Integration between data from register and sample surveys: enterprises classified by use of ICT and economic indicators

Methodological Note

The Statistical information system for estimating structural economic variables on business accounts (Turnover, Value Added, …) FRAME SBS, released in 2012-2013, is based on the primary use of integrated administrative/fiscal data, combined with survey data. FRAME SBS is currently the pillar of the new system of business statistics in Italy. As further development, the use of new opportunities in economic analysis by integrating FRAME SBS with other sources of data stemming from sample surveys has been particularly fostered. Recently the Italian National Statistic institute (Istat) has been developing towards an integrated production system of SBS statistics. In this framework, the core of the information is represented by administrative sources while sample surveys are conducted in order to estimate only not directly available specific sub-population information.

With the purpose of defining a methodological approach for the production of indicators, a Task Force was established by involving variables stemming from both FRAME SBS and sample surveys such as ICT usage and e-Commerce in Enterprises survey (ICT).

ICT survey is under the European Regulations requiring estimates for given domains of the target population, i.e. enterprises with at least 10 persons employed and belonging to given NACE codes. The sampling design is one-stage stratified random sampling. The strata are defined by combining the economic activities (NACE classification), size classes (Number of persons employed) and regions (NUTS classification) by the domains of interest. Enterprises belonging to the same stratum had equal probabilities to be sampled. The sample size in each stratum is mainly defined according to the Bethel procedure (Bethel, 1989) as the minimum sample size ensuring that the coefficient of variation of estimates in predefined domains does not exceed a given threshold. Estimates are then calculated through calibration methodology (Särndal et al., 1992) to compensate non-response and to match known population totals (benchmarks) of selected auxiliary variables. The population totals are computed using the Italian Statistical Business Register (ASIA). Considering the reference period, the survey is conducted within the first 6 months of the year t, while the ASIA business Register data used for sampling and calibration is referred to the year t-2. Enterprises belonging to the sample that cannot be linked to FRAME SBS are considered as "non-response" units.

The objective of Task Force is to produce indicators integrating the information gathered by a sampling (ICT) and exhaustive (FRAME SBS) sources through methods qualifying the results in terms of comparability and consistency according Istat methodological standards.

Two different approaches have been considered: macro approach (balancing, Iterative proportional fitting) and micro approach (consistent repeated weighting, calibration estimators and statistical matching). The selected method guarantying data quality and processing is the calibration estimators, adopted also to produce estimations in the ICT sample survey. However, several differences in applying method are listed as follows:

- The statistical archives for estimations is FRAME SBS referred to year t-1, while in the ICT sample survey is ASIA, referred to the year t-2. The given population are identified by updated information in the reference year t-1;

- A share, small percentage, of enterprises interviewed by the survey are no longer part of the target population (most of the times because the number of persons employed is less than 10);

- The above mentioned units have been excluded from the sample: the data set used to produce the indicator estimations is based on matching between data set units for the ICT survey estimation and the units in the FRAME SBS;

- The estimation domains are redefined by matching the domains required from the survey (small details) in compliance with objective of the study (tables of indicators with certain characteristics) and the quality of the obtained results. Particularly, the territorial level has been reduced, excluding the regional information;

- With regards to the model used for the calibration of the weights in the ICT survey (totals for the variables Number of enterprises and Number of persons employed by NACE and geographical level), the use of FRAME SBS was crucial to consider within the known totals those related to Value Added, Turnover and Gross Operating Surplus.

The methodological framework applied is the same used for ICT survey. Hence, in order to estimate the accuracy and the reliability of produced estimation the same criteria can be adopted. With regards to the data processing, the software ReGenesees (Zardetto, 2015), which performs methods for business statistics produced in Istat, has been used. Furthermore, as the adopted strategy creates microdata file with estimations of weights, the estimations of new ICT indicators have been produced. These estimations are coherent with those already released and this output ensures the results consistency (the consistency with FRAME SBS is ensured by creating the new estimation domains). The objective of the study is to define several tables of indicators which combine information from FRAME SBS and ICT survey. Only with this purpose, the dataset with the elementary units (and with the calibration weights) can be used accordingly.

References

[1]   AA.VV. (2016), "Integrazione del Frame con altre indagini e fonti amministrative ai fini della produzione di indicatori complessi", *Istat Working Papers*, n. 17/2016.

[2]   AA.VV. (2014), "L'integrazione dei risultati delle indagini sulla tecnologia e l'innovazione nelle imprese: una sperimentazione", Rivista di statistica ufficiale, n.3/2014.

[3]   Accetturo A., Bassanetti A., Bugamelli M., Faiella I., Finaldi Russo P., Franco D., Giacomelli S., Omiccioli M. (2013), "Il sistema industriale italiano tra globalizzazione e crisi", *Questioni di Economia e Finanza*, Banca d'Italia.

[4]   Airaksinen A. (2004), "Impacts of ICT usage on business organisation and business processes" – Final report. *NESIS work package* 5.4.

[5]   Andrews, D., Criscuolo C., Gal P. (2016), "The Best versus the Rest: The Global Productivity Slowdown, Divergence across Firms and the Role of Public Policy", OECD Productivity Working Papers, 2016-05, OECD Publishing, Paris.
 https://www.oecd.org/eco/growth/Frontier-Firms-Technology-Diffusion-and-Public-Policy-Micro-Evidence-from-OECD-Countries.pdf.

[6]   Brynjolfsson E., McAfee A. (2011), *Race Against The Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy*, Digital Frontier Press.

[7]   Bugamelli M., Pagano P. (2004), "Barriers to investment in ICT", Applied Economics, 36(20), pp. 2275-2286.

[8]   Biagi F. (2013), "ICT and Productivity: A Review of the Literature", Institute for Prospective Technological Studies, *Digital Economy Working Paper* 2013/09, JRC. http://publications.jrc.ec.europa.eu/repository/bitstream/JRC84470/jrc84470%20final%2011111 3.pdf.

[9]   Bronzini R., Piselli P. (2016), "The impact of R&D subsidies on firm innovation", *Research policy*, Vol. 45 (2): pp. 442-457.

[10]   Cette G., Fernald J.G., Mojon B. (2016), "The Pre-Great Recession Slowdown in Productivity", Federal Reserve Bank of San Francisco Working Paper No. 2016-08. http://www.frbsf.org/economic-research/publications/working-papers/wp2016-08.pdf.

[11]   Costa S., De Santis S., Nurra N., Salamone S., Vicarelli C. (2018), Indifferenza, sensibilità, compiutezza: una mappa della propensione alla trasformazione digitale nel sistema produttivo italiano, AISRe, XXXIX Conferenza scientifica annuale Bolzano, 17-19 Settembre 2018.

[12]   Eurostat (2013), ESSnet on Linking of Microdata to Analyse ICT Impact, ESSLait project. https://ec.europa.eu/eurostat/cros/content/esslait_en.

[13]   Eurostat (2010-2012), ESSnet on Linking of Microdata on ICT usage, ESSLimit project. https://ec.europa.eu/eurostat/cros/content/esslimit-finished_en.

[14]   Fernald J.G. (2014). "Productivity and Potential Output Before, During, and After the Great Recession." *NBER Macroeconomics Annual*.

[15]   Gal P. et al. (2019), "Digitalisation and productivity: In search of the holy grail – Firm-level empirical evidence from EU countries", OECD Economics Department Working Papers, No. 1533, OECD Publishing, Paris, https://doi.org/10.1787/5080f4b6-en.

[16] Gordon R.J. (2014). "The Demise of U.S. Economic Growth: Restatement, Rebuttal, and Reflections." *NBER Working Papers 19895*, February.

[17] Istat (2018), Rapporto sulla competitività dei settori produttivi – Edizione 2018. https://www.istat.it/storage/settori-produttivi/2018/Capitolo-3.pdf.

[18] Istat (2017), Cittadini, imprese e ICT - Statistica Report, 21 dicembre 2017. https://www.istat.it/it/files//2018/06/a5_ICT2017_Testo-integrale-e-nota-metodologica.pdf.

[19] Mokyr, J., Vickers C., Ziebarth N.L. (2015), "The history of technological anxiety and the future of economic growth: Is this time different?", *Journal of Economic Perspectives*, Vol. 29 (3): 31-50.

[20] OECD (2019), Measuring the Digital Transformation: A Roadmap for the Future, OECD Publishing, Paris. https://doi.org/10.1787/9789264311992-en.

[21] Pellegrino B., Zingales L. (2017), "Diagnosing the Italian disease", *NBER Working Paper* n. w23964, ottobre.

[22] Spiezia V. (2012), "ICT investments and productivity: Measuring the contribution of ICTS to growth", *OECD Journal: Economic Studies*, Vol. 2012/1. https://www.oecd-ilibrary.org/economics/oecd-journal-economic-studies_19952856.

[23] Zardetto D. (2015), "ReGenesees: an Advanced R System for Calibration, Estimation and Sampling Error Assessment in Complex Sample Surveys", *Journal of Official Statistics*, 31(2):177-203.

Working Group:

■ Alessandro Faramondi, Alessandra Nurra, Giovanni Seri, Valeria Tomeo.