AGRODEP Household survey data course Dakar, 8-10 October 2012

Overview of household survey data









Acknowledgement

Kathleen Beegle^a, Gero Carletto^a, Kristen Himelein^a, Juan Muñoz^b, Talip Kilic^a, Kinnon Scott^a, Diane Steele^a

^a World Bank ^b Sistemas Integrales

Context

- Part of increased emphasis on performance-based management. *Examples*:
 - Performance of public administrations
 - Effectiveness of aid
 - □Impact of country social policies
- 2. Information is needed on inputs, outputs, outcomes and impacts (information on outcomes and impacts has to come from households)
- 3. Demand for good poverty data and poverty monitoring systems has greatly increased

1. Enhancing social policies

- Detect issue/problem
- Identify determinants of observed outcomes
- Simulate changes resulting from alternative policies
- □ Monitor performance
- Evaluate impact

2. Assessing information needs



Data Sources

- National accounts
- Current public expenditure statistics
- Program of Price collection (cons./prod.)
- Administrative Records (from line ministries)
- Qualitative Work
- Surveys:

-Household/Community

- Enterprise
- Facilities

The Demand for Data

- Performance-based management
 - Is the public sector delivering good services? Are they properly targeted?
 - Are country policies/poverty reduction strategies reducing poverty?
 - Is aid supporting poverty reduction?
 - In the World Bank: e.g. "Results-based"
 Country Partnership Strategies

The Demand for Data

- Performance-based management
- Millennium Development Goals (MDGs)
 - MDG 1: Eradicate extreme poverty and hunger
 - MDG 2: Achieve universal primary education
 - MDG 3: Promote gender awareness, empower women
 - MDG 4: Reduce child mortality
 - MDG 5: Improve maternal health
 - MDG 6: Combat HIV/AIDS, malaria and others
 - MDG 7: Ensure environmental sustainability
 - MDG 8: Develop a global partnership for development

MDGs 1 - 3

		LSMS/IS
GOALS	INDICATORS	(usually)
1. Eradicate extreme	 Proportion of population below \$1 a day 	Yes
poverty and hunger	 Poverty gap ratio (incidence x depth of poverty) 	Yes
	Share of poorest quintile in national consumption	Yes
	 Prevalence of underweight in children (under five years of age) 	Yes
	 Proportion of population below minimum level of dietary energy consumption 	No
2. Achieve universal primary education	 Net enrollment ratio in primary education 	Yes
	 Proportion of pupils starting grade 1 who reach grade 5 	Yes
	 Literacy rate of 15 to 24-year-olds 	Yes
3. Promote gender equality and	 Ratio of girls to boys in primary, secondary, and tertiary education 	Yes
empower women	Ratio of literate females to males among 15- to 24-year-olds	Yes
	 Share of women in wage employment in the nonagricultural sector 	Yes
	 Proportion of seats held by women in national parliament 	No

MDGs 4 - 8

		LSMS/IS
GOALS	INDICATORS	(usually)
4. Reduce child mortality	IMR and immunizations: usually not, but possible	
5. Improve maternal	 Maternal mortality ratio 	No
health	 Proportion of births attended by skilled health personnel 	Yes (births in last
		2 years)
6.Combat	 HIV prevalence among 15- to 24-year-old pregnant women 	n/a
HIV/AIDS, malaria,	 Contraceptive prevalence rate 	No
and other diseases	 Number of children orphaned by HIV/AIDS 	No
	 Prevalence and death rates associated with malaria 	No
	 Proportion of population in malaria-risk areas using effective 	prevention: Yes
	malaria prevention and treatment measures	treatment: No
	TB indicators	No
7. Ensure	 Land use, GDP per unit of energy use, Carbon dioxide emissions 	No
environmental	 Proportion of population with sustainable access to an improved 	Yes
sustainability	water source, access to improved sanitation, access to secure	
	tenure	
8. Develop a global partnership for development	No for most (with exception of unemployment rate of 15- to 24-year-olds))

The Demand for Data

- Performance-based management
- Millennium Development Goals (MDGs)
- Poverty Reduction Strategies (PRSP)
 - Measure welfare/poverty
 - Identify problems--magnitude, causes
 - Alternative policies
 - Cost/benefit
 - Monitor
 - Evaluate

The Demand for Data

- Performance-based management
- Millennium Development Goals (MDGs)
- Poverty Reduction Strategies (PRSP)
- General Demand
 - Poverty and Inequality
 - Poverty Mapping
 - Benefit Incidence Analysis
 - Public services
 - Determinants of observed outcomes
 - <u>Targeting of programs</u>
 - Inputs to Program Design
 - Impact Evaluation
 - RCT vs. quasi-experimental design
 - Research

Household Data

- Variety of types of data about and from households/individuals:
 - Administrative data
 - Case studies
 - Qualitative/participatory assessments
 - Censuses



Heterogeneity in Surveys

- Initial purpose of the survey drives the way survey is designed and implemented
 - Different agenda \rightarrow Different instrument
- An increasingly crowded field...

Instrument	Sponsor
Censuses	UNFPA
Income Expenditure /Budget Surveys (IES/HBS)	Central Banks, IMF, NSOs
Labor Force Surveys (LFS)	ILO
Demographic and Health Surveys (DHS)	USAID
Multiple Indicator Cluster Surveys (MICS)	UNICEF
Core Welfare Indicator Questionnaires (CWIQ)	UNDP, DfID WB Africa Reg.
Welfare Monitoring Survey (WMS)	Stat Norway
Statistics on Income and Living Conditions (SILC)	Eurostat
Comprehensive Food Security and Vulnerability Analysis (CFSVA)	WFP
Integrated, Multi-Topic Surveys [Living Standards Measurement Study (LSMS), Integrated Surveys (IS), Family Life Surveys (FLS)]	World Bank RAND NSOs

Heterogeneity in Surveys

- Dimensions of a possible typology ...
 - 1. "Representativeness" (sampling)
 - 2. "Directness" of measurement
 - 3. Analytic complexity
 - 4. Respondent Burden
 - 5. Methods

Degree of Representativeness

Case	Purposive	Quota	Small prob.	Large prob.	Census
study	selection	sampling	sample	sample	

Subjective/Objective Dimension



Tools to gather information from households



Tools to gather information from households



Household Budget Surveys (HBS)

- Purpose: collect information on household expenditures to produce or update the weights for consumer price indices as well as to provide inputs for national accounts.
- Countries often add modules on income to their HBS in order to facilitate the measurement of national income as well. (then IES)
- Restricted set of questions that often mimic what is captured in the decennial population and housing census.
- Topics can include:
 - basic demographic information
 - education levels
 - and employment status
 - agricultural module (rare)
- Supported by Central Banks, IMF, EUROSTAT, WB

Labor Force Survey (LFS)

- Purpose: Measure and monitor indicators of a country's economic situation; for planning and evaluating many government programs.
- Done monthly in many developed countries; quarterly or annually or less in most developing countries.
- Topics include those related to labor:
 - employment,
 - unemployment,
 - Earnings,
 - hours of work,
 - occupation, industry, and class of worker.
 - Supplemental questions-- income, previous work experience, health, employee benefits, and work schedules
 - May ask other sources of income/poverty measurement
- Supported by Ministry of Labor, NSOs, ILO definitions

Demographic and Health Surveys (DHS)

- Purpose: collect data on health, primarily maternal and infant health, but not limited to this, and demography.
- Started in 1984 (continuation of the World Fertility Survey and the Contraceptive Prevalence Surveys that had been done previously.)
- Done in > 80 countries (> 210 standard DHS done)
- Women in reproductive age
- Topics usually covered by the surveys include:
 - basic characteristics of the household and the respondents,
 - child health, education
 - family planning, fertility and fertility preferences
 - HIV/AIDS knowledge, attitudes and behavior,
 - infant and child mortality,
 - maternal health,
 - nutrition
 - socio-economic indicators based on asset ownership
- Supported by USAID through Macro Int'l.

The Multiple Indicator Cluster Surveys (MICS)

- Purpose: Monitor progress on the 1990 World Summit for Children Goals
- Assessing progress on HIV/AIDS and malaria reduction
- Four waves so far, 62 countries in MICS IV, starting MICS V in 2012
- Main topics covered
 - MDGs
 - nutrition
 - child health and mortality
 - water and sanitation
 - housing
 - reproductive health and contraceptive use
 - literacy, child protection
 - labor
 - domestic violence
- Supported by UNICEF

Core Welfare Indicator Questionnaire (CWIQ)

- Purpose: Measure and monitor a limited range of human development indicators, on access, utilization and satisfaction with social services
- Mainly done in Africa region
- In conjunction with IHS-type baseline?
- Topics- indicators:
 - Roster
 - Education- use
 - Health-use
 - Sanitation
 - Correlates of poverty ... consumption?
- Supported by World Bank

Living Standards Measurement Study (LSMS) Surveys

- Long tradition, started in 1980s
- Purpose: Measure poverty plus study household behavior, welfare, interactions with government policies: determinants of outcomes, and linkages among assets/ characteristics of households/livelihood sources/government interventions.
- Topics include (inter alia)
 - HH composition Consumption
 - Education

 - Labor
 - Migration

- Agriculture
- Health/Anthro HH enterprises
 - Other
 - Community characteristics, prices
- Credit Use
 Facility characteristics
- Supported by World Bank, UN agencies, IADB, bilateral agencies, governments

Survey	Sample - hhlds	Geographic desegregation	Freq. data collection	Period of data collection	No., visits	Interview Duration
Censuses	All hhlds in country	Any level	10 years	1 day to 1 month	1	¹ ∕₂ hour
Income / Expenditure Surveys (IES)	2,000-20,000	3-10 regions Urban/rura ¹	1-5-10 years	12 months	5-10	1-2 hours per visit
Labor Force Surveys (LFS)	5,000-50,000	5-20 regions Urban/rural	Month5 yrs	3 months	1	30 minutes per active hh member
Demographic and Health Surveys (DHS)	5,000-20,000	5-20 regions Urban/rural	5-10 years	3-4 months	1	2-4 hours
Multiple Indicator Cluster Surveys (MICS)	2,000-15,000	<5 regions Urban/rural	3-5 years	3 months or less	1	1 hour
Core Welfare Indicator Questionnaires (CWIQ)	5,000-15,000	5-20 regions Urban/rural	Once or twice	1 month	1	< 1 hour
Integrated, Multi- Topic Surveys (LSMS/IS/FLS)	2,000-5,000	3-8 regions Urban/rural	3-5years	2-12 months	1 or more	1-3 hours per visit

Multi-topic Surveys

	Single-Topic (e.g. LFS)	Multi-Topic (e.g. LSMS)
Questionnaire	Small	Large
Sample	Large	Small
Frequency	High	Low

Surveys and Policy Analysis



Implications for Survey Design Individual/Household level information critical Multi-topic needed **Community**/spatial level data supplements

The thinking behind multi-topic surveys

- Need to <u>understand</u> living standards, poverty, inequality and the correlates and determinants of these- not just monitor.
- Unit of analysis is the household, as both a consuming and producing unit
- One survey collecting data on a range of topics is a more powerful tool for policy formulation than a series of single purpose surveys: the sum is greater than the parts
 - Farmers are diversified
 - Poverty and FS are multidimensional

The thinking behind the LSMS survey

- Demand driven: implemented in a specific country as needed
- Priority given to meeting the policy needs of each country, but an eye to x-country comparability
- Implications
 - no standard set of LSMS questionnaires: content, length and complexity varies by country and, at times, over time within a given country.
 - Questionnaire development- lengthy process linking data users, stakeholders and data producers
 - Capacity building, sustainability

What is an LSMS Survey?

DISTINCTIVE FEATURES:

Multi-topic Questionnaire

Modules LSMS Questionnaires

- □ Food expenditures
- Home production
- □Non-food expend.
- Housing
- Durable goods

- Household roster
- Housing
- Education
- Mental Health

- □Non-farm Self-Empl.
- Agric. Activities
- Labor activities
- Other income
- Savings and credit
- Health, fertility
 Migration
 Anthropometric
 Social capital
 Subjective poverty

No two LSMS are exactly the same

- □ Special purpose topics in the questionnaire:
 - Tanzania: contingent valuation questions (willingness to pay)
 - Guatemala (2000): social capital module
 - Bosnia: mental health module
 - Kagera region, Tanz.: extensive module on adult deaths

□ Special purpose samples:

- Northeast China: focus on agricultural activities in rural households
- Northeast and Southeast Brazil
- Kagera region, Tanz: focus on HIV/AIDS

What is an LSMS Survey?

□DISTINCTIVE FEATURES: ⇒Multi-Topic Questionnaire

Multiple Instruments
Multiple Instruments

- Household Questionnaire
- Community Questionnaire
- Price Questionnaire (regional differences)
- □ Facility Questionnaire

What is an LSMS Survey?

□ DISTINCTIVE FEATURES:
 ⇒ Multi-Topic Questionnaire
 ⇒ Multiple Instruments
 → Quality Control

Quality Control

Small Sample

- Pre-coding, closed ended questions
- Direct/multiple informants
- Germal pilot(s)
- Training: in-depth
- □Supervision: formal (1 to 2-3)
- Data access policy
- Two-round format
- Concurrent Data Entry and editing (tworound format)

Two Round Interview

First Round

- →Household roster
- →Education
- →Health
- ➔Income and Employment
- ➔ Migration
- ⇔[Community level]
- ⇔[Food diary]

Second Round

- ➔ Agricultural Activities
- →Non-Farm Self-Empl.
- ➔ Household Expenditure
- Home Production
- →Fertility
- ➔Credit, Savings
- →Anthropometrics

Quality Control (cont'd)

Missing data
Internal consistency
Inaccuracies
Omission of key issues by analyst

Missing data

Country	Survey	% missing income data for:			% direct
		Salaried Workers	Self- Employed	Employers	informants
Ecuador	LFS, 1997	6.3	6.7	13.2	n.a.
	LSMS, 1998	3.6	8.5	6.5	96.5
Nicaragua	Urban LFS, 1997	1.0	1.4	5.7	n.a.
	LSMS, 1998	1.1	1.0	4.7	84.6
Panama	LFS 1997	2.9	36.2	26.0	n.a.
	LSMS, 1996	1.0	3.5	8.4	98.7

Missing data (cont'd)

Country	Year	Final Sample Size	Households with complete consumption aggregate		
			Number	Percent	
Bosnia and Herzegovina	2001	5,402	5,395	99.9	
Ghana	1998/ 99	5,998	5,258	87.7	
Guatemala	2000	7,940	7,276	91.6	
Jamaica	1999	1,879	1,876	99.8	
FRY: Kosovo	2000	2,880	2,880	100.0	
Kyrgyz Republic	1998	2,979	2,962	99.4	
Nicaragua	1998- 99	4,209	4,040	96.0	
Tajikistan	1999	2,000	2,000	100.0	
Viet Nam	1997/ 98	5,999	5,999	100.0	

Consistency Checks

	1985	Cote	d'Ivoire	1985	Peru	LSS
Consistency	No. of	No.	%	No. of	No.	%
Check	observ.	Failing	Failing	observ.	Failing	Failing
		Check	Check		Check	Check
Non-hhld members correctly classified Date of birth consistent, verbal-	1,135	8	0.7	1,122	6	0.5
certificate	7,485	45	0.6	27,120	153	0.6
Valid id of father	6,084	17	0.3	12,581	32	0.3
Car and gas						
expenditure	143	31	21.7	355	88	24.8

What is an LSMS Survey?

□DISTINCTIVE FEATURES: ⇒Multi-Topic Questionnaire

- ⇒Multiple Instruments
- ⇒Quality Control
- →Welfare Measure

Welfare Measure

Consumption vs. Income

Income from LFS vs. LSMS

USE OF MULTI-TOPIC SURVEYS

Use for Social Policies

	Multi-topic	IES	LFS	Ag survey	CQIW / PS
Measure welfare			$\overline{\mathbf{i}}$	$\overline{\mathbf{i}}$	$\overline{\mathbf{i}}$
Service utilization		$\overline{\bigcirc}$	$(\dot{\mathbf{c}})$		\odot
Relationship poverty-service utilization		$\overline{\mathbf{i}}$	$\overline{\mathbf{\dot{o}}}$	$(\dot{\mathbf{c}})$	$\overline{\mathbf{i}}$
Simulate alternative policies					$\overline{\mathbf{i}}$

Multi-topic surveys bring an added dimension

Social indicators become more meaningful when disaggregated, so that comparisons can be made between different population groups







• ...possibly, official statistics can add the gender dimension

•...the figures show that, in urban areas, there is no gender differential but a large gap in rural areas.

•But we still don't know much about who sends their children to school



Common Analytic Applications

Poverty Profiles

- Incidence of Commodity Tax/Subsidy
- Targeting of Large Programs
- Response of Household to User Fees
- Impact of Education on Earnings
- Impact Evaluation

Panama: Poverty Profile



Bulgaria: Poverty Update



Cumulative percentage

Who Benefits from Food Subsidy Programs in Jamaica?



➔ Food stamps are more pro-poor than food subsidies

Simulated Impact of Raising Hospital Fees in Côte d'Ivoire

Percentage of ill children seeking care in clinics and hospitals



➔ Increased hospital fees shift demand from hospitals to clinics

Nicaragua: FISE Evaluation

Financing institution that provides small-scale grants for social sector projects identified by communities

... but little known about targeting and impact

Evaluation Objectives

□ Focus of Impact Evaluation

- Poverty targeting
- □Household impact on human capital formation
- □Supply and utilization of FISE investments

Poverty Targeting - Findings

Distribution of Investments by Poverty Levels of Beneficiary Households*

	Latrines	Educ.	Health	Water	Sewer.
Bottom 20%	33.5	26.3	19.3	12.6	8.3
Bottom 40%	63.8	43.3	58.1	42.5	9
Тор 40%	17.3	30.8	27.9	42.7	71

* households ordered by consumption deciles

- □ Most Progressive Latrines (poorest 20%)
- □ Education investments are slightly progressive
- Health investments are neutral for poorest 20%, but quite progressive for poorest 40%
- Targeting for water is neutral, except for poorest 20% where it is regressive
- ☐ Most Regressive -- Sewerage

Policy Impact of Study

□Suspension of new sewerage projects

New demand-side conditional cash transfer program piloted for extreme poor

Combining Census and Survey Data

NEED:

 Large data sets, representative at small geographical units

PROBLEM:

Household surveys satisfy 2., not 1.

2. Data on consumption expenditure/income

Census data satisfies 1., but not 2.

Combining Census and Survey Data

- Select all variables which exist in both the survey and the census data set (pay attention to variable definitions - best if chance to plan in advance)
- Use the household survey (LSMS) to run a linear regression explaining household consumption in each region that is designed to be representative
- Use the parameter estimates from the regression models to impute household consumption for each household in the census
- Construct poverty maps at the level of spatial aggregation desired (based on average probability of being poor in area)

Poverty mapping



Poverty mapping





Ratio of the poor persons to the total population (in %)





Ratio of the number of poor persons to the total population (in %)

< than 24
24 - 28.6
28.6 - 36
36 - 43
> than 43

Example: Yunnan Province (China)



Poverty and Social Impact Analysis: PSIA

- Analysis of consequences and distributional impacts of policy interventions/reforms, such as:
 - Utilities
 - Pension reforms
 - Civil service reform
 - Ag reform
 - Education/health (fees, decentralization)
 - Fiscal (VAT, other taxes)
 - Land reforms
 - Etc...
- http://www.worldbank.org/psia

Tools for PSIA

Types Direct impact analysis

Behavior models

Partial equilibrium tools

General equilibrium tools

Examples

Incidence tools

Poverty mapping

Supply and demand analysis

Household models

Multi-market models

- CGEs
- SAM-IO

Macro-micro models

- 1-2-3 PRSP
- PAMS

Volume of case studies (Coudouel, Dani and Paternostro 2006)

Example: Malawi ADMARC

- Restructuring marketing functions of ADMARC (closing loss-making markets for inputs and outputs)
- Objective: Investigate the importance of ADMARC services for various groups
- Data: 1997/98 Malawi Integrated Household Survey, merged with location of ADMARC markets and roads network



Malawi ADMARC reforms

- Proximity has a larger positive effect in remote areas:
 - Impact of markets on maize yields, demand for fertilizer farm profits and consumption is significant only in remote areas.
- Policy recommendations:
 - In areas where the private sector operates and which are close to a main road, loss-making markets could be closed without major distributional impacts.
 - In areas where the private sector does not operate and where households are isolated, subsidy to lossmaking markets could be justified for their social role.

Proxy Means Testing for Programs

• Who should be beneficiaries? How to identify these people?

(Other uses of household survey data that influence program design: Geographic coverage; level of benefits people receive)

- Using household survey data to develop short list of simple indicators that can be collected in the field to "proxy" the household income/consumption.
- Compile long list of possible indicators, then use econometrics to determine which indicators are useful and the weight to place on these indicators.
- Analysis can be made more accurate by using more specific geographic regions (urban/rural, districts, etc.) but this depends on the level at which results can be generalized from household data.
Proxy Means Testing: Examples

- KIHBS 2007 data being used to create targeting system for OVC CCT program that targets poorest 20%.
- Panama Red de Oportunidades CCT program, developed with input from the 2003 Panama Living Standards Survey (Encuesta de Niveles de Vida, ENV)



Tools

- Comparative Living Standards Project (CLSP)
 - Survey Finder
 - Harmonized Data for x-country analysis
- ADePT-Agriculture
- ADePT-Livestock
- CAPI
- Source books/best practice docs:
 - Migration
 - Climate Change Adaptation
 - Tracking
 - Use of GPS (in progress)

- Fisheries
- Livestock (in progress)
- CAPI

Tradeoffs to Consider When Planning a Survey as Part of a System of Surveys

Overall scope

- Single vs. Multi-topic
- Probability vs. Purposive Sampling
- Sampling vs. Non-Sampling Errors
- Time vs. Cost
- Data vs. Capacity Building
- Surveys over time: repeated cross sections, panels, rotating

Summary

- Surveys are one source of information among many (system of information)
 - Consider all the key elements of a National Statistical System

Summary

- Surveys are one source of information among many (system of information)
- No one survey can meet all data needs: System of Household Surveys

System of Household Surveys

- **Goal**: System able to respond to evolving needs: not produce data X or survey Y
 - Determine data needs before they are URGENT
 - Identify appropriate instruments
 - Implement them properly, timely fashion
 - Analyze the resulting data

Improving the SHS

- Linking Users and Producers
- Providing adequate resources
- <u>Continuous</u> Survey Program
 - Not necessarily permanent survey
 - Benefits
 - Avoid loss of capacity
 - Create greater levels of capacity (building on existing)
 - Economies of scale
 - Policy makers know when data will be available
 - Protects NSO from pressures for ad hoc surveys
 - Ongoing system actually allows more flexibility and responsiveness

Summary

- Expanding demand for timely, relevant data
- Need to determine the range of data needs to begin to define a system of information
- Surveys are one, important, source of information among many
- No one survey can meet all data needs: System of Household Surveys

Further Information on HH Surveys

- LSMS:
 - http://www.worldbank.org/lsms
- LSMS-ISA:
 - http:/www.worldbank.org/lsms-isa
- DHS
 - http://www.measuredhs.com
- MICS
 - http://www.unicef.org/statistics/index_24303.html
 - http://www.childinfo.org
- LFS
 - http://www.statistics.gov.uk/statbase/Product.asp?vlnk=1537
 - http://www.census.gov
- IES/HBS
 - http://www.bls.gov/cex/home.htm
 - http://europa.eu.int/estatref/info/sdds/en/hbs/hbs_base.htm
- CWIQ
 - http://www.worldbank.org/afr/stat

References

- International Monetary Fund and International Development Association (1999a). "Building Poverty Reduction Strategies in Developing Countries." Report to the Board of Directors, International Monetary Fund and International Development Association, Washington, D.C.
- International Monetary Fund and International Development Association (1999b). "Heavily Indebted Poor Countries (HIPC) Initiative: Strengthening the Link between Debt Relief and Poverty Reduction." International Monetary Fund and International Development Association, Washington, D.C.
- United Nations (2000). "United Nations Millennium Declaration." United Nations' General Assembly, Fifty-fifth Session, New York, New York.
- Marrakech action plan for Statistics (2004). "Better Data for Better Results An Action Plan for Improving Development Statistics".
- Muñoz, Juan and Kinnon Scott (2005). "Household Surveys and the Millennium Development Goals." Paris21, processed.