

Statistical Commission
Fifty-second session
1–3 and 5 March 2021
Item 3(1) of the provisional agenda
Item for discussion and decision: Household surveys

Background document
Available in English only

Positioning Household Surveys for the Next Decade
Annotated Outline

Prepared by Inter-Secretariat Working Group on Household Surveys

Positioning Household Surveys for the Next Decade

Annotated Outline

February 2021

1. Introduction

Household surveys are a vital component of national statistical systems and a key source of social and economic statistics. They provide data for a wide range of research efforts that inform the design and evaluation of development policies. They are critical for tracking progress towards national and international goals. In the context of the Sustainable Development Goals (SDGs), household surveys are the source of one-third (80) of all (232) indicators, cutting across 13 out of 17 SDGs. Furthermore, household surveys are important for improving and complementing administrative sources. They are also critical for validating and calibrating machine learning models that combine household surveys with alternative data sources, providing insights with accuracy and precision that otherwise cannot be achieved by using these data sources alone. The need for household surveys is now greater than ever, given the widespread socio-economic and health impacts of the COVID-19 pandemic that have resulted in an increase in global poverty for the first time in two decades.

Despite their importance for development, weaknesses persist regarding the availability, coverage, accuracy, timeliness, cost-effectiveness, policy-relevance and use of household surveys, particularly in the low-income countries that stand to benefit most from survey data. Household surveys also face challenges such as diminishing response rates brought on by urbanization and higher income levels; coordination failures within overburdened statistical systems; and lengthy questionnaires that can trigger respondent fatigue with negative consequences for data quality.

This document provides an outline of a paper being prepared by the Inter-Secretariat Working Group of Household Surveys (ISWGHS) which aims to identify priority areas and the enabling environment, both at the national and international level, for household surveys to better perform their foundational roles, meet new data demands and increase their development policy and research impact in the remaining decade for the 2030 Agenda for Sustainable Development. The paper will articulate these priorities as a blueprint for action for the ISWGHS, given its mandate of supporting countries as they work towards achieving the SDGs.

2. Re-imagining household surveys for the next decade

Recent methodological and technological advances offer opportunities to improve household surveys. The COVID-19 pandemic has accelerated the adoption of new modes of data collection, leading to a surge in phone and web surveys, as well as alternative data sources that can meet the demand for timely data on the impacts of the pandemic. In the post-pandemic era, stronger and more responsive household survey systems will be needed that can enable mixed-mode data collection efforts, complement other data sources and validate insights from non-traditional data sources.

For the decade ahead, we propose to focus on:

- Increasing the policy-relevance and use of household survey data to better inform policy and research;
- Maximizing the efficiency and coverage of household surveys programmes through coordination within the country and at the regional and international level;
- Improving the quality and cost-effectiveness of household survey data production; and
- Enhancing the value addition of household surveys not only for primary data collection, but also as a source for data integration and for validating data from non-traditional data sources.

In pursuit of these objectives, work in the following areas will be proposed: (1) enhancing the policy-relevance and use of household survey data for policymaking and research; (2) pursuing an ambitious agenda on the adoption of innovative survey methods; (3) investing in capacity for CAPI, phone, web, and mixed-mode surveys; and (4) improving interoperability and integration of household surveys with other data sources.

2.1. Enhance the policy-relevance and use of household survey data for policymaking and research

The value of household surveys is realized only when data are used to inform policy. To promote more effective data use, it is critical for household surveys to be aligned with policy data needs and to involve users at all stages of survey planning, data collection, analysis and dissemination. Efforts should also be made to promote the full use of data through better dissemination of survey metadata and microdata as well as on building capacity on data literacy and use of household survey data for policymaking.

2.1.1. Align household survey programmes with policy data needs and improve coordination

Data programmes in the country should be fully aligned with policy data needs. Household survey programmes are integral part of the overall data collection programme that responds to these needs. Countries need to design a comprehensive system of surveys to produce data that responds to policy needs, complements the other data sources and is well-funded. At the regional and international level, technical and financial support should be coordinated when working with countries to help maximize the efficiency and coverage of household surveys programmes at the national level.

Lack of a clear comprehensive plans for an integrated household programme at the national level and lack of coordination of international efforts, as well as unpredictable, short-lived and ad-hoc funding for household surveys remain major challenges for an efficient and well-coordinated national household survey system.

2.1.2. Enhance engagement with policymakers and key stakeholders at every stage of household survey operations

Policymakers and all relevant stakeholders should be a key partner at all stages of survey planning, data collection, analysis and dissemination. Such engagement helps build co-ownership of produced data and the entire household survey process with policymakers and in turn helps secure financial support for household survey operations in the country.

2.1.3. Respond to the needs for SDG data disaggregation

The unprecedented requirement for more disaggregated data for SDG indicators has raised a great challenge for countries. The flexibility and customizable nature of household survey design allow inclusion of specific sub-groups of the population to help ensure no one is left behind. Providing more disaggregated data for SDGs requires a thorough consideration of topics covered, sampling approaches (including sub-sampling approaches) and other aspects of survey methodology.

For example, improvements in the availability and quality of individual-disaggregated survey data on key topics are necessary to further understanding of the various dimensions of gender equality. Critical areas of focus, for instance, may include asset ownership, work and employment, food consumption and food away from home, anthropometrics, biomarkers, and agriculture. Similar approach should be taken to ensure data for all other vulnerable groups, as identified by the 2030 Agenda, are properly obtained through household surveys. Another example includes minimizing the reliance of proxy respondents in collecting personal information and undertaking methodological research to develop and validate survey methods for individual-level data collection.

2.1.4. Increase data access and data use

Any improvements to household surveys must include effective strategies for documentation and dissemination, to leverage the full analytical potential of collected data and maximize the return on investing in household surveys. Efforts to improve data availability must be accompanied by the development of strict data privacy and anonymization protocols.

Looking forward, data producers should aim for timely data dissemination and accelerate gains in the public availability of de-identified household survey datasets, focusing on regions with higher levels of data deprivation. It is also important to strengthen the dissemination of household survey metadata that will increase transparency of survey operations and improve the ability of users to interpret and use the data appropriately.

Increasing data use in country contexts requires investments in local technical capacity, including increasing data analysis skills among statistics officials, supporting data dissemination platforms with innovative data visualization capabilities and addressing country-level political impediments.

2.2. Pursue an ambitious agenda on the adoption of innovative survey methods

Household surveys have been sometimes criticized for being expensive and providing less timely data. Low unit and item response rates across countries and measurement errors have raised concerns over quality of household survey data. It is therefore important to improve the planning and adopt innovative methods to improve the efficiency of survey operations, reduce respondent burden, and enhance the quality and timeliness of survey data.

2.2.1. Adopt smarter sampling approaches to maximize efficiency and coverage

Effective household survey programmes rely on comprehensive and up-to-date sampling frames. A comprehensive sampling frame not only covers all population in the country but also provides adequate contact information for survey organisers to approach respondents. This is particularly relevant for COVID-19 fieldwork protocols and the need to reduce travel footprint in the post-pandemic era. Some key approaches to be considered include:

- Multi-frame designs: These can better capture groups underrepresented in census frames such as refugees, internally displaced people and homeless people
- Spatial sampling: Using spatial sampling to establish the main sampling frame for the primary statistical units can improve the efficiency of the sample in meeting fieldwork requirements and increasing accuracy.
- Responsive and adaptive sampling design: Assists in overcoming challenges in the increasing complexity of surveys and declining response rates, while taking advantage of technological advances such as geospatial information and paradata from digital data collection.
- Sampling for specialized surveys anchored in large national surveys: Survey costs can be lowered by using large surveys to assist in the development of sampling frames for specialized surveys such as learning assessments or surveys targeting specific age groups within the national survey.
- Innovative sampling methods for rare events and sensitive topics: The use of certain methods such as the Network Scale-Up Method can maximize coverage when collecting data on rare events or sensitive topics.

2.2.2. Implement effective questionnaire and data collection design

Questionnaires have become increasingly lengthy, leading to higher levels of respondent fatigue with potential adverse effects on data quality. This reflects both increasing demand for data and the limitations in the scale of household surveys in countries, creating pressure to increase content on existing or planned surveys potentially to the serious detriment of data quality.

Going forward, a more integrated approach through a thorough assessment of national policy priorities and data availability within the country, along with the development of capacity in multiple modes of data collection can help achieve the delicate balance between burden, cost and range of data collected.

For example, data producers can experiment with shorter face-to-face interviews paired with follow-up complementary data collection, either through phone surveys, in-person follow-up interviews relying on local enumerators, or a combination of the two. Implementing shorter cross-sectional face-to-face surveys and leveraging survey-to-survey imputation to fill data gaps between “longer” face-to-face survey rounds is another example of effective survey design.

Well-designed and tested questionnaires are key to survey data quality. With increasingly limited resources for extensive questionnaire testing, new methods should be used and experimented to generalize knowledge about characteristics of questions that influence response times or item non-responses.

During the questionnaire design stage, a pool of survey questionnaire modules should be established and consulted for common variables across different surveys. This improves efficiency greatly by saving the cost of building a questionnaire and testing from scratch. The exercise also helps improve interoperability of data collected from the survey.

2.2.3. Increase direct measurements to reduce measurement error in self-reported data

Research has shown that scalable technologies that permit direct measurement of development outcomes, with proper quality control, could eliminate measurement errors that are otherwise inherent in self-reported survey data. Examples include GPS for area measurement, crop cutting for crop yield estimation, DNA fingerprinting for crop variety identification, accelerometers for physical activity tracking, and mobile phones for diary-keeping on consumption or labor data. Once collected, direct measures have the potential to reduce data collection burden and improve accuracy of data.

More experiments should be carried out to gain better understanding of the reliability of direct measurements and feasibility of scaling up the approach in countries.

2.2.4. Improve efficiency of data processing

Data processing is a resource intensive and error prone part of the statistical life cycle. This is particularly true for data collected through open-ended questions (for example, on occupation, industry, and time use activities, among others). Innovative approaches such as machine learning on improving efficiency of data processing should be documented and further developed. Good practices in properly designing household surveys to reduce data processing burden should also be documented and compiled.

2.3. Invest in capacity for CAPI, phone, web, and mixed-mode surveys

In the past decade, many countries have moved from paper-and-pencil interviewing (PAPI) to computer-assisted personal interviewing (CAPI) for their household survey data collection. While it represents a significant technological advancement to move from PAPI to CAPI, the COVID-19 pandemic and the social distancing measures that halted face-to-face surveys have revealed the need to build technical capacity and technological infrastructure to also implement phone, web, and mixed-mode surveys, which can be launched either as part of deliberate and routine data collection activities or in response to shocks and rapid assessment needs as in the case of the COVID-19 pandemic.

2.3.1. Systematically collect survey data to build sampling frames for phone and web surveys

To successfully implement phone and web surveys that use face-to-face surveys or population censuses as sampling frames, survey implementing agencies must collect phone numbers not only for household members but also for reference individuals, including neighbors and friends, in face-to-face surveys and population censuses. Similarly, email addresses, information on online activity, and Twitter handles should be obtained for individuals interviewed face-to-face.

These efforts should also be coupled with (a) revisions to privacy and consent agreements with face-to-face household survey respondents, given the potential for the individuals to be contacted for other

surveys down the line; and (b) collaboration with private data providers such as mobile telecommunication and social media companies to obtain access to data needed in building relevant sampling frames.

2.3.2. Develop phone and web survey tools and protocols

Strong phone and web survey infrastructure should be coupled with required survey tools and protocols. For example, phone questionnaires need to be significantly shorter and simpler given the shorter attention span of respondents, potentially poor mobile network connectivity, and concerns for respondent fatigue. The flow of questions and visual cues for phone and web questionnaires also vary significantly from face-to-face questionnaires. Meanwhile, protocols should be established for respondent selection, incentive provision, phone and e-mail contact attempts to recruit respondents, and the formulation of scripted introductions and transitions during the interviews. These will be critical for successful survey implementation, as well as to ensure the representativeness of the data to the general population.

2.3.3. Build capacity in the use of paradata for data quality assurance

With the increasing use of CAPI, phone, and web modes for censuses and household surveys, more paradata are being collected as a byproduct of the data collection process. Paradata can be used to monitor data quality throughout data collection, reduce non-response bias, and help assess measurement errors in household survey data.

The format, layout, and content of paradata may vary considerably in accordance with the software platform used for data collection, ranging from data on interviewer and respondent keystrokes to data from GPS-based monitoring of interviewer travel patterns. As statistics offices intensify the use of CAPI, phone, and web modes in survey data collection, technical capacity needs to be built to manage and analyze large volumes of paradata as an inherent part of the survey implementation process.

2.3.4. Conduct more systematic analysis of mode effects

Phone surveys and other data sources like crowdsourcing and social media data raise many concerns of quality, selectivity, and coverage, inter alia. Given the increased reliance on phone and web surveys, users need to better understand the relative accuracy, reliability, and affordability of these surveys vis-à-vis their face-to-face counterparts, preferably through survey experiments that would randomize the mode of interview for a common questionnaire instrument. Rigorous randomized experiments have been increasingly used to inform survey design choices and assess their impact; however, coordination and capacity building for such experiments in low-income countries is needed.

2.4. Improve interoperability and integration of household surveys with other data sources

Household surveys become exponentially more useful when integrated other surveys carried out in the country and with other types of data, such as geospatial data, administrative data, call detail records, and social media data. Enabling seamless integration of household surveys with other data sources can

counteract criticisms regarding respondent burden, costs of survey implementation, and limits to coverage, granularity, and timeliness.

For example, including common questions for integration with other existing surveys and censuses enables survey-to-survey imputation to achieve greater spatial resolution of the estimates. Common questions can link household surveys with administrative data and Big Data and thus help validate data collected with non-probabilistic approaches, such as surveys run on social media platforms or citizen-generated data. Meanwhile, collecting geo-referenced household, plot, and community locations can link household surveys with a host of geospatial data from satellites, helping to validate and calibrate high-resolution mapping of poverty, asset, wealth, and child nutrition, among others.

Data integration methodologies need to be further developed, tested and standardized to facilitate better integration of survey data and other data sources. In addition, coordinated efforts should be made on all data sources within the national statistical system to enable data integration.

3. Fostering a stronger enabling environment for household surveys

The proposed actions to improve household surveys require a stronger enabling environment. The following are some of the prerequisites, both at the national and international level, for household surveys to maintain an influential role in countries' data systems and meet the challenges of the next decade.

3.1. Quantify the benefits and communicate the value of surveys

The household survey community needs to invest in data visualization and data journalism to better communicate – relative to other data sources and based on empirical evidence – the “value addition” of household surveys, both in and of themselves and through integration. This is critical for boosting the understanding of the critical importance of household survey data, and in turn, securing political commitment to and predictable financing for household survey programme.

3.2. Promote legislative support to facilitate data integration

Successful integration of survey data with other data sources within the National Statistical Offices as well as those from the private sector requires strong legislative backing that allows access to other data for official statistical purposes and guards the confidentiality of personal information in the integrated data, as well as in the original data sources.

3.3. Invest in ICT infrastructure

A recent survey of NSOs conducted by the UNSD and the World Bank, in collaboration with the UN Regional Commissions, revealed how national statistical offices currently lack adequate equipment, software, and ICT infrastructure to support the continued regular production of statistics, including national surveys that increasingly rely on computer-assisted personal and phone interviewing. The lack of cloud computing services for data storage and exchange as well as of suitable facilities for remote training are common challenges highlighted by the UNSD-WB survey. Stronger and smarter technological infrastructure should be built at the national level for implementing CAPI, phone, web, and mixed-mode

surveys. These measures should be coupled with enhancements in data storage, data dissemination platforms, and steps to address hardware and software needs on the part of NSOs for the analysis of household surveys and their integration with alternative data sources.

3.4. Move towards a more diverse skills mix at national statistics offices

The skills mix of national statistics offices must diversify to increase the development impact of household surveys. Capacity must be built in experimental statistics, particularly on topics related to data integration and the use of alternative data sources as complements to household surveys. To build capacity, rigorous training should be held for statistics staff on all relevant topics including survey data collection, compilation and dissemination, data integration, new data sources, innovative methods, data visualization and data journalism.

3.5. Invest in capabilities for remote delivery of technical assistance

The COVID-19 pandemic made remote training and technical assistance a necessity, which may be an accelerator for the adoption of remote training by national statistics offices for cost-saving purposes. Materials and guidance on remote training delivery mechanisms should be developed. Countries could also consider establishing training centers at regional level. The abovementioned investment in the ICT infrastructure is a must for this aspect to be implemented successfully.

3.6. Pursue a coordinated and systemic approach to supporting national statistics

To unleash its full potential, household surveys should be positioned strategically within the national data ecosystem while recognizing its strength and weaknesses and adequately funded. Meanwhile the critical role of a functional household survey programme needs to be emphasized for pursuing the above recommendations.

For example, during COVID-19, countries were able to rapidly launch and successfully implement phone surveys on COVID-19 that used existing household surveys as sampling frames, taking advantage of investments in capacity building over a decade prior to the pandemic. We can learn from this that investment is needed now to improve the future responsiveness and resilience of data collection systems.

Collaborative efforts at the regional and international level are key to a coordinated household survey programme at the country level, especially for countries that rely technically or financially on international agencies and donor community. The ISWGHS should be the forum to foster such collaboration.

3.7. Sustain financing for household surveys

Sustainable financing for household surveys is crucial at both national and international level. At the national level, systematic funding mechanisms should be in place to support household survey operations as an integral part of national statistical system. At the global level, donor investments for global public goods must be sustained towards carrying out survey research, developing and testing survey tools for national scalability, producing guidelines for data collection, use, and dissemination and curating and

disseminating good national practices. Other innovative financing mechanisms that integrate data capacity as a key dimension of development should be encouraged.

3.8. Support a stronger role for the Inter-Secretariat Working Group on Household Surveys

Implementing these priority areas will require a strong coalition of international agencies and countries to support a shared agenda. The ISWGHS can bring together this coalition by complying with its mandate in three areas of work:

- Coordination: The ISWGHS will foster coordination of national household surveys towards building a comprehensive household survey programme that responds to policy needs; as well as collaboration at the international level in the planning, funding and implementing household surveys. The ISWGHS will broaden its outreach and collaboration with partners including NSOs, civil society and academia, and with other data communities such as those focused on administrative data, geospatial data, and citizen-generated data. It will serve as an intermediary to link innovative methods with national data needs, as well as in curating and disseminating good practices to promote exchange of experiences.
- Methodological development: The ISWGHS will focus on innovative topics for its methodological work, following the areas outlined above. Its methodological work will ensure that the solutions provided are scalable and easy to adapt for countries at different levels of statistical development. The group will also encourage and support experimentation with innovative approaches to establish systematic empirical evidence within countries. Finally, it will support the scaling up of innovative survey methods in countries.
- Advocacy and communication: The ISWGHS will promote the documentation and exchange of national experiences in carrying out different components of the proposed ambitious agenda. It will also continue advocating for the better financing of household survey programs as an integral part of a national data ecosystem.