ISWGHS Side event at the 52nd United Nations Statistical Commission

Positioning Household Surveys for the Next Decade, 19 February 2021

Summary

COVID-19 has brought tremendous challenges to national household survey programmes. A paper on how to position household surveys for the next decade is being prepared by ISWGHS and the <u>annotated</u> <u>outline</u> is available as a background document for the 52nd United Nations Statistical Commission. A virtual consultation meeting took place on 19 February 2021. The meeting was attended by 125 participants and panelists were from National Statistical Offices of Canada, Palestine and Uganda, an UN Regional Commission, the Demographic and Health Survey (DHS) and academia. The concept note, including the work programme of the side event is available in Annex 1.

During the meeting, an introduction <u>presentation</u> was made by ISWGHS members. Panelists and participants responded positively to the timely preparation of the paper, provided input to the annotated outline and brainstormed on how to help countries in operationalizing the proposals. Recording of the side event is available <u>here</u>. Statements made by the panelists are included in Annex 2. A consolidated list of questions raised by participants and by those who registered to the event is available in Annex 3.

Summary of key points

The meeting acknowledged the continuing and increasing importance of household surveys in supporting the implementation of the 2030 Agenda for Sustainable Development, especially in reaching the vulnerable population groups. The importance of household surveys has been further demonstrated following the COVID-19 pandemic when surveys, mostly by phone, produced the very urgently needed timely, high quality and disaggregated data to address the challenges posed by the pandemic.

Participants of the meeting praised ISWGHS for taking on the task to identify priorities and the enabling environment for household surveys in the next decade, as covered by the outline of the position paper. Participants emphasized that collective and concerted efforts should be made to accelerate the adoption of innovative approaches in countries to better position household surveys for the next decade.

The participants however noted that workable innovative solutions differ by statistical capacity and infrastructure readiness at the country level; as well as by survey topics, types of surveys and population covered. For example, while national household surveys have adapted quickly to the challenge of COVID-19 by adopting different data collection modes such as phone and web, it is important to ensure that survey tools do not leave out the vulnerable population groups who do not have access to those media.

Innovative approaches should also be thoroughly tested within each country before being fully adopted for official statistics. Countries are also encouraged to regularly incorporate small-scale experiments when implementing their surveys.

The discussants were also supportive of the emphasis of the position paper on the critical importance of integration of household survey data with other sources for more timely and more disaggregated data and to increase the value of household surveys. The role of household surveys as a benchmarking tool for non-probability data was emphasized.

While more data are being integrated that contain geospatial information and personal contact information, participants noted that the privacy issue brought by the richness of data should not be overlooked. Methods to protect confidentiality of individual level data should also be developed.

The need for more research on tools to monitor and evaluate the quality of interviewer training and for survey data quality assessment was raised. In this regard, collecting and mining paradata using machine learning algorithms should be explored as cost-effective ways to improve quality of survey data.

The importance of experiences sharing and peer-to-peer learning on survey methodologies and innovative methods to the adoption of innovations was emphasized.

Lastly, the participants reiterated the importance of survey coordination within the country and at the international level; as well as broad collaboration with all stakeholders. The importance of strengthening communication and advocacy for the importance of household surveys was also raised.

Annex 1. Concept note and work programme

Positioning Household Surveys for the Next Decade February 19th, 2021 8:00 – 9:30 am (EST)

Household surveys play a critical role in meeting national data needs. A recent mapping exercise found that approximately one-third of all Sustainable Development Goal global indicators (80 out of 232 indicators), covering 13 different goals, can be sourced from household surveys. Despite their fundamental role in national statistical systems over the past decades, household surveys are facing funding challenges and skepticism on their continued utility within the changing data landscape.

Within this context, the COVID-19 pandemic exacerbated existing challenges for household surveys, given limitations on face-to-face survey operations as a result of social distancing and other restrictions in many countries. A survey conducted by the UN Statistical Division and the World Bank in May 2020 found that 96 percent of national statistical offices either partially or fully stopped face-to-face data collection at some point during the pandemic.¹The increased need for timely, high-quality data on the impacts of the pandemic heightened demands for more efficient household survey systems relative to the status quo.

In response to these challenges, national statistical offices have acted quickly to respond to new data needs introduced by the pandemic, while maintaining the continuity of existing data collection programmes. To do so, countries have adopted innovative approaches and/or turned to alternative data sources², with most countries implementing one or more rounds of phone or web surveys measuring the impact of COVID-19 in the past year.

Looking ahead, the global data community faces several key questions on the future of household surveys. How will COVID-19 reshape national household survey programs in the future? How can we establish sustainable household survey programs that are resilient and versatile to future shocks like COVID-19? Will COVID-19 be a catalyst for innovation in the field of household surveys? What are the priority actions for countries and for the international community in the next decade for household surveys?

Organized by the Inter-Secretariat Working Group on Household Surveys (ISWGHS), the session provides a platform for national statistical offices, international organizations, the research community, and civil society organizations to share their views on how household surveys should be best positioned for the next decade. The discussion will contribute to the drafting of a position paper being prepared by the ISWGHS. Following a short (10 min) presentation by the ISWGHS outlining the content of the position paper, discussants from national statistical offices, development partners and academia will be asked to make brief (5 min) remarks on the document and the proposed role of the ISWGHS, followed by a Q&A session. For further reading, an annotated outline for the paper is available as a background document under agenda item 3(I) at https://unstats.un.org/unsd/statcom/52nd-session/documents/BG-3l-Positioning_household_surveys-E.pdf.

¹https://covid-19-response.unstatshub.org/statistical-programmes/covid19-nso-survey/

 $^{^2\} https://unstats.un.org/iswghs/task-forces/covid-19-and-household-surveys/national-responses-to-COVID-19/$

Livestream video available at: <u>https://youtu.be/F2t_GJu2sYI</u>

Work programme

Opening: Mr. Stefan Schweinfest, Director, UN Statistics Division

Moderator: Mr. Gero Carletto, co-Chair, ISWGHS and Manager, Living Standards Measurement Study, World Bank

Introduction presentation: Ms. Haoyi Chen, Coordinator, ISWGHS and Mr. Talip Kilic, Senior Economist, World Bank

Discussants:

Ms. Ola Awad, President, Palestinian Central Bureau of Statistics

Mr. Jack Gambino, Former Director, Household Survey Methods Division, Statistics Canada

Mr. James Muwonge, Director for Professional Services and Project Management, former Director Socioeconomic Surveys, Uganda Bureau of Statistics

Ms. Sunita Kishor, Director, The Demographic and Health Surveys (DHS) program

Ms. Frauke Kreuter, Professor, Ludwig-Maximilians-Universität München and University of Maryland

Mr. Andrés Gutiérrez, Regional Adviser on Social Statistics, UN Economic Commission for Latin America and the Caribbean (ECLAC)

Q&A

Closing: Mr. Rafael Diez de Medina, Director, ILO Department of Statistics

Annex 2. Statements made by panelists

Ms. Ola Awad, President, Palestinian Central Bureau of Statistics

Coordinate, development of methodology, communication and advocacy are 3 important aspects. More specifically, it is very important

- to integrate household surveys and administrative data and other existing data sources.
- focus on data at local and regional level. It is important to develop new methodologies;
- encourage the exchange of experiences
- strengthen the communication and partnership with stakeholders, including civil society, private sectors and academia, in additional to line ministries and government institutions
- provide support to CSOs and manage other types of information from only collecting data through household surveys

Mr. Jack Gambino, Former Director, Household Survey Methods Division, Statistics Canada

I knew I would like the document when I read in the very first paragraph that "*The need for household surveys is now greater than ever*...". Just a few years ago, some people – not statisticians, but some non-statisticians – gave the impression that big data would solve all our information problems and might even make surveys obsolete. Of course, that was far too unrealistic.

As I read further into the document, I found nothing that I disagreed with, but I did find that I would have emphasized some things more than others, and I would have added some other ideas as well. I realize that the document is an outline for a lengthier document, so it's likely that some of my comments would have been addressed in that version, but let me go through some ideas just in case.

Before giving some specific comments, I have one general comment, namely that the document doesn't say enough about the many challenges faced by NSOs when they try to implement such recommendations. In particular, I would like to have seen mention of the fact that there are different challenges for different NSOs. For example, at the beginning of section 2 it says *"stronger and more responsive household survey systems will be needed"*; this is true for any country, but what it means for Statistics Sweden is very different from what it means for an NSO from a developing country. Obviously the Working Group is aware of this, but it should be noted explicitly, and so I hope that the longer document will address this.

Moving on to specific comments, section 2.1.3 is on "*data disaggregation*". I agree that there is such a need – it's a constant demand in Canada. This has become an even bigger issue in Canada due to the way that COVID-19 has impacted different subpopulations differently. Measuring this differential impact is a major challenge, and Statistics Canada is working actively on this issue.

Section 2.2.1. on "*smarter sampling approaches*": The approaches mentioned in this section are important and useful, but they are not a panacea. For example, in my experience, responsive/adaptive design certainly helps improve the data collection process, but the gains are not dramatic – of course, results may be different in other contexts.

Section 2.2.2. on "*effective questionnaire and data collection design*": This is a tremendous challenge (e.g., balancing questionnaire length and response burden with the need for detailed information from respondents). The section mentions some approaches to deal with this challenge, such as "*leveraging survey-to-survey imputation to fill data gaps*". For this particular point, there is some nice theory to support such approaches but practical success stories are not common (at least as of a couple of years ago when I last looked at this).

Section 2.3.1 on building *"sampling frames for phone and web surveys"*: A distinction should be made between the use of phone numbers and email addresses (1) as a means of **contacting** a person, (2) as a means of **selecting** people for a survey and (3) as a means of **collecting** data (in the email case, this could be to direct the person to a website that has an online questionnaire) Some of this is addressed in item 2.3.2 and it should also be covered more explicitly in the longer document.

2.3.4 on "systematic analysis of mode effects": My only comment here is that, of course, there has been a lot of country-specific work on this topic, but there may be non-trivial differences among countries, so it's good that the document stresses the importance of conducting additional experiments.

2.4. on "*integration of household surveys with other data sources*": This is an important area to develop. In my experience (which I admit is limited), integrating data from different sources works well when data for a person or a household are combined using record linkage but only moderately well when using other methods (such as statistical matching).

One important issue is the legal framework that allows (or makes difficult) combining data from different sources. It helps if a country has a single NSO, but even that happy situation is not sufficient; for example, can the NSO get data from other government departments? and transaction data from the private sector? and data from social media companies? This issue is addressed briefly in section 3.2 of the document, but perhaps more can be said in the longer document.

Finally, I was very happy to see the statement in section 3.8 on "*experimentation … to establish systematic empirical evidence within countries.*" The words "within countries" are important here, because for certain issues, there will be differences among countries. In fact, not only will the results differ, but even the nature of the experiments will differ (e.g., a test that is possible in one country may be a major challenge in another country because of differences in data collection infrastructure).

References added after the meeting:

This document gives a good overview of experiences with responsive design; it is related to my comment on section 2.2.1:

Tourangeau, R., Brick, J.M., Lohr, S., and Li, J. (2017). Adaptive and responsive survey designs: a review and assessment. *Journal of the Royal Statistical Society, Series A*, 180, 203-223.

This document was literally published just days ago. It provides an overview of recent developments in survey-nonsurvey data integration:

Beaumont, Jean-François and J. N. K. Rao (2021). Pitfalls of making inferences from nonprobability samples: Can data integration through probability samples provide remedies? *The Survey Statistician*, January 2021 issue. URL: <u>http://isi-iass.org/home/wp-</u> <u>content/uploads/Survey Statistician 2021 January N83 02.pdf</u>

Mr. James Muwonge, Director for Professional Services and Project Management, former Director Socioeconomic Surveys, Uganda Bureau of Statistics

I just wanted to add the following points to the need for establishing partnerships with the private sector (telecommunication company, and other key players in developing countries):

(a) the paper should take note of the fact that phone surveys are good for specific types of studies and face-to-face should continue alongside phone surveys

(b) the Censuses (population and business) could be improved to capture information that is useful in constructing the different sample frames (that would increase reliability of estimates from phone surveys and for other specific studies)

Finally in developing countries, we may not achieve much in terms of increasing utilisation of data and information if we preach to the same audience (planners etc). The paper should propose how to segment the audience to ensure increased use of data

Ms. Sunita Kishor, Director, The Demographic and Health Surveys (DHS) program

Thanks Gero. First, congratulations on the paper—it certainly covers a lot of ground covering many of the most important issues. As you can imagine, for us at The Demographic and Health Surveys Program much of the action items in the paper are our daily bread and butter: including such questions as how do we improve our methodologies for data collection; how do we balance the policy and programmatic needs for data with the survey and questionnaire burden; and what tools can we use to make sampling and data processing a shared responsibility with our country partners. Also, data collection is not the only key objective of the DHS Program—other objectives include an emphasis on data use, capacity strengthening and dissemination all of which the paper touches on. Against this context of similar challenges, let me restrict my comments to a few key areas:

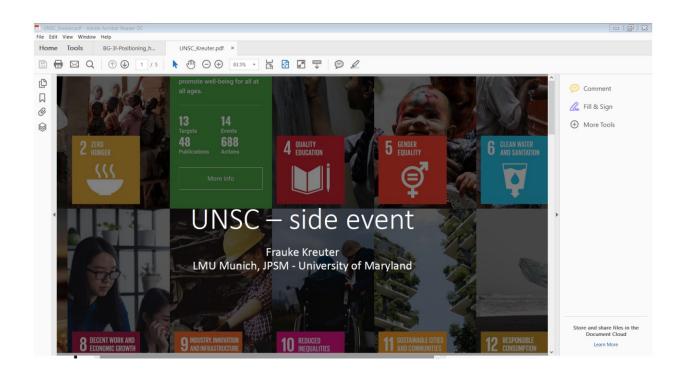
- First the continued relevance of surveys: I do not believe that there is any situation in the foreseeable future, and definitely not by 2030, that surveys will not be central to meeting the data needs of countries. I say this not just because of all the reasons laid out in the paper, but also because of the fact that as long as inequalities exist, particularly in education and literacy and in access to information, and information tools, not just surveys, but face-to-face surveys will continue to be necessary. Without them, we will surely be unable to include the most vulnerable and marginalized populations—the inclusion of whom is a major objective of the SDGs. Additionally, a key part of the paper mentions the need to verify and 'proof' self-reported survey findings. For health surveys in particular, biomarker collection and testing does exactly that: selfreported health status can be checked with biomarker data collection. For this too, face-to-face surveys continue to be essential.
- 2. The second issue I want to comment on is a discussion of alternatives to face-to-face surveys. With the advent of the COVID era there are plenty of examples of how quickly and effectively various survey platforms have been able to shift to on-line or cell phone data capture. However, in low and middle income countries, there is widespread acknowledgement that complex surveys with multiple modules need face-to-face interaction. Additionally, not all surveys are made equal: in fact, while some data can be effectively collected on cell phones or other remote means, we need special care for surveys that contain sensitive information. Can we, should we, attempt to get data on say the experience of domestic violence from women over the phone? The phone may well be controlled by the very person who is abusing her. Some cell-phone based survey research indicates that many women have their speaker phone during phone interviews. So, we need careful research on the types of information that can safely be collected over the phone.
- 3. This leads to my third point regarding PII security and confidentiality. There are many innovations proposed in the paper—all of which I agree need to be pursued. However, not too much attention has been paid to questions regarding confidentiality of information and the ethical collection of data. Here I refer not only to the work the DHS and the World Bank are doing together to find more efficient ways of displacing GPS locations of the survey households to ensure respondent confidentiality, but also to the issue of how to safeguard other PII. While the collection of phone numbers and emails for respondents as part of a regular survey can indeed prove very useful, the maintenance of this information runs counter to the need to anonymize the data so that respondent privacy is maintained. While this is relevant for all types of surveys, it

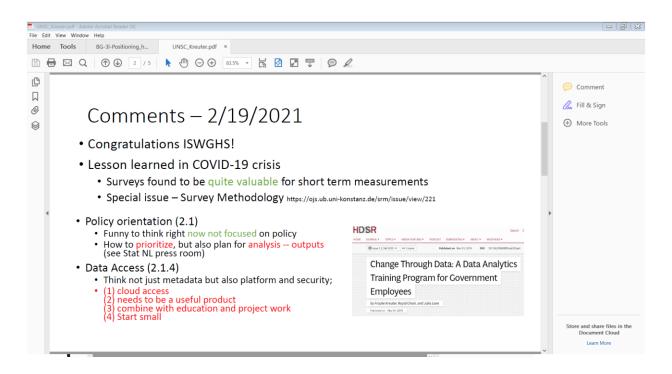
is a particular challenge for surveys like the DHS that collect highly sensitive information including on HIV status, sexual behavior and domestic violence. The data belong to the country, and there is a need for fool proof systems to be set us so that the telephone data base cannot be used to violate respondent confidentiality.

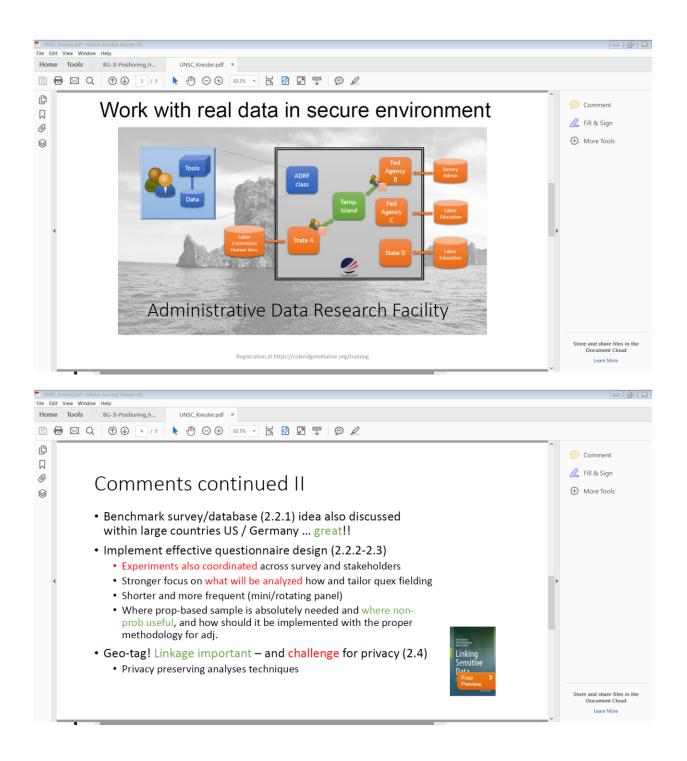
4. I am almost out of time, but let me mention one last issue regarding the need for more research on tools to monitor and evaluate the quality of interviewer training and for survey data quality assessment. This is particularly relevant if we want to explore further the idea of linking surveys to mine for complementary information. We must have ways to make quality assessments so that when we try to link surveys, we are linking apples and apples and not apple and rotten fruit. In this context, the collecting and mining of paradata I believe will prove to be one of the most important innovations in survey research. Developing machine learning algorithms that cut down on human monitoring of data quality as it is being collected holds the promise of not just improving data quality but also doing it in a timely and potentially cost effective way.

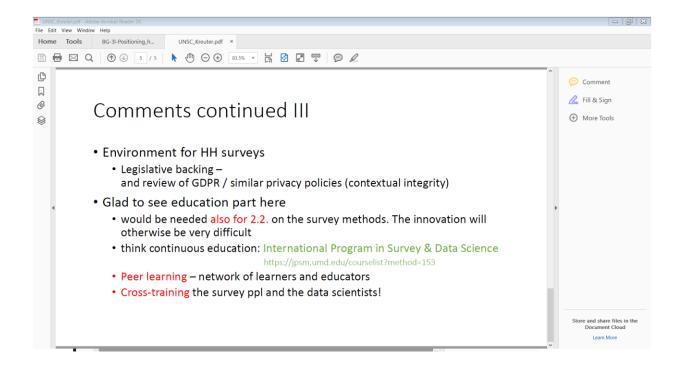
There is more I could have said, but let me stop here.

Ms. Frauke Kreuter, Professor, Ludwig-Maximilians-Universität München and University of Maryland









Mr. Andrés Gutiérrez, Regional Adviser on Social Statistics, UN Economic Commission for Latin America and the Caribbean (ECLAC)

1. The Outline gives a useful and comprehensive overview of the desired position that household surveys should assume in the next decade. These statistical processes could play a central role in the integration of statistical systems. As a survey statistician, I am convinced that no other collection method allows having such gains in accuracy and precision while keeping a reduced sample size . If we could summarize the Outline in one word, it should be *integration*. However, in Latin America and the Caribbean, household surveys have failed to completely incorporate external auxiliary information from censuses, administrative records, and other important statistical sources.

2. At first, I want to focus on a single concern and how this issue is related to the recommendations we have found in the outline. It looks like the society demands for more data, but at the same time respondents are more reluctant to answer interviews. When looking at trends over time, it is found that response rates are still declining over the years. Literature has shown that nonresponse trends differ over countries, but not over surveys. Some countries show a steeper decline in response than others, but all types of surveys show the same downward trend. The differences in (non)response trends over countries can be partly explained by differences in survey design between the countries.

3. In Latin America and the Caribbean, the pandemic brought many changes in household surveys, and perhaps many of them should be adopted in the medium term by national statistical offices. For example, in many countries, the collection mode went from face-to-face to telephone, although the sample continued to be based on an area framework. With declining response rates and challenges of using RDD sampling for telephone surveys, collecting data from address-based samples has become more attractive. In fact, it was shown that face-to-face interviews are not mandatory to produce social statistics in this region.

4. Household survey samples should continue to be planned to represent the whole population and major disaggregations. However, the sampling design, more specifically the sample allocation, should be defined so that in a second stage (or step), the sample can be combined with auxiliary information at the unit or area level to produce disaggregated official statistics. There are recent papers that explore the advantages of defining this kind of sample design.

5. Household surveys cannot exist without censuses. This is important because, in Latin America and the Caribbean, censuses define the sampling area frame and the corresponding stratification. However, the use of censuses in household surveys can be extended, not only to calibrate the population totals but also to include their information in a small area estimation setup to produce again disaggregated statistics from household surveys.

6. The change of mode collection and the integration of household surveys with other statistical processes should consider the comparability within the time series. It is essential that national statistical offices can assess the impact that these innovations can have. For example, some countries in the region have started to define the assessment of the mode effect given to the pandemic and have used randomized groups taking advantage of rotative sampling designs.

7. In Latin America and the Caribbean, this Outline can become a useful guide to countries and national statistical offices to take specific actions and decisions to transform the way we think household surveys

and how they can interact with other sources. In particular, countries can be supported by the regional commissions to adopt and adapt new methodologies to their specific concerns and needs.

8. Finally, I think that regional commissions could play a central role on positioning the recommendations defined on the outline by providing support and assistance to countries in order to create capacity building into the national statistical offices regarding the design and analysis of household surveys, and most importantly their integration with other sources of information, and the estimation of disaggregated statistics by fitting small area estimation models.

ISWG session – positioning household surveys for the next decade:

Concluding Remarks

- Probably everyone participating in this session understands very well how important household surveys are and I think also recognise that to meet data demands we actually need many more household surveys, not less.
- However, we are standing on a burning platform and know the resources needed to complete HHS are under threat, even before COVID-19. For example recently we commented on a national statistical strategy that set a target to replace the majority of data collected from household surveys by other sources in less than 10 years. Whether or not that target is realistic it demonstrates that for many HHS may look like an expensive luxury.
- COVID-19 has shown that we have the possibility to adapt HHS to be fit for purpose, but we also know how challenging that process has been for countries. With that very fresh, and ongoing, experience, this is a good time for us to reflect on how we can envision the future of HHS. Thankfully I believe we have many good practices to refer to, such as some of those shared today.
- I would like to focus in particular on three words that I think summarise what we need to do as highlighted in the earlier presentations. They are **co-ordination**, innovation and advocacy.

COORDINATION

- When I talk about **co-ordination** I mean this **at various levels**. Certainly, co-ordination **among international agencies** is something we need to be concerned about and is an important part of the mandate of the ISWG. I believe this is an area where progress has certainly be made but we need to follow this through strongly to ensure that countries feel the benefits.
- Of course the other important area is co-ordination within countries, which can also operate on multiple levels, across ministries, within an NSO etc. We know well the cost of poor co-ordination – duplication of efforts, lack of coherence and consistency, sub-optimal use of resources and so on.
- However, we also know that countries have learned much about how to co-ordinate their household survey programmes to good effect so we need to **learn from this and spread these good experiences to the benefit of all**. I know this is an important focus of the current work of the ISWG so I really encourage that we see this through.

INNOVATION

• Talking about innovation the COVID-19 pandemic has already shown us some of the ways we can innovate, for example by updating modes of data collection, improving or changing sampling frames etc. We need to build on this, even if in the short-term there can be a need to revert to traditional approaches because that is what the capacity and conditions allow (fully agree with Sunita's points). The proposed document points to a variety of the innovations we can target all towards the idea of making surveys more efficient in delivering high quality data that meets user needs.

• By successfully innovating we create the potential to significantly the availability of high quality data and more granular data (as Andrés has highlighted), even with the same resources as before, but what we need to guard against is the disappearance of those resources.

ADVOCACY

- The final word I think is key is **advocacy**. No matter how well statisticians do their technical work, **we often fall short in convincing others how important this work is**. In the case of household surveys I think we have **a very clear risk that even those who recognise their use don't see them as good investments, particularly in the tight budgetary environment we will undoubtedly face,** so there is a major task for all of us to raise the profile of household surveys and their value. We need to do this at the same time as modernising in order to avoid the risk of looking like we are stuck in the past.
- So in summary I think this **discussion is timely**, no less so because of the massive disruptions created by the COVID-19 pandemic. Household surveys are increasingly necessary, but also increasingly under threat. We have the **opportunity to improve them and the risk that if we don't they could become more rare, particularly in less developed countries where data gaps are largest.**
- We do need to take a longer term view as we know all too well that the achievement of significant innovation and good co-ordination take time, even if we can hope to accelerate development by building on good country practices. With this in mind I find the reference to the 'Next Decade' appropriate as a way to give a sense of the scale of the task ahead. However, of course this does not mean we have the luxury of time to act. A genuine sense of urgency is needed in this and I hope the document can act as a platform to focus our efforts.
- While we have to take a longer term view we still cannot forget that for many countries thoughts of innovation and better co-ordination are something of a luxury and what they require now is strong support from international agencies to enable household surveys to take place at all, particularly in these challenging times.
- Finally, as noted by the colleagues who presented earlier, there is a key role for the ISWG HS to play here. The role of international organisations in supporting the household survey system varies substantially across countries. However, for those most in need of support there is a clear demand for strong and co-ordinated action from the community of international organisations. The ISWG HS is a key mechanism to co-ordinate and prioritise our work and, only by combining our efforts can we maximise our outreach and impact. However, we need to deliver on the promise that our mandate provides. It is fair to say it has taken time for the ISWG to find its feet but I am increasingly confident that it will play a critical role in advancing the cause of co-ordination, innovation and advocacy over the coming years, with the tangible results our constituents demand of us.