

From experimental statistics to official statistics: state of the art and prospects in Istat

Arianna Carciotto¹, Orietta Luzi¹

¹Istat – Italian National Statistical Institute, Italy

Abstract

The Italian National Statistical Institute (Istat) has been producing Experimental Statistics (ES) since 2018, following the initiatives carried out by Eurostat and other National Statistical Institutes (NSIs) in this field. These statistics are defined “experimental” as they use new data sources and/or new methods and/or new tools to meet users’ needs in a more effective and/or timely way. Often, ES are considered not fully “mature” since they need to be further refined and validated in terms of compliance with quality requirements. In fact, quality assessment is one of the main challenges with ES, especially when evaluating transition to official statistics. However, developing a quality framework for ES is a complex task as they are characterised by a wide heterogeneity. Relevance seems the quality dimension which deserves the greatest attention. In fact, users’ information need is one of the main driver to produce ES while their feedback is deemed as essential both to further improve experimental outputs and to shift them to official statistics. In any case, ES can always be seen as the result of research activities aimed at improving efficiency and boosting official statistics. Comparing some NSIs within the European Statistical System (ESS) that produce ES, neither a common definition nor shared rules to produce, manage and disseminate experimental products have been agreed upon. However, ES originate from the same need to innovate both products and processes, trying to meet stakeholders’ needs in a timely way. In the paper, some examples of ES particularly appreciated by external users, and others that have shifted to official statistics will be reported, along with an overall description of their strengths in terms of relevance and potential usefulness for stakeholders. Some general considerations, at the ESS level, about challenges on the production of ES, possible shared rules, standardisation and exchange of best practices, will also be outlined.

Keywords: experimental statistics, innovation, quality assessment, users’ feedback

1. Introduction

Defining precisely experimental statistics (ES hereafter) is a very demanding task. We can say that ES are based upon some key elements: innovation, the use of new data sources, new methods, new tools, the production of new analyses, production processes and/or outputs not fully “mature”.

Innovation is the main driver to develop ES given the massive availability of IT tools more and more sophisticated and high performing. As well, the recent experience of the increasing need for data and analyses of new phenomena in emergency conditions (Covid pandemic) have urged National Statistical Institutes (NSIs hereafter) to re-think their statistical production strategy. At the same time, the use of new data sources (in particular, Big Data) to enhance survey data is constantly growing. In a world of increasing survey costs, falling response rates and rising potential for survey biases, Big Data often offer a less expensive and timely way for producing a variety of statistics, than may either supplement or replace “traditional” survey data (Japiec, et al. 2015).

ES are mainly aimed at finding innovative solutions for both products and processes in order to produce data to properly and timely meet new and emerging users' needs. In fact, users play a crucial role for producing ES since they can define a new product relevance and its possible shifting to OS. Of course, ES are generally produced more quickly with respect to official statistics (OS hereafter), and can provide new information and analyses often complementing and enriching OS. In this way, ES can result in gains in terms of efficiency and can promote OS further development. On the other hand, experimental data may not be useful as expected and can be affected by quality problems. For example, ES based on the use of Big Data pose specific quality issues due to the use of data characterised by specific sources of error, access challenges and confidentiality concerns. The need to have a framework to assess ES quality is also fundamental for their transition to OS. To this aim, NSIs need to invest in the development of quality criteria, consistent with the OS requirements (Eurostat, 2007), and adequately adapted to the specific ES context. However, it is worth underlining that, even if not all the ES evolve in OS, their development always provides a valuable contribution to research and to the methodological, technological or thematic innovation of statistical products and processes.

This paper is structured as follows: Section 2 discusses general aspects related to the ES production in NSIs, including definition, main benefits and possible drawbacks. Section 3 focuses on Istat experience in developing ES. Section 4 is devoted to ES quality evaluation and the transition from experimental to OS in Istat. Section 5 contains general conclusions and suggestions for future work.

2. Experimental statistics in NSIs

In the last years, Eurostat has put an increasing attention to ES development, promoting their production in the NSIs. In fact, a section on its website is devoted to ES produced by Eurostat itself and to the European Statistical System (ESS hereafter) – Experimental Statistics Hub¹. This last collects links to the NSIs that produce ES. It is worth mentioning that not all the ESS NSIs produce ES since it is not mandatory. Moreover, the ES definition, production processes, quality measures, as well as procedures to shift them into OS are characterised by a great heterogeneity.

The ES importance is underlined by the forthcoming European legislation too. In fact, the amending Regulation 223/2009 on European statistics, not yet finalised, states that statistics can be developed, in accordance with users' needs and in specific areas, in the form of statistics under development or ES. Those statistics should be treated as European statistics and their release should be accompanied by information on their quality.

It is important to remind that many NSIs developing ES deem extremely important and valuable to collaborate with Academia and other research institutions as well as with private/public bodies owning data for other purposes (e.g. administrative data and Big Data).

¹ <https://ec.europa.eu/eurostat/web/experimental-statistics/overview/european-statistical-system>

2.1 An attempt to define experimental statistics

At the moment, there is neither a univocal definition for ES nor shared rules for producing and managing them at the ESS level. As already said, ES are mainly related to innovation: given the massive IT development, many NSIs are investing towards this direction. ES are broadly defined as those statistics that are produced using new data sources, new methods and new tools in order to investigate, through innovative products and/or processes, new phenomena in a more rapid way.

In any case, these outputs/processes are considered not fully “mature” and still under development. For this reason, users have to be carefully informed about the possible uses of these statistics: ES are usually published with supporting methodological notes clarifying the output “limits”. Moreover, this issue raises quality problems, which will be analysed in the following sections.

In some cases, ES are defined in terms of differences with regard to OS. Specifically, while the latter is mainly based on “direct” measures (data from surveys, in some cases administrative data), ES are produced through “indirect” measurements and carried out with innovative methods (multi-source statistical models, data integration tools, big data analysis, etc). An important aspect of OS is related to the “closeness” between what has to be measured and what is actually (directly) measured. This is particularly true when the information is extracted from accurately designed (possibly sample) surveys (Carciotto, 2023), allowing statisticians to tailor surveys exactly to the specific research question and to measure sampling and measurement error components. On the other hand, data based on traditional statistical surveys as well as on administrative sources often do not ensure adequate timeliness, while the availability of immediate, real-time digital data generally characterising ES could much faster respond to users' information needs.

2.2 Main benefits and critical issues

The main ES positive aspects is that they are produced more quickly than OS, they are able to provide information that may not be observable through traditional surveys, and they may support new and/or more granular analyses. Furthermore, they can generate gains in terms of efficiency and give an impulse to OS for a wider information supply.

Although the timeliness is an important strength characterising ES, the length of the process for publishing them may have a negative effect on timeliness itself. This is an issue NSIs have to cope with if they aim at increasing users' satisfaction, also considering the high importance users and stakeholders have in determining ES relevance.

Another important aspect to consider is that in many cases ES are based on the use of data sources held by private subjects. In this case, it is fundamental that the ES production does not rely solely on voluntary and occasional collaboration by data providers, but that a formalised model for data access is established, based on clear agreements between data holders and NSIs to ensure continuity of collaboration and stable accessibility to data. For example, there is a risk that data change beyond NSI's control - e.g. in their structure or in terms of frequency of data collection - with negative effects on validity, stability and quality of the estimates.

Furthermore, privately held data should be accessed for statistical purposes in a sustainable way, considering also the increasing attention paid to privacy issues and the recent legislation on data protection. The forthcoming European statistical law is expected to ensure an easier and more sustainable access to privately held data, based on fair, clear and predictable rules.

As for costs, the development of ES may result in reduction of data collection costs and burden on respondents. However, producing ES often implies large investments with regard to human resources, IT tools, organisation, use of data sources held by private subjects, etc. How to quantify and how to distribute costs among NSIs and other participant subjects is still an open issue.

3. Experimental statistics at Istat

Istat has been producing ES since 2018. On the institutional website,² ES are classified according to 4 different categories:

- 1) *Non-standard classifications*: new experimental classifications derived from official taxonomies or from analysis and research activities.
- 2) *New indicators*: produced by integrating many official and non-official data sources.
- 3) *Interpretation frameworks and analyses* of complex phenomena obtained through the integration of official data sources.
- 4) Results of *experiments on Big Data*, characterised by the use of non-official data sources.

An internal procedure has been established to select the ES proposals that will be published in a dedicated section of the institutional website. The Director responsible for the product, who is in charge of a prior evaluation, can submit a proposal at any time of the year. The proposal has to be accompanied by supporting documents providing clear and complete information on the new output: the data used, the methodological and IT solutions adopted, any useful information to evaluate quality, its innovative aspects. The proposal is then submitted to the Istat Research Committee – which is in charge to monitor and address Istat research activity – that assesses it with the support of its scientific secretariat and experts identified *ad hoc*. The evaluation takes into account the new output soundness, innovative aspects and relevance. An important aspect considered is the ES sustainability that is if it can be regularly produced and updated in terms of workload, data availability and organisational problems. The Istat legal office is involved as well to assess the compliance with privacy issues. The proposal can be rejected or amended based on the experts' recommendations. In case of positive evaluation, the proposal is then submitted to the Board of Directors for the final decision. If approved, it is published according to defined graphic standards to distinguish the experimental output form OS. The whole process is coordinated and monitored by a senior officer who has to make it as smooth as possible, considering the relations among the actors involved in the evaluation process and the importance of timeliness.

² <https://www.istat.it/en/experimental-statistics>

Since the first release in 2018, Istat has published 20 ES, with the participation of 75 colleagues working on them. Some ES are updated whenever more recent data are available and when it is sustainable considering workload. Some others are updated regularly, quarterly or yearly. Finally, 4 ES have been shifted to OS, having reached a proper level of “maturity” and being of interest for users. In fact, as already mentioned, the relevance of an experimental output is strictly related to users’ satisfaction, as discussed in next section. In order to evaluate which experimental products are mostly of interest for users, the accesses to the website section devoted to ES have been analysed. The main users’ interest is focused on the following ES:

- *Daily population for study and work reasons*, an experimental approach using administrative data;
- *Use of the Open Street Map to calculate indicators for road accidents*, a calculation of new indicators on road accidents;
- *Social Mood on Economy Index*, a daily measure of the Italian sentiment on the economy based on X data;
- *Municipal demographic projections*, aimed at studying the probable future of a population in terms of total size and structural components;
- *Municipality-tailored indicators*, a selection of indicators at Municipality, Province and Region level from a multi-source system that uses experimental and other more consolidated sources.

Analysing comments provided by users, their feedback mainly focuses on generic data requests, in particular an updated version of the experimental output. In other cases, requests concern the information systems functionality or clarifications on the adopted methodology. Sometimes specialised users provide some suggestions for ES further development or ask for collaboration.

Finally, considering the report on user satisfaction about Istat products, concerning ES in general, users declare to be not fully satisfied mainly with reference to timeliness and comparability, while they are satisfied enough in terms of accuracy and reliability. It is worth mentioning that, from the analysis of accesses to the ES website, it is evident that the users’ level of interest is strictly related to the specific ES theme.

4. Experimental Statistics quality at Istat: from Experimental to Official Statistics

As already mentioned, statistics coming from experimentations could not exactly fit the pre-defined statistical objectives, in other words, they may be affected by quality problems, such as accuracy, reliability and coherence. Besides information on the used data and on the methodologies and technical solutions adopted for producing ES, NSIs should accompany the release of experimental outputs with quality reports to make users aware of the strengths and limitations in using experimental figures. Especially in case of ES based on new data sources, such as Big Data, an open research issue is defining an appropriate “quality framework” for each type of source (social data, images, text data, sensor data, etc.). Such framework has to take into account the peculiarities

of both data generating process and data production process, identifying the different sources of errors and the statistical measures quantifying them.

Currently, at Istat the proposal for a new ES should include the level of fulfilment of quality dimensions and possible quality indicators to be monitored over time. In order to develop a standard framework for evaluating ES quality, a research activity is ongoing, focusing, for the time being, on ES based on the use of Big Data, as described in next session.

4.1 Quality assessment for Experimental Statistics based on Big Data: the case of Trusted Smart Statistics

The term Big Data includes data characterised by their large volume, their variety and velocity, the way how they are created, and the new types of processes needed to treat them and make inference from them. It is well known that, in order to produce the so-called Trusted Smart Statistics (TSS hereafter), i.e. statistics based on Big Data, these sources are to be used responsibly, and their limitations are to be understood and accounted for. At present, assessing TSS quality is a challenging issue at both National and ESS level. At Istat, a research project is ongoing to address this issue in a structured and systematic way. The main objective is to define a methodological framework that allows the measurement and documentation of quality for statistics based on the use of new data sources (Big and other non-probabilistic data).

The first step for developing the framework was identifying the scope of the framework itself. Thus, the research team defined a template to collect structured information on the different TSS typologies, exploiting the research team members' work experiences. Based on the information gathered with the template, together with a review of specialised literature and the results of international projects on quality of statistics based on new data sources (e.g. Daas, 2023; ESSnet Big Data II, 2020; Multi-MNO project³), three broad categories of new data sources were identified: sensor data, social media, electronic transactions, each one with specific sub-categories. Smart Surveys are then a further category with largely different characteristics.

The analysis of the template's information is now supporting the identification of the main quality issues for the different categories, finding commonalities and main differences among them.

It is worth noting that some common issues arose at Institutional level as requirements for the usability of the data source. These aspects, like the mandate for accessing such data, cooperation with private data holders, issues connected to confidentiality, etc. will be mentioned in Istat quality framework. This approach will comply with the existing legislation and the principles and indicators of the ESS Code of Practice (CoP, Eurostat, 2017), focusing on statistical processes and outputs.

Another common trait of the production process of many TSS is that, differently from traditional surveys and also from administrative data, not only data generation but also part of data processing (Ricciato et al., 2020) is not managed by Statistical Authorities but externally (e.g. by the private data

³ <https://cros.ec.europa.eu/multi-mno-project> , <https://cros.ec.europa.eu/mno-minds>

holder). This poses further challenges to NSIs that cannot monitor directly the quality of these complex processes (Ricciato, 2022).

The need for transforming Big Data into statistical information is the main methodological common feature as these data are generated for different purposes. The process of transforming/processing Big Data can lead to the creation of proxy variables other than the actual statistical concepts but also to the introduction of new errors due to the use of innovative methods and techniques. The statistical results obtained may be affected by coverage and measurement errors, and efforts should be made to define statistical methods to compensate for these errors. Based on preliminary analyses, the research team is working to develop a methodological framework, e.g. for selecting web data sources as well as processing quality and performance measures for the applied methods.

4.2 From Experimental to Official Statistics

As already mentioned, in the context of OS, the principles of the ESS CoP require an assessment of all the product quality components like relevance, accuracy (sampling and non-sampling errors), timeliness and punctuality, accessibility, comparability and coherence. The CoP also requires systematic assessments of all the production processes phases. For this reason, as mentioned before, shifting from ES to OS requires putting in place a number of enabling conditions and a transition process based on solid elements (methodological, organisational, legal, infrastructural). The objective is ensuring data generation process transparency, production process sustainability (in terms of costs, secured data access, continuity of the collaboration with data owners, etc.), statistical outputs efficiency and quality. In particular, adopting transparent and reproducible methodologies would ensure that the transition from ES to OS complies with ESS quality standards and principles. As an example, in case of shifting an experimental TSS to OS, the output must properly reflect the object of observation, strengths and weaknesses in terms of methodology, and quality must be transparently and properly described and communicated to users. Therefore, open, transparent and reproducible methods for data treatment and estimation are to be ensured, and appropriate measures for assessing estimates accuracy, reliability and coherence are to be adopted. A common effort at ESS level would allow the development of a shared solution to the problem.

At Istat, the shifting procedure is similar to the one followed to publish ES. The request for shifting has to be accompanied by supporting documents stating that the output meets some fixed criteria, such as improving methodologies, estimates' stability in time, stabilising and securing the production process, updating and improving data sources, spatial coverage, relevance and timeliness, positive feedback from users, coherence with the official statistics framework. Of course, these criteria imply the need to evaluate the risk that the new statistics might have an impact in terms of competitiveness or overlapping with regard to official ones already published (Carciotto, Signore, 2021).

Until now, at Istat 4 experimental products have been shifted to official ones:

- *Classification of Municipalities based on Italian ecoregions*, a new classification according to the ecoregions based on homogeneity with respect to climatic, biogeographical, physiographic and hydrographic factors.
- *Integrated economic and environmental accounts for tourism*, statistics on environmental pressures related to tourism industries, obtained by integrating the Tourism and the Environmental Satellite Accounts.
- *Integration between data from register and sample surveys: enterprises classified by use of ICT and economic indicators*, new indicators obtained by integrating register data on structural business statistics with sample survey data.
- *Businesses behaviours and sustainable development*, new indicators to analyse business characteristics in terms of environmental and social sustainability.

The decision of shifting these ES to OS has been based on different considerations: i) relevance of the experimental outputs in terms of ability to either complement current OS or to allow more in-depth analyses; ii) interest from either “wide-range” or “specialised” users; iii) ES ability to provide information on new/emerging phenomena of public interest.

5. Conclusions and future work

The production of ES gives NSIs the opportunity of exploring new areas for innovating traditional processes, increasing their efficiency, and providing users with timely and/or new statistical products. However, the regular production of ES implies some new issues for NSIs. The first one is the trade-off between the possible reduction of data collection costs and respondents’ burden, and the required investments in terms of (human and IT) resources and development of new methodologies. In order to strengthen technological and methodological innovations, collaboration with research institutions and Academia should be promoted at both National and ESS level, taking into account the legal issue on privacy.

Another important problem relates to the need of assessing quality of these new outputs and/or production processes. For example, in the context of ES based on the use of new data sources, Istat is carrying out a specific project which is expected to produce first valuable results within this year, and which will also provide useful elements for defining criteria to shift ES produced by using Big Data to OS. This implies the more general issue of defining a quality framework for ES, considering the European CoP requirements and the fact that, in most cases, ES are based on new, non-probabilistic and non-traditional data sources (e.g. Big Data). Relating to this last aspect, it worth mentioning that the forthcoming European statistical law is expected to ensure an easier access to privately held data, provided that requested data are strictly necessary for the production of European statistics as well as for statistics under development.

In any case, a common effort at ESS level is needed to develop a new and shared quality framework for ES, also taking into account key experiences gained within European research projects on the

matter, for example in the context of the use of new data sources for statistical purposes. Moreover, Eurostat could promote, possibly through financed projects, some strategies for adopting a minimum set of shared criteria to standardise both the production of ES (definitions, quality standards, production processes, domains, etc.) and the transition process from ES to OS. This could also represent a basis for promoting comparability among NSIs that produces ES on similar topics. Concerning users' perspective, the role of stakeholders is crucial, as their information need is one of the main driver to produce new ES, and their feedback is an important impulse to improve published products, also in the perspective of shifting them to OS. To this end, Istat regularly analyses the accesses to the published ES and users' feedback. However, it is worth investigating the possibility not only to exchange best practices within the ESS, but also to gather feedback on new experimental outputs from a dedicated ESS network, before their release to the public.

References

- Carciotto, A. (2023). Experimental Statistics: a New Common Language for Uncommon Data. *Rivista Italiana di Economia Demografia e Statistica*, Volume LXXVII – N. 4 October-December 2023. ISSN: 0035-6832
- Carciotto, A., Signore, M. (2021). Improving relevance: Istat experience on experimental statistics. *Statistical Journal of the IAOS*, Vol. 37, No.2, pp. 593-601, IOS Press. DOI 10.3233/SJI-200764
- Daas, P. (2023). Big data in official statistics. Technische Universiteit Eindhoven
- ESSnet Big data II (2020). Work package K Methodology and quality. Deliverable K3: Revised Version of the Quality Guidelines for the Acquisition and Usage of Big Data
- Eurostat (2017). European Statistics Code of Practice
<https://ec.europa.eu/eurostat/documents/4031688/8971242/KS-02-18-142-EN-N.pdf/e7f85f07-91db-4312-8118-f729c75878c7?t=1528447068000>
- Eurostat (2007). Handbook on Data Quality Assessment Methods and Tools
<https://unstats.un.org/unsd/dnss/docs-nqaf/Eurostat-HANDBOOK%20ON%20DATA%20QUALITY%20ASSESSMENT%20METHODS%20AND%20TOOLS%20%20I.pdf>
- Japac L., Kreuter F., Berg M., Biemer P., Decker P., Lampe C., Lane J., O'Neil C., Usher A. (2015). Big data in survey research: Aapor task force report. *Public Opinion Quarterly*, 79(4), 839-880. <https://doi.org/10.1093/poq/nfv039>
- Ricciato F, Wirthmann A, Hahn M. (2020). Trusted Smart Statistics: How new data will change official statistics. *Data & Policy*, Volume 2, e7. doi:10.1017/dap.2020.7
- Ricciato F (2022). A reflection on methodological sensitivity, quality and transparency in the processing of new "big" data sources, European Conference on quality in Official statistics Q2022, Vilnius Lithuania, June 2022