

Data in Development: an Overview of Microdata on Developing Countries

Abstract

Finding quality microdata on developing countries can seem problematic as their national infrastructures may not support large-scale surveys. In fact a variety of organizations are collecting and distributing data, though the types of data and reasons for collection often differ from those in the most developed countries. Much of the data is collected by groups involved in, or interested in researching, the field of international development. This paper provides an introduction to the different groups involved with collecting data on developing countries and to the data they collect, including population health and welfare surveys, program assessments, finance, and opinion data.

Keywords: Developing Countries; Development; International; Foreign Aid, Survey data

As all data librarians know, a good datum can be hard to find. When the data being sought deal with countries generally referred to as developing, less developed, or low and middle income, finding good data can seem like even more of a challenge. The less developed countries form an extremely heterogeneous group, and from a data perspective, they are united primarily by what they lack. In developed countries, the librarian can usually expect national governments and related institutions such as central banks to collect demographic and business data, and national archives or major research institutions to archive surveys. In less developed countries, national governments and institutions may not collect much detailed data beyond the census and other material strictly required for internal use, or may not have developed the infrastructure to process and distribute the data they do have for public consumption. And yet, "there has been a spectacular increase in the availability and quality of data from developing countries in recent years." (Bureau for Research in Economic Analysis of Development). Data is being collected and distributed, but in many cases the groups collecting the data, and the purpose behind the collection, are specific to the field of international development. This paper looks at data on developing countries and the field of international development which concerns itself with them. For the sake of keeping this overview to a manageable length and also of highlighting the sources will be of

by Kristi Thompson¹

use to the most people, the discussion will be limited to microdata available from sources that do not require a paid subscription.

Development is a process, not a permanent state. Developing countries are not united only by their lack of the economic advantages that distinguished the most developed countries. They are

united by the fact that they are subject to development. "Developing countries is an international practice. The essence of this practice is the mobilization and allocation of resources, and the design of institutions, to transform national economies and societies, in an orderly way, from a state and status of being less developed to one of being more developed." (Gore) This practice of developing countries has developed itself into a highly diverse international field, with its own sets of standards and practices and key players large and small. People and institutions active in the practice of development collect data, both to assist with their own operations and for fundraising purposes, to "demonstrate that they can perform effectively and are accountable for their actions." (Degomme and Guha-Sapir)

An initial challenge is simply to define the countries under consideration. "There is no established convention for the designation of 'developed' and 'developing' countries or areas in the United Nations system." However, "in common practice, Japan in Asia, Canada and the United States in northern America, Australia and New Zealand in Oceania, and Europe are considered 'developed' regions or areas," (United Nations Statistics Division) leaving most of the planet still under development. A more fine-grained categorization is the United Nations' Human Development Index ranking, which divides member states into Very High, High, Medium and Low human development. The World Bank uses Gross National Income to divide countries into High Income, Upper Middle, Lower Middle and Low Income. The UN measure takes into account a broader range of factors, including life expectancy and education as well as income. For purposes of this paper, and where relevant, I will use the groupings from the UN 2009 Human Development Report² to distinguish level of development.

The different groups active in collecting data on developing

countries include national governments, intergovernmental organizations (IGO's), non-governmental organizations (NGO's) and other – this last being a catchall category containing such groups as academics and for-profit private sector firms. The bulk of the data considered in this paper comes from intergovernmental and nongovernmental organizations. In the interest of brevity, data from individual governments is considered here only insofar as it appears in other compilations. Intergovernmental organizations include such familiar large and well-funded organizations as the World Bank, the United Nations and the World Health Organization. These are valuable sources both because they conduct surveys and collect data, and because they serve as compilers and standardizers of country-level macroeconomic data. Intergovernmental organizations tend to conduct large-scale, nationally representative surveys that are standardized across a number of countries. Non-governmental organizations are more heterogeneous, and include larger and better funded organizations such as Demographic and Health Surveys and the International Food Policy Research Group along with a myriad of smaller and less known organizations. The data they collect is similarly heterogeneous, but non-governmental organizations are more likely to provide subnational surveys targeted towards a particular context or issue. “NGO surveys aim at assessing a local situation for needs and programming... (while) UN surveys tend to be large scale snap shots of a situation that serves as a point of reference”. (Degomme and Guha-Sapir) Surveys conducted by academics, or collaborating groups of academics, range from the large-scale, standardized, cross-national World Values Survey to the small, very specific local assessments available from MIT's Poverty Action Lab.

Survey Catalogs

There are two databases that compile surveys on developing countries, the International Household Survey Network catalog, and the Bureau for Research and Economic Analysis of Development (BREAD) / McArthur survey database. The International Household Survey Network catalog, which is maintained by the World Bank Data Group, contains 4147 surveys at the time of writing. Surveys include major population welfare surveys, censuses, firm-level economic surveys, and others, and the catalog can be browsed by country or survey series. The other database, which is available from the BREAD website, is smaller, claiming to hold only about 500 surveys, and lists as its focus surveys on poverty and health. It can be searched by survey location and module. Both of these databases provide information and links to access survey microdata where available.

A third database, the Complex Emergency Database (CE-DAT), has information about specialized surveys carried out by non-governmental organizations in smaller populations such as refugee camps, often under crisis conditions. At the time of writing it contained 2713

surveys. While microdata is not available from the database, the citations will assist in finding the appropriate contact person to inquire about availability, and the database includes summary statistics for the indicators and has an interface for mapping and charting them.

Major Population Health and Welfare Surveys

There are three major, long-running international data collection programs that focus on surveying developing countries: the Demographic and Health Surveys (DHS), the World Bank's Living Standards Measurement Study (LMS) surveys, and UNICEF's Multiple Indicator Cluster Surveys (MICS). The World Health Organization's World Health Surveys cover both developing and developed countries and are also worth a look. Each has advantages and drawbacks in terms of geographic coverage and topical focus. I have limited this comparison to these four series because each provides relatively uniform data collection across a number of countries, enabling cross-national comparisons. I have excluded survey programs such as the Food Policy Research Institute household surveys, which are not uniform enough to consider as a group, and the World Fertility Surveys, which were last conducted in the 1980's and are now quite dated. The data from all these surveys is largely available to researchers as microdata, though individual surveys may be unavailable and registration or application may be required.

The Demographic and Health Surveys (DHS) program is the largest of the major population welfare survey series. The DHS program is primarily funded by the U.S. Agency for International Development, and as such has a broadly defined focus on aid, development and policy improvement. They have conducted over 240 surveys in about 90 High, Medium and Low human development index countries in Africa, Asia, Latin America, the Middle East and Europe. The earliest surveys date to the mid-1980's, and multiple waves have been done in many countries. The primary surveys focus on fertility, family planning, maternal and child health, gender issues, and nutrition, and the surveys also cover household and respondent characteristics including education and school attendance, employment, and income and family wealth, making them useful for a range of analyses. DHS also conducts special modules on key topics such as AIDS and Malaria, and scaled-down surveys called the Key Indicators Surveys that are used to assess smaller sub-national populations that may be targeted by special initiatives. The focus is on women and children and men are often excluded.

The Living Standards Measurements Study surveys are conducted by the World Bank, and perhaps not surprisingly, it has excellent coverage of consumption and income. Around 90 surveys have been completed in over 40 countries between 1985 and the present. Most of the surveys cover countries with High or Medium Development, with only one (Malawi) classified as Low,

and another, Iraq, that is currently not classified. While the focus is on economic measures ranging from income and employment to debt and purchasing behavior, there are a number of health, demographic and social measures. Some of the surveys are available directly from the World Bank

but do not include the usual economic and demographic variables such as income and employment.

Comparing the Major Population Health and Welfare Survey Series

	Time Period	Focus	Demographics	Labour and Income variables	Countries Covered
DHS	1985 - present	Health, particularly reproductive	Good	Some	90+ low, medium and high development
LSMS	1985 - present	Consumption and income	Good	Good	40+ mostly high and medium development
WHS	2002	Health and health systems	Good	Some, no income	70 ranging from low to very high development
MIC	Every 5 years from 1995	Child health and welfare, reproductive health	Limited, mostly household head	Limited	65+ low and medium development

web site, others are distributed by the government of the country in which they were conducted.

The World Health Surveys were done in 2002, and follow-up studies such as the WHO Study on Global Ageing and Adult Health are being conducted. Microdata has been released for all the countries in the original surveys and is available upon signing a data use agreement. There are detailed questions on health, mortality, health insurance and use of health services along with some basic demographic and wealth variables. Employment is included, but not income. A different questionnaire is used for countries in the highest income group.

The Multiple Indicator Cluster Surveys, done by UNICEF, are currently on their fourth wave, and have been done every five years since 1995. The MICS are particularly notable for their coverage of the poorest countries; in the 2005 wave, 11 of the 55 countries they surveyed were in the low development group. They also surveyed some sub-national populations. However, the surveys are limited in their usefulness for general analysis due to the selection of variables available. The MICS were developed specifically to track progress on the World Fit for Children Plan of Action and the Millennium Development Goals, and cover indicators relating to child health and welfare, women's reproductive health, and some basic household variables,

Data for Assessing, Modeling and Targeting Development

While much of the data collected on developing countries is used to provide evidence to assess developmental progress, data focusing on the evaluation of specific approaches is rarer. The Abdul Latif Jameel Poverty Action Lab was formed at MIT as a network of professors around the world who use Randomized Evaluations to answer questions on poverty alleviation. "What makes J-PAL's work innovative is that such randomized studies haven't typically been used in evaluating poverty-alleviation programs, or even in the wider field of economics." (Standish) They offer training on randomized evaluation and maintain a database that currently holds information about 234 randomized assessments, with microdata available for 13 of them, on topics ranging from textbook provision to microfinance. The International Food Policy Research Institute (IFPRI) is another source for random evaluations and has released data for studies such as the Comparing Food versus Cash for Education Program, along with their more standard household welfare datasets.

The International Food Policy Research Institute also collects data to construct Social Accounting Matrices, which are currently available for about 30 countries. Social Accounting Matrices are used as the basis for various economic models that can be used to estimate the effects that different policy or aid approaches will have before

actually carrying them out.

Finance

The World Bank claims to provide “the world's most comprehensive company-level data in emerging markets and developing economies” (World Bank Group) and I would not attempt to dispute that claim. They have conducted surveys in over 125 countries, including survey projects such as the World Business Environment Surveys, as well as more specialised projects such as the recent three-round Financial Crisis Surveys and the Management, Organisation and Innovation survey work. Firm-level microdata is available to researchers at no cost from these surveys, in contrast to the International Monetary Fund, which also conducts surveys but only releases macrodata, and some of that only to paying customers. The World Bank’s Enterprise Survey Portal also allows users to construct charts online.

A couple of regional sources for financial microdata are the Economic Research Forum, focusing on the Middle East and North Africa, and Oxford University’s Centre for the Study of African Economies, which focuses on Sub-Saharan Africa. The Economic Research Forum has released data on micro and small enterprises in Egypt, Lebanon, Morocco and Turkey, as well as an Egypt labour market panel study. The methodology documents for the micro and small enterprises data provides interesting insights into some of the difficulties in collecting and analysing data on developing economies; for example, in Turkey the researchers were unable to weight the rural sample because no more authoritative data on rural enterprises existed.³ The Centre for the Study of African Economies has comparative cross-national firm-level datasets primarily of manufacturing firms, as well as a couple of household panel surveys, and some more specialized datasets from working papers.

Opinions, Attitudes, and Values

The data collections discussed so far have been factual, consisting of theoretically objective measures of demographic and economic variables that can be used to implement and assess development programs. Opinion and values surveys may not appear to have the same practical, development-oriented application that most of the data discussed so far have. Development is done, in the end, for people, to improve their lives as they actually experience them, not merely to increase GNI or some other measure. Opinion data can “delve deeper than, for example, official poverty data, by reporting on people’s experiences in obtaining basic human needs and their own perceptions of whether or not they feel poor.”(Corporacion Latinobarometro) In addition, development happens in a political as well as an economic context. It is ideally done in cooperation with and with the support of governments and the people in the area being developed.

Two of the largest and best-known cross-national opinion surveys are the World Values Surveys and the Global Barometers. The World Values Surveys are an international collaboration among academics who are attempting to survey the “basic values and beliefs of the publics of more than 80 societies.” (World Values Survey) Five waves have been completed, conducted between 1981 and 2008, with a sixth currently in progress. “Since each national group funded its own survey, its first wave was largely limited to relatively developed societies.” (World Values Survey) By the second wave, the researchers behind the project had decided that “it was important to include societies across the entire range of development, from low income societies to rich societies,” (World Values Survey) and additional researchers and sources of funding were found.

The Global Barometers are done more modularly, with separate Afrobarometer, Arab Barometer, Asian Barometer, East Asian Barometer, and Latinobarometro surveys, and a Eurasia Barometer under development. (Another one, America’s Barometer, covers some of the small Latin American nations and is available for subscription through the Latin American Public Opinion Project (LAPOP).) While the Global Barometers were inspired by Eurobarometer, there is no formal association.⁴ The Barometers are designed to be a comparative survey of attitudes and values toward politics, power, reform, democracy and citizens’ political actions. They are repeated at approximately three year intervals, and include a core module of questions asked across regions, plus additional region-specific questions. “Whereas the WVS addresses deep-seated, semi-permanent cultural values, the GB is concerned with tracking emerging political and economic attitudes, which are often subject to rapid change.”(Corporacion Latinobarometro)

The Pew Global Attitudes Project is conducted by the Pew Research Centre, a U.S. based non-governmental organization. It is a series of opinion surveys, each covering anywhere between five and 50 countries, that have been conducted between 2001 and the present. Key areas of interest include “attitudes toward the U.S. and American foreign policy, globalization, terrorism, and democracy.”(Pew Research Center) While there is a decided focus on American foreign policy (for example, the data formed a basis for the book *America Against the World: How We Are Different and Why We Are Disliked*⁵), the data also includes questions of broader interest such as global reactions to issues in the news.

The multinational surveys mentioned above are generally too large and slow to implement to track responses to local issues under rapidly changing circumstances. Local polls, whether done by local media organizations, academics, or others are more likely to track these highly situational opinions, but the decentralised nature of local polling as well as language barriers make them particularly difficult to

access. Worldpublicopinion.org, which is managed by the Program on International Policy Attitudes at the University of Maryland, compiles and analyzes some of these opinion polls, and they also conduct their own local polls. Their own studies are performed by a network of research centres in 25 countries, and many of their datasets are available to download.

Conclusion

This overview has only touched on some of the sources of data available on the developing world. The choice to focus on microdata means that macroeconomic and many finance sources have been neglected; limiting myself to freely available sources means that resources in archives such as ICPSR or subscription services such as Polling the Nations were left out; most smaller projects covering only one or a few countries are included only insofar as they appear in one of the survey databanks. Still, it should be clear by now that a great variety of data on developing countries and development is available, whether it is collected by practitioners of development trying to improve their outcomes, researchers studying developing countries and various types of comparative social science, or other interested parties from both within and outside the developing world. While much of this data is gathered to serve relatively narrow purposes, collectively it can provide the researcher with a window, however imperfect, into the minds and lives of a large and often unheard portion of the world's population

Appendix: list of data sources in the order they were mentioned.

Survey catalogs

BREAD/Mcarthur/CCPR Survey Database: <http://ipl.econ.duke.edu:8080/survey/>

International Household Survey Network's central survey catalog: <http://www.internationalsurveynetwork.org/home/index.php?q=activities/catalog/surveys>

Complex Emergency Database (CE-DAT): <http://www.cedat.be/>

Major Population Health and Welfare Surveys

Demographic and Health Surveys: <http://www.measuredhs.com/>

Living Standards Measurement Study: <http://www.worldbank.org/lsms/>

World Health Surveys: <http://www.who.int/healthinfo/survey/en/index.html>

Multiple Indicator Cluster Surveys: <http://www.childinfo.org/mics.html>

Data for Assessing, Modeling and Targeting Development
MIT's Abdul Latif Jameel Poverty Action Lab :<http://www.povertyactionlab.org/evaluations?filters=type:evaluation>

International Food Policy Research Institute (IFPRI)
Surveys: <http://www.ifpri.org/datasets>

Finance

World Bank Enterprise Surveys:<http://www.enterprisesurveys.org/>

Economic Research Forum Data:<http://www.erf.org.eg/cms.php?id=Datasets>

Oxford University's Centre for the Study of African Economies:<http://www.csae.ox.ac.uk/>

Opinions, Attitudes, and Values

World Values Surveys: <http://www.worldvaluessurvey.org/>

Global Barometer:<http://www.globalbarometer.net/>

Worldpublicopinion.org: <http://www.worldpublicopinion.org/>

Pew Global Attitudes Project: <http://pewglobal.org/>

References

B R E A D Bureau for Research in Economic Analysis of Development. "Data from developing countries." Available online: http://ipl.econ.duke.edu/dthomas/dev_data/index.html, accessed 18 Oct 2010.

Corporacion Latinobarometro. "Global Barometer – Strategy." Available online: <http://www.globalbarometer.net/strategy.htm>. Accessed 18 Oct 2010.

Degomme, Olivier and Debarati Guha-Sapir. "Mortality and nutrition surveys by Non-Governmental organisations. Perspectives from the CE-DAT database" *Emerging Themes in Epidemiology* 2007 Vol. 4, No. 11. Available online: <http://www.ete-online.com/content/4/1/11>. Accessed 18 Oct 2010.

Gore, Charles. "The Rise and Fall of the Washington Consensus as a Paradigm for Developing Countries." *World Development* Vol. 28, No. 5, pp. 789-804, 2000

Pew Research Center. "About the project." Available online: <http://pewglobal.org/about/>. Accessed 18 Oct 2010.

Standish, Sarah. "Researching Better Ways to End Poverty" (blog post). 10 Dec. 2009.

Available online: <http://www.globalenvision.org/2009/12/03/rigor-science-now-economics-too>

United Nations Statistics Division. "Standard Country and Area Codes Classifications." (See footnote c.) Available online: <http://unstats.un.org/unsd/methods/m49/m49regin.htm>

The World Bank Group. "Enterprise Surveys." Available online: <http://www.enterprisesurveys.org/>

World Values Survey. "Building A Worldwide Network Of Social Scientists." Available online: http://www.worldvaluessurvey.org/wvs/articles/folder_published/article_base_51. Accessed 18 Oct 2010.

Endnotes

1. Contact: Kristi Thompson, Data Librarian, Leddy Library, University of Windsor, Windsor, Ontario, N9B 3P4. Phone: 1 519 253 3000 x3858

Email: kathomps@uwindsor.ca

This paper is based on the presentation Data in Development given by the author, IASSIST 2010.

2 United Nations Development Programme. "Human Development Report 2009." Available online: http://hdr.undp.org/en/media/HDR_2009_EN_Complete.pdf

3: Ozar, Semsal. "Micro and Small Enterprises in Turkey: Uneasy Development". Available from the data download area of the Economic Research Forum web site; contact http://www.erf.org.eg/cms.php?id=MSE_Database for access.

4 See <http://www.globalbarometer.net/background.htm>

5 See <http://pewglobal.org/americaagainsttheworld/>