, West Lafayette, Indiana.
(113) Hummels, D. (1999), Toward a Geography of Trade Costs, *GTAP Working Paper No. 17*, Purdue University, West Lafayette, IN: Global Trade Analysis Project (GTAP). Retrieved from https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=1162.
(114) Jomini, P., Zeitsch, J.F., McDougall, R., Welsh, A., Brown, S., Hambley, J., and J. Kelly (1991), *SALTER: A General Equilibrium Model of the World Economy, Vol. 1. Model Structure, Data Base, and Parameters*, Canberra, Australia: Industry Commission..

is associated with difficulties. Analysts have to assess what the overall value or worth is of a particular FTA, including the reduction of specific NTBs, which is challenging from both a data and methodology point of view. There are also specific choices to be made as far as the implementation of the trade policy scenarios are concerned such as e.g. how the labour market and the current account balance are treated in the model simulations. In addition, the incorporation of productivity effects in trade models based on heterogeneous firm theory has by and large only started.

The Commission has actively participated in the development of the underlying workhorse GTAP database in the past to ensure the relevance of this global modelling project for the assessment of EU policies and it is constantly working to provide new and more detailed data to be able to better assess the outcome of trade policy initiatives. Recently, efforts have been launched to split the services sectors in the workhorse GTAP database and to provide data on the modes of supply of services trade.([115]) Work is further on-going to gather data on procurement flows and procurement barriers in order to assess the impact of procurement liberalisation. This work will feed into a GTAP CGE module which the Commission has sponsored.

The Commission is currently considering the need to carry out work to update the Armington elasticities in the GTAP database which are crucial for the results of FTA simulations. It is also looking into the possibility of achieving additional sectoral detail in the analyses by potentially splitting some economically and/or politically important GTAP sectors.

Regardless, of the methodology chosen, the level of detail of the data and how recent the data is, one should always remember that the modelling results should be seen as a rough indication of the magnitude of the effects, which sector that are likely to be affected and whether the impact is expected to be positive or negative. Modelling results should thus not be taken at face value, but should be interpreted with a certain degree of caution while always recalling that the impact should be seen against the counterfactual, i.e., what the situation would have been in absence of the policy initiative.

(115) Mode 1 - Cross border trade, Mode 2 - Consumption abroad, Mode 3 - Commercial presence and Mode 4 - Presence of natural persons.

12 Trade asymmetries and Consistency between National Accounts and Balance of Payments

ROBERT OBRZUT[116]

Table of content

([116]) Eurostat Unit C5 – Integrated Global Accounts and Balance of Payments

The author thanks Mr. Ferdinando Biscosi (Eurostat) for the preparation of the data from International Trade in Services Statistics and for his valuable comments to the chapter on bilateral trade asymmetries.

1. Introduction

In 2014 the process of convergence in the methodological standards for the compiling of European national accounts and balance of payments statistics (BOP) was finally concluded. Hence, in applying both methodologies respectively, the European System of Accounts 2010 (ESA2010) and the Balance of Payments and International Investment Position Manual in its 6th edition (BPM6) ensure a high degree of comparability and consistency between BOP and the external account of national accounts (rest of the world account)(117). The essential question in this context remains how this methodological consistency is reflected in the statistical data.

This paper presents the results of a consistency analysis between European balance of payments statistics and the rest of the world account of national accounts, and is based on available statistical data, Eurostat surveys, quality reports, ad-hoc data confrontations and discussions with compilers in the Eurostat Balance of Payments Working Group (BOPWG). Possible reasons for inconsistencies, where applicable, are discussed, although the debate about the causes is still ongoing. This paper is further complemented by an analysis in bilateral trade flows in services between the EU Member States, trying to establish a second dimension to the above issue – the geographical asymmetries in BOP/international trade in services statistics (ITSS). Perfectly consistent Balance of Payments and National Accounts at national level may not be very credible in the presence of huge bilateral external asymmetries with partner countries. These bilateral asymmetries, where occurring, deserve therefore additional attention to be analysed and dealt with in parallel to consistency considerations, thus commending reconciliation of both statistics to a playing field of several quality dimensions.

([117]) BPM6 Appendix 7, ESA2010 Chapter 18

2. Measuring BOP-ROW consistency in the EU-28

2.1. Time frame and methodology of Eurostat's regular data comparisons

Regular data comparisons of quarterly statistics in BOP and the rest of the world (ROW) sector account are conducted by Eurostat since 2015 after the introduction of the BPM6 standard in European BOP statistics. Since then we are able to assess the evolution of consistency over time between the two statistics with a particular interest in the nonfinancial accounts. Although available time series are reported by some countries even back to 1999, reliable data comparisons across all EU-28 Member States appear currently feasible and meaningful from 2010 to 2015. Data are compared from quarterly statistics([118]), thus effectively reflecting back data revisions during the compilation year. Discrepancies are measured in absolute differences on gross transactions in the underlying nonfinancial accounts, as patterns could be different for export or import transactions. This appears instrumental to avoid offsetting effects, which could blur the picture. For this purpose the respective transactions in BOP are compared with those of the Sector Accounts (QSA), and annualised for more convenient reading (Table 1).

Table 1: Reconciling the nonfinancial accounts of BOP and the ROW sector

BOP component	ROW item	Description
Goods	P61	Exports of goods
	P71	Imports of goods
Services	P62	Exports of services
	P72	Imports of services
Primary income	D1	Compensation of employees
	D2	Taxes on production and imports
	D3	Subsidies
	D4	Property income
Secondary income	D5	Current taxes on income & wealth
	D6	Social contributions and benefits
	D7	Other current transfers
	D8	Adjustment for the change in pension entitlements
Capital account	D9	Capital transfers
	NP	Acquisition less disposal of nonfinancial nonproduced assets

BOP and ROW items according to BPM6 and ESA2010

The prevailing data comparison refers to the latest data releases of January 2017. It compares the releases of QSA which are published about 3 weeks after QBOP in European statistics. Eurostat has conducted this comparison systematically since October 2015([119]). A bias for revision and vintage effects however cannot be completely excluded in such analyses.

2.2. Recent results on BOP-ROW consistency in a nutshell

Against the methodological consistency of the standards current measures confirm a continued and persistent overall exposure to discrepancies in some particular components of the European nonfinancial accounts. In Table 2 we calculated total absolute discrepancies of the EU-28 as a sum of absolute discrepancies occurring in all 28 Member States. The extent of discrepancies measured on average around EUR 206 billion over the observed period (1.5% of average EU-28 GDP 2010-2015). In 2015 it culminated to EUR 272 billion (1.9% of GDP) in the EU.

([118]) Quarterly BOP (QBOP) and the ROW sector of the Quarterly Sector Accounts (QSA)
([119]) The latest report can be download from the Eurostat website: http://ec.europa.eu/eurostat/statistics-explained/index.php/Consistency_between_national_accounts_and_balance_of_payments_statistics

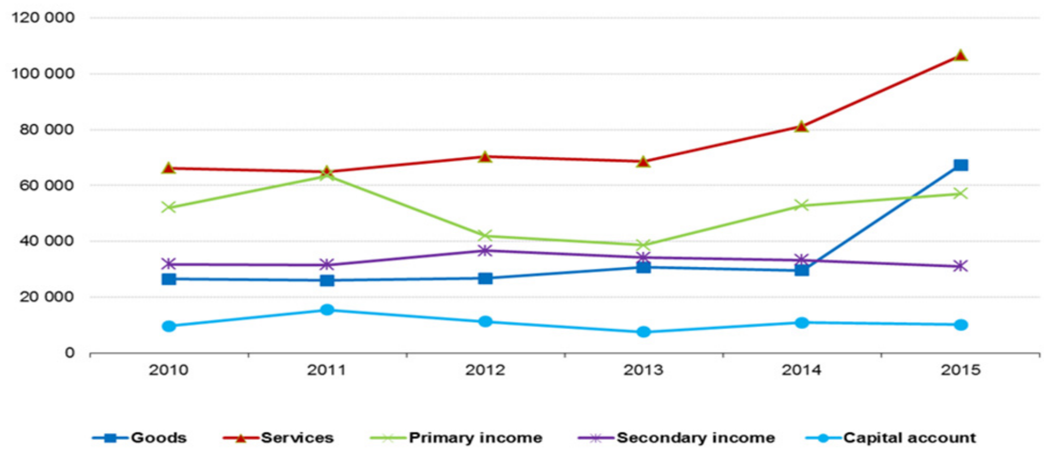
Table 2: Absolute BOP-ROW discrepancies in the nonfinancial accounts, sum of EU-28 Member States, by BOP item, 2010-2015 (EUR million; percentage of GDP)

	2010	2011	2012	2013	2014	2015
Goods	26 478	26 029	26 781	30 799	29 688	67 390
Services	66 245	64 995	70 294	68 591	81 271	106 554
Primary income	52 149	63 479	42 033	38 678	52 831	57 112
Secondary income	31 896	31 651	36 710	34 232	33 324	31 085
Capital account	9 727	15 466	11 264	7 600	10 918	10 264
Total	186 495	201 620	187 082	179 900	208 031	272 404
% EU-28 GDP	1.5	1.5	1.4	1.3	1.5	1.9

Source: Eurostat

The measured discrepancies affect in particular the **goods, services** and **primary income accounts** in European statistics, where elevated discrepancy levels were observed over the past years with usually higher measures for 2015 in these components (Figure 1).

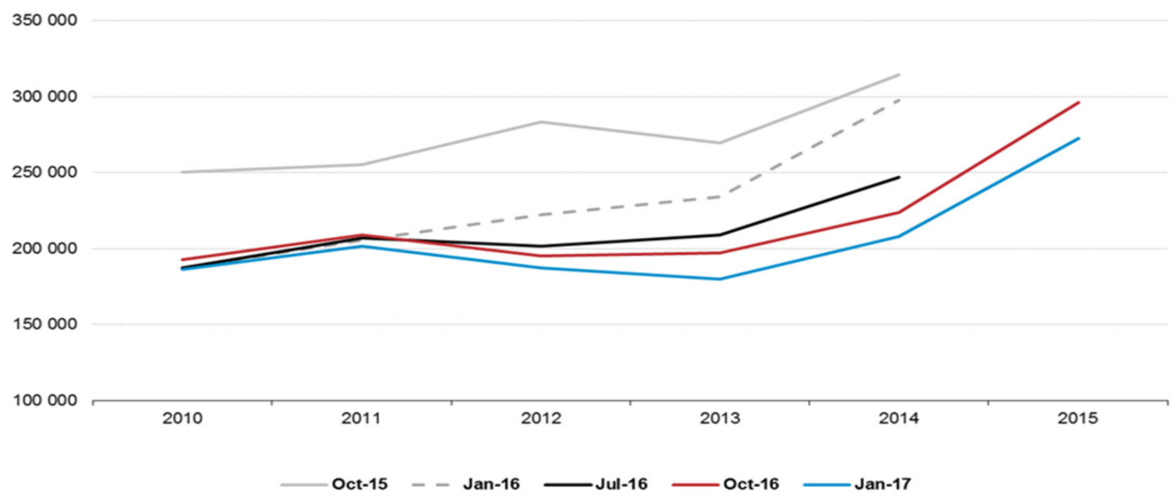
Figure 1: **Absolute discrepancies by components of nonfinancial accounts as per January 2017, sum of EU-28 Member States, 2010–2015 (EUR million)**



Source: Eurostat

Services and goods show to some extent a parallel evolution, as the underlying discrepancies also reveal a systematic bias due to different classification practices of component items between goods and services (e.g. treatment of goods acquired by households abroad and/or travelers) in the National Accounts and BOP. However, in regard to total transaction volumes the exposure in services appears much more prominent than in goods. Differences in the primary income accounts relate mostly to the component "property income" (D4), which due to its heterogeneous character and high incidence of estimations (e.g. on reinvested earnings, direct investment income) gives sufficient causes to divergent compilation practices among compilers.

Figure 2: **Absolute discrepancies in the European nonfinancial accounts over time, sum of EU-28 Member States, 2010–2015 (EUR million)**



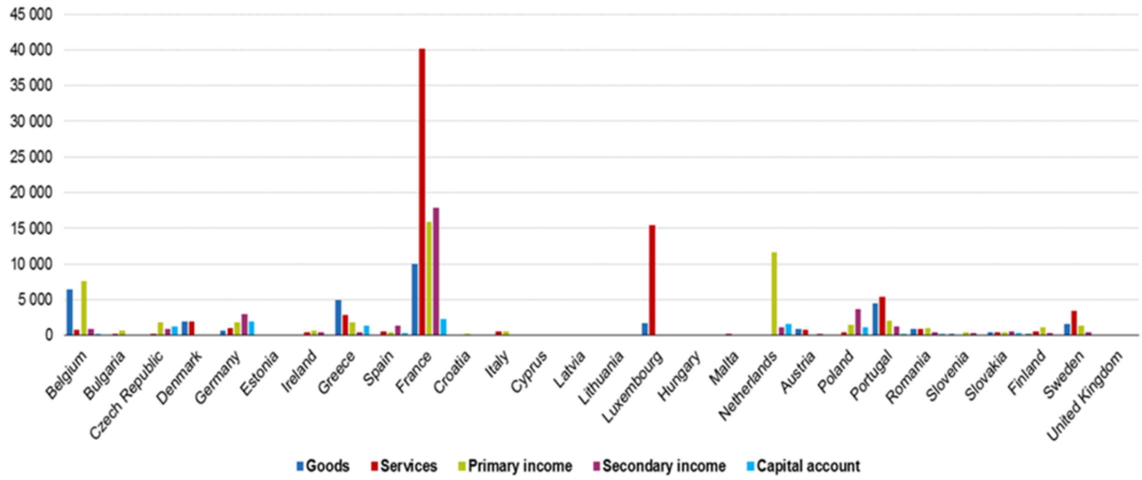
Source: Eurostat

Despite the above evidence of persistent discrepancies, we may also conclude on the beneficial impact of data revisions during the past (Figure 2). Since October 2015, when Eurostat first started to monitor consistency in the nonfinancial accounts after the convergence of the methodological standards, a considerable downward shift in discrepancy levels has been noticed, thus effectively reflecting most European compilers' ambitions to reconcile their statistics.

Overall total absolute discrepancies of the EU-28 fell from a multiannual average of EUR 274 billion in October 2015 to around EUR 206 billion in January 2017 for the observed period 2010-2015. This is an improvement of ca. 25%, bringing down relative exposure to discrepancies from earlier above 3% to below 2% of total EU-28 GDP. Between October 2015 and January 2016 most comprehensive revisions took place for the period 2010-2013, where compilers concentrated particularly on improving back data consistency. Later in 2016 compilers dedicated their attention to the more recent data, where considerable improvement was measured for 2012-2015. The recent data of January 2017 (blue line) confirm a continuation of this trend.

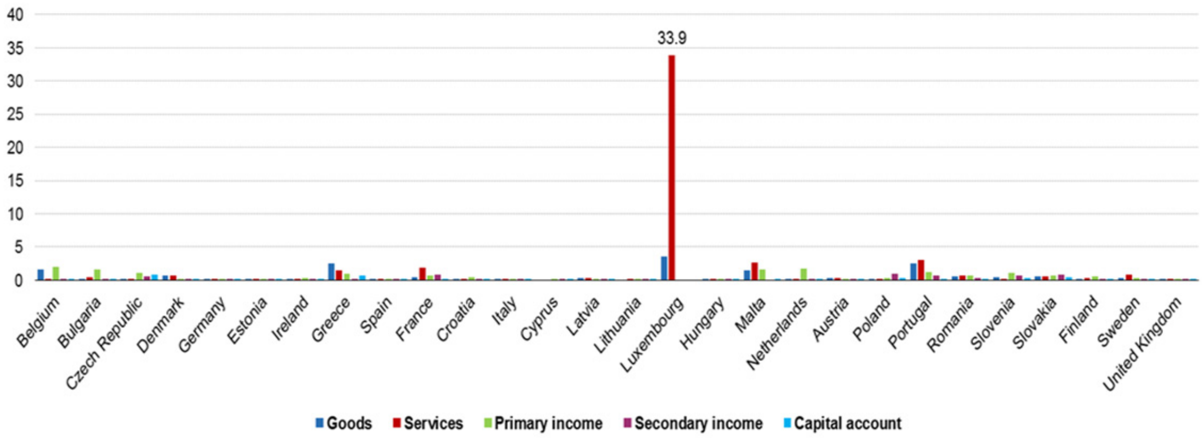
In view of the underlying country data of the EU-28, the geographical image of discrepancies in the EU-28 appears however highly dispersed. Major discrepancies originate from a group of 6 Member States only (Figure 3a).

Figure 3a: **Absolute discrepancies in the European nonfinancial accounts by EU-28 Member States, mean 2010–2015 (EUR million)**



Source: Eurostat – Absolute discrepancies = BOP minus ROW items

Figure 3b: **Relative discrepancies in the European nonfinancial accounts by EU-28 Member States, mean 2010–2015 (percentage of GDP)**



Source: Eurostat – Relative discrepancies in % of GDP (mean 2010-2015)

Depending on their exposure to the components of the nonfinancial accounts France, Luxembourg, Belgium, the Netherlands, Portugal and Greece show absolute discrepancies higher than EUR 10 billion in their multiannual means 2010-2015. These countries contribute currently to more than 76% of all discrepancies in the EU-28. 42% of mean annual discrepancies during 2010–2015 alone are attributed to

France, which is currently the main contributor to BOP-ROW discrepancies in Europe. On the other hand the statistics of the United Kingdom, Cyprus or Ireland appear highly consistent.

In regard to discrepancies in the components the profiles of the 6 mentioned countries differ significantly from each other: while France, Luxembourg and Portugal see high levels of discrepancies in services, the Netherlands and Belgium face most of their discrepancies in the primary income accounts. France, Belgium, Greece and Portugal also have elevated discrepancies in their goods account, with the first two contributing prominently to the outlier in 2015.

However, in relative terms to the countries' GDP, discrepancies appear significantly downsized, except for Luxembourg (Figure 3b). Luxembourg's discrepancies related to services amount to 34% of its GDP, and to goods ca. 4% of GDP. All other countries measure lower relative discrepancies to their GDPs, thus indicating little statistical significance.

3. Reasons for inconsistencies — findings and ongoing work

Due to the potential impact on comparability, a clearer view on the causes for inconsistent statistics is deemed instrumental. Since the introduction of the BPM6 to European BOP statistics, Eurostat has launched two surveys among European compilers, which allowed them to give explanations for the measured discrepancies in both statistics([120]). This feedback from the compilers helped to establish a few patterns about discrepancies.

- The organisational setup of national compilation processes plays a prominent role in explaining the occurrence of discrepancies. Decentralised statistical compilation systems lead more likely to institutional coordination and thus consistency issues.
- Different access to (micro) data sources or source statistics could generate discrepancies, in particular for items that can be measured from a heterogeneous spectrum of data sources. These data sources come to the compiler also at different frequencies in BOP and National Accounts statistics. Further, "contagion effects" arising from different (vintages of) source data, could import discrepancies to the final statistical product (e.g. financial data for the calculation of investment income)([121]).
- Items that are difficult to measure by surveys or administrative data sources are naturally subject to estimations or extrapolations (e.g. FISIM([122]), CIF/FOB adjustment([123])). Their uncoordinated use could pave the way for discrepancies.
- The methodological standards serve different statistical purposes. As a consequence the manuals are not always specific as regards thematic issues in the mirror statistics (e.g. the concepts of tourism and travel, delineation of goods and services). This leaves room for interpretation when applied by more than one compiler, resulting in different compilation practices due to consistency aspects with other macroeconomic statistics.
- Due to the specific objectives in each statistics and the foregone investment in IT infrastructure, (automatic) compilation systems are less flexible for being redesigned or adapted to new needs. As a consequence compilers generally appear less inclined to challenge already established and effectively working operational processes, even when their statistical products diverge from each other to some extent (low relative discrepancies).
- Institutional peculiarities foster discrepancies arising from different delineations of economic sectors (e.g. captive financial institutions, government-owned banks) or the economic territory([124]).

([120]) Results were published in a dedicated Working Paper, see Eurostat (2016), which were confirmed by the recent survey in 2017.
([121]) Obrzut (2016), p. 118
([122]) Financial Intermediation Services Indirectly Measured (FISIM)
([123]) Cost Insurance Freight (CIF) and Free On Board (FOB). The standards require an adjustment in order to make export and import transactions comparable.
([124]) For example, Swiss BOP and National Accounts treat the principality of Liechtenstein differently.

- Different institutional progress in fully adopting the corresponding statistical standards BPM6 and ESA2010 also explained to some extent the occurred discrepancies in the past (e.g. the inclusion of FISIM or illegal economic activities).
- Revision and vintage effects always persist as "statistical noise" due to different publication calendars and revision practices. Consequently, zero absolute discrepancies appear only achievable from fully integrated production systems (e.g. United Kingdom).

4. Bridging consistency and asymmetry aspects for international trade in services

While the analysis of BOP-ROW consistency is based upon the presumed identity of concepts in National Accounts and BOP statistics, the analysis of asymmetries refers to the compilation of geographical breakdowns in BOP statistics. Bilateral asymmetries have received particular attention during the past years and were subject to interesting analytical studies in causalities that appear also relevant to BOP([125]).

4.1. General assumptions and evidence

When establishing a direct relationship between BOP-ROW consistency and bilateral asymmetries in BOP, it is assumed that BOP statistics is the main data source for compiling the ROW account (or vice versa), thus the geographical image of cross-border transactions of one statistics is imputed to the other. Through this a direct transmission of consistency issues to asymmetry issues (and vice versa) appears justified. We have evidence that this indeed applies to the majority of Member States. Only in 6 Member States the ROW sector is not systematically sourced from the corresponding BOP data due to different compilation practices, horizontal balancing, different revision practices, or different (and assumed more reliable) data sources. However, for the compilation of goods and services BOP is the predominant source statistics in most EU-28 countries, thus we can consider the above assumptions for those countries justified. Given its high exposure to discrepancies with the ROW in European statistics, we would like to particularly select international transactions in the services accounts for our analysis. We compare the EU Member States' trade in services available from annual statistics with the corresponding mirror data of their EU partners (intra-EU asymmetries)([126]).

4.2. BOP-ROW discrepancies and bilateral asymmetries in services in the EU

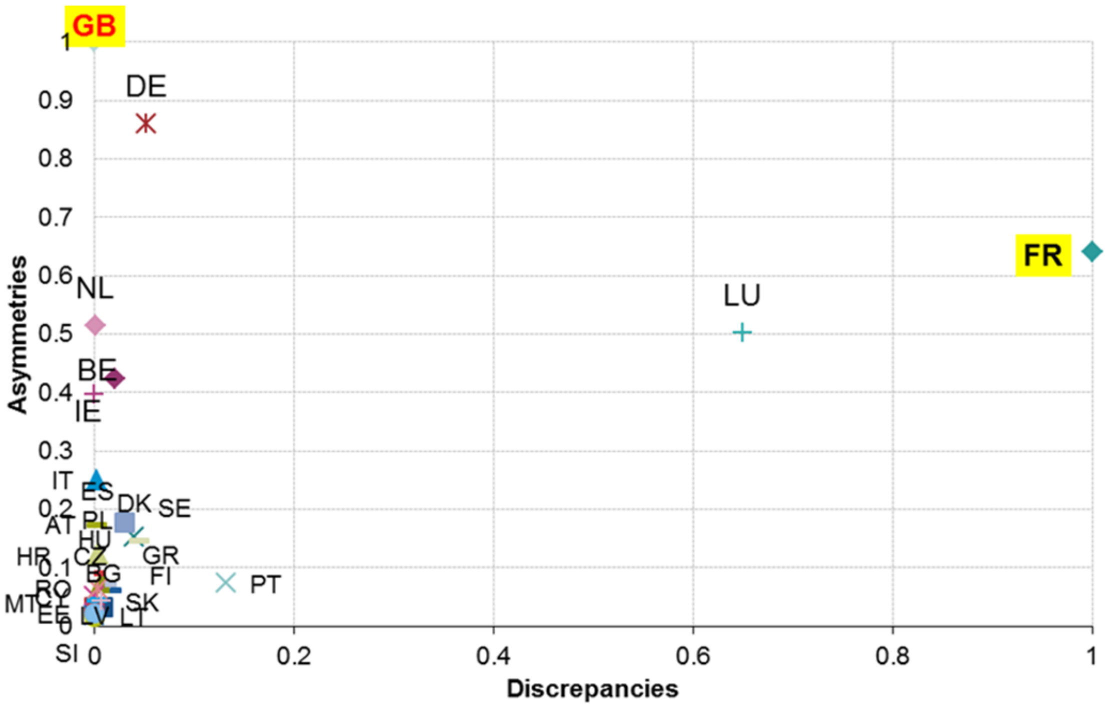
We extend our analysis to a second risk dimensions next to the risk of discrepancies (showing high absolute/relative discrepancies between BOP services and ROW services in one country) – the risk of asymmetries (showing high absolute/relative differences between BOP services of one country with the mirror BOP services of its EU partners. While the risk of discrepancies compares data from quarterly BOP and sector accounts, we measure the risk of asymmetries from data of annual International Trade in Services Statistics (ITSS)([127]). In order to do so, we calculate the extent of absolute differences between one country's exports and imports with the mirror data of its EU partners, and sum them up to a total. As a consequence the measure does not identify in detail those EU partners which trigger asymmetries, but depicts the extent of one country's asymmetries with all its EU partners.

Figure 4a combines the two measures on BOP-ROW discrepancies and bilateral asymmetries for the EU Member States in services. Each Member States is positioned on a two-dimensional risk map, showing its unique exposure to both risk dimensions based on absolute values in 2015. For comparison purposes, the

([125]) Fortanier, et al. (2016)
([126]) International Trade in Services Statistics (ITSS) for 2015 – ITSS provides for a more comprehensive country breakdown than the quarterly BOP time series (QBOP).
([127]) With some minor exception in the Netherlands quarterly BOP and annual ITSS time series are fully comparable in the EU-28.

measures have been normalised to the respective “outlier” values([128]). By means of normalising the data, we make them mutually comparable without the blurring effects arising from absolute discrepancy/asymmetry levels([129]). For the purposes of the analysis a two-dimensional playing field of risk exposure is created, on which each country is uniquely positioned according to its exposure to the risk of discrepancies and asymmetries. However, the reading of this presentation deserves some further clarifications: while exposure to the risk of discrepancies is directed at the counterparts of a specific country, exposure to asymmetries due to its nature has to be interpreted in a bilateral context. A country showing high exposure to asymmetries is not necessarily the originator of asymmetric data – this could also be the respective partner. However, both countries are held for initiating reconciliation efforts.

Figure 4a: **Absolute discrepancies/asymmetries, services, by EU-28 Member States, 2014 (normalised presentation)**

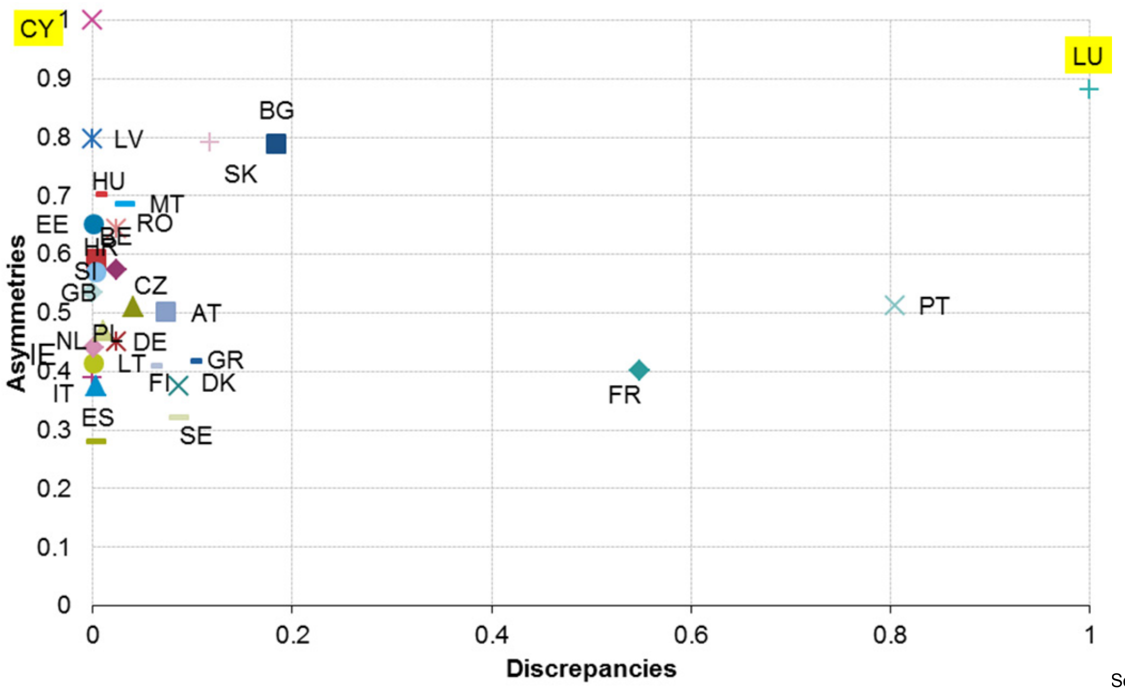


Source: Eurostat

With this in mind, France is the only country showing elevated levels in both dimensions (for absolute discrepancies it represents the outlier value), while noteworthy United Kingdom ranked highest in bilateral asymmetries with its EU partners, but shows no discrepancy with its ROW services. Similarly Germany ranks with little exposure to discrepancies, but high exposure to asymmetries with its EU partners. Luxembourg ranks relatively high in both dimensions, while the Netherlands, Belgium and Ireland have moderate exposure to asymmetries but little exposure to discrepancies (in services). Overall, EU countries appear more exposed to bilateral asymmetries than discrepancies, which is illustrated by the majority of countries’ positioning on the left side of the diagram.

Figure 4b: **Relative discrepancies/asymmetries, services, by EU-28 Member States, 2015 (normalised presentation)**

([128]) Outlier values relate to countries with highest exposure to the dimensions “Discrepancies” and “Asymmetries” respectively.
([129]) In absolute measures Member States are much more exposed to asymmetries than to BOP-ROW discrepancies. This would optically blur the image in favour of asymmetries. However, we assume both risks equally relevant to the analysis.



Source: Eurostat – Normalised presentation of exposure to discrepancies as percentage of total transactions in services (sum of exports and imports flows)

Depicting absolute measures in both risk dimensions ignores however countries' general exposure to transactions in international services. As a consequence, countries with high transactions in services are more likely to observe high risk exposure, while countries with little international transactions in services observe little risk exposure, respectively. Relative measures could accommodate this bias. Figure 4b illustrates EU Member States' relative exposure to both risk dimensions, showing Cyprus as outlier in asymmetries but no discrepancy, and Luxembourg as outlier in discrepancies and high exposure to asymmetries. Compared to the diagram on absolute risk exposure most countries move up in their relative exposure to asymmetries, i.e. in relative terms the risk of bilateral asymmetries becomes a quality concern in the majority of the EU countries, while BOP-ROW discrepancies in services remains a concern only for three countries (Luxembourg, Portugal and France).

This presentation is an attempt to illustrate the multidimensionality in which the BOP-ROW consistency debate is embedded. Countries find themselves in a playing field between at least two quality risk dimensions where they have to find a strategy in organising their compilation processes to such an extent that exposure to one or the other dimension (or both) is minimised. In the light of the earlier findings, it is argued that tackling the risk of BOP-ROW discrepancies can be achieved by improved national coordination measures, tackling the risk of asymmetries by international coordination with the respective EU partners. Usually the compiler is facing a trade-off between the two, and will have to choose a path related to other systemic parameters applying (budget restrictions, institutional autonomy, resource limitations in IT and/or human capital, geographical distances, language barriers, etc.). However, it is argued that through a higher degree of coordination both at national and international level, a move towards lower exposure to both asymmetries and discrepancies appears feasible.

4.3. Impact on data comparability and credibility

From the above findings we can deduce that exposure to quality risks (inconsistent and/or asymmetric statistics) has unfortunate consequences, if remaining unexplained:

1. It casts severe doubts on the accuracy of the statistical product and impact its compiler's credibility (e.g. questioning the correctness of intra- and extra-EU breakdowns);

2. It brings up the question about the actually correct figures: Who is right – BOP or National Accounts? Is a weighted average between asymmetric countries closer to reality? Does the involvement of third

countries in the case of asymmetric data blur the reported bilateral data (If yes, to what extent)? This in turn impacts other macroeconomic statistics (Input-Output Tables);

3. They may involve the disclosing of confidential company-related data for studying asymmetries, which cannot be shared with partner countries. On the other hand, partner countries need this information in order to identify potential gaps in their data sources.

4. Asymmetries may be significant in one country, but insignificant in the partner country, thus impeding any further bilateral investigations and follow-up; BOP-ROW discrepancies may be relatively insignificant in the originating country, although contributing most prominently to overall inconsistencies in Europe;

Table 4 shows some selected significant pairs in asymmetric data in the EU-28, involving a few Member States. While asymmetries for example in Financial Services are prominent for Luxembourg (in relation to its GDP), for its partner countries (e.g. Germany and United Kingdom) they are not significant (0.2% and 0.5% of their GDP respectively). Similarly, while bilateral asymmetries in charges for the use of intellectual property are relatively significant for Luxembourg (11.3% of its GDP), for Ireland they have less significance (2.3% of its GDP).

Table 4: Absolute asymmetries in selected EU Member States and services components, 2015

(EUR million; percentage of GDP)

Reference country	Partner country	Asymmetry (EUR million)	Services component	Reference country (% of GDP)	Partner country (% of GDP)
DE	**LU**	5 069	SG - Financial services	0.2	9.9
LU	**GB**	13 549	SG - Financial services	26.5	0.5
DE	**AT**	4 823	SC -Transport	0.2	1.4
GB	**IE**	3 788	SC -Transport	0.1	1.5
GB	**IE**	5 088	SI - Telecommunications	0.2	2.0
LU	**IE**	5 774	SH - Charges for the use of intell. property	11.3	2.3
FR	**NL**	6 054	SJ - Other business services	0.3	0.9

Source: Eurostat – Asymmetry = sum of absolute differences of export and import flows with their mirror transactions in partner countries

We would also like to illustrate this situation by some EU Member States' exposure to discrepancies in cross-border services for the year 2015 – the most exposed component to discrepancies in the EU-28 as shown earlier. The below mentioned 6 countries have a common pattern in their discrepancies for services – the measured discrepancies appear to these countries negligible when related to their GDP, although they contribute prominently to overall discrepancies of the EU-28 (assuming more than 60% of total EU discrepancies in services for 2015). Most dramatically France contributes to almost 50% of total discrepancies in services, but measures its discrepancies for services only by 2.4% of GDP. From a national point of view it could be argued that the consistency issue is statistically not significant, and appears somewhat exaggerated.

Table 5: Selected EU Member States with high relative impact to discrepancies in services, 2015 (EUR million; percentage of GDP)

	Discrepancy (EUR million)	Country share to total EU discrepancies	Share in % of countries' GDP
Belgium	1 031	1.0	0.3
Denmark	2 107	2.0	0.8
Greece	871	0.8	0.5
France	52 370	49.1	2.4
Portugal	6 894	6.5	3.8
Sweden	2 341	2.2	0.5

Source: Eurostat – Discrepancy = sum of absolute differences of credit and debit flows of reference country with their mirror transactions in the ROW account of the reference country

Disregarding whether these quality risks involve national or international counterparts, the above situation is described in scientific literature as “Nash equilibrium”. Nash equilibrium is a stable state of a system that involves several interacting participants in which no participant can gain by a change of strategy as long as all other participants keep their strategy unchanged([130]). In the above context this characterises a situation, where Member States with relatively insignificant discrepancies/asymmetries will not pursue any other strategies than conserving the status quo, while at the same time these discrepancies/asymmetries remain significant to their counterparts’ (mirror) statistics. These Member States will therefore not prioritise a move towards higher consistency/symmetry levels due to this lack of significance as it makes unilateral initiatives appear uneconomical, and accept the (second best) status quo as their dominant strategy. As a way out cooperative strategies are suggested, moving participants towards an overall optimal situation (characterised by fully consistent and symmetric data). Such endeavours cannot be launched unilaterally, but require a neutral coordinating body. This emphasises international institutions’ role as motor in order to pave the way towards more consistent and symmetric statistics among Member States, where these Member States would otherwise consider exposure to quality risks insignificant. Due to the high concentration of inconsistencies/asymmetries around a few countries quality initiatives involving these countries could already create a high overall impact.

5. Conclusions and outlook

In this paper we illustrated the multidimensional character of statistical quality, and put the concept of fully consistent National Accounts and BOP statistics into a broader perspective including geographical asymmetries. It is clearly suggested that Member States submit their statistical products more actively to national and international coordination. As in a majority of EU Member States compilation processes for National Accounts and BOP statistics are decentralised, an EU/global perspective is required in order to consistently implement a higher degree of coordination across European (and worldwide) counterparts. This does not only include coordinated compilation practices, but also a common reading of methodological standards that were written under different statistical objectives, common access to micro data sources, and a higher degree of standardisation in the underlying micro data and business registers. Successful examples of the past have shown that the provision of common access to micro databases, e.g. Centralised Securities Database (CSDB) of the European System of Central Banks can contribute to more harmonised financial statistics([131]). Current initiatives to establish an international identification standard (e.g. Legal Entity Identifier) into international business registers or the establishment of an FDI Network([132]) have emphasised the prominence of international coordination for the sake of more symmetric statistics on international transactions.

It has been argued that this situation cannot be successfully changed without a cooperative approach (“Nash equilibrium”). International organisations are challenged with “closing the gaps” in the standards which still foster diverging compilation practices in BOP and National Accounts on one side, while

([130]) Princeton University, http://www.princeton.edu/main/tools/search/?q=Nash%20Equilibrium
([131]) Pérez / Huerga (2015)
([132]) https://www.imf.org/external/pubs/ft/bop/2014/pdf/14-20.pdf

counteracting consistency requirements with other macroeconomic statistics on the other. As a consequence the consistency debate has to be put into perspective and submitted to overall priorities in a multidimensional context, incorporating aspects such as trade asymmetries, etc. as well.

In the light of the above experience the following obstacles seem to hamper progress and will prominently remain on statisticians' agenda during the oncoming years:

- Institutional autonomies and strategic rationales;
- Resource restrictions on human and IT capacities;
- Need for flexibility in adjustment practices among compilers;
- Strict confidentiality regimes;

This list of structural obstacles finally suggests that a paradigm shift is necessary in order to tackle them, by moving away from autonomous statistical compilation systems and rationales towards more systematic cross-border sharing of information, access to common databases, coordinated cross-border profiling (e.g. "early warning systems", "large cases units"), as well as unique and linked identifiers in coordinated registers, in order to tackle the increasing complexities of compiling international statistics in a globalised world([133]).

References

Central Statistics Office, *Ireland's Economic Growth Figures 2015*, Press Release 12th July 2016, Corcaigh, Éire, 2016.

Eurostat, *Consistency between national accounts and balance of payments*, Working Paper, Luxembourg 2016.

Fortanier F., et al, *Asymmetries in Merchandise trade and trade in services statistics: Introduction to the Country Notes*, OECD, Paris, 2016.

Obrzut R., 'Consistency between national accounts and balance of payments statistics', in: EURONA, Luxembourg, 2016.

Pérez Cornejo A./ Huerga J., *The Centralised Securities Database (CSDB): Standardised micro data for financial stability purposes*, BIS Working Papers, Basle, 2015.

Stapel-Weber S./ Verrinder J., *Globalisation at work in statistics: Questions arising from the 'Irish case'*, in: EURONA, Luxembourg, 2016.

BPM6, *Balance of Payments and International Investment Position Manual*, Sixth edition, IMF, Washington DC, 2009.

ESA2010, *European System of Accounts*, Eurostat, Luxembourg 2013.

([133]) Stapel &Verrinder (2016), p. 30ff.

13 An innovative webGIS system for dissemination and visualization of official statistics and geospatial analysis

EMANUELA RECCHINI, ALESSANDRO CAPEZZUOLI[134]

Table of Content

([134]) Italian National Institute of Statistics (Istat).
The views expressed in this paper are solely those of the authors and do not involve the responsibility of Istat. Addresses for correspondence: emanuela.recchini@istat.it; and alessandro.capezzuoli@istat.it

1. Introduction

Statistical information is essential for decision makers, politicians, researchers, journalists and the general public. They need relevant, reliable and even more timely data for their work.

Official statistics are best suited to meet this need. In fact, data that are produced under the statistical programme of a given country or an international organization are characterized by the highest quality possible. Not only are quality criteria met for the data themselves but also strict conditions concerning processes are in place, following the United Nations Fundamental Principles of Official Statistics and the European Statistics Code of Practice.
ESS Vision 2020 emphasizes the importance to complement existing data derived from traditional sources with newer emerging ones, including geospatial and where possible big data. It is also stressed the relevance and the opportunity of sharing tools within the ESS and the usefulness of proper dissemination channels that meet the needs of as many people as possible.
In this respect, it is evident how crucial the role played by National Statistical Offices (NSOs) may be in developing suitable tools for the integration of different types of information and the dissemination of statistical data.

In the present chapter, after an overview of main formats, standards and methods used for data dissemination, data visualization issues and webGIS systems are considered, with particular reference to Istat experience. A specific focus is put on StatVIEW, a tool which has been created for purposes linked to this topic and can be conveniently shared among NSOs; concrete examples are shown to highlight its possible uses. Some concluding remarks emphasize the main features of StatVIEW.

2. Dissemination and visualization of official statistics

2.1. Main formats, standard and methods for data dissemination

At present, one challenge for NSOs is the digital transformation going on world-wide. The ever increasing availability of data – facilitated by advances in information technology and the Internet – requires the implementation of methods and tools for the proper assessment, interpretation and dissemination of the same data.

Dissemination of high quality data gives users the opportunity to compare and integrate statistical information from different sources to best analyze the phenomena under investigation.

Before the creation of web-sites, the main channel of data dissemination was publication on paper. Now, thanks to Internet, the speed at which information travels and the evolution of tools available to users (PCs, tablets, smartphones, apps) have redefined the way of data dissemination and utilization. The web regularly provides up-to-date data: this is a crucial point that distinguishes this type of dissemination from paper publication.

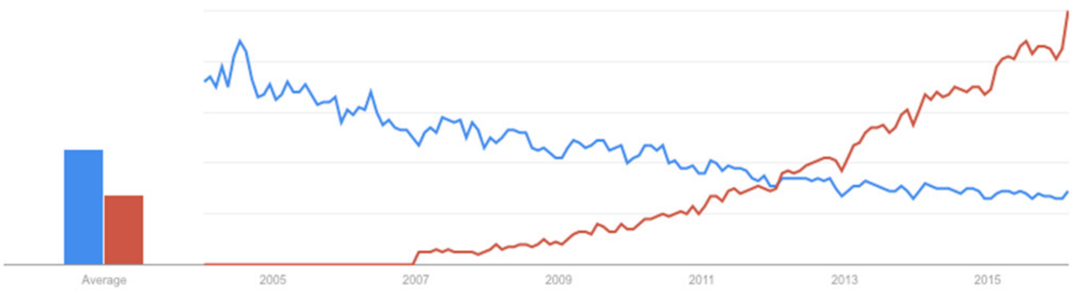
Solicited by ESS, NSOs have started exploring new technology-driven areas when dealing with very large amount of data and are developing suitable tools for the integration of different types of data and its dissemination, i.e. platforms for data storage, analysis and visualization.

There are several formats, standards and methods used for data dissemination on the web. In this paragraph, two open formats regularly in use are considered: XML (eXtensible Markup Language) and JSON (JavaScript Object Notation). They are both defined as technical standard by the World Wide Web Consortium (W3C), an international community intended to realize the full potential of the Web.

XML defines a set of rules for encoding documents in a format that is both human-readable and machine-readable. JSON provides a format that uses human-readable text to transmit data objects consisting of attribute–value pairs. It is speedy and easy to understand and use. In recent years, thanks also to web 2.0

and the considerably increasing popularity of the JavaScript language, the use of JSON format has largely replaced XML. The graph below shows the interest of users in textual search of 'Json API' (red line) versus 'XML' (blue line) on Google (Figure 1).

Figure 1: XML vs JSON (users' interest) – Years 2005-2015



Source: Google Trend.

Being a JavaScript Object Notation, JSON is the natural environment for the construction of web applications using programming languages (like Javascript, Ajax, PHP, JSP, PYTHON) combined with *ad hoc* frameworks as appropriate (e.g., Node.js, OpenLayers, Leaflet, Raphael JS).

Different purposes may be served by the available frameworks: examples are graph construction, database management, asynchronous calls to external resources, cartographic applications, etc.

OpenLayers and Leaftlet are the main frameworks used for cartographic applications. The latter, in particular, provides a set of powerful and complete tools and plugins for data integration and the use of web resources.

JSON and XML differ in the type of architecture used: the former uses a resource-oriented architecture, while the latter uses a service-oriented one.

It is worth noting that the protocol used to uniquely identify (through a URL, Uniform Resource Locator) and then accessing web resources is the HTTP (HyperText Transfer Protocol). REST (Representational State Transfer) defines a set of guidelines to use the HTTP protocol in order to perform four operations summarized in the acronym CRUD (Create, Read, Update, Delete), by means of API (Application Programming Interface). See figures 2 and 3.

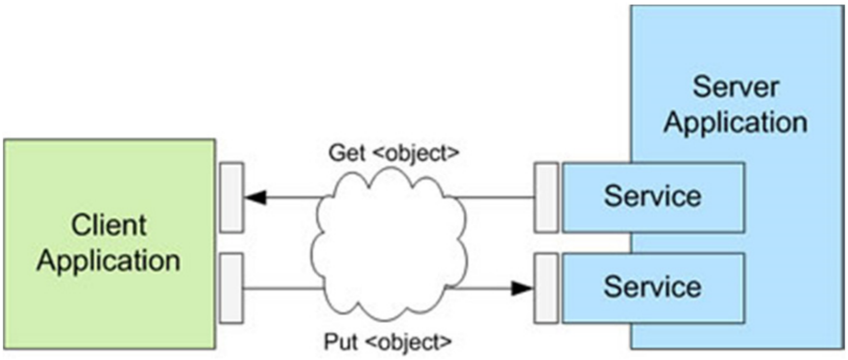
Figure 2: HTTP methods vs CRUD operations

HTTP METHODS	CRUD OPERATIONS	DESCRIPTION
POST	CREATE	Create or add a new resource
GET	READ	Read, retrieve, search, or view existing resources
PUT	UPDATE	Update or edit existing resources
DELETE	DELETE	Delete/deactivate/remove existing resources

An API is a web resource that is based on a client server architecture. It provides a service when called. In this case, the service is represented by the data obtained. Figure 3 suggests how an API interface with REST architecture allows users to work in a simple and well-

ordered way.

Figure 3: REST API Architecture



2.2. Data visualization and webGIS systems

In the last few years, there has been a rapid increase of languages, frameworks and tools for the web, allowing users wide opportunities for cartographic visualization and geospatial analysis of statistical data. In order to produce simple density maps or dashboard containing tables and graphs, many commercial and open source tools have been developed (for example, Tableau, Microstrategy, Spago BI, etc.). However, these tools do not often take advantage of the potential of geo-referenced data, simply providing users with little more than infographics.
WebGIS systems are more advanced tools, allowing users to satisfy their different needs. This is because they enable not only data dissemination and cartographic visualization, but also dynamic territorial analyses. WebGIS systems are very useful to produce *ad hoc* applications for real-time and time series analyses, monitoring of specific phenomena such as earthquakes, urban mobility, environmental pollution, as well as data integration using REST (or SOAP) web services and standards like JSON-stat.

2.3. Istat experience

The usability of data is related to the format (XML, JSON, CSV, etc.) and the model adopted for their representation. At present, the Italian National Institute of Statistics (Istat) uses different channels for data dissemination and visualization. The main ones are:

- The single exit point for exchanging and delivering data to Eurostat and other Institutions through SDMX format;
- I.Stat, the data warehouse of statistics currently produced by Istat, where data are organized in a coherent and homogeneous way and are constantly upgraded. I.Stat permits data downloading in several formats (csv, excel, SDMX, etc);
- A set of API REST to disseminate data in the JSON-stat format.

SDMX (Statistical Data and Metadata eXchange) is an international standard providing: a logical model to describe statistical data, together with guidelines on how to structure the content; a standard for automated communication machine to machine; a technology supporting standardized IT tools that can be used by all parties involved in the data exchange and processing([135]). In recent years, other models have appeared, in addition to SDMX (for example JSON-stat, DDI, etc.).

Istat data warehouse is conceived on the basis of the multi-dimensional array of data OLAP Cube (On-Line Analytical Processing). An OLAP cube is a method of storing data in a multidimensional form, generally for reporting purposes. The faces of the cube contain the dimensions – i.e. the dimensional spreadsheets containing the descriptive attributes of the measures (region, province, gender, date of birth, etc.) –, while each cell contains the value related to the intersection of dimensions. The arrangement of data into cubes overcomes a limitation of relational databases, which are not well suited for near instantaneous analysis and display of large amounts of data.

([135]) SDMX is an international initiative that aims at standardizing and modernizing the mechanisms and processes for the exchange of statistical data and metadata among international organizations and their member countries. For further details, see: https://sdmx.org/?page_id=3425 and http://ec.europa.eu/eurostat/data/sdmx-data-metadata-exchange.

The cube model is at the basis of the JSON-stat format, which is a JSON-based data cube packaging format. The JSON-stat format is a simple lightweight JSON format for data dissemination. It is based on a cube model that arises from the evidence that the most common form of data dissemination is the tabular form. In this cube model, datasets are organized in dimensions, which in turn are organized in categories([136]).

3. StatVIEW: a webGIS system for data visualization, dissemination, monitoring and geospatial analysis

3.1. What is StatVIEW?

StatVIEW (http://www.statview.eu) is an innovative tool which can conveniently support whatever analysis focused on one or more subject areas covered by official statistics (social, demographic, economic, environmental, etc.). StatVIEW uses JSON-stat API for the dynamic construction of density maps based on data from different sources. It is a user-friendly webGIS system, developed by making use of open source technologies; it permits to gather, link, standardize, disseminate and visualize statistical data in different formats.

Similarly to data on population, health, education, labour, environment, national accounts and social security and welfare produced and provided by Istat currently visualized through StatVIEW, data produced by NSOs, research institutes and government agencies, available according to different geographical units, can be disseminated through the web using web services complying with the JSON-stat standard or other data exchange standards (i.e., SDMX, OData); also data other than official statistics can easily enter the webGIS system. The release of geospatial data in the form of Web Map Service (WMS) and Web Features Service (WFS) is managed by Geoserver, which provides data on territorial boundaries and georeferenced point data. StatVIEW may be seen as a JSON-stat hub that can be used as a web service to further disseminate data in a machine-readable format.

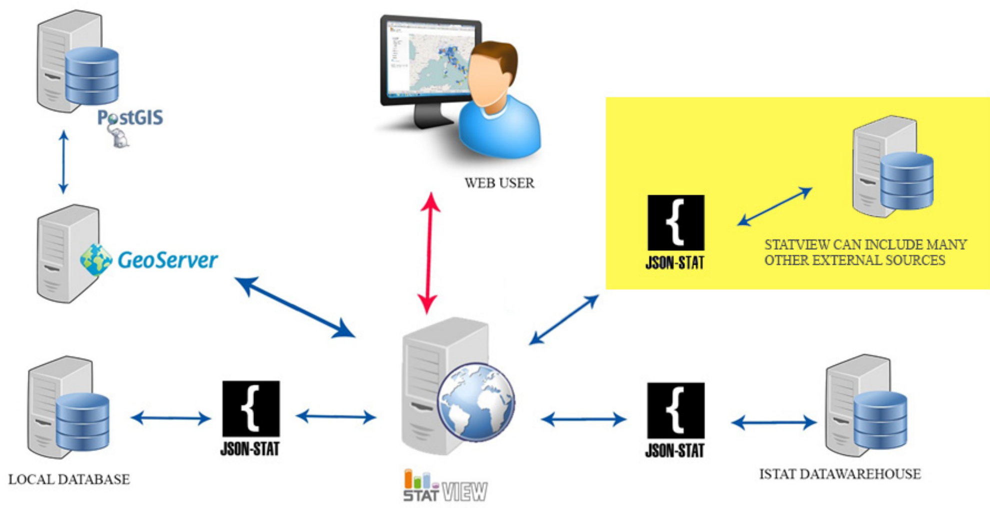
3.2. StatVIEW architecture

StatVIEW, based on a microservices architecture pattern, is composed of a map server and a geospatial database (Geoserver, Postgres, PostGIS), different data sources (data warehouse, web services, database, etc.), interrogation engine and user interface (Figure 4). By using a map server you can import shapefiles in geospatial database containing geographic data and disseminate them through the HTTP protocol, in compliance with the standards defined by the OGC Consortium (WMS, WFS, WCS), in a format that suits the web – GeoJSON, TopoJSON, etc. – and existing frameworks (OpenLayers, Leaflet, D3.js).

Though StatVIEW is a highly scalable system that can support different data sources (databases, files, web services), different models (SDMX, DDI, JSON-stat) and different formats (XML, JSON), it was chosen to standardize the source output in the JSON-stat format. This choice, in addition to standardizing data from different sources, permits to integrate the frameworks to work with (JSON-stat framework toolkit, OpenLayers, Leaflet and Raphael JS) and permits to unify the programming language (Javascript). Unlike many tools that enable users only density maps visualization, StatVIEW allows them to perform operations normally included in GIS Desktop and WebGIS systems (Topology Overlay; Network analysis; Buffering).

([136]) For further details on JSON-stat, see: https://json-stat.org/.

Figure 4: StatVIEW architecture



When multidimensional cubes are updated in the Istat data warehouse, updated data are automatically visualized in the user interface. Similarly to what happens in the case of Istat data warehouse, it is possible to visualize or monitor data from NSOs, research institutes and government agencies.

3.3. Graphical representations and geospatial analysis through StatVIEW

3.3.1. DENSITY MAPS, GRAPHS AND TABLES

Among the different possible types of graphical representations, StatVIEW permits to visualize statistical data in the form of density map, graph or table (table visualization is achieved through JSON-stat Table Browser). The box below shows the case of Istat data related to residence permits of non-EU citizens. By clicking on the menu available on StatVIEW, it is possible to:

- select the dimensions associated to the dataset;
- explore the entire content of the dataset;
- choose the type of data representation: cartographic (Figure 5); graphic (Figure 6); tabular (Figure 7);
- customize the research (for instance, by choosing dimension, territorial level, type of visualization, or by splitting density areas; several other possibilities are also available). Gradients and subdivisions can be easily customized, adding extra layers in the database management. The customization depends on the layers available in the chosen map. In the case of residence permits of non-EU citizens, the available layers are: the layer containing the georeferenced organizations providing assistance to migrants, the layer that, via geospatial query, creates a bubble chart representing the number of the georeferenced organizations active in the territory (Figure 8). By selecting the layers and, for example, different base map or colour shade or partition (quartile, quintile), it is possible to get a different visualization, conducing to a different interpretation of the phenomenon;
- visualize data according to different geographical units.

A legenda allows users to know the threshold values associated with colour shades.

Residence permits data visualization

Figure 5: Density map

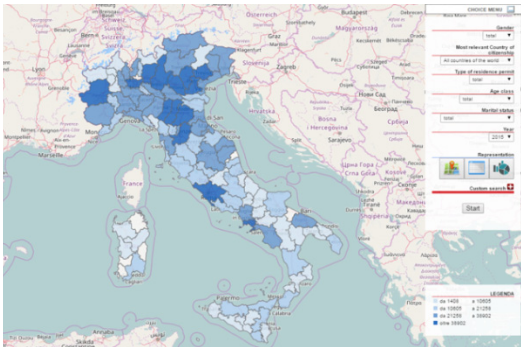


Figure 6: Graph

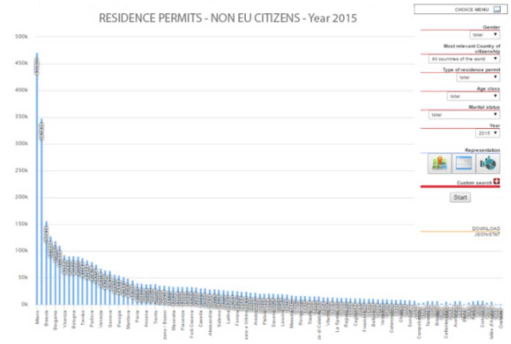


Figure 7: Table

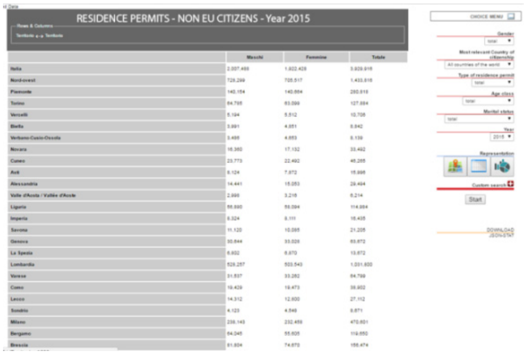
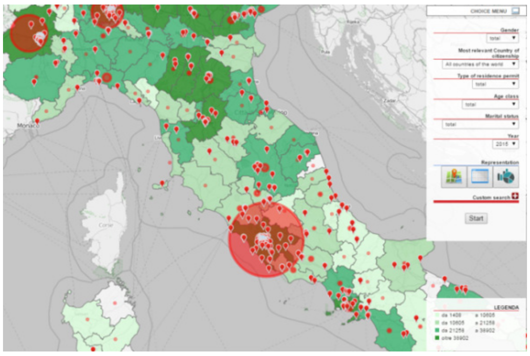
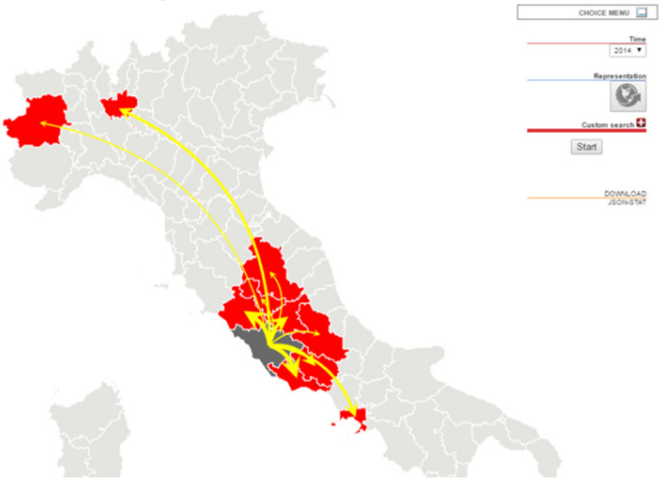


Figure 8: Topology overlay



3.3.2. INFLOWS AND OUTFLOWS

Figure 9: Outflows from Rome



Transfers of residence

Through StatVIEW it is possible to represent inflows and outflows to and from a given territorial border.

Figure 9 shows outflows of foreign citizens cancelled from the population register of Rome and registered in the population register of another Italian municipality.

It is possible to set the maximum distance or the distance range for the desired flows.

For example, you may be interested in visualizing the flows within 100 km from a given province or within a distance ranging 100-300 km. The dimension of the arrows gives an idea of volumes: the thicker the arrow, the greater the flow.

3.3.3. ROUTING

Figure 10: Routing

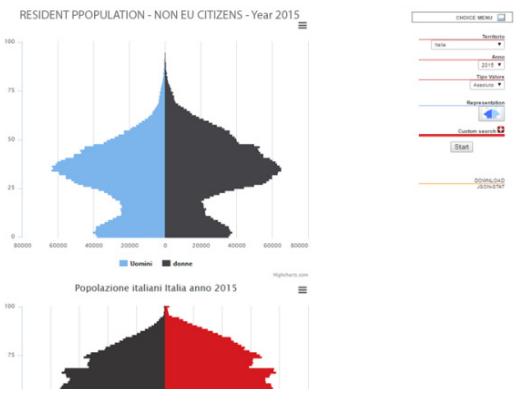
OSRM Routing machine

StatVIEW uses the OSRM (Open Source Routing Machine) routing algorithm.

By using this algorithm, it is possible to obtain the shortest route to a geo-localised point in the map server used by StatVIEW (e.g., hospital, police station, etc.) from any point in the map, as well as the respective travel time (Figure 10).

3.3.4. CUSTOM GRAPHS

Figure 11: Pyramid graph



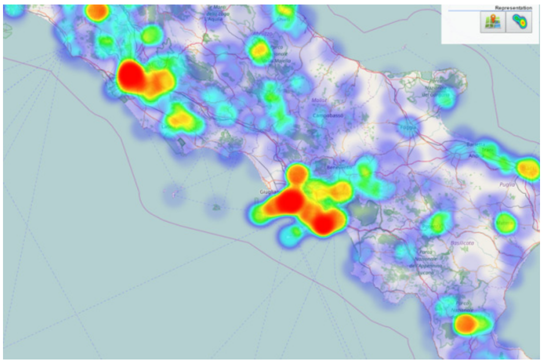
Pyramid graph

The standardization of sources' output allows users the easily creation of *ad hoc* plugins for the generalized visualization of multi-source graphs or dashboards.

Figure 11 presents a pyramid graph showing the age of foreign citizens resident in Italy, compared to that of resident Italian citizens.

3.3.5. HEATMAPS AND CLUSTERS

Figure 12: Heatmap of private museum



Private museums

StatVIEW permits to create heatmaps and clusters of geolocated points to represent their density and distribution on the territory.

Figure 12 represents the heatmap of private museums resulting from the census survey carried out annually by Istat, the Ministry for Cultural Heritage and Activities and Tourism, the Regions and Autonomous Provinces.

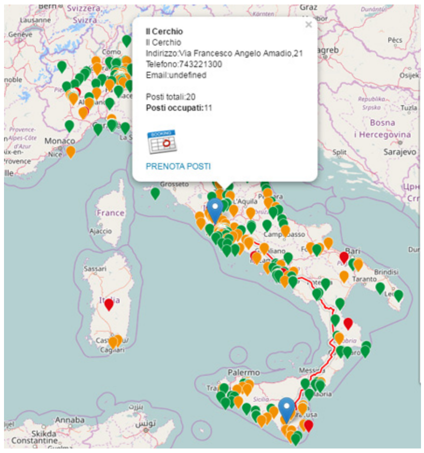
3.3.6. AD HOC APPLICATIONS

Management of immigration centres

Thanks to the support of geospatial database, StatVIEW permits the development of *ad hoc* applications.

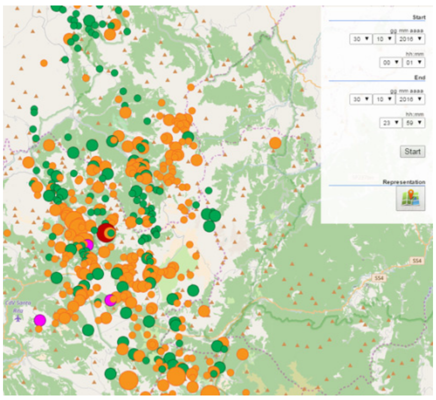
The map in Figure 13 indicates the position of immigration centres. The colour indicates the number of available sleeping accommodations. By clicking on the marker of an immigration centre, an interested user can book a certain number of accommodations, in order to equally distribute immigrants. In a reserved area, it is possible to manage the immigration centre and the booking.

Figure 13: Ad hoc application



3.3.7. MONITORING

Figure 14: Earthquakes real-time monitoring

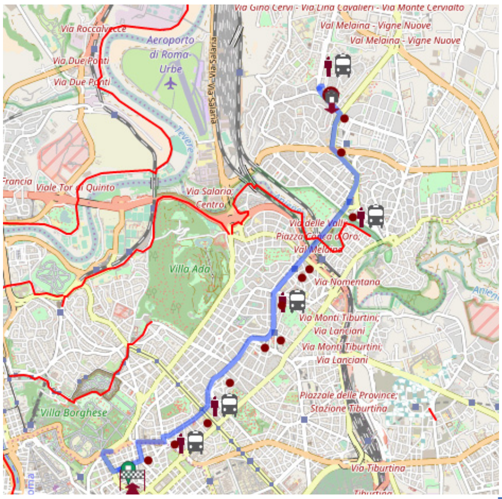


Real-time monitoring

Micro services architecture at the basis of StatVIEW permits the creation of different monitoring system (real-time and time-series) to study and control several phenomena.

Figure 14 is an example of how StatVIEW enable users to real-time visualize and monitor earthquakes and their main features (location – with the precise indication of geographic coordinates –, magnitude, depth, day and time of occurrence). In the case of earthquakes, the data source is INGV, the Italian National Institute of Geophysics and Volcanology, devoted to 24-hour countrywide seismic surveillance, real-time volcanic monitoring, early warning and forecast activities.

Figure 15: Urban mobility real-time monitoring



Urban mobility

StatVIEW permits the exploitation of urban mobility data to real-time control and monitor road traffic, timetables and public transports routes.

In Figure 15 the case of collective mobility in the Metropolitan Area of Rome is represented. The website of the Municipality of Rome provides a set of APIs through which it is possible to monitor the positions of public transports (buses, trolleybuses, trams, metro lines, regional railways, etc.), know routes and route planners to follow to get from one point to another of the city using public transport.

The overlay, by combining for example routes and bus positions to the tracks of the bike paths and car sharing areas, facilitates urban mobility.

4. Concluding remarks

In the Internet Era, the ever increasing availability of data, essential in supporting research and decision making, necessarily involves the use of proper tools to best analyze, assess and interpret their meaning in an interactive fashion. Solicited by ESS, NSOs, playing an important role in developing, producing and disseminating statistics that respect standardized quality criteria, have started developing suitable tools for the integration of different types of data and its dissemination, i.e. platforms for data storage, analysis and visualization. A particularly interesting way to integrate different sets of data is to bring in a spatial dimension; in this way, the complexity of phenomena under investigation is still there but their analysis can be facilitated and early warning is made more at hand.

StatVIEW, the proposed webGIS system, is an innovative tool providing a wide range of solutions to easily analyze the many phenomena for which statistical data exist as well as to quickly monitor the evolution of different phenomena at the same time. It is currently in an experimental phase, consequently a web analytic tool to keep track of visitors, although already in place in a basic form, does not permit at the moment a deeper analysis on traffic and insight on visitors' behavior in order to create strategies according to it. What is certain is that the platform, already presented in numerous national and international conferences, is receiving tangible and more than positive feedback from academics, researchers and statistics stakeholders, mostly interested in the use and replicability of it.

One of its significant features is represented by the opportunity of having constantly updated statistical information, given the fact that the webGIS system does not imply the loading and transferring of data, while, instead, it provides a dynamic link.

Overall, this open source tool can be conveniently shared among NSOs and extended to any institution that performs data dissemination through API REST.

Furthermore, a webGIS system such as StatVIEW is along the lines of ESS Vision 2020, in the sense that it helps official statistics to engage users proactively and to meet their demands in a cost-effective and responsive manner.

References

De Smith, M.J., Goodchild, M.F. and Longley, P.A., *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools,* Troubador Publishing Ltd, 2007.

Henderson, C., *Mastering Geoserver*, Packt Publishing Ltd, 2014.

Kalkhan, M.A., *Spatial Statistics: GeoSpatial Information Modeling and Thematic Mapping*, CRC Press, 2011.

Obe R.O. and Hsu L.S., *PostGIS in Action, 2nd Edition,* Manning Publications, 2015.

Richardson, L., and Ruby, S., *RESTful Web Services. Web services for the real world*, O'Reilly Media, 2007.

Schabenberger, O. and Gotway, C.A., *Statistical Methods for Spatial Data Analysis*, Chapman & Hall/CRC Press, 2005.

Webber, J., Parastatidis, S. and Robinson, I., *REST in Practice: Hypermedia and Systems Architecture*, O'Reilly Media, 2010.

14 Spontaneous recognition: an unnecessary control on data access?

FELIX RITCHIE[137]

Table of Content

([137]) Bristol Business School. A version of this paper has also been published in the ECB Statistics Paper Series (no. 24 / August 2017; http://www.ecb.europa.eu/pub/pdf/scpsps/ecb.sps24.en.pdf?4b7c709063a3b704e1582c9b29ef77c2)

Abstract

Social scientists increasingly expect to have access to detailed source microdata for research purposes. As the level of detail increases, data owners worry about 'spontaneous recognition', the likelihood that a microdata user believes that he or she has accidentally identified one of the data subjects in the dataset, and may share that information. This concern, particularly in respect of microdata on businesses, leads to excessive restrictions on data use.

We argue that spontaneous recognition presents no meaningful risk to confidentiality. The standard 'intruder' model covers re-identification risk to an acceptable standard under most current legislation. If a spontaneous re-identification did occur, the user is very unlikely to be in breach of any law or condition of access. Any breach would only occur as a result of further actions by the user to confirm or assert identity, and these should be seen as a managerial problem.

Nevertheless, a consideration of spontaneous recognition does highlight some of the implicit assumptions make in data access decisions. It also shows the importance of the data owner's institutional culture: for a default-open data owner, spontaneous recognition is a useful check on whether all relevant risks have been addressed, but for a default-closed data owner spontaneous recognition provides a way to place insurmountable barriers in front of those wanting to increase data access.

1. Introduction

Social scientists increasingly expect to have access to detailed source microdata. The twenty-first century has seen major advances in the availability of detailed social science microdata for research purposes. Two elements have combined to make much more of the data collected by national statistics institutes (NSIs) available to researchers:

- secure, remote access to detailed data with few limitations on researcher use
- external pressure on NSIs to make data available

These have driven massive expansion in the use of NSI data sources; and despite the potential of 'big data', NSI and other government data remains the most important for much research, particularly in economics.

However, this growth in data access has not always been actively driven by the data owners. As Ritchie (2016) notes, data owners, particularly in government, are often reluctant to release data. This arises from institutional cultures and incentive structures which encourage a risk-averse attitude to decision-making (Ritchie, 2014b). Decision-makers are supported in this attitude by the academic literature; this overwhelmingly focuses on hypotheticals and the risk to the data owner, rather than any gain to the researcher (Hafner et al, 2015a).

The issue of 'spontaneous recognition' (SR) illustrates the difficulties facing those trying to improve access. Spontaneous recognition occurs when a microdata user believes that he or she has identified, without trying, one of the data subjects in the dataset: a neighbour, a co-worker, an organisation, or a group of patients, for example. The identification need not even be correct: the perceived breach of confidentiality can be as important as an actual breach. This concerns data providers wanting to allow researchers to use confidential data: no matter how trustworthy researchers are, they are still human and the recognition of an individual might lead to the disclosure of information about an identified data subject. Hence, data providers often insist on minimising the risk of spontaneous recognition.

Despite this, very little attention is paid to the topic in the literature. Almost every manual on statistical disclosure control (SDC) or data owner's guide to data handling mentions it, but largely as a pedagogical device before moving on to more sophisticated models.

Nevertheless, it is an important topic. Spontaneous recognition creates an additional hurdle to be addressed by those requesting access to data. It is not possible, in general, to show that there is no risk of spontaneous recognition, and so this gives great scope to those unwilling to release data for re-use scientific researchers or the public. This is particularly true in the case of business data, where the 'obvious' identifiability of businesses by one or two characteristics such as size and industry has been used in the past to restrict research access to the data.

This paper argues that the hurdle is an irrelevant distraction: spontaneous recognition has little or no practical contribution to make in the question of whether or not a dataset should be released. It is unlikely that there is any lawful or ethical basis for the concept; other SDC methods cover the reasonable requirements of law and access conditions. If any re-identification did occur spontaneously, it is the actions of the data user which should be governed, not the recognition itself which is an unconscious human act. These are better governed by management procedures.

The structure of the paper is as follows. The next section describes the fleeting appearance of spontaneous recognition in the literature. Section three considers whether spontaneous recognition is a useful statistical concept, in terms of building risk models. Section four reflects on whether there is any legal requirement to address spontaneous recognition. Section five shows how the real problem arising from spontaneous recognition comes from the actions of the user, not from the recognition itself. Section six considers the institutional impact of the concept, and shows how it can be helpful or obstructive depending on the organisation's attitude. Section seven concludes.

2. SR in the literature

The OECD Glossary of Statistical Terms defines 'spontaneous recognition' as

> *… the recognition of an individual within the dataset. This may occur by accident or because a data intruder is searching for a particular individual. This is more likely to be successful if the individual has a rare combination of characteristics which is known to the intruder. (OECD, 2005)*

This glossary was defined and adopted to encourage the consistent use of statistical language across countries, academics and NSIs. However, the definition given above is not widely used, largely because it includes deliberate searching (and therefore includes 'intruder' models).

Duncan et al's (2011) definition is closer to the more commonly understood definition:

> *You know of person X who has an unusual combination of attribute values. You are working on a data set and observe that a record within that data set also has those same attribute values. You infer that the record must be that of person X. In order to be truly spontaneous you must have no intent to identify. Otherwise this is just a specific form of deliberate linkage. (Duncan et al, p35)*

Duncan et al (2011, p29) explicitly distinguish spontaneous recognition from 'snooping' or 'intruder' models where the user takes actions specifically to identify a data subject. We use the more common Duncan et al definition here.

In one sense, there is a large literature discussing spontaneous recognition: most of the general statistical works on SDC (eg Hundepool et al, 2012) cover it, as do the guidelines produced by NSIs. It is often used to give examples of extreme values; for example, in GSS(2014):

> *An intruder may spontaneously recognise an individual in the microdata by means of published information. This can occur for instance when a respondent has unusual characteristics and is either an acquaintance or a well-known public figure such as a politician, an entertainer or a very successful business person. An example is the "Rich List" which publishes annual salaries of high-earning individuals. (GSS, 2014, p18)*

Spontaneous recognition is then typically used to explain why extreme values or population uniques are problematic and may need to be removed from the data.

However, this is the full extent of the discussion in the SDC literature. As Hafner et al (2015a) have noted, almost all of the confidentiality literature is focused on deliberate attempts to re-identify the data, the so-called ‘intruder model’. In this context, spontaneous recognition disappears as an uninteresting special case.

In summary, in the confidentiality literature, spontaneous recognition is a useful teaching tool but otherwise not considered.

3. SR as a statistical problem

3.1. Defining spontaneous recognition

Following Duncan et al (2011) we define spontaneous recognition as the “accidental identification of a data subject (that is, without actively searching), whether that identification is accurate or not.”

The second condition differs from most of the SDC literature, which assumes that identification is only a problem is the true identity is uncovered([138]). This is clearly consistent with legal requirements to keep data confidential, but it ignores the institutional impact. Asserting that confidentiality can be breached can have a substantial effect on the reputation of the data owner, whether that assertion is true or not: one NSI had to undertake a substantial public relations operation after it was (falsely) claimed that a multinational supermarket used confidential Census data to send out mailshots.

An additional complication is what ‘identity’ is being uncovered. Some data subjects, such as organisations, can have complex structures which makes identification a much more difficult concept. For example, assume that there is one university on the Isle of Wight, in the southern UK([139]), and that a researcher using business data from the island comes across a single entity whose industrial classification describes it as ‘university’. There are three possibilities.

- This is the whole university
- This is part of the university
- This is a branch office of a mainland university; the reporting units of the Isle of Wight university are not classified as ‘university’ for some reason

Clearly with complex data subjects the ‘accuracy’ of the identification has more room for interpretation. In line with the above definition, we treat spontaneous recognition as occurring when the data users thinks ‘I have found a data subject that I can put a name to’, irrespective of whether that name relates to an accurately identified unit or not.

A related concept is ‘identity confirmation’: a user spontaneously recognising a data subject takes active steps to confirm his or her suspicions. For example, the researcher could cross-check with other information in the dataset, or external information. This differs from the intruder model in that the researcher has no specific interest in attacking the dataset; the researcher’s curiosity has been aroused, and the purpose of further investigation is to satisfy that curiosity.

Finally, ‘identity assertion’ is where a user who spontaneously recognises a data subject reports his or her suspicions to another, again without deliberate intent to reveal information but as a human response to an interesting finding.

Because these two concepts both require action by the user after the initial suspicion has been aroused, we combine them as ‘identity confirmation or assertion’.

([138]) This is done for pedagogical purposes. The literature does of course recognise that false inferences may occur. However, the implications of this are rarely discussed; only accurate inference is considered.
([139]) This example is for illustration only. We are not aware of any university on the Isle of Wight at present.

3.2. Spontaneous recognition as a theoretical risk

Two obvious risks are widely described in the literature:

- *Population uniques on one or two characteristics or extreme values.*

For example, in 2014 the first female bishop in the UK Anglican church was elected. This was a high profile event, with wide newspaper coverage. Female clerics in the UK are comparatively rare, and a detailed job description or a salary range (indicating the highest paid) combined with gender could prompt a memory in the data user. Alternatively, a small geographical area might have one well-known high-value celebrity resident. Salary or wealth data may be enough to identify that individual. For business data, this is the most significant problem. Detailed industrial classification and size of business (typically employment or turnover) are assumed to be enough to identify well-known large players, such as in telecoms or aerospace.

- *Sample uniques where the sample is known.*

Sample uniques which are not population uniques on the key characteristics are not normally of concern; by definition they represent at least two indistinguishable data subjects. However, we can consider cases where the data user might have additional information about the sample, making the sample uniques into population uniques. For example, a neighbour learning that a research is using a particular data set may tell the researcher that she was included in a particular wave.

A third risk, less commonly documented, is that an unsophisticated user may mistake a sample unique for a population unique, and draw an inappropriate inference.

To avoid these risks, SDC good practice normally requires that population uniques are disguised or removed. For example, Statistics New Zealand's old Confidentiality Protocol explicitly identified spontaneous recognition with population uniques (Statistics NZ, 2000, appendix B). This is why business data is commonly described as being impossible to anonymise: the variables of interest (industrial classification, size) are essential components in research, and so cannot be removed while maintaining value in the data.

Sample uniques are also avoided where the underlying number of population uniques is small; for example, Schulte-Nordholt (2013) describe Dutch public use Census files as having a minimum 1000 observations on trivariate categorisation, and minimum five observations on all household characteristics.

This practice is uncontroversial, and allows SDC advisors to concentrate on the seemingly more important problem of active, intruder, attacks on confidentiality. It is assumed that re-identiification through deliberate action must have a success rate which is no lower than that of accidental discovery. Spontaneous recognition is intruder matching without a match model; therefore it can be treated as the less interesting special case. Mackey (2013), for example, treats spontaneous recognition as the starting point for a formal intruder-based review.

Nevertheless, examining spontaneous recognition in its own right can throw light on some of the underlying assumptions.

First, intruder models are designed to give risk measures based on assumptions about behaviour. However, the risk of spontaneous recognition cannot be estimated and cannot be proved not to exist, because the personal information that leads to it is unknowable. A dataset will contain population uniques unless it is K-anonymous on all variables (that is, any combination of variables, including continuous variables, must give at least K duplicate observations); for spontaneous recognition, K=2 as no active searching is considered. A fully K-anonymous dataset has limited research value, and so in practice removal of uniques is carried out using a subset of variables which is determined subjectively (Skinner, 2012). Therefore spontaneous recognition based on an unpredicted combination of variables must be possible. For example, wages are not normally considered identifying except for extreme values; but a researcher looking at a dataset on employees might notice a promotion from the wage data, and link that to personal knowledge of a specific employee.

Second, protection measures designed to stop intruders may not be relevant for identity confirmation. By construction, spontaneous recognition arises from a combination of knowledge not foreseen in the protection algorithm. It follows that cross-checking of information in the dataset to confirm a suspicion does not need to use the scenario used to create the protection algorithm.

Third, it is not possible to test for spontaneous recognition. By construction, spontaneous recognition arises from the accidental linkage of personal information with specific data. Testing by asking users to try re-identifying data subjects (as for example in Spicer et al, 2013) cannot formally replicate the conditions for accidental re-identification, although it would provide useful evidence as to whether the data protection is 'good enough'.

Fourth, spontaneous recognition is implicitly accepted in research files. For public use files (PUFs; those with no restrictions on use), stringent precautions are taken against direct attack, and hence spontaneous recognition. For scientific use files (SUFs; access limited to verified researchers) and secure use files (SecUFs; access limited to controlled environments), precautions are more complex as more controls are available. For example, Spicer et al (2013) discuss how the access restrictions on SUFs enable a more relaxed approach to intruder attacks. However, this implies a simultaneous increase in the probability of spontaneous recognition. In other words, for anything other than PUFs, a non-negligible level of spontaneous recognition is implicitly accepted; see Table 1.

Table 1: Subjective expectation of SDC controls

File type	SDC controls	Acceptable intruder risk	Acceptable spontaneous recognition risk
PUF	Many	None	None
SUF	Fewer than PUF	More than PUF	More than PUF
SecUF	Fewer than SUF	More than SUF	More than SUF

Finally, the likelihood of inaccurate spontaneous recognition is not covered in the SDC literature for the simple reason that there is no meaningful way to address the problem. What is the probability that a user looking at a dataset will make assertions based on an inaccurate identification? There is good empirical evidence that humans are over-confident in their ability to re-identify data subjects, but this evidence is based on tests under known conditions where test subjects are required to express their confidence in their predictions. It is not clear how one would test whether this applies in non-test conditions where data users are under no pressure to express an opinion.

3.3. Practical implications of the statistical approach

In summary, spontaneous recognition and its consequences are unpredictable, untestable and unprovable. This means that protection based around the notion of the predictable intruder might be ineffective.

In practice, spontaneous recognition in PUFs is ignored. The evidence of a half century of anonymisation suggests that focusing on intruders seems to provide adequate protection. Inaccurate assertions about individuals do not seem problematic.

In files created for researchers (SUFs and SecUFs), a non-negligible risk of spontaneous recognition is implicitly accepted. However, in terms of behaviour, spontaneous recognition is the exact opposite of active, intruder, re-identification. To many data owners this suggests that it should be treated as a separate problem and tackled independently.

4. SR as a legal problem

If spontaneous recognition occurs, it is not clear that any breach of confidentiality has occurred.

In PUFs, spontaneous recognition would imply that the anonymisation procedure has failed. While the data owners would be expected to review the anonymisation procedures, it is unlikely that this would lead to legal consequences. Most data protection laws (for example, the regulation covering data management in the EU) require data owners to take <u>all reasonable</u> protection measures, not <u>all</u> measures; some laws explicitly absolve the data owner of any legal responsibility in the case of a mistake. It would be difficult to argue that an intruder-protected PUF is inadequately protected against spontaneous recognition (assuming, of course, that the intruder protection is carried out to an accepted standard).

For researcher files, a non-negligible risk of spontaneous recognition is implicitly if not explicitly approved, as noted above. Are there consequences if a researcher recognises a data subject? No; the researcher has been granted lawful access to the data in that state, and nothing has changed.

What happens next does matter. The researcher has four options:

8. Identity confirmation: cross-checking the data with any other information
9. Identity assertion: mentioning the fact to another researcher or a non-user
10. Identity assertion: mentioning the fact to the data owner
11. Taking no further action

Action (1) may or may not be a breach of confidentiality, but it is almost certainly a breach of the access terms to the data, as the researcher is now trying to actively re-identify a data subject (in many jurisdictions, this would also be a breach of the law).

Action (2) is also likely to be a breach of access terms: mentioning something discovered about a data subject could be taken as seeking confirmation of the identity of the data subject, seeking to provide another with identifying information, or both. It is not clear whether an offence has been committed if the identification is inaccurate, but most data access agreements ban any information being shared about data subjects, whether that information is true or not.

The consequences of action (3) depend on the attitudes of the data owner. Data owners following the 'active researcher management' principles (Desai and Ritchie, 2010) should welcome information about easily recognisable subjects as an opportunity to review protection measures in the light of new information. However, the authors have observed data facilities where any speculation about the identity of data subjects, even to the data owners and irrespective of intent, is strictly forbidden and liable to penalties.

Some data owners require users to report any suspected identification, and so action (4) may be a breach of access conditions. However, it is not clear how a data owner could prove that spontaneous recognition has happened, unless one of actions (1) to (3) was also taken.

In summary, spontaneous recognition by itself does not seem a breach of confidentiality on behalf of the data user. For PUFs, the fault lies with the creator of the file. For research files, any breach of law or access agreements arises from additional actions taken by the researcher. In other words, the problem arises from the actions of the users, and not from a failure of statistical protection.

5. SR as a management problem

A non-negligible possibility of spontaneous recognition is implicitly accepted in research data files, and poses no legal problems. Breaches of confidentiality or procedures occur when the data user takes some follow-up action: identity confirmation or assertion. This clearly identifies the risk associated with spontaneous recognition as a user management problem.

This perspective offers several advantages over seeing a statistical or legal problem.

First, it focuses on the unlawful activity: searching for identity, or speculating on identity with another. It does not criminalise users for an automatic response (recognition) to some information presented to them. It penalises behaviour, not thoughts.

Second, it is likely to be easier to detect actions to confirm or assert an identity, whereas detecting whether someone has identified a data subject is impossible to know until they share that knowledge.

Third, it requires no assumptions to be made about what personal knowledge a data user might have that could lead to spontaneous recognition. It is only the outcomes that matter, not the inputs.

Fourth, it reduces incentives to damage data as protection against something which is explicitly acceptable, at least in research files.

Fifth, the protection measures are already covered to some degree. Providers of research files usually give users training or written guidelines, or both, which state that attempts to re-identify data subjects, or discuss characteristics of the data with unauthorised users, is prohibited. Data access agreements may have similar wording, although there is little evidence to suggest that users read these.

Six, the management approach can be applied to PUFs as well as research files. Thus, recognition of an individual in a PUF is no longer a failure of statistical technique, but the manifestation of a known and accepted managerial risk. The change in emphasis, from blaming individuals to corporate learning, should discourage SDC advisors from worst-case risk avoidance strategies to protect themselves.

Finally, as this is a management issue, and not a legal one, then the data owner can choose to promote positive behaviours. For example, reporting of suspected identification without penalty can be encouraged.

Best practice user training and communication encourages the development of a community of interest between researchers and the data owners (Desai and Ritchie, 2010; Eurostat, 2016). Training about spontaneous recognition versus identity confirmation or assertion can be used to reinforce messages of trust in the training. The unavoidability of human failings can be contrasted with working with the support team to ensure that no-one gets into trouble. Messages about inaccurate identification can also be pushed in training. The fictional example of the Isle of Wight university, given above, could be used to emphasise the scope for error in any assertions.

In summary, viewing spontaneous recognition as an irrelevance, and seeing identity confirmation or assertion as a problem of user management,

- focuses attention on unlawful activity only
- allows data owners more flexibility to deal with problems, even for PUFs
- encourages the community of interest amongst data owners and users
- is entirely consistent with best user training practice

6. Culture, attitudes and default perspectives

Although it may have no statistical value, the concept of spontaneous recognition can have a practical impact because of the data owner's cultural perspective and resulting attitudes. These reflect the way that decisions are taken.

The institutional culture of data owners can be simplified to one of the following (Hafner et al, 2015a):

- Default-open: release data unless the release is shown to be unsafe

- Default-closed: do not release data unless the release is shown to be safe

In theory these two positions are identical, but Ritchie (2014a) shows that the phrasing generates very different outcomes. Most NSIs notionally claim to be default-open; in practice, almost every organisation is default-closed. Ritchie (2014a, 2016) and Hafner et al (2015a) note that the default-closed culture arises from three sources

- institutional incentives focus on the probability of individuals making bad decisions, rather than the overall loss to society of good decisions foregone
- the statistical literature emphasises extreme scenarios, hypothetical examples, and worst-case risk avoidance, rather than evidence-based modelling, encouraging risk aversion in data owners who are not experts in assessing disclosure risk analyses
- the public choice literature which encourages data owners to view users as self-interested and so inherently untrustworthy

In these circumstances, it is not surprising that NSIs develop risk-averse cultures. In this context, statistical analysis is essential for allowing an NSI to claim an objective rationale for its decisions, despite these decision being highly subjective (Skinner, 2012).

For a default-closed data owner, spontaneous recognition offers an unbeatable hand. As noted above, spontaneous recognition arises from an unexpected combination of luck and unpredictable knowledge. It is not possible to demonstrate that this cannot happen, nor is there any evidence that it cannot happen (such evidence would have been incorporated into predictable risk). Hence, a data owner unwilling to release data can call spontaneous recognition as a risk without fear of losing the argument.

This can be done even though, for all practical purposes, the intruder model and management strategies make spontaneous recognition irrelevant. For a default-closed data owner, little expected practical impact may not be enough; it is the <u>potential</u> for spontaneous recognition to occur that must be demonstrably negligible.

For a default-open data owner, the reverse is true: spontaneous recognition can be a very useful check on the validity of one's risk scenarios. Considering spontaneous recognition encourages one to treat and eliminate the foreseeable risks; when the only remaining untreated risk is spontaneous recognition (ie entirely unpredictable risk), then the data owner can be satisfied that the release is now no longer 'shown to be unsafe' to any limit of reasonableness.

This issue raises important questions about the way data owners are persuaded to allow their data to be used. Data owners should be concerned about identity confirmation or assertion, but they may be unable to articulate it as the literature focuses on spontaneous recognition. Those advocating greater use therefore have a role to play in making data providers aware of the specific risk being raised.

The creation of 2010 Community Innovation Survey SUF (described in detail in Hafner et al, 2015b) illustrate the impact of the default-open/default-closed attitude. The creation of an SUF for business data is unusual, because of the assumed identifiability of the businesses. Prior to the 2010 survey, all continuous variables for all observations in the SUF were perturbed to reduce the attribution risk associated with a successful identification. However, a review taking a default-open approach argued that the only risk not handled by practical controls was the chance of careless assertions of identity. Accordingly, risk management focused on user training, and less than 1% of observations, the cases most likely to prompt speculation, were perturbed.

In short, for the default-open data owner spontaneous recognition offers a handy rule-of-thumb to determine whether there are any remaining untreated risks in a dataset; for a default-closed owner, it offers unlimited potential to place an unfeasible burden of proof on those wanting to release data.

7. Conclusion

The public good is best served by making data available for research with as little damage as possible. Data owners, facing institutional pressures which encourage them to place the needs of the organisation over the wider public good, raise concerns about both the deliberate re-identification of individuals (the intruder model) and accidental re-identification through spontaneous recognition.

The intruder model is the workhorse of statistical data protection; almost the entire literature and most practice is based on it. In contrast, spontaneous recognition has had negligible formal examination. It is used for pedagogical purposes, to demonstrate simple examples, before the intruder model takes over in formal modelling.

One reason for this is that spontaneous recognition is not easily amenable to formal modelling. By its nature, it arises from unpredictable knowledge. If that knowledge were predictable, then intruder models would be able to incorporate it. This leads to the second reason for ignoring spontaneous recognition: it is subsumed into the intruder model as a special case.

Despite this lack of theoretical or practical value, data owners do raise concerns about spontaneous recognition. The authors observe this most often in relation to business data, where it seems 'obvious' that any useful microdataset is going to include many pieces of information that would prompt a researcher to speculate on the identity of the business. Concerns about spontaneous recognition have in the past led to restrictions on research.

Hence the conclusion of this paper, that the concept of spontaneous recognition has no place in data protection, is important. This conclusion derives from two observations.

The first observation is that the intruder model, despite its flaws, does effectively encompass the spontaneous recognition problem: it focuses on predictable risks and allows for active searching, meaning that any remaining risk arises from luck and a complete unpredictable set of events. This is likely to meet the test of 'reasonableness' embodied in most data protection laws. Because the intruder model can be applied to different environments (PUFs, SUFs, SecUFs), it can incorporate spontaneous recognition in those different environments.

The second observation is that any re-identification from spontaneous recognition leads to a managerial problem. Breach of confidentiality arises from the follow-on actions of the user, the identity confirmation or assertion, not the spontaneous recognition itself. Hence any management plan must focus on the training of users and the relationship between users and data owners, not on predicting the unpredictable.

Nevertheless, an examination of spontaneous recognition can usefully highlight implicit assumptions being made in data release decisions; and it demonstrates the importance of the data owner's institutional culture. With the default-closed attitude, spontaneous recognition is an unplayable hand whose only value is to block access. With a default-open attitude, spontaneous recognition becomes a useful sounding board to explore the limits of our knowledge and develop non-statistical risk models to cover for the 'unknown unknowns'.

In summary, the statistical problem of spontaneous recognition is an unhelpful chimaera encouraging the underutilisation of valuable data. The problem that should be addressed is one of identity confirmation, which is management issue. A change in both language and attitudes, a focus on the exact nature of the problem being raised, and the use of evidence can generate substantial dividends for both data providers and users.

References

Desai, T. and Ritchie, F., 'Effective researcher management', *Work session on statistical data confidentiality 2009*; Eurostat.

Desai T., Ritchie F., and Welpton, R., 'The Five Safes: designing data access for research', Working papers in Economics, No. 1601, University of the West of England, Bristol, January 2016.

Duncan, G T, Elliot, M, Salazar-Gonzalez, J, 'Statistical Confidentiality', *Principles and Practice*, Springer: New York Dordrecht Heidelberg London, 2011.

Eurostat, *Self-study material for the users of Eurostat microdata sets*, 2016, http://ec.europa.eu/eurostat/web/microdata/overview/self-study-material-for-microdata-users

GSS, GSS/GSR Disclosure Control Guidance for Microdata Produced from Social Surveys. Office for National Statistics/Government Statistica,l Service, 2014 https://gss.civilservice.gov.uk/wp-content/uploads/2014/11/Guidance-for-microdata-produced-from-social-surveys.pdf .

Hafner H-P., Lenz R., Ritchie F., and Welpton R., 'Evidence-based, context-sensitive, user-centred, risk-managed SDC planning: designing data access solutions for scientific use', in UNECE/Eurostat Worksession on Statistical Data Confidentiality, Helsinki, 2015a.

Hafner H.-P., Ritchie F. and Lenz R., 'User-centred threat identification for anonymized microdata', *Working papers in Economics,* no. 1503, University of the West of England, Bristol. March 2015b.

Hundepool A, Domingo-Ferrer J, Franconi L, Giessing S, Schulte Nordholt E, Spicer K, de Wolf P., Statistical Disclosure Control. Wiley, 2012.

Mackey E., European Union Statistics on Income and Living Conditions (EU-SILC). case study. Mimeo, UK Anonymisation network, 2013, http://ukanon.net/wp-content/uploads/2015/09/EUROSTAT-EU-SILC-DATA-Nov-2013-pdf.pdf .

OECD, Glossary of statistical terms. November 2005,. https://stats.oecd.org/glossary.

Ritchie F., 'Access to sensitive data: satisfying objectives, not constraints', *J. Official Statistics,* v30:3, September 2014a,. DOI: 10.2478/jos-2014-0033, pp. 533-545.

Ritchie F.,'Resistance to change in government: risk, inertia and incentives', *Working papers in Economics* no. 1412, University of the West of England, Bristol, December 2014b.

Ritchie F., "Can a change in attitudes improve effective access to administrative data for research?", Working papers in economics no. 1607. University of the West of England, Bristol, 2016.

Schulte-Nordholt E, *Access to microdata in the Netherlands: from a cold war to co-operation projects,* Work session on statistical data confidentiality 2013; Eurostat, 2013.

Skinner C., *Statistical Disclosure Risk: Separating Potential and Harm*, Int. Stat. Rev. v80:38, 2012.

Spicer K., Tudor C. and Cornish G., *Intruder Testing: Demonstrating practical evidence of disclosure protection in 2011 UK Census,* Worksession on statistical data confidentiality 2013; Eurostat, 2013.

Statistics New Zealand, Confidentiality Protocol, 2002 http://unstats.un.org/unsd/dnss/print.aspx?docID=141.

PDF: KS-TC-17-006-EN-N
Print: KS-TC-17-006-EN-C

Selected papers from the 2016 Conference of European Statistics Stakeholders

The Conference of European Statistics Stakeholders (CESS) brings together European stakeholders such as official statisticians, methodologists, private data producers/providers and data users. The conference, held biennially, gives participants an opportunity to discuss user needs and exchange good practices in statistical production and outreach. It is also an opportunity to illustrate innovative ways of visualising and communicating statistics and promote new methodological ideas in data conception, collection and analysis.

This special Statistical Working Paper (SWP) issue presents selected papers from CESS 2016, held in Budapest on 20-21 October 2016. The papers highlight many of the current and future themes for ensuring good-quality statistics that are fit for use.

For more information
http://ec.europa.eu/eurostat/

Publications Office

PDF: ISBN 978-92-79-73628-5
Print: ISBN 978-92-79-73629-2