Editorial

Statistics on difficult to measure population groups: Leaving no-one not-included

The report '*Realizing the SDGs for All: Ensuring Inclusiveness and Equality for Every Person, Everywhere; Together 2030 written inputs to the UN High-Level Political Forum on Sustainable Development (HLPF) 2019*'[1] reminds in its introduction that leaving no one behind lies at the heart of the 2030 Agenda for Sustainable Development. The 2030 agenda makes a clear commitment to inclusiveness, and many of the SDG indicators aim to include data on the most vulnerable and marginalized groups, several indicators are explicitly focussing on these groups. The document however concludes that '..., *in spite of the frequent use and reference to this principle, focused efforts to leave no one behind remain insufficient, in terms of policy design, implementation and review*' (see 1, 2019, p. 1). A main reason for this is the lack of sufficient data, which again is a result of problems of measurement due to the accessibility of these groups for the collection of information or of hampering register and administrative systems. Examples of such difficult to measure groups are homeless and stateless people, refugees, IDP's, but also other groups that for a variety of cultural or political reasons are not included in registers (like the population or birth register) or surveys. Beyond national publications based on very specific and often a-typical national data, international comparable statistical information is missing for many of these marginal and vulnerable groups.

Inclusiveness concerns all types of population groups, in all countries and regions. This is not only an issue in developing countries and not only concerns the marginalized and vulnerable, but also groups that are

excluded by administrative or for measurement reasons.[2] Collecting the relevant information on these groups in the context of the SDG's and more in general for official statistics, calls for action from the member states, but also from civil society and the international organizations. The role for the international organizations, is, alongside their mandate as custodian of specific SDG's, to develop and maintain clear guidelines and standards for producing and disseminating harmonized data on these issues and guarantee the integrated approach of statistical information of these groups with all other groups and the total population. Such guidelines support, beyond the comparability, also the other quality aspects of statistics on these population groups.

It requires a special effort by the international statistical community to develop guidelines and start collecting comparable information on the difficult to measure groups. A number of actors under the umbrella of the Expert Group on Refugee and IDP Statistics (EGRIS) are collaborating to produce draft International Recommendations on Statelessness Statistics (IROSS).[3] Statelessness is a situation that causes severe disadvantages for the people involved; importantly the lack of access to the regular socio-economic and political rights. As a result, people without citizenship are also particularly vulnerable to severe forms of exploitation and abuse, such as human trafficking.

The starting point for preparing these guidelines is the fact that there are no definitive or reliable global

---

[1] Ensuring Inclusiveness and Equality for Every Person, Everywhere; Together 2030 written inputs to the UN High-Level Political Forum on Sustainable Development (HLPF) 2019': https://sustainabledevelopment.un.org/content/documents/23216Together_2030__Position_Paper__HLPF_2019.pdf.

[2] The manuscript in this issue by Avni Kastrati and Nico Keilman 'Culture, tradition, and the registration of deaths: The case of Kosovo', In: Statistical Journal of the IAOS, Vol 37 (2021). illustrates how culture can influence register information.

[3] The Expert Group on Refugee and IDP Statistics works under the auspices of the UN High Commissioner for Refugees (UNHCR).

estimates for these groups, only some but very different in quality and coverage national figures drawn from a variety of quantitative and qualitative sources from both official and non-official sources.

One of the strategic objectives of the IAOS is to give via official statistics people a voice in decisions that affect them; such statistics make the invisible visible. Official statistics are fundamental to democracy, helping society to leave no one behind. Promoting inclusiveness in official statistics and supporting developments that help to produce and disseminate statistics on difficult to measure marginalized and vulnerable groups is also a way for the Statistical Journal of the IAOS to support this strategic objective.

In this issue statelessness is the topic of the core manuscript and therefore also the theme for the 10th discussion on the SJIAOS discussion platform (www.officialstatistics.com). The manuscript summarizes the progress made to date by the EGRIS in preparing this new set of recommendations for consideration by the UN Statistical Commission (UNSC). It is expected that in the near future also progress on guidelines for the measurement of other difficult to measure groups can be reported in this journal.

It is all about high quality and especially timely data...

As argued in several manuscripts in the last four issues of this journal, the COVID-19 pandemic, the related economic downturn and the current fast recovery demands for superior data, new methodologies and procedures and agile statistical systems. The opening article in the Economist '*Data and the economy*' [4] argues about the need for good data and statistics and that, taking the example of Central Bank statistics,[5] (but of course this need for good data holds true for all kind of statistics) much of the required data still are not adequately available and the traditional statistical systems lag behind with survey and even with integrated data.

Presentations in recent conferences, like the ISI WSC[6] and the UN WDF,[7] showcased how national statistical offices and international organizations are –

following these demands – rapidly taking initiatives on procedures and methodologies. Many of these very needed and newly developed methods and applications use new types of data like big data, also instant data resulting from sensor technology. Such instant data are more and more required to respond to the demands from governments and the international business world for timely an often real time data. The culture of 'production on demand' appears like in the business world also in official statistics. Historical data and traditional statistics are still required but the dynamics in society do no longer allow time lags for using data for the forecast and management of the processes in society.

New methodologies also include new forms of integration of survey data with all types of other official and non official data as well as data generation technologies as Artificial Intelligence and Machine Learning. In this issue several manuscripts deal with such new data sources and methodologies. The characteristics of estimation and missing data in new and existing data sources are dealt with in three further manuscripts. In this issue the value added by combining different data sources as well as relevance of high quality guidelines is illustrated via an application for the Oslo region with Economic and Environmental Accounts.[8] Integration is one of the key words also in the context of accounting.

... in a world with irresistible challenges for young statisticians

In a recent seminar I reflected on the dynamics in official statistics during the last 40 years: the period that my generation (born in the fifties of last millennium) matured from young statistician to senior expert. Starting the career 40 years ago in a survey based, very local oriented national statistical office, known and respected by only a relatively small number of policy and business world stakeholders, we were dazzling about new survey and sampling techniques and methodologies (hand held computers were still only on the drawing table) without having the slightest idea how statistics would look like some 40 years later.

The period 1980–2020 has been a revolutionary period for official statistics, offering a fast changing, chal-

---

[4]The Economist, October 23rd-29th 2021.

[5]See also in this issue: New developments in central bank statistics around the world, by Bruno Tissot, Alfonso Rosolia and Silke Stapel-Weber; Statistical Journal of the IAOS, Vol 37 (2021).

[6]ISI World Statistics Conference, 10–16 July 2021, The Hague, The Netherlands.

[7]World Data Forum, 3–6 October, 2021, Bern, Switzerland.

[8]Per Arild Garnåsjordet, Margrete Steines, Sofie Cimburova, Megan Nowell, David Barton, Iulie Aslaksen: Urban Green. Integrating ecosystem extent and condition as a basis for ecosystem accounts. Examples from the Oslo region. Statistical Journal of the IAOS, Vol 37 (2021).

lenging and due to these developments a never boring working environment. Data sources, the production and dissemination methodologies and publication cultures have fundamentally changed; the organization structure of the statistical offices and moreover the growth of statistical systems and statistical ecosystems beyond the national statistical office has been enormous; the cooperation with academia and civil society has opened new venues for coordination and research and development; no longer national but much more international policies require statistical information. The agenda for statistics is no longer national or regional, but is dominated by global challenges and formal at UN level agreed obligations, for example to fulfill the SDG objectives. The domains that are covered by official statistics have consequently changed from pure national issues to global issues like digitization, globalization and climate change. The world for official statisticians has substantially changed to a data science environment, even when the norms and values for official statistics are relatively unchanged.

For the generation of the 1950's, entering into the world of official statistics as a young statistician was challenging. For the current generation of young statisticians, this step into a career is, due to the enormous growth in the importance of statistical information, the complexity, diversity in sources, methodologies and IT applications, as well as the more forefront position in society, of an unbeatable and irresistible nature. A nice illustration of the world of statistics that today's young statisticians steps in, is given in this issue in four manuscripts by the winners of the 2021 IAOS Young Statisticians Prize. The contributions show the variety of their projects, ranging from domain specific applications of new methodologies to the forefront position in developing accurate COVID-19 statistics as well as their high expertise in new techniques like Machine Learning and statistical integration methodologies. Elham Sirag and Gautier Gissler (Statistics Canada) won the first prize with '*Estimating excess mortality in Canada during the COVID-19 pandemic: Statistical methods adapted for rapid response in an evolving crisis*'. Kevin Kloos (Statistics Netherlands (CBS)): with '*A new generic method to improve machine learning applications in official statistics*' was awarded the second prize and the third prize was awarded to Caio Gonçalves (João Pinheiro Foundation) and Ms. Luna Hidalgo (Brazilian Institute of Geography and Statistics (IBGE)): with '*Model-based single-month unemployment rate estimates for the Brazilian Labour Force Survey*'. A special commendation for a paper from a

developing nation was awarded to Muhammad Fajar and Zelani Nurfalah (Badan Pusat Statistik – Statistics Indonesia): '*Hybrid Fourier Regression-Multilayer Perceptron Neural Network for Forecasting*'.

Next to the already mentioned topics this issue covers a lot of other themes, a first paper stemming for the recent ISI World Statistics Conference, a set of Covid related articles, some articles on population and social statistics, agricultural statistics and some remaining papers from the cancelled 2020 European Statistics Quality conference.

1. The manuscripts in this issue in some more detail

New developments in central bank statistics

Bruno Tissot, Alfonso Rosolia and Silke Stapel-Weber report in their manuscript '*New developments in central bank statistics around the world*' on the presentations and discussion of the ISI WSC session focussing on this theme. The session concluded that the need to broaden the ability of central bank statistics to face future shocks that can test the resilience of today's economies in unexpected ways as the key lesson from central banks' experience during the COVID-19 pandemic. To achieve this challenge higher-frequency, more granular and timelier indicators, leveraging on the growing availability of alternative data sources have to be developed. The new types of information caused by in particular, increased digitalization can complement and expand traditional analysis and statistical measurements. Yet, a key issue is that reaping the full the benefits of such new and alternative data sources can face several important challenges.

The impact of COVID-19

In '*Monitoring the Newly Infected Cases of COVID19 Data Weekly: A Survival Data Analysis (SDA) Perspective*' Ramachandran Ramasamy and Maniam Kaliannan describe their attempt to fit the best survival model distribution for the Malaysian COVID-19 new infections experience of Wave I/II and Wave III using t Survival Data Analysis (SDA) procedures. Based on their analysis of a set of distribution functions of the conditional probability of infection in a short time interval, they confirm the Weibull as the best statistical fit for the COVID-19 new infection data. They produce weekly estimates of scale and shape parameters relating the

COVID-19 new infections experiences by pandemic waves in Malaysia. Succinctly put, the SDA procedures have a distinct merit of reducing numerous frequency counts characterized by high fluctuations, skew and kurtosis into a single measure of scale and shape parameters, which in turn used in gauging and monitoring the virality trends by weekly and undertaking short-term forecasts.

The second COVID-19 Pandemic related manuscript in this issue is the earlier as a blog on www.officialstatistics.com published '*Robust official business statistics methodology during COVID-19-related and other economic downturns*' by Paul Smith and Boris Lorenc. They argue that official statistics has not properly researched and understood how its methods and models behave at times of downturns (and potentially in the corresponding situation of similarly paced (unpredictable and fast) growths. As shown by many analyst during the last 18 months the production of official statistics requested and caused radical changes in both data collection and statistical methods. In anticipating poor measurement in times of future downturns Smith and Lorenc discuss the issues of the robustness of statistical methods.

Statistics on difficult to measure population groups

The manuscript '*Improving official statistics on stateless people: challenges, solutions, and the road ahead*' by Melanie Khanna and Mary Strode, illustrates the very important work of the EGRIS working group and constitutes the base for the 10th discussion on the SJIAOS discussion platform and is also the front-cover topic for this issue. The manuscript reports on the important work of the Expert Group on Refugee and IDP Statistics (EGRIS) in producing draft International Recommendations on Statelessness Statistics (IROSS). One of the biggest challenges for describing and analyzing and consequently guiding policies on this issue is the lack of reliable statistics about statelessness. There is no definitive or reliable global estimate, national figures are scant in most regions of the world, and where they exist they are often of questionable quality. The figures are currently drawn from a variety of quantitative and qualitative sources from both official and non-official sources. This manuscript summarizes the progress made to date by the EGRIS in preparing this new set of recommendations for consideration by the UN Statistical Commission.

Population and housing census/social statistics

In '*Accuracy of French Census Population Estimates*' Gwennaël Solard, Lionel Espinasse, Vincent Le Palud, Julie Prévot, Lucile Vanotti, describe in detail the quality of the population estimates produced by the French rolling census. The French census is subject to numerous quality controls throughout the process: development of a housing register, preparation of the collection, the collection itself and the post collection, adjustment and estimation operations. They argue that the many checks carried out throughout the process guarantee that the estimates produced are of a high quality. They also show that, there are some weak points, for example even when many instructions are included in questionnaires, the answers given by enumerated persons are imperfect due to misunderstandings, an inability to adapt questions to real-life situations, or deliberately incorrect answers.

Avni Kastrati and Nico Keilman in their manuscript titled '*Culture, tradition, and the registration of deaths: The case of Kosovo*' show how the patriarchal culture in traditional parts in Kosovo, explains an unusually high share of male deaths (SMD) among all deaths. In this culture the so-called Kanun of Lekë Dukagjini sets behavioral rules implying a very low status for woman, for example a woman cannot own immovable property. Consequently, to register the death of a family member at the office for civil registration is less urgent for women than for men. They also mention other factors that could explain the high SMD: under-registration of deaths among Serbs in Kosovo, violent deaths and smoking among men, and bad physical and mental health among veterans of the war of 1999.

Mohamed El Vilaly, Maureen Jones, Mahamadou Tankari, Gil Make, Sabrina Juran aim in their manuscript '*Mauritania's Internal Migration Dynamics and Trends in Response to Rainfall Variability and Change*' to determine and examine internal migration flows to analyze the relationship between long-term rainfall changes and dynamic spatial demographic shifts in terms of movements toward urban centers. The study area, the northwest African country of Mauritania is a vast, desert territory, which was historically been dominated by pastoral nomads. A vast sedentarization movement since the 60ties of last century coupled with internal and inter-regional migration has resulted in the growth of Mauritania's urban population from less than 10 percent of the total population in 1965 to nearly 90 percent in 2013. The factors that have caused this rapid urbanization, include the droughts that spanned the late 1960s through

to the early 1980s. Combining the decennial census and rainfall data, with available socioeconomic variables, they demonstrate distinct interactions between climate variability and interregional migration in Mauritania throughout the past four decades.

The establishment of the Sustainable Development Goals in 2015 entails generating relevant and timely statistics for monitoring and policymaking. In this context the Philippine Statistics Authority (PSA) generates poverty statistics using the three-yearly Family Income and Expenditure Survey (FIES). Manuel Leonard Albis, Jesaa Lopez, Roxanne Jean Elumbre, Daniela Jann Galias in '*Estimating Official Poverty Statistics in Non-FIES Years*' present a method of filling in the gaps by interpolating annual poverty statistics on state level, particularly the poverty incidence, using macroeconomic indicators and demographic and employment information from the Labor Force Survey (LFS). Relatively high forecast accuracy was observed for the predicted values of poverty incidence.

Young Statisticians Prize 2021

The Young Statisticians Price (YSP) is a competition that promotes younger official statisticians to publish and present their work. The prize winners are awarded beyond a monetary award and participation to an international conference, the publication of their winning manuscript in the Statistical Journal. In this issue we present the winning manuscripts from the YSP 2021 competition.

The first prize was awarded to Elham Sirag and Gautier Gissler from Statistics Canada for their article '*Excess mortality in Canada during the COVID-19 pandemic: Statistical methods adapted for rapid response in an evolving crisis*'. In this paper they present the approach adopted at Statistics Canada to produce timely and accurate estimates of excess mortality during the ongoing COVID-19 pandemic. They describe the two models involved in the estimation of excess mortality: the model used to estimate the expected number of deaths in the absence of the pandemic (baseline mortality), and the model used to adjust provisional death counts for under-coverage. The manuscript concludes by presenting selected results from Statistics Canada's official release of excess mortality estimates from February 8th, 2021.

The second prize was awarded to Kevin Kloos (Statistics Netherlands) for his manuscript titled '*A new generic method to improve machine learning applications in official statistics*'. Mr. Kloos states presents a new generic method to correct misclassification bias for time series and its statistical properties, starting from the premises that the use of machine learning in official statistics always introduces a misclassification bias. The importance of this new method is numerically shown by its lower mean squared error than the existing alternatives in a wide variety of settings.

Caio Cesar Goncalves and Luna Hidalgo (IBGE Brazil) were awarded the third prize for their manuscript '*Model-based unemployment rate estimates for the Brazilian Labour Force Survey*'. Monthly unemployment rate estimates in Brazil are regularly produced based on a three-month average of direct estimates based on the Brazilian Labour Force Survey (BLFS). The COVID-19 pandemic and its effects in the economy and labour market, forced the IBGE to investigate model-based estimation procedures to obtain unemployment rate single-month estimates. They present structural time series models developed to produce model-based single month estimates at national level as well as small area (state-level) estimates at a higher frequency than those currently being published. The new improved model-based estimates were proposed as experimental statistics for the Brazilian national statistical office (IBGE).[9]

The winners of a special YSP 2021 commendation for a paper from a developing nation are Muhammad Fajar and Zelani Nurfalah (Badan Pusat Statistik – Statistics Indonesia). In '*Hybrid Fourier Regression-Multilayer Perceptron Neural Network for Forecasting*' they propose the Multilayer Perceptrons Neural Networks Model as a new forecasting method using hybrid Fourier Regression. Applying to the forecast of the production of big chili their results show hybrid Fourier Regression – Multilayer Perceptrons Neural Networks Model is more accurate than Fourier Regression and Multilayer Perceptrons (MPNN). They conclude that by using the hybrid Fourier Regression-MPNN method the government can be helped to find out the potential production of big chili in the next few quarters and considering some government policies about big chili needs.

Quality in statistics

Rudi Seljak and Tina Steenvoorden, in '*Centralised management of reference metadata and its application*',

---

[9]As the authors of this manuscript were invited to present their paper already in the Statistical Journal of the Royal Statistical Society, the SJIAOS only publishes the abstract of their paper.

report on the basic methodological fundaments of the project in the Statistical Office of the Republic of Slovenia (SURS) to develop a new, multipurpose application that enables easier and more effective usage of reference metadata produced through the statistical process and supports the evaluation phase of the statistical business model. It describes the individual steps in the design of the application, details on the functionalities of the application and points out the main challenges that had to be met during the development.

Sofie de Broe, Olav ten Bosch, Piet Daas, Gert Buiten, Ben Laevens, Bert Kroese (all statistics Netherlands) discuss in '*The need for timely official statistics. The pandemic as a driver for innovation*' how Statistics Netherlands in the context of the innovation process already in place, as well as the innovations in response to the pandemic, could respond rather quickly with a range of new outputs to the sudden increase in the need for statistical information following the outbreak of the COVID-19 pandemic. Interesting in their article is the discussion of what made speedy innovation and implementation possible and the lessons drawn in order to maintain the ability to react quickly to future policy questions. One important success factor is the combination of new data sources with already existing statistics for calibration.

With the release of seasonal and calendar adjusted series statisticians have to face which revision policy to use. In '*Fixing the model for the seasonal component: a new revision policy*', Maria Novás Filgueira, Felix Aparicio, Rafael López, Soledad Saldaña, David Salgado and Louis Sanguiao describe how INE Spain used to apply the policy of Partial Concurrent Adjustment: ARIMA Parameters in JDemetra+, but huge revisions from the beginning of the series were occasionally observed. Analyzing model changes happening several times in a year, leading to a quite unstable seasonal adjustment, they choose to apply a new revision policy, which may be considered as a compromise between the Partial Concurrent Adjustment: ARIMA Parameters policy and the Partial Concurrent Adjustment: Fixed Model, both implemented in JDemetra+. This policy avoids model changes by: (i) fixing the last estimation of the model with admissible decomposition when a model change is triggered and (ii) adjusting root assignment parameters to make sure autoregressive roots remain in the same component. This way, the authors state, the estimation of the model parameters is improved with the new data, while avoiding big revisions.

Data sources and methodology

Madior Fall's (Afristat) manuscript '*Food balance sheets provide information on food security, indicators of the prevalence of undernourishment and losses in the cases of Benin, Guinea and Mali*', fits within the Monitoring of the SDGs in Africa (SODDA) project. Analysis of the self-sufficiency rate over the 2010–2015 period shows that Mali has higher food self-sufficiency than Benin and Guinea. In Guinea, overall, 43.2% of domestic product supplies are on average imports. Plant products are the most dependent on imports with an average annual IDR of 48.2% compared to 12.5% for animal products. In the three countries, plant products are the most dependent on imports. The manuscript illustrates how the use of FAO methodologies for calculating the prevalence of undernourishment under SDG 2 and the food loss index under SDG 12 allows to estimate these two indicators and other related indicators.

Environmental-Economic Ecosystem Accounting (SEEA EA) has recently been adopted as an international statistical standard. The SEEA EA is based on spatial extent accounts (area of ecosystems) and biophysical condition accounts (ecological state of ecosystems). In the manuscript '*Urban Green. Integrating ecosystem extent and condition as a basis for ecosystem accounts. Examples from the Oslo region*', Per Arild Garnåsjordet, Margrete Steinnes, Zofia Cimburova, Megan Nowell, David Burton and Iulie Aslaken, explore case studies for the Oslo region, combining land use/land cover maps from Statistics Norway with satellite data. They argue, based on their results, that especially in an urban context, extent and condition accounts are not separate approaches as suggested by SEEA EA but should be integrated for ecosystem accounting. Moreover, the basic spatial unit should not be fixed, as suggested by SEEA EA, but should reflect that modeling of different ecosystem services, as basis for trade-offs in urban planning, requires different spatial units to capture urban green elements. The article enhances the knowledge base for assessment of urban ecosystem services within the SEEA EA.

Ahmed Youssef, Amr Kamel and Abonazel Mohammed proposed in their paper titled '*Robust SURE Estimates of Profitability in the Egyptian Insurance Market*' three robust estimators (M-estimation, S-estimation, and MM-estimation) for handling the problem of outlier values in seemingly unrelated regression equations (SURE) models. In their paper, they compare the performance of these three estimations with the traditional Ordinary Least Squares (OLS), Zellner estima-

tions based on a real dataset on the Egyptian insurance market during the financial year from 1999 to 2018 containing important indicators issued by insurance corporations on the Net Profit for the Year (NPY). The results of their application shows that robust estimations greatly improve the efficiency of the SURE estimation, and the best robust estimation is M-estimation.

Jonas Klingworth, Joep Burger, Bart Buelens and Rainer Schnell argue that capture-recapture (CRC) is currently considered a promising method to integrate big data in official statistics. In applying CRC to estimate road freight transport with survey data (as the first capture) and road sensor data (as the second capture), using license plate and time-stamp to identify re-captured vehicles they found a considerable difference found between the single-source, design-based survey estimate, and the multiple-source, model-based CRC estimate. A possible explanation for this is underreporting in the survey. In this paper, '*Transition from survey to sensor-enhanced official statistics: Road freight transport as an example*' the authors report on the effects of 1) reporting errors, 2) measurement errors, 3) considering vehicles reported not owned as nonresponse error instead of frame error, and 4) response mode. They conclude that alternative hypotheses are unlikely to fully explain the difference between the survey estimate and the CRC estimate. Underreporting, therefore, remains a likely explanation, illustrating the power of combining survey and sensor data.

Ger Snijkers, Tim Punt, José Gómez Pérez and Sofie de Broe, in '*Exploring sensor data for agricultural statistics: the fruit is not hanging so low as we thought*' argue that in theory sensor data could be a valuable new source for official statistics, for example in agricultural statistics, where new business processes are heavily data-driven with farmers using machines with sensors for precision agriculture. The assumption is that generated sensor data from agriculture might also be useful for official statistics. The manuscript reports on a small-scale data use case study in collaboration with an innovative farmer. The aim of the study was to obtain insights in the available data and the data structure in order to take informed decisions concerning next steps in innovating primary data collection for agricultural statistics. The authors conclude that though this data source may be valuable there is still time needed to prepare for using it in full and to be ready for the future.

A third manuscript on the use of new data sources is '*Scanner data in inflation measurement: from raw data to price indices*' by Jacek Bialek and Maciej Beresewicz. They illustrate the new opportunities offered by scanner data (barcode information) for CPI or HICP calculation. In their article Bialek and Beresewicz present a proposal for the implementation of individual stages of handling scanner data as well as describe potential problems during scanner data processing and their solutions. The proposal is based on the insights from comparing a large number of price index methods based on real scanner data sets, not only from a methodological perspective but also from the perspective of how time-consuming they are. As an additional criteria in the selection of the appropriate price index. They present an approach based on distances between these indices and the theoretical, expected value of the price share when prices are log-normally distributed.

David Marker describes in his article '*Suppression Criteria for Inaccurate Estimates*' that statistical offices regularly have to decide at what level of aggregation to publish results of their data collection. These decisions are typically driven by two separate concerns: first, they do not want to publish estimates with large amounts of uncertainty; second, they do not want to provide potentially-identifying information that could disclose an individual person or company. This article focuses instead on the first concern, when are data so uncertain that an agency should not publish the results? We focus on policies adopted by 16 statistical offices around the world.

The article '*Big Data in the Philippines: How Do We Actually Use Them?*' Reports on ways in which the government of the Philippines can recognize the use of Big Data for official statistics. Lisa Grace Bersales, Josefina V Almeda, Sabrina O Romasoc, Maria Nadeed R Martinez and Daniela Jann B Galias in this useful overview article start with describing gathering and presenting Big Data-related initiatives and projects across the globe for various types and sources of Big Data. Next, they discuss the opportunities, challenges, and risks associated with using Big Data, particularly in official statistics. Via an assessment of the current utilization of Big Data in the country through focus group discussions and key informant interviews, desk review, discussions, and interviews, the paper then concludes with a proposed framework that provides ways in which Big Data may be utilized by the government to augment official statistics.

Maria Thurow, Florian Dumpert, Burim Ramosaj and Markus Pauly use in their manuscript '*Imputing missings in official statistics for general tasks – our vote for distributional accuracy*' the German Structure of Earnings data from the Federal Statistical Office of Germany (DESTATIS), to investigate various imputation meth-

ods regarding their accuracy and impact on parameter estimates in the analysis phase after imputation. With the aim to deliver guidelines for correctly assessing distributional accuracy after imputation and the potential effect on parameter estimates such as the mean gross income they studied different measures for assessing imputation accuracy, beyond the most common measures, the normalized-root mean squared error (NRMSE) and the proportion of false classification (PFC), they put a special focus on (distribution) distance measures for assessing imputation accuracy.

The manuscript '*Data Integration Using Statistical Matching Techniques: A Review*' by Israa Lewaa, Mai Hafez and Mohammad Ali Ismael, summarizes the main elements of Statistical Data Integration (SDI) as tool undertaking research that involves integrating data from multiple sources in order to make the best use out of it. This paper aims at giving a complete overview of existing Statistical Matching methods, both classical and recent, in order to provide a unified summary of various SM techniques along with their drawbacks. Some points for future research are suggested at the end of this paper.

I wish you pleasant readings of these interesting articles.

2. SJIAOS discussion platform, the 10th discussion 'Statistics on difficult to measure population groups: Challenges to leave no-one not-included'

In August 2019 the Statistical Journal of the IAOS launched the on-line platform for discussion on topics of significant relevance for official statistics (www.officialstatistics.com) as part of the SJIAOS website. The discussion platform invites to contribute to important discussions at a time of own choosing. With each release of an issue of the Statistical Journal, a new discussion topic is launched via a leading article or based on a section in the Journal Each discussion runs for about a year and is closed with a concluding commentary by the article author(s).

With the release of this issue of the Journal (December 2021), also the 10th discussion will be opened. This discussion will be asking your comments/opinion on the importance to be in official statistics as inclusive as possible and especially to statistically cover the – often very vulnerable – population groups that are difficult to measure because of a lack of administrative or register data and/or difficulties in approaching or accessing these groups.

This 10th discussion is triggered by the manuscript '*Improving official statistics on stateless people: challenges, solutions, and the road ahead*', by Mary Strode (Independent Consultant to UNHCR) and Melanie Khanna (Former Chief of the Statelessness Section, UNHCR), in: Statistical Journal of the IAOS, Volume 37 (2021).

Several other discussions are still also on line on the SJIAOS Discussion platform (www.officialstatistics.com). For more information about the statements and how to react see the introduction into the 'SJIAOS Discussion platform' at the end of this issue.

3. Two issues that ask for your special attention

The December 2021 issue (Volume 37, Issue 4) is a full open access issue that contains 22 high quality contributions focussing on 'New Developments in Training in Official Statistics'. It describes the recent trends in the training in official statistics of those producing and those using results of official statistics, with the aim to develop respectively their specific knowledge, skills and competencies and to increase the 'statistical thinking'. In this context I especially would like your attention for the ninth discussion on the SJIAOS discussion platform, that is is based on seven statements on 'New Developments in Training in Official Statistics'. https://officialstatistics.com/news-blog/demand-and-format-training-official-statistics.

Full open access is now also granted on The Supplement to Volume 36: Extra issue in 2020 "Official statistics in Africa'. While the 2020 Zambia conference has been delayed, it was decided to go ahead with this extra issue to not lose the momentum of the extra attention on statistics in Africa. This extra issue with 15 contribution on official statistics in Africa and authored by statisticians from the region, has become full open access due to sponsoring by the UN Economic Commission for Africa.

4. Some words about the next issues (Volume 38 (2022), Nr. 1, Vol38 (2022), Nr. 2)

The next two issues of the journal are already in full preparation.

The March 2022 issue (Vol38 (2022), Nr. 1), will start with an interview with Mario Palma the former IAOS president (2017–2019) about his recent book on the history of INEGI, the Mexican Statistical Institute.

An article by Mario Palma based on his book will also be in this issue. The new president of the IAOS, Misha Belkindas will via an interview present the IAOS strategy for the period 2021–2023. The March issue is further expecting a selection of articles on the FAO and World Bank 50 by 2030 project on agricultural and rural statistics. The issue will further contain manuscripts from the 2021 ISI World Statistics Conference. Also the 2021 Bern UN World Data Forum (UN WDF) might result in some contributions. The June 2022 issue will be mainly dedicated to papers from these two recent conferences (see footnotes 6 and 7).

Beyond these issues with a diversity of manuscripts, there are several special issues and sections in preparation. I especially ask your attention for the planned special on the 'History of Official Statistics'. The guest editorial team for this issue is in search for authors and relevant manuscripts, so, do not hesitate to inform me when you have a manuscript or idea for a manuscript for this special. (pevssjiaos@gmail.com). It is also expected that in the upcoming IAOS conference in Krakow in April 2022, this will be an important topic for discussion. This conference is expected to result in many manuscripts for the September 2022 issue of the Journal. Participants are invited to reflect from the start of their preparation for the conference on a possible submission of their manuscript for consideration in the Statistical Journal.

Of course there are always slots for other manuscripts; authors are kindly invited to submit their manuscript to: https://www.iospress.nl/journal/statistical-journal-of-the-iaos/?tab=submission-of-manuscripts.

5. The COVID-19 pandemic and new ways of soliciting manuscripts

The COVID-19 pandemic has in 2020 and 2021 substantially changed the international conference agenda. Conferences are canceled or postponed (or organized virtually). As for many other research fields the cancelation or change of format of the international conferences has an important impact. Many Journals (also SJIAOS) are partly based on the active soliciting by the editors of articles on important and relevant new developments via the participation in conferences, networking and observing presentations listening to peers etc.

Virtual and recently also hybride conferences have proven to be a good alternative. In general it is easier to participate in a virtual conference (from home, no traveling costs, etc.). However the oversight and flexibility for the editor in chief will be substantially restricted compared to walking around and switching sessions in physical conference, and this risk that Journals will – to a lesser extent than before – be able to catch at an early stage important developments. New ways to solicit manuscripts are experienced. The editorial board of SJIAOS is inviting all readers, the editors and reviewers and other interested not to hesitate to send important papers and manuscripts for review.

https://www.iospress.nl/journal/statistical-journal-of-the-iaos/?tab=submission-of-manuscripts

Pieter Everaers
Editor-in-Chief
October 2021
Statistical Journal of the IAOS
E-mail: pevssjiaos@gmail.com