

Achieving the Full Potential of Household Surveys: the work of Inter-Secretariat Working Group on Household Surveys

Haoyi Chen Coordinator, Inter-Secretariat Working Group on Household Surveys



Outline

- ☐ What is the Inter-Secretariat Working Group on Household Surveys
- ☐ Highlights of our work
 - Coordination
 - Methodological development
 - Communication and advocacy
- ☐ Position paper the future of household surveys



The Inter-Secretariat Working Group on Household Surveys (ISWGHS): a primer

- ☐ Established in 2015 under the aegis of the UNSC
- Objectives:
 - ☐ Improve coordination of household surveys
 - ☐ Advance cross-cutting survey methodology
 - ☐ Enhance communication and advocacy

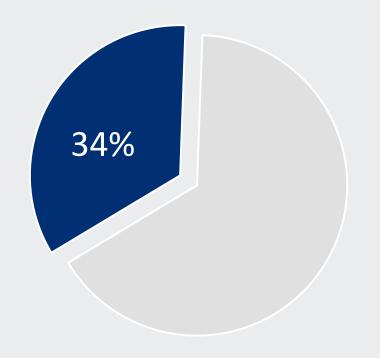
□Governance

- Membership: 11 international agencies + 8 (rotating) member states
- Secretariat: UN Statistics Division
- Current co-chairs: WB and UNW
- ☐ Work through time-bound Task Forces, led by and with contribution from members and non-member experts.



Household surveys for SDG monitoring and national policymaking

SDG indicators can be obtained from household surveys



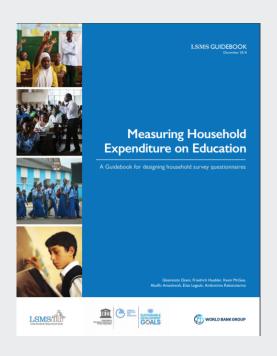
Covering 13 goals



Source: SDG indicator mapping

Developing thematic survey methodology

- ☐ Standards for **education spending** (UIS and WB)
- ☐ Measuring **food consumption** in household surveys (FAO and WB)
- ☐ A survey module on **SDG 16 indicators** (UNODC and UNDP)
- ☐ Consistent measurement of paid and unpaid work activities (ILO)









Developing cross-cutting survey methodologies

- ☐ Moving towards more **coordinated** national household survey programmes
 - Collecting and sharing survey information, a survey calendar?
 - Developing recommendations on a comprehensive national household survey programme
- ☐ Dissemination of household survey microdata
 - Household survey microdata dissemination
 - Spatial anonymization in public-use household survey datasets (completed)
 - Standards and best practices for survey data documentation
- □COVID-19 and household surveys (UIS, UNSD, WB)



Survey co-ordination task force - Terms of Reference

□Objective

 Work with a selected number of countries in deriving comprehensive, practical and efficient survey plans that fit various data needs and statistical capacity.

□Output

- A set of required information for a comprehensive and efficient survey plan, to be collected from the focus countries
- leading the work with countries in assessing their data needs and household survey programmes
 - Assisting the drafting of a comprehensive, efficient and practical multi-year survey plan for each of the participating country, led by participation of national statistical offices
 - Providing recommendations on steps for the focus countries to move towards a coordinated household survey programmes



How much have we achieved?

- □ Toolkit to assess coordination of national household survey prorgrammes, completed
- ☐ Worked with 5 countries and 1 sub-regional office
 - Canada, Costa Rica, Ghana, Ireland, Samoa, SPC
 - In collaboration with ECLAC and ESCAP
- ☐ For each country (and SPC)
 - Desk research, meetings and email communications with countries
 - Individual reports
- ☐ Consolidated report



Toolkit to assess coordination of national household survey programmes

INTER SECRETARIAT MICREING GROUP ON HOUSEHOLD SURVEYS Light to an founder

DRAFT, AS OF FEB 2021

Table of contents

	•
Background	3
Study scope	3
Conduct of the study.	4
Online research	4
Meeting	4
Questionnaire	4
Lessons learned and general recommendations	6
Broad overview of the household surveys infrastructure	6
mirastructure	0
Assessing the degree of coordination, integration,	
and harmonisation	8
Summary	8
Identification of obstacles and challenges towards more coordinated, integrated and harmonized household	9
surveys	9
Current and anticipated impacts of COVID-19	9
References	9
Glossary	10
Appendix	11
Appendix: Requirements	11
Background and Objectives	11
Work Assignment	12



Case studies (1)

Samoa – centralised system with periodic surveys – strong demand for support from International Organisations, co-ordination between Government agencies on funding to support surveys, also emphasis on efficiency/integration where possible – e.g. MICS/DHS integration

Ghana – centralised system with periodic surveys – <u>strong central role for NSO</u> driven by legislation, NSO documents and upholds standards on methodologies to promote harmonisation, provides technical support (e.g. sampling support) to other agencies

Costa Rica – centralised system with multiple regular surveys – ongoing major project to <u>redevelop household</u> <u>survey system</u> – greater integration of processes and harmonisation of practices and systems – process was more resource intensive than expected, key concern of breaks in series



Case studies (2)

Ireland – incremental process of development of system – significant project completed in 2018 to integrate household surveys, introduce telephone interviewing, **greater centralisation of processes** within the NSO (process approach), demands substantially set by **European legislation** but significant national demand for additional info (particularly related to COVID-19)

Canada – long path of development to current system – <u>highly centralised and integrated processes</u> for all household surveys run by statistics Canada, significant focus on metadata and related IT systems, dissemination by theme across sources – recent increased used of web-interviewing – particular focus on trying to improve reactiveness, key importance of senior level buy in to

SPC – regional organisation promoting increased <u>co-ordination and harmonisation</u> including promoting a common survey programme and harmonised tools, e.g. Census of Population, HIES - also co-ordination in funding activities – significant challenges to achieve harmonisation due to varying capacity of PICS and external demands which can disrupt planned activities



Key findings

- ☐ No single best fit system or approach
- □ Countries continually adapting their approaches based on technological or methodological advancement
- ☐ More advanced systems with regular activities
 - emphasis on greater integration of processes across surveys (e.g. centralised sampling, data collection, testing, estimation)
 - modernisation of methodologies (e.g. modes of data collection)
 - Integration of data sources
 - However, flexibility an issue and change processes are very resource intensive
- ☐ Systems with periodic collection
 - Insufficient resources to target wide integration of processes and instability of funding a key constraint
 - Importance to promote key co-ordination role of NSO to promote co-ordinated planning and apply consistent and harmonised standards
 - Important to have some documentation of standards
 - Higher importance of support from International and regional organisations and improved co-ordination requested



Key messages (cont.)

- ☐ Projects to modernise/intergrate are highly resource intensive
- □ Need to tackle internal resistance, external resistance (breaks in series) and high level buy-in critical
- □Common elements maintenance of clear standards and focus on harmonisation relevant in all cases
- □All countries were asked about COVID-19 impacts and lessons learned for household surveys
 - Unsurprising messages
 - Need to modernise data collection (mode) develop sampling frames with contact details (e.g. through Census), develop survey infrastructure to increase efficiency
 - Need for greater flexibility ability to activate supplementary data collection at short notice but with good standards and practices

COVID-19 Task Force

TECHNICAL GUIDANCE NOTE

Task Force on COVID-19 and household surveys

The COVID-19 pandemic has presented itself as both a tremendous challenge and an opportunity to household surv programmes in countries.

The overall objectives of the Task Force are to (a) support the coordination of the COVID-19 impact surveys in provide guidance to countries both on maintaining the continuity of regular survey programmes and various r aspects of COVID-19 related surveys; and (c) establish a collective vision on the implications of COVID-19 for renational household survey programmes.

National responses to COVID-19

Based on a survey carried out by UN Statistics Division and the World Bank, many national statistical offices have responded through innovative approaches.

COVID-19 impact surveys

Members of the Intersecretariat Working Group on Household Surveys are supporting countries to measure the impact of COVID-19 through sample surveys. As of 8 July 2020, we have supported about 180 countries through 480 surveys.

Methodology on (surveys

ISWGHS members h guidelines to help co out surveys to asses COVID-19.

Planning and Implementing Household Surveys

Annex 1. Checklist for planning and carrying out household surveys under COVID-19

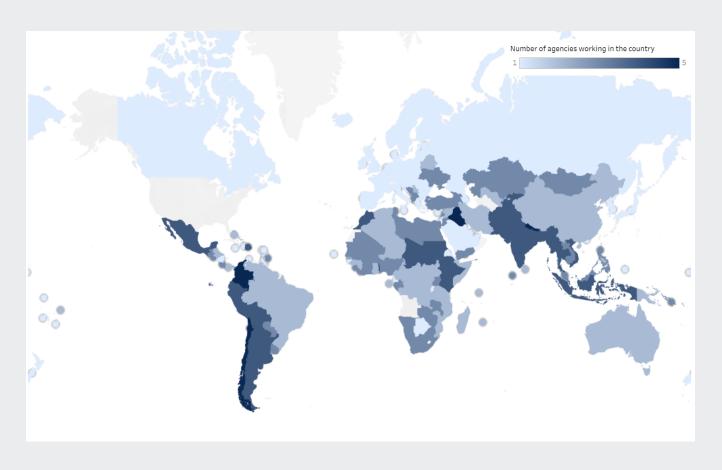
The checklist is created to facilitate the planning and organization of household surveys. More detailed explanation of each item is available in the text in its corresponding section.

1. General principles

- Minimizing the risk of COVID-19 transmission among household survey field staff and survey respondents
- Limiting the field data collection activity to the minimum that is necessary
- Ensuring high quality and timely data are available for policymaking at the national and local level.

Phase	Sub-phase	Task	Status
2. Planning	2.1. Setting/	Develop/revise survey objectives with all	
data collection	Revisiting survey	stakeholders	
	objectives	Prioritization of modules/questions to be	
		included	
	2.2. Assessing	Gather data on government restrictions, number	
	COVID-19 situation	of cases, overall transmission rate, transmission	
		hot spots	
	\	Assess whether face-to-face interview is plausible	
		at the national, regional and local level	
		Develop safety protocols to minimize the risk of	
		COVID-19 transmission	
		Assess the need for personal protective	
		equipment	
	2.2 Duilding the	Involve health authorities and experts in the	

Virtually all countries are now conducting phone surveys to monitor impact of pandemic and beyond, most with support from ISWGHS members ...





COVID-19 task force: Assessing and Minimizing the COVID Impact on Survey Quality

Table of Contents
Section

- 1. Introduction
- 2. Household survey operations pre- and during COVID: Potential impact of the pandemic on survey quality
 - 2.1 Survey error frameworks
 - 2.2 Potential ways the pandemic impacted major sources of error
- 3. Assessing quality and methods to reduce and correct errors
 - 3.1 Accuracy and reliability
 - 3.2 Timeliness and punctuality
 - 3.3 Coherence and comparability
- 4. Communicating Data Quality
- 5. Acknowledgements
- 6. References



Pag

SDG data disaggregation

TOOLRIT | GENDER DATA PRODUCTION AND COLLECTION

Counted and Visible Toolkit

05 Mar 2021

The <u>Counted and Visible. Toolkit to Better Utilize Existing Data from Household Surveys to Generate Disaggregated Gender Statistics</u> (Counted and Visible Toolkit) provides a compilation of tools and mechanisms used by several countries to produce evidence to inform gender-responsive policies and catalyze actions to leave no one behind. The selected countries are linked to UN Weema's global gender data programme, <u>Women Count.</u>

The toolkit was developed by UN Women, in collaboration with the <u>Intersecretariat Working Group on Household Surveys</u> (ISWGHS), and benefited from the outcomes of the <u>Counted and Visible qlobol conference</u> in 2020.

Explore the toolkit

SPACE SHORTCUTS

How to use Confluence

PAGE TREE

- Introduction
- > Chapter 1. Policy relevance for
- Chapter 2. Population groups (
- Chapter 3. An overview of chal
- Chapter 4. An overview of sam
- > Chapter 5. Guidance on sampli
- Chapter 6. Sampling versus sur
- Chapter 7. Documenting samp
- How to use Confluence
- Documents and literature

SAE4SDG
Pages
Pages
Pages
Pages
Pages
Pages
Page SHORTCUTS
Documents and case stu...
First all member meetin...
Months articles
PAGE TREE
How to use Confluence
Motivating SAE
Producing SAE
Communicating SAE
SAE by SDGs
SAE for vulnerable populatic
From SAE experiment to pro

SAE practices

Australia

Asian Development Bank

Dashboard 🔓 🖉 🛆 475 views

SAE4SDG

Created by UNSD Clarence Lio, last modified by Haoyi Chen on Jun 05, 2021

Welcome to the Toolkit for using Small Area Estimation for the SDGs!

In committing to the realization of the 2030 Agenda for Sustainable Development, Member States recognized that the dignity of the individuals is fundamental and that the Agenda's Goals and targets should be met for all nations and people and for all segments of society. Ensuring that these commitments are translated into effective action requires a precise understanding of the target populations and progress made in addressing their particular priorities.

To properly measure this, statistics need to be presented for different population groups and geographical areas. The Sustainable Development Goal (SDG) indicator framework has included an overarching principle of data disaggregation: SDG indicators should be disaggregated, where relevant, by income, sex, age, race, ethnicity, migratory status, disability and geographic location, or other characteristics, in accordance with the Fundamental Principles of Official Statistics.

To enable national statistical offices to estimate disaggregated indicators, guidelines are needed to support the process. The idea of writing guidelines on how to use statistical methods and, in particular small area estimation (SAE), to receive disaggregated statistical indicators is not new. Some focus on methodological aspects, others provide methodology in a specific program language or focus on a specific topic as poverty mapping. Usability of SAE for official statistics has also been carried out over the past 10 years. So how do these guidelines differ from the existing work?

The **SAE4SDG Toolkit** in Wiki is to provide a space to provide information on methods to produce disaggregated data through small area estimation. It aims to complement and use the existing methodological work and case studies to encourage and enable national statistical offices to use SAE for the monitoring of the SDGs. The Toolkit will be an evolving project/document that will incorporate newly available methods, case studies and practical examples in future versions. The Toolkit also focuses on key steps to help countries in moving from SAE experiment to official data production. Finally, the Toolkit aims to be a space for partners to document and include references for their work on small area estimation.

What to expec

The SAE4SDG Toolkit targets practitioners and technical staff in National Statistical Offices and other institutions within the National Statistical System that are interested in using! While the Toolkit provides information on SAE models and the process around building the models, it also offers discussions around elements that help countries make the transit production. This is to respond to the challenges that the use in official statistics is still rather limited even though the method has been around for a long time.

Sampling to Leave No One Behind

Created by Paul Pacheco (UNSD), last modified by Haoyi Chen just a moment ago

Introduction

Chapter 1. Policy relevance for SDG data disaggregation and implication on sampling

Chapter 2. Population groups covered by the Guidance

Chapter 3. An overview of challenges in sampling certain population groups

Chapter 4. An overview of sampling strategies

Chapter 5. Guidance on sampling various population groups: concepts and definitions, sampling methods and country practices

Income, poor and extreme poor

Sex (gender)

Age (including children and older persons)

Race and ethnicity (including people with African descendants)

☐ Migratory status (including migrants, forced displaced)

Disability

☑ Geographic location (including urban/rural)

✓ LGBT (gender identity and/or sexual orientation)

Drug users

Homeless

✓ Prisoners

Chapter 6. Sampling versus survey design - taking a holistic approach

Chapter 7 Documenting sampling methods



Communication and advocacy







REVAMPING WEBSITE
INCLUDING PORTAL FOR SDGS

WEBINARS AND CONFERENCES

BLOGS, NEWSLETTER, ANNUAL REPORT



Connect and serve our stakeholders

- ☐ Establish the network of survey focal points from NSOs
- ■Assess national needs
- ☐ Foster peer-to-peer learning

Please suggest household survey-related area(s) that we should be focusing on

Gridded Population Sampling; sampling hard-to-survey population Mobile phone data collection and quality

Survey integration, especially with remote sensing and mobile phone data; model validation

Methods and standard on microdata dissemination Guidance on ensuring "state of preparedness" of NSOs when

planning, conducting or analyzing household surveys

Guidance on gathering sensitive information (HR-based principles, ethical and safety considerations) - e.g., corruption, governance, violence

Inclusion of forcibly displaced populations in national HH surveys Measurement of migration intention

Better measurement poverty, and combining/distinguishing it from vulnerability

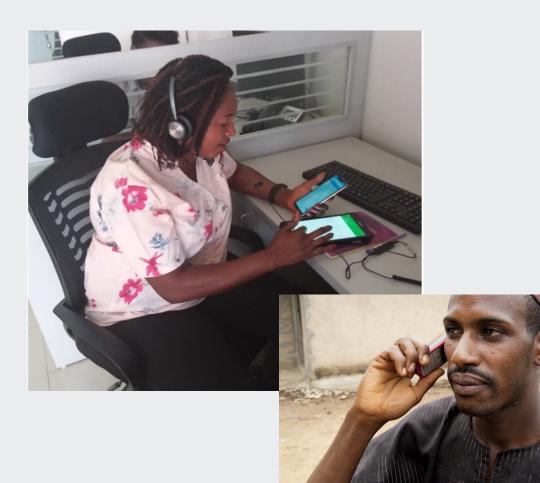


In less than a year from KL, we went ...

... from this ...



... to this!

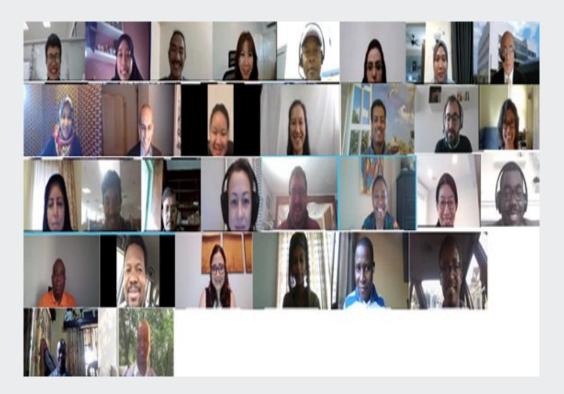




From this ...



... to this!





As of May 2020, virtually all countries had fully/partly stopped F2F surveys ...

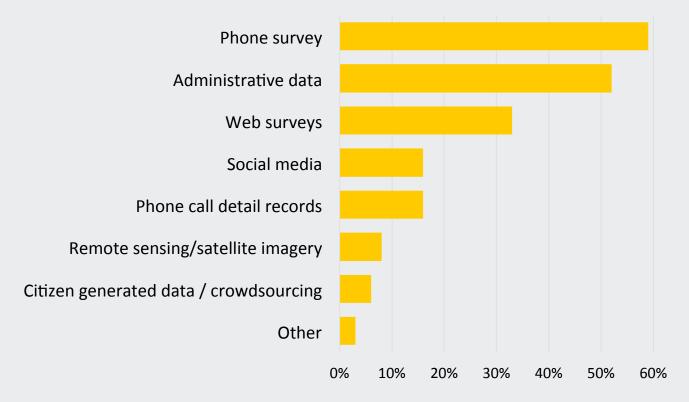
Stopped face-to-face data collection





... but many NSOs responded rapidly by adopting new data sources/modes!

Is your institution using alternative /nontraditional data sources/ approaches to analyze or monitor aspects of the COVID-19 pandemic?

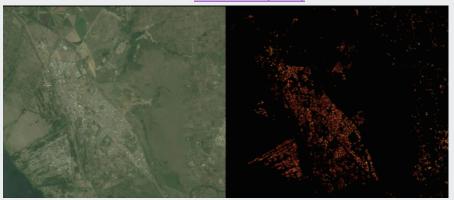




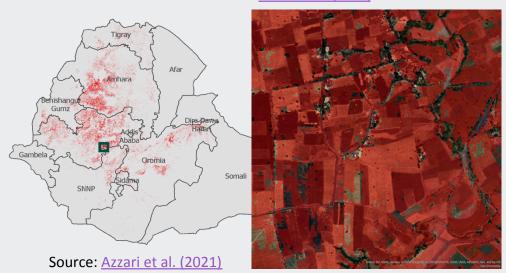
... and many are integrating EO and survey data and leveraging ML to improve spatial disaggregation and timeliness of development data ...



Source: Lobell et al. (2020)



Source: Tiecke et al. (2017)





... meanwhile, affordable sensors and other digital instruments are increasingly used in survey operations.







Source: Carletto et al. (2017)



Source: Sinha et al. (2020)

However ...

□ Lack of sound survey infrastructure: only 43% of 180 countries implementing phone surveys used a recent, updated sampling frame ☐ Institutional barriers: NSS not set up to foster interoperability □ Lack of financial and technical capacity: main concern voiced by (L/MICs) NSOs in UNSD/WB survey ■ More often, data integration and correction for bias done ex-post with data which are unfit-for-purpose ■EO/modelled estimates of crop production and other remote-sensed applications reveal huge differences with ground-based, "gold standards" measurements ☐ Huge demand from EO and ML communities for better ground-truthing data ☐ Technology availability and adoption unevenly distributed across countries and constrained by analytical capacity



Against this backdrop ...

Positioning household surveys for the next decade

- ☐ The document is being developed in response to the rapid changing landscape and the perceived need by member countries and DPs for a shared vision
- □ Disclaimer: document is still being developed. The plan is to continue consultation to get inputs and share draft at next UNSC
- □ 9 priority areas, including enabling environment to accelerate realization of vision
- ☐ Strike balance between foundational and frontier
 - x-country equity is strong consideration with focus on LICs and LMICs where needs are greatest and both foundational and frontier features are weakest
- □Goal not to be comprehensive but ensure that key priorities are included, in light of recent developments and potential for medium-term impact
 - This webinar is part of our consultations and we hope to get your feedback



9 priority areas

- 1. Enhance interoperability and integration of survey data
- 2. Improve sampling efficiency and coverage
- 3. Scale up adoption of improved methods and affordable technologies
- 4. Invest in capacity and research on CAPI, CATI, CAWI and mixed-mode
- 5. Understand and address non-response
- 6. Systematize collection, storage and use of meta/paradata
- 7. Expand capacity and use of machine learning and AI
- 8. Improve access, discoverability and dissemination of microdata
- 9. Foster a stronger operating environment



Enhance interoperability/integration of survey data

- □Interoperability as one key attribute for (survey) data to generate value for development (Jolliffe et al, 2021)
 - Improve timeliness, cost-effectiveness, accuracy, and granularity of insights
 - Address issues of coverage and respondent burden
- Interoperable surveys as a calibration/validation instrument for other data sources; examples in agriculture, poverty measurement and monitoring nutritional outcomes
 - Achieve economies of scale and scope through systematic collection of ground-truthing layers with multiple applications/uses
- ☐ To ensure interoperability, must act at design stage (e.g., "poverty mapping"); it requires "interoperable institutions/actors"
- ☐ To maximize benefits of interoperability, must promote data access while ensuring privacy
 - Need more research on spatial anonymization of public use datasets and comparative assessment of trade-offs between risk of disclosure and losses in empirical utility
 - Tailor to different users: spatially-anonymized public use datasets (w/ random offsets) versus confidential data accessed securely in data enclaves
- □ Need to enhance statistical capacity in LICs on data integration and design, implementation and analysis of interoperable surveys



Improve sampling efficiency and coverage

- Lacking/outdated censuses and under-coverage of important populations of interest major limitation of population-based survey frames
- Increase reliance on satellite imagery/geospatial data (e.g., <u>HRSL</u>) and multiple frames, particularly in conflict settings and to target hard-to-reach areas and population groups
- □ Increase piloting (and eventually scale up) of responsive/adaptive sampling techniques to meet specific needs
- □COVID-19 phone surveys clearly show advantages of using recent nationally-representative F2F surveys as sampling frames
 - Requires a systematic approach to collecting contact info and tracing F2F survey respondents
 - Evidence on leveraging F2F survey data on bias adjustment at the household-level, but...
 - o Limits to bias-adjustment at the individual-level (Brubaker et al., 2021)
 - Coverage concerns remain as a function of aging F2F survey samples used as frames
- □Clear demand from countries to enhance statistical capacity and promote sustainable use of advanced sampling techniques



Scale up adoption of improved methods and affordable technologies

- □ Non-classical measurement errors (NCME) in survey data have been shown to bias empirical analyses and policy recommendations
- Adopting improved survey methods, both "high-tech" sensors (e.g., handheld GPS for land area measurement, accelerometers for physical activity tracking) and "low-tech" objective measures (e.g., crop cutting for crop yield measurement), can help eliminate NCME
 - Cost may still be an issue. Multiple use/sharing? Imputation? Optimal size of sub-samples?
- High-frequency phone surveys in mixed-mode can reduce recall bias and improve timeliness/frequency
 - There are, however, unanswered operational questions (e.g., assign cell phones? Set incentives and how much?)
- □ A business line on experimental statistics *cum* capacity strengthening can help streamline methodological research and development into NSO work program and help promote scaling up of validated methods



Invest in capacity and research on CAPI, CATI, CAWI and mixed-mode surveys

- □ Rapid transition from PAPI to CAPI over the past decade; pandemic-induced move to CATI and CAWI in L/MICs; significant scope for building on this momentum to ensure more sustainable reliance on mixed-mode for recurrent events and crisis monitoring
- □ Capacity and infrastructure remain a constrain across NSOs in many countries (UNSD/WB, 2021)
- ☐ Need to invest in apps and tools for facilitating CATI/CAWI/mixed-mode surveys, including protocols for respondent selection, incentive provision and tools for data management
- □Also, more randomized experiments and research needed on mode effect to understand data quality and comparability, likely to vary by variable



Understand and address non-response

- ☐ With rapid urbanization and increasing income levels, non-response rates rising, even in LICs
 - Increasing complexity of instruments, waning trust in public institutions and privacy concerns making things worse
 - Shifts to new modes of data collection e.g. phone, web further accentuating problem
- ☐ More research and harmonization needed to guide ex-ante survey design choices to reduce response burden and increase trust/cooperation (e.g. contact protocols, questionnaire content and length, training protocols, ...)
- □ Ability for ex-post corrections depends on availability of data, thus focus on design stage and metadata



Systematize collection, storage and use of meta/ paradata

- □CAPI/CATI/CAWI generating enormous amount of paradata which, if analyzed in real time, can be a game changer in fieldwork supervision and quality control.
- ☐ Possible uses include assignment of survey features based on previous responses and respondent's profile, gather info on contact burden, evaluate adoption of new data collection modes, etc.
- □Also, systematic collection and use of metadata based on harmonized protocols extremely helpful for quality assurance as well as for treatment of non-response.
- □ Investments in protocols, apps and statistical tools for collection, storage and use/sharing of meta/paradata should be prioritized.



Expand capacity and use of ML and Al

- □AI, ML and predictive analytics have great potential for virtually every step in the survey value chain, from questionnaire design to data collection and processing, to data use and dissemination
 - Predicting attrition in longitudinal surveys
 - Auto-coding of open-ended questions
 - Fast-tracking data editing and imputation
 - High-resolution mapping of development outcomes through data integration
 - Use of paradata for quality assurance
- ☐ Efforts still quite scattered and concentrated in HICs.



Improve data access, discoverability and dissemination of microdata

- □Address issue of underutilization of data data use often constrained by data access and fitness-for-purpose of existing data
- □ Long-standing efforts e.g. IHSN, Microdata Library, IPUMS ... but plagued by underinvestment and disincentives
- □ Data integration rest on improving access of both surveys (eg georeferencing) and other data sources (eg admin data)
- New types of data (e.g. georeferenced, rare populations, social media and other private data, ...) create additional hurdles, raising additional privacy and ethical concerns; need new protocols and methods
- □ Capacity and IT infrastructure, particularly in LICs, remain constraint
- ☐ Suite of options by type of user/data? Institutional setting?



Fostering a stronger operating environment at national, regional and global level

- ☐ Strengthen engagement with data users and policy makers ☐ Make research on survey methods easier to conduct and more "appealing" to researchers □ Invest in ICT infrastructure for remote work, training, data collection □Support (sub-regional?) hubs to strengthen capacity in use of new data sources and frontier applications/methods □ Sustain financing and new financing models ☐ Foster more coordinated and systemic approach to support NSOs • Sustain coordinated investments and research on global standards and data public
 - Support a stronger role of ISWGHS

goods



Some final thoughts ...

available resources. The ISWGHS can help!

☐For household surveys to remain relevant and grow, we must strike the right balance between foundational and frontier work
□Alternative data sources provide an opportunity to add value to household surveys while also help increase spatial and temporal granularity
☐The rapid diffusion of new, affordable technology a game changer to improve accuracy, coverage and use of household surveys
☐ The Covid-19 pandemic —potentially a major blow to the future of F2F surveys — has instead accelerated the process of modernization and innovation already in place
☐ More attention should go on quantifying benefits and communicating the value of survey data
☐Rigorous methodological research addressing key measurement issues should be more systematic and systemic, as well as context-specific
☐No technical solution will succeed and achieve scale without a proper enabling

environment. Thus, the need for greater coordination and a more efficient use of





Positioning Household Surveys for the Next Decade

Haoyi Chen chen9@un.org

