Title and author
Fisher, Nicholas I. and Trewin, Dennis J. (2021) A Proposal to Enhance National Capability to Manage Epidemics: The Critical Importance of Expert Statistical Input Including Official Statistics. *Statistical Journal of the IAOS*, Vol. 37: 465 – 481.

Type of manuscript: Policy proposal meant for (high-level) decision making.

Structure
The IMRAD structure does not fit as good as it does for research papers. Still here it is.
Introduction. Management hand-out. The main organizational structure that should be in place is described.
Methods. Inventarization of the information needed on basis of a) evaluation of handling of the current outbreak, b) description of information needed in models, c) in epidemiological considerations heterogeneity and dispersion are important issues that must be translated into the right levels of granularity.
Results. Description of the needed information in the various phases of the pandemic, with emphasis on the statistical input needed to obtain reliable quantitative data. The plan itself is described in the Appendix.
Discussion. Some extra aspects of the proposal.

Short description:
The paper describes a plan to manage and acquire all information necessary during the next outbreak of a pandemic, the Pandemic Information Plan. This includes a description of the data in the various fields that are needed in de different phases of the pandemic. It is emphasized that statistical skills are needed to provide reliable data.

The proposal is based on an analysis of the handling of the COVID-19 pandemic in Australia. What went well, what could have been done better and what was missing?
Many of the missing information could have been supplied if the data acquisition already in place to support operational goals had been founded on a solid statistical basis allowing estimates of the relevant parameters for the population and for relevant subgroups.
Since the spread of the pandemic was heterogenic, data should allow for various des-aggregations.
Most of the parameters to model the course of the pandemic were unknown, of not sufficiently known due to poor data management (and random surveys).

Main message:
The paper stresses the critical importance of reliable information to manage the crisis of a new pandemic. To be reliable, statistical expertise is needed, especially the expertise available in NSO's.

Title and author
Diaz-Bone, Rainer and Horvath, Kenneth. (2021) Official Statistics, Big Data and Civil Society. Introducing the Approach of 'economics of Convention' for Understanding the Rise of New Data Worlds and Their Implications. *Statistical Journal of the IAOS*, Vol. 37.

Type of manuscript and the main audience:
Expository paper on a sociological approach to the role of OS in society and its relation to other important providers of data and statistics. Paper is intended for a broad audience of official statisticians and other producers of social statistics.

Structure
Intro. The paper fits into a discussion on the role of OS.
Methods. Sociological theory of Economics of Convention (EC, a French theory on social role of statistics and quantification, with people of the INSEE involved.)
Results. Characterization of the OS data-world in terms of EC-theory and sketch of data-world of (commercial) Big Data and Civic Data in the same framework.
Discussion. Remarks on the future evolution of the relation of OS to the two other data-worlds.

Conceptual model
With the help of the theory of conventions different data worlds are described. The theory allows us to discuss, on a more abstract level, the role of social statistics in society and to put recent discussions in a more general perspective.

Short description:
The paper introduces several new and abstract notions. The most important one is Data World. Three data worlds are described and discussed. The world of official statistics, The big Data world and the civic data world. These worlds are seen as social institutions that produce social statistics. These statistics are produced along statistical chains that rest on division of labor. All actors (these may be organizations) in the chain must coordinate with each other and this is established by conventions. The main theory of this article is the Economy of Conventions. Conventions are complexes of values and attitudes that are associated with spheres of activity in society. The authors give 8 different conventions and give some of the qualitative features of these. Conventions are used to coordinate between different actors in the statistical chain and this gives the results of the statistical chain social and political dimensions.
Three data worlds are contrasted. De data world of official statistics, that of big data and finally the world of civic data. The last data world is emerging and connected with citizen science and open science and is critical of official and state data, for example GDP. The three data worlds are structured by different mixtures of conventions, even the conventions of quality are related to different spheres. Official statistics is related to the industrial and civic conventions, where big data follows primarily market convention and the industrial conventions and finally the civic data world follows primarily civic convention. The theory of Economy of Conventions provides a way of discussing the different roles, but also the different organization of the data worlds. These worlds are not static, but change and adapt to each other, for example around the SDG's.

Main message:
The official statistics data world will lose their dominant position and have to move in the direction of the civic data world in order to strengthen their public role.

Title and author(s)
Unit Nonresponse and Weighting Adjustments: A Critical Review, J. Michael Brick;
Journal of Official Statistics, Vol. 29, No. 3, 2013, pp. 329–353,

Type of manuscript and the main audience:
Review;
Statisticians/researchers dealing with cross sectional household sample surveys with unit non-response. Especially those involved in developing data collection protocols, estimation strategies, and the resulting nonresponse bias in the estimates will be interested in this manuscript.

Structure
Paragraph 1 Introduction, describes the issue and characteristics of unit-non response. In paragraph 2 a historical and thematic extensive background of the issues of response and non-response. Paragraph 3 focusses on the mathematical representation of the non-response bias under different circumstances, stochastic and deterministic, as well as based on characteristics of the non-response population.In paragraph 4 the discussion concentrates on response propensity modeling in situations with an at-random non response. In the following paragraphs the aspects of calculation of weights to adjust for non-response (5), the choice and use of variables and metrics that can be used to weight the non-response (6) and the calculation of the non-response propensity (7) . In paragraph 8 the non-response propensity is related to characteristics in the survey and sample procedure and protocols. Finally in paragraph 9 (Discussion) the author reflects on new approaches, and the need to develop a better understanding of the effects of non-response strategies that include changing modes of data collection, providing incentives, and converting reluctant respondents. When these strategies reduce nonresponse bias. A very lengthy literature overview complements the manuscript.
Reflection on the structure: as an overview article, very useful for a researcher to orientate/specialize into the issue of unit-non response. Though, using parts of the manuscript as reference is relatively useless, considering the logical structure as a historical and thematic overview.

The conceptual model/ the theoretical background/ the theory the manuscript is building on
The manuscript builds on a set of related conceptual models of the effect of person and household characteristics on survey unit-non responses and characteristics of sample and surveys (protocols).

The methodology used/ when relevant also the data used
The manuscript is based on desk research. There is no direct application or use of specific research data.

Main message:
Based on the extensive overview of existing literature, the manuscript concludes that in the analysis of unit-non response time is ripe for new approaches to the vexing and important question of why people do and do not respond to surveys. The author addresses the need that further interdisciplinary and basic research may prove profitable if the structural issues can be addressed and that substantive progress in this research cannot be cannot be guaranteed by any one single approach.

Title and author
MacFeely, S.(2020) Measuring the Sustainable Development Goal Indicators: An Unprecedented Statistical Challenge. *Journal of Official Statistics,* 36(2) 361-378.
https://doi.org/10.2478/jos-2020-0019

Type of manuscript and the main audience
Policy paper providing background information on Global Indicator Framework.
General public wider than Official Statisticians.

Structure (IMRAD is not applicable)
Section 1 gives the structure of the paper. Section 2 gives the key differences between MDG's and SDG's. Compared to the MDG's the SDG's are wide-ranging and the result of broad process of consultation and negotiation. Section 3 gives some challenges in measuring SDG's. Different challenges are given:
-   Complexity and lack of clarity
-   Lack of priority
-   Differences between national/ regional/ global indicators
-   Changes in compositions of groups e.g., LDC (Least Developed Countries)
Section 4 The process revealed tensions between national and global indicators. This may result in many indicators remaining unpopulated. Section 5 gives the classification of the indicators into three Tiers. With Tiers 1 containing well understood and widely measured indicators, Tier 2 well understood, but not regularly produced indicators, while Tier 3 indicators lack methodology and standards. Only 45% of the indicators is in Tier 1. Sections 6 stresses that the costs of measurement are enormous. Section 6 gives some unanticipated consequences:
-   The definitions of the statisticians have real-world political consequences
-   Many aspects remain unmeasured and maybe unattended
-   Broad targets with narrow indicators
Section 7 shows some unanticipated consequences:
-   The definitions of the statisticians have real-world political consequences
-   Many aspects remain unmeasured and maybe unattended
-   Broad targets with narrow indicators
Section 8: Discussion points:
-   Should Statistical Organizations be more inclusive
-   Big data how can these be integrated
-   There is the risk of reputation damage because of the too ambitious measurement goals
As general remark the indicators are performance metrics and less suitable for diagnostic and informational role for policy. Section 9 gives some conclusions and recommendations
-   NSO's should engage more with other statistical organizations to broaden the scope of OS
-   More international cooperation is needed to produce international statistics.
-   National statistical systems should be strengthened
-   Prepare for post-SDG.

The conceptual model/ the theoretical background
As policy paper there is no clear theoretical/statistical theory. The paper addresses a wide audience and has a broad focus.

Title and author(s)
Sample survey theory and methods: Past, present, and future directions
J.N.K. Rao and Wayne A. Fuller;
Survey Methodology, December 2017 145 Vol. 43, No. 2, pp. 145-160

Type of manuscript and the main audience:
Review article ;
The audience for this manuscript are statisticians/researchers interested and/or working with samples surveys. The manuscripts provides a very extensive overview, describing 100 years of development in sample survey theory and methods with the aim to set the scene for future use of sampling surveys in a time of fast changing availability of data sources and an increasing demand for empirical results and reliable statistical information for policy making. The article can be considered an introduction and reflection on the use of sample surveys.

Structure
The article follows overall a historical approach. After defining its context in paragraph 1 (Introduction) it describes in paragraph 2 the development during the period 1920-1960. Paragraph 3 describes the characteristics of the integration since the fifties of last century of sample survey theory with statistical inference theory and especially the likelihood approaches. From that perspective the subparagraphs in paragraph 3 describe the theoretical foundation of statistical inference, the analytical use of sample data, missing data as well as small area sampling as specific examples of application of modern sampling theory and approaches. In the final paragraph the authors discuss several developments in the light of future expectations concerning new data sources and new methodologies, as well as the impact of social developments concerning IT use, societal participation and respondents attitudes.

The conceptual model/ the theoretical background/ the theory the manuscript is building on
The manuscript reviews developments in the domain of sample surveys over period of circa 100 years and refers to a selected group of as important contributions labelled books/ manuscripts.

The methodology used/ when relevant also the data used
The manuscript is based on desk research. There is no direct application or use of specific research data.

Main message:
The manuscript - beyond giving a synthetic overview of a century of developments in sampling theory and application — aims to raise the question and invites for a reflection on the future of statistical research based on sampling. It asks the question on how the field will be impacted by social and cultural changes, factors external to our discipline and that this all will   require to adapt in data collection, data processing, and data presentation-dissemination.

Title and author
Van den Brakel, J., Söhler, E., Daas, P. and Buelens, B. (2017). Social media as a data source for official statistics; the Dutch Consumer Confidence Index. *Survey Methodology*, Statistics Canada, Catalogue No. 12-001-X, Vol. 43, No. 2. Paper available at http://www.statcan.gc.ca/pub/12-001-x/2017002/article/54871-eng.htm.

Type of manuscript: Expository. Assessing different data sources using advanced techniques.

Structure
Introduction. New, non-sampled data are compared with survey data. Time series techniques are used to reach a) more precise estimates and b) get an indication whether both data sets reflect the same underlying trend.
Methods. Structural time series techniques are applied to periodic survey data to improve the efficiency in a bivariate context. The concept of cointegration is used get an indication whether both time series reflect the same trends.
Results. The various analyses are given. In the end, the use of the social media data needs extra attention before it can be used in full. The modelling approach did not give in general the extra precision.
Discussion. The discussion is used to asses some theoretical aspects of the data analysis in the context of the current application.

Short description
Two data sources are combined, one consists of survey data and one is constructed from social media data, both measuring (economic) sentiment in the Netherland. The two data sets are combined and analyzed with time series methodology. Time series methods can provide more precision. The use of alternative data sources can supply extra precision. Social media data can provide early predictions. A new concept of cointegration is used to evaluate whether both series represent the same underlying process.
Some theory of (bivariate) time series is given. The variance reduction of the time series model approach is mixed. Sometimes the noise in the process is high and does not reduce the sampling error, sometimes the time series approach increases precision.

Title and author

De Waal, T., van Delden, A., and Scholtus, S. (2020) Multi-source Statistics: Basic Situations and Methods. *International Statistical Review*, 88: 203– 228.

Type of manuscript: Review article with emphasis on applications.

Structure

Introduction. The chosen approach to multi-source statistics is given.

Methods. The various ways to address multi-source statistics are classified according to the properties of the data sources. Combining these properties leads to 8 situations. This provides the organization of the rest of the paper.

Results. The main body gives some general considerations applicable for all situations. Each situation is covered with an example, possible problems and approaches in the literature.

Discussion provides some further reading.

Theory/ conceptual model

The careful classification of basic situations provides the access to the existing literature and examples.

It is a hands-on guide, between theory and technical elaboration.

Short description

The article is a survey of methods that can be applied and have been applied in the analysis of multi-source data. The emphasis is on the multi-source character, less on the different methods of data acquisition.

The survey is organized into 8 situations that have been derived from the properties of the different datasets that are combined. The organization into data-situations makes the article more accessible and helps to find relevant literature and applications.

Title and author(s)
Das, S., and Haslett, S. (2019) A Comparison of Methods for Poverty Estimation in Developing Countries. *International Statistical Review*, 87: 368– 392. https://doi.org/10.1111/insr.12314.

Type of manuscript and the main audience:
Combining review with technical note comparing three methods by desk research and an extensive simulation study.
Statisticians/researchers working on poverty estimation using small area estimation technique. The paper assumes a technical understanding of Small Area Estimation.

Structure
Introduction (Section 1, Section2) After a short introduction to the poverty measures (FGT), three methods are introduced, ELL (World Bank method), EBP (empirical best prediction) and MQ (M-quantile). All three methods use regression under various assumptions. ELL has been applied in many studies in developing countries, the other two in poverty studies in Europe (contrary to the title). The models of the methods are introduced. Later on a three level version of ELL is introduced.
Methods Desk research (Section 3) A table of different requirements and assumptions is given and results from studies comparing the methods are given. Simulation study (Section 4). The Bangladesh Household Income and Expenditure Survey (2000) and a sample from Population and Housing Census (2001) are combined via three-way sampling. Based on these data 500 censuses are simulated plus additional data.
Results. (Section 5) The results of the simulations show that a three-level version of ELL performs best, but this is partly thanks to the structure of the data. ELL may be difficult to apply surveys in developed countries.
Discussion Final consideration of the three models results in tentative conclusions, with pros and cons for all methods.

Theory/ conceptual model
Simulation study based on empirical data exploring the sensitivity of each method to errors in the data and comparing how well the methods estimate the errors in the (simulated) data.

Main message:
The three methods differ in the handling of census data in combination with a survey. Especially the demarcation of the small areas in the census and survey data raises problems. The data in the European studies are organized in a different way, making the application of EBP and MQ more suitable. All models are different, in assumptions on distributions, robustness to outliers and in robustness to misspecification. There is not a model that is uniformly the best in all situations.

Title and author(s)
Exploring developments in population size estimation; James Brown, Christine Bycroft, Davide Di Cecco, Joane Elleouet, Gareth Powell, Viktor Račinskij, Paul Smith, Siu Ming Tam, Tiziana Tuoto, and Li-Chun Zhang ;The Survey Statistician, 2020, Vol. 82, 27-39.

Type of manuscript and the main audience:
Combining review with Technical note;
Statisticians/researchers working on population census and aiming to utilize administrative sources via linking to either check and correct the traditional census or to replace them. (Dual-system approach). Readers experienced in this methodology and involved in tackling the issue of both under-coverage and over-coverage and trying to estimate the coverage errors might especially be interested in this manuscript. It especially focuses on countries where there is not a strong history of population registers and administrative data use in the compilation of population statistics.

Structure
Paragraph 1 Introduction; a few words about the general situation concerning the 2021 population census round, the influence of the pandemic on collecting data for the census and the change to work with administrative based population counts. Under and over coverage are then main challenges to tackle to result in reliable population size estimates. Working with combining different data sets, both administrative as well as surveys leads to different types of coverage problems. In paragraph 2 some brief examples of country experiences and the influence of coverage errors is given. Paragraph 3 goes into the main body of the note, describing methods to deal with linkage errors when one or more sources are available to estimate/correct for the under and over coverage in other sources (dual system).
Paragraph 4 aims to give examples of how in several countries administrative data are utilized throughout the census process: assessing the coverage and improving the census returns by combining the data. A concluding paragraph (5) with some general remarks, repeating the need and importance of the research as described in the manuscript.
Reflection on the structure: rather a typical manuscript, considering that it also aims to report on the papers as presented in an expert-workshop.

The conceptual model/ the theoretical background/ the theory the manuscript is building on
There is no extensive theoretical background given in this manuscript. Examples are given from experiences in several countries. For the dual-system approaches and the handling of the under and over coverage it refers to a growing body of work looking at both the estimation of linkage errors as well as developing adjustments for estimated linkage errors. There is the implicit assumption of the relation between several person/household and dwelling variables in use to estimate other population variables.

The methodology used/ when relevant also the data used
The manuscript is giving an overview of a few methods to estimate and calculate effects and risk of over and under coverage using different types of (administrative and sample survey) data sets.

Main message:
In census taking more and more use is made of administrative sources to complement and replace traditional census taking. This process of complementing and replacing (based on linking methodologies) demands methods to calculate and handle under and over coverage. New methods are quickly becoming available that help researchers in improving the quality of the resulting population size estimators.

Title and author
Puts, M. and Daas, P. (2021) Machine Learning from the Perspective of Official Statistics, *The Survey Statistician,* Vol. 84, 12-17.

Type of manuscript and the main audience
Expository note that explores some challenges for applying Machine Learning (ML) in OS.
Audience: Official Statisticians and statisticians applying Machine Learning.

Structure
Introduction (Section 1) Very short intro in AI
Messages and Results (section 2)
Machine Learning is considered along the five dimensions of quality of the EES. Two dimensions are not relevant here:
- Relevance
- Timeliness and Punctuality

The other three are considered:
- Accessibility and Clarity.
  Machine learning is a black box. It requires lots of research to gain more insight.
- Coherence and Comparability
  Problems here:
    o Concept drift, for example change in use of words that measure a concept on social media (retraining can help)
    o Correlation and causation, this is a problem for many statistical applications. In big data the dataset is usually fixed.
- Accuracy and Reliability
    o Annotated data and representativity. Annotated data that are used to train the algorithm can be compared to designing sample survey with issues as inter-comparability, bias (perceptual and perception) of the people annotating the training set. Machine Learning works best is a set that is equal to the training set. Sometimes it is necessary to over represent examples (stratified sampling).
    o Misclassification will introduce bias if the training data and actual data do not match in for example the ratio of Type I error and Type II error.

Discussion (Section 3)
Research in the problems above is needed before this form of ML can be used in OS.
In general, the expectations of direct application of Machine Learning are lower than some years ago. (It has been applied in areas as outlier detection.)

The conceptual model/ the theoretical background
There is no extensive theoretical background given in this manuscript

Main message:
Research in the problems above is needed before this form of Machine Learning can be used in Official Statistics.