A Thing of the Past? Household Surveys in the New Global Data Landscape

Gero Carletto Development Data Group The World Bank





A few take-home messages

- No single data source will do the (SDG) job!
 - Household surveys remain critical, but ...
 - More emphasis on data integration
 - Adding value to instruments
- Go beyond indicators
 - For evidence-based policy, need RHS as well, preferably in same instrument
 - Also, better sampling and imputation methods
- How you measure it matters!
 - Need for common standards, rigorously validated
- Importance of equitable adoption of improved, cost-effective methodologies and technologies
 - Implications for development and choice of best, <u>scalable</u> methodology
 - NSO's involvement



 Demand for data and evidence has increased

 SDG provides a unique opportunity

Country-driven











- •17 goals
- •169 targets
- •232 indicators

Themes

- People
- Planet
- Prosperity
- Peace
- Partnerships





Where will the data come from?

- No single data source will do the job
 - Improve individual data sources
 - Integrate different, new data sources
- New data sources need validation
 - Calibration/groundtruthing of remote sensing
 - Citizen-generated data
 - Machine learning/AI

- Data sources
 - Administrative data
 - e.g. CRVS, ag routine data, ...
 - Geospatial data
 - Big data
 - Censuses
 - Household surveys





Household Survey Data & the SDGs





1/3

of





Household Survey Data & the SDGs

- 77 indicators in total identified as coming from household surveys
- Goal 3 with highest number followed by goals 16, 8, 5, 7, 1 and 2
- About 80% are either Tier I or Tier II, 13 of the indicators are Tier III

By Goal:	Tier I	Tier II	Tier III	Mixed	Total
Goal 1: Poverty	2	2	2	0	6
Goal 2. Hunger	4	0	1	0	5
Goal 3. Health	8	9	1	0	18
Goal 4. Education	1	3	1	2	7
Goal 5. Gender equality	2	7	0	0	9
Goal 6. Water and sanitation	2	0	0	0	2
Goal 7. Energy	2	0	0	0	2
Goal 8. Decent work	6	2	1	0	9
Goal 9. Infrastructure	1	0	0	0	1
Goal 10. Inequality	1	0	3	0	4
Goal 11. Cities	1	1	1	0	3
Goal 16. Justice	1	6	3	0	10
Goal 17. Partnership	1	0	0	0	1
Total	32	30	13	2	77

Mitra and Walsh, 2017





Moving beyond indicators!



To understand, not only monitor, we need an integrated approach involving ...

- Integration within same instrument
 - Cost saving
 - Analytical advantages ... but also drawbacks!
- Integration <u>across</u> data sources
 - Need better methods (survey to survey imputation, smarter sampling...)





Zooming In On ...





1 Poverty

Data Availability

Much progress...

Figure 1: Number of Poverty Data since 1976





Serajuddin et al. (2016)

Data Availability

...but large gaps remain



92 low/middle income countries do not have a multi-topic survey every 3 years, as per the WB President's commitment

- No data: mainly in EAP and LAC small countries
- Only 1 point: mainly in AFR
- **77** with "extreme" deprivation (> 5-year interval)
- Irregular (ad hoc) survey implementation

But also, beyond data deprivation, issues with:

- Uncertainty of funding: many more (IDA) countries "at risk"
- Data **quality** (reliability, comparability) and accessibility
 - E.g., only 27 of 48 SSA countries have at least two comparable surveys between 1990-2012

Note: number of data deprived countries estimated based on surveys conducted during 2002-2011



Poverty-related surveys, 1994

Data deprivation: Number of poverty surveys per decade available via the World Bank, 1994 All values refer to the decade that ends at the shown year (e.g. 2013 refers to 2004 to 2013).





1 NO POVERTY

Poverty-related surveys, 2014

Data deprivation: Number of poverty surveys per decade available via the World Bank, 2014 All values refer to the decade that ends at the shown year (e.g. 2013 refers to 2004 to 2013).





NO Poverty



Income vs. consumption





Pros and cons ...







East Asia & Pacific



Surveys	Consumption vs Income	Household Size
Cambodia 2011	Consumption	Per Capita
Indonesia 2016	Consumption	Per Capita
Lao PDR 2012	Consumption	Per Capita
Malaysia 2016	Income	Per Capita
Mongolia 2016	Consumption	Per Capita
Myanmar 2015	Consumption	Per Capita Per Adult Equivalent
Philippines 2015	Income	Per Capita
Timor-Leste 2014	Consumption	Per Capita
Vietnam 2016	Consumption	Per Capita



Europe and Central Asia



Surveys	Consumption vs Income	Household Size
Armenia 2015	Consumption	Per Adult Equivalent
Bosnia and Herzegovina 2004	Income	-
Kosovo 2015	Consumption	Per Adult Equivalent
Kyrgyz Republic 2013	Consumption	Per Capita
Macedonia 2017	Income	Per Adult Equivalent
Moldova 2013	Consumption	Per Adult Equivalent
Russian Federation 2008	Consumption	Per Adult Equivalent
Tajikistan 2014	Consumption	Per Capita

Latin America & Caribbean



Surveys	Consumption vs Income	Household Size
Argentina 2016	Income	Per Capita Per Adult Equivalent
Bolivia 2015	Income	Per Capita
Colombia 2017	Income	Per Capita
Ecuador 2013	Consumption	Per Capita
Ecuador 2018	Income	Per Capita
El Salvador 2015	Income	Per Capita
Guatemala 2014	Consumption	Per Capita
Haiti 2012	Consumption	Per Capita
Honduras 2016	Income	Per Capita
Mexico 2016	Income	Per Capita
Nicaragua 2014	Consumption	Per Capita
Panama 2008	Consumption	Per Capita
Paraguay 2017	Income	Per Capita
Peru 2017	Consumption	Per Capita





Middle East & North Africa



Surveys	Consumption vs Income	Household Size
Egypt 2008	Consumption	Per Capita
Iraq 2012	Consumption	Per Capita
Jordan 2010	Consumption	Per Capita
Lebanon 2011	Consumption	Per Capita
Djibouti 2017	Consumption	Per Adult Equivalent
Morocco 2013	Consumption	Per Capita
West Bank and Gaza 2011	Consumption	Per Adult Equivalent
Yemen 2005	Consumption	Per Capita

South Asia



Surveys	Consumption vs Income	Household Size
Afghanistan	Consumption	Per Capita
Bangladesh 2016	Consumption	Per Capita
Bhutan 2017	Consumption	Per Capita
Sri Lanka 2016	Consumption	Per Capita
India 2011	Consumption	Per Capita
Pakistan 2013	Consumption	Per Adult Equivalent
Maldives 2016	Consumption	Per Capita

1 NO Poverty

Ňŧ**Ť**ŧ



Sub-Saharan Africa



Surveys	Consumption vs Income	Household Size
Côte d'Ivoire 2015	Consumption	Per Capita
Kenya 2015	Consumption	Per Adult Equivalent
Malawi 2010	Consumption	Per Capita
Mozambique 2014	Consumption	Per Capita
Nigeria 2010	Consumption	Per Adult Equivalent
South Africa 2014	Consumption	Per Capita
Tanzania 2014	Consumption	Per Adult Equivalent
Uganda 2011	Consumption	Per Adult Equivalent
Zambia 2015	Consumption	Per Adult Equivalent
Zimbabwe 2011	Consumption	Per Capita

How You Measure it Matters!





Measuring Consumption

Lisa C. Smith, Olivier Dupriez and Nathalie Troubat. Assessment of the Reliability and Relevance of the Food Data Collected in National Household Consumption and Expenditure Surveys. International Household Survey Network Working Paper No. 008, February 2014.



INTERNATIONAL HOUSEHOLD SURVEY NETWORK

Assessment of the Reliability and Relevance of the Food Data **Collected in National Household Consumption and Expenditure** Surveys

Lisa C. Smith Olivier Dubriez Nathalie Troubat

> **IHSN Working Paper No. 008** February 2014

www.ihsnorg

IHSN

Poor Harmonization Across All Criteria



1 NO POVERTY

Note: The percent meeting all criteria is based on the 93 countries with no missing data.

* Food consumed away from home.

Poor Harmonization of Recall Periods





1 Poverty

Methods Matter!



Beegle et al. (2012). Methods of household consumption measurement through surveys: Experimental results from Tanzania. Journal of Development Economics Volume 98, Issue 1, May 2012, Pages 3-18.



Snacked on gromperekichelcher lately? Food Away from Home



% of food consumption coming from FAFH



% from energy intake coming from FAFH



Food Away From Home – Needs Harmonization

1 NO POVERTY

 90 out of 100 surveys collect some information on FAFH (Smith et al. 2014), of those:



Food Away From Home – Needs Harmonization

1 NO POVERTY

 90 out of 100 surveys collect some information on FAFH (Smith et al. 2014), of those:



Food Away From Home – Needs Harmonization

1 NO POVERTY

 90 out of 100 surveys collect some information on FAFH (Smith et al. 2014), of those:



Food Away From Home

Poverty rates over time, Peru 2010-2013



Farfán, G., Genoni, M. E., & Vakis, R. (2017). You are what (and where) you eat: Capturing food away from home in welfare measures. Food Policy. 72, 146-156.





Food Away From Home

Poverty rates over time, Peru 2010-2013



Gini coefficient: Peru 2010-2013











LEMST

The Use of Non-Standard Units for the Collection of Food Quantity

Classificant for large unit of the terms of the second sec

C THE MORE A BANK

LSMS





Special Yeave, Food Perksy

second to a laman plot

Food Counts: Measuring Food Consumption and Expenditures in Household Consumption and Expenditure Surveys

Touling Toulin

Their York of social top can be very bold for anyon for an easily of much internation of prively. Two intensity, with infilter top boards or which advanced carry of disrupt interface, of these provides of the operating their theorem allocs a final construction and apport boards or infold any path someone internation of a source of any path someone provide top of a source of any path someone.

indexending the welfare of a service proprogram. Next of the consequence point. The opposite terms of means indexes on the second consequence and the undexes in Next of Second Second 2019 address in Next Oct (Next Oct (Next

protificient worklands and also protific policy-views with transactions with a letter rentermenting of powers for rescaling proceedings for which a contraction of





More methodological guidelines ...

Published or forthcoming LSMS Guidebooks on:

- Food Consumption
- Education expenditures
- Use of non-standard units
- Land area measurement
- Soil quality measurement
- Livestock
- Fisheries
- Forestry

- Labor
- Disability
- Conflict
- Migration
- Climate change adaptation
- Justice
- Service delivery
- Energy





How you measure it (in Ag) matters!

A look at yields...





Gourlay, S., Kilic, T., & Lobell, D. (2017). *Could the debate be over? errors in farmer-reported production and their implications for the inverse scale-productivity relationship in Uganda*. The World Bank. [*Forthcoming in JDE*]

Data Quality - Methods Matter

Over-Estimation in Self-Reported Yields Varies by GPS-Based Plot Size = Non-Random Error



LSMS

2 ZERO HUNGER

Gourlay, S., Kilic, T., & Lobell, D. (2017). *Could the debate be over? errors in farmer-reported production and their implications for the inverse scale-productivity relationship in Uganda*. The World Bank. [*Forthcoming in JDE*]



Data Quality - Methods Matter

Bias in Self-Reported Land Area = Non-Random Error





Carletto, C., Gourlay, S., Murray, S., & Zezza, A. (2017, October). Cheaper, Faster, and More Than Good Enough: Is GPS the New Gold Standard in Land Area Measurement?. In *Survey Research Methods* (Vol. 11, No. 3, pp. 235-265).



On new data sources ... and the need for validation!



Every company has big data in its future and every company will eventually be in the data business.

— Thomas H. Davenport —



Big Data is like teenage sex: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it.

— Dan Ariely —





New data sources – Use, Validation & Integration

- New data sources offer tremendous opportunities, e.g.
 - Night light for poverty estimation
 - Crowdsourcing, citizen-generated data
 - Remote sensing for yield estimation
- Use of nationally representative HH surveys for correcting for selection bias of citizen-generated data
- HH surveys can be instrumental in ground truthing and calibrating of remote sensing models



Source: Strobl. forthcom



Big Data, Big Errors



- Integration of data sources to address biases and maximize accuracy, efficiency, and value
- For example:
 - Zillow, the online database of housing units, holds more than 200 million records BUT sample survey data with ~6000 records may provide better estimates
 - Survey quality often trumps Big Data quantity, as was the case for Zillow data. (Paul Biemer)
- Institutional constraints of data integration

Measuring yields using remote sensing

- Status quo: self reporting
- Gold standard: Crop cutting (full plot)
- Scalable solution: crop cut on subplots (cum imputation)
- Innovation: high-resolution remote sensing combined with crop modeling based on groundtruthing (hh survey) data







New data sources – Use, Validation & Integration







Uganda



Brazil







Measuring yields using remote sensing

- Evidence from Uganda



Source: Lobell, D. B., Azzari, G., Burke, M., Gourlay, S., Jin, Z., Kilic, T., and Murray, S. (2018). <u>"Eyes in the sky, boots on the</u> <u>ground: assessing satellite- and ground-based approaches to crop yield measurement and analysis in Uganda."</u> World Bank Policy Research Working Paper No. 8374. *[Forthcoming in AJAE]*



Measuring yields using remote sensing - Evidence from Uganda





Measured from space, but validated on the ground...







Lower accuracy, high precision

Source: Lobell, D. B., Azzari, G., Burke, M., Gourlay, S., Jin, Z., Kilic, T., and Murray, S. (2018). "Eves in the sky, boots on the ground: assessing satellite- and ground-based approaches to crop yield measurement and analysis in Uganda." World Bank Policy Research Working Paper No. 8374.

- Evidence from Uganda



Measuring yields using remote sensing





Higher accuracy, less precision

Lower accuracy, high precision

Source: Lobell, D. B., Azzari, G., Burke, M., Gourlay, S., Jin, Z., Kilic, T., and Murray, S. (2018). <u>"Eyes in the sky, boots on the ground:</u> <u>assessing satellite- and ground-based approaches to crop yield measurement and analysis in Uganda.</u>" World Bank Policy Research Working Paper No. 8374.







Measuring yields using remote sensing - Evidence from Uganda



Higher accuracy, higher precision

Higher accuracy, less precision

Lower accuracy, high precision

Source: Lobell, D. B., Azzari, G., Burke, M., Gourlay, S., Jin, Z., Kilic, T., and Murray, S. (2018). <u>"Eyes in the sky, boots on the ground:</u> assessing satellite- and ground-based approaches to crop yield measurement and analysis in Uganda." World Bank Policy Research Working Paper No. 8374. *[Forthcoming in AJAE]*



"Without data you're just another person with an opinion."

– W. Edwards Deming



"TORTURE THE DATA, AND IT WILL CONFESS TO ANYTHING."

-RONALD COASE, ECONOMICS, NOBEL PRIZE LAUREATE



Home

About the Survey Unit

Living Standards Measurement Study (LSMS)

About LSMS

• Our Work

Programs

News

Events

Publications

Newsletter

Multimedia

Microdata Library

Software



FEATURED

Study

< 4/4 >

Exploring Youth's Role and Engagement in African Rural Economies

Generating high-quality data, improving survey methods, and building capacity

How do young Africans engage in the rural economy? Learn how an interdisciplinary research project is answering this question with LSMS data. Read More $\!$



surveys.worldbank.org/lsms



