WPS7252

Data Deprivation

Another Deprivation to End

Umar Serajuddin
Hiroki Uematsu
Christina Wieser
Nobuo Yoshida
Andrew Dabalen

Abstract

The Millennium Development Goal of halving the incidence of extreme poverty from its 1990 level will be achieved in 2015, and the international development community is now moving to a new goal of "ending extreme poverty." However, the data needed to monitor progress remain severely limited. During the 10 year period between 2002 and 2011, as many as 57 countries have zero or only one poverty estimate. This paper refers to such lack of poverty data as "data deprivation," because the poor are often socially marginalized and voiceless, and the collection of objective and quantitative data is crucial in locating them and formulating policy to help them exit extreme deprivation. This paper studies the extent of data deprivation and proposes targets for ending data deprivation by 2030—the year by when the international community aims to end extreme poverty. According to the analysis in this paper, this target is ambitious but possible, and achieving it is necessary to be able to declare the end of extreme poverty with confidence.

Data Deprivation: Another Deprivation to End

Umar Serajuddin, Hiroki Uematsu, Christina Wieser, Nobuo Yoshida, and Andrew Dabalen*

* World Bank, Washington DC. Email: userajuddin@worldbank.org, huematsu@worldbank.org, cwieser@worldbank.org, nyoshida@worldbank.org, adabalen@worldbank.org

"Nothing will come of nothing" – *King Lear,* William Shakespeare

1.   Motivation

Data are crucial in knowing where a country stands in terms of social and human development and can help explain which policies work and which do not work in promoting growth and eradicating poverty and other social deprivations. A very concrete feature of data is that they can make the situation of the poor visible to policy makers. The poor, who often lack political representation and agency, may remain invisible unless data reveal where they are and how they are. In this sense, lack of data on key dimensions of human and social development can be seen as a form of 'deprivation'. This understanding of the role of data is consistent with the recent United Nations Report titled *A World That Counts* (IEAG 2014), which argues that a lack of data can lead to a "denial of basic rights" and discusses how to mobilize a 'data revolution' for sustainable development.[1] The objective of this paper is to take stock of data on poverty and discuss goals and targets to address this data deprivation.

The availability of data is not only important to citizens and policy makers in countries, but also for the international development community. With the Millennium Development Goals (MDGs) expiring in 2015 and the world preparing to begin monitoring the post-MDG development agenda – expected to be reflected in the Sustainable Development Goals (SDGs) – the role of data in development has received strong attention, as the call for a 'data revolution' to monitor the SDGs indicates (IEAG 2014). Poverty related goals were among the most prominent of the MDGs and are expected to feature prominently in the post-2015 development agenda as well. The World Bank Group (WBG) recently adopted two principal goals to guide its work: (i) to end extreme poverty at the global level by 2030; and (ii) to promote "shared prosperity": a sustainable increase in the income and well-being of the poorer segments of the population – namely the bottom 40 percent – in every country. A "world free of poverty" has been the WBG's stated mission since its inception; however, the goal of ending extreme poverty (defined as reducing the share of the global poor to 3 percent[2]) is the first time the WBG's efforts focus around a specific target. Similarly, the goal of shared prosperity incorporates the notion of 'growth with equality' into the WBG's operations quite explicitly. The frequency and quality of data on poverty and shared prosperity are central to satisfactorily monitoring these twin goals, and more significantly, to set effective policies for a country's poverty reduction program. A recent study by Joliffe *et al* 2014 that devotes itself entirely to the methodological and empirical issues surrounding monitoring these goals, illustrates the importance of improved data, both in terms of regularity and quality.

The starting point in monitoring progress in poverty reduction and enhancing shared prosperity is to have household consumption survey data that are not only available at reasonably frequent intervals, but are also comparable over time. The availability of such data has expanded rapidly in recent times, starting from a very low coverage of only 13 countries in the early 1990s, growing to 40 countries in 2001, and further increasing to 62 countries in 2011.

---

[1] The report was prepared by the Independent Expert Advisory Group (IEAG) on the Data Revolution for Sustainable Development at the request of the United Nations Secretary General.
[2] For more details on this see Jolliffe *et al* (2014).

All the same, substantial data gaps persist. During the ten year period between 2002 and 2011, among the 155 countries[3] for which the World Bank monitors poverty data using the WDI database, 29 countries do not have any poverty data point and 28 countries have only one poverty data point.[4] Thus, in over a third of the world's developing or middle income countries there is essentially no meaningful way of monitoring poverty or shared prosperity for that specific period. Moreover, among countries that have poverty estimates for two years over the ten year period, often large time gaps exist. Of 35 such countries, 20 have two poverty estimates with a larger than five year interval, which resulted in poverty monitoring efforts being dated. Therefore, a total of 77 countries – about half of the 155 countries – faced challenges in producing timely or any poverty estimates during the 2002-11 period. If one considers intervals shorter than 5 years, the picture becomes worse.

The analysis in this note is conducted using a World Bank database of poverty monitoring data for 155 countries and territories. By examining the availability of data across countries over the last two decades, the note identifies improvements that have taken place over time and provides a richer understanding of the current state of data deprivation, which hinders effective poverty monitoring. This discussion can then be used as a benchmark to set targets for addressing poverty data deprivation in the medium and long term.

While this note focuses on 'poverty' data, it should be noted outright that surveys collecting consumption data collect information on many human development and social indicators, and addressing poverty data deprivation can go a long way – if not all the way – in addressing data deprivation regarding many other indicators as well. Thus the usefulness of consumption surveys goes far beyond that of poverty monitoring. By gathering a host of statistics on individual and household attributes, as well as on locational attributes, such surveys are the pivotal source of data for governments and policy makers in understanding linkages between poverty and other socioeconomic outcomes and policy. At the same time as informing key priorities of governments, such as the design of social programs and the impacts of policy interventions or shocks, such data also provide vital input to constructing national accounts and the CPI (consumer price index).

Alongside household consumption survey data, complementary data are also needed for tracking poverty (Jolliffe *et al* 2014). For example, if population census data are unavailable, outdated or unreliable, it is problematic to make inferences about a country's population from a sample-based household survey. Cross-country comparisons place additional data demands. Cost of living adjustments are necessary to count the poor across countries using a common currency and a global poverty line. At present, this comes from data on purchasing power parity (PPP) collected by the International Comparison Program (ICP) every few years. Moreover, to compare poverty across countries in any given year, inflation data and real GDP growth data (or private/household consumption data) would be needed to account for changes in prices and in real economic activity between the survey year and the reference year.

However, in reality, the main bottleneck for poverty data is lack of household surveys. Usually if household survey data are available, the aforementioned complementary data are available. Therefore, it is safe to say lack of poverty data can be attributed to lack of household survey data.

---

[3] When using the term countries, we refer to countries and territories.

[4] A poverty data point consists of a poverty estimate in a certain year at either the international poverty line of $1.25 a day or the national poverty line compiled from household consumption survey data.

This note is organized as follows. Following a brief description of the database used, the note examines trends in poverty data availability across the world. It then moves on to propose approaches to think about targets for improving poverty data availability.

2.   Data Source

For a comprehensive assessment of available poverty data, we use poverty statistics from the World Development Indicators (WDI) database (http://data.worldbank.org/data-catalog/world-development-indicators). This database identifies 1,101 different poverty data points across developing countries between 1976 and 2013, for which poverty measures were computed at national or international poverty lines.[5]

In this note, we define the developing world as a total of 155 countries and territories, based on a union of the three popular country groupings (defined as of July 1, 2014). First, from the WDI database we take the 139 countries that are Low, Lower Middle or Upper Middle Income countries as classified by the World Bank (this is determined by GNI per capita based on the Atlas method). The WDI's definition excludes some countries that recently graduated into High Income countries (e.g., Russian Federation, Chile, Uruguay), but where the World Bank is active. Consequently, we also draw 10 additional countries from the 144 countries listed as belonging to the World Bank's different lending groups (classified as IDA, IBRD or Blend). Finally, we add another six countries that are in the PovcalNet database[6], an online tool for poverty measurement developed by the World Bank's Development Research Group. We exclude the wealthier countries where extreme poverty – as defined by the population living on less than $1.25 a day, measured at 2005 international prices adjusted for purchasing power parity (PPP) – is presumed to be zero, and for which poverty trends have not been tracked by the World Bank.

The poverty data points in the WDI database are computed from household surveys with full-fledged consumption modules that enable countries to produce reliable poverty estimates. In reality, many household surveys include consumption or income modules, but most are not suitable for producing official poverty statistics because they include too few items in the consumption or income module or they do not have wide enough geographic coverage to produce a nationally representative statistic. Moreover, the poverty estimates published in the WDI database have been scrutinized by World Bank staff and can be considered 'fit for purpose' for poverty monitoring.

The analysis in this note attempts to focus on poverty data points that are comparable over time, and exclude additional non-comparable poverty data points included in sources such as the World Bank's Poverty and Equity Database (http://povertydata.worldbank.org/poverty/home) or in WDI's non-comparable poverty estimate series. Even if poverty data are frequent and timely, if they are not comparable over time they cannot be used in poverty or shared prosperity monitoring. Non-comparability essentially arises when either the welfare aggregate (consumption or income) is measured differently across surveys from different years due to changes in survey design, or if a poverty line cannot be fixed in real terms across years due to flawed

---

[5] The WDI database was last accessed on October, 2014.

[6] All the six countries – Czech Republic, Estonia, Latvia, Lithuania, Slovak Republic and Slovenia – have very small but still positive levels of extreme poverty as defined by the $1.25 (2005 PPP) International Poverty Line. With the October 2014 update, PovcalNet now includes many high income and OECD countries. What we refer to as "countries included in PovcalNet" does not include these newly added high income countries as they have never been used in the calculation of extreme poverty rate for the developing world.

price adjustment data. An assessment on the comparability of poverty data has been conducted by the Global Poverty Working Group (GPWG), a group of poverty measurement experts in the World Bank. Given the large volume of the database, the GPWG started this assessment from the most recent data. When this note was prepared, the assessment was not completed across all years. Therefore, our assessment on data availability, particularly on poverty data before the 2000s, should be seen as provisional.

It is worth noting that some of the data sets underlying poverty estimates published in the WDI database are not directly accessible to the World Bank's poverty economists. However, even in such atypical cases, the poverty economists still have sufficient confidence in the robustness of the poverty estimates to include them in the WDI database. We decided to include these estimates in our notion of available poverty data as the challenge for them is not availability *per se* but access, which is a different challenge.

A candidate database that we could have drawn from is the International Household Survey Network or IHSN database, a repository of various types of household surveys including LSMS (Living Standard Measurement Surveys), DHS (Demographic and Health Surveys), LFS (Labor Force Surveys), MICS (Multiple Indicator Cluster Survey), etc.. A recent paper by Demombynes and Sandefur (2014) utilizes the IHSN database to understand data production and access across countries and estimate the financial costs of filling data gaps. The IHSN database has an advantage of being able to identify whether surveys are 'open, i.e., accessible on-line.' However, since the focus is on poverty data, the WDI database was the preferred choice. That said, comparisons between these two databases give us very useful insights on how surveys are planned and carried out in the field. We will come back to this point in section 6, where we discuss practical concerns in addressing the "Data Deprivation".

3.   Key Findings

The availability of poverty data has increased sharply since the early 1990s (Figure 1). A first rapid increase occurred between 1990 and 1996 when data points for those particular years increased from 13 to 36, driven largely by the Europe and Central Asia (ECA), Latin America and Caribbean (LAC), and Sub-Saharan Africa (SSA) regions. At the beginning of the 1990s, poverty data from the ECA region were not available in the WDI database, mainly due to the fact that many post-Soviet states did not yet exist or data were not accessible to the World Bank. The number of poverty estimates in the ECA region increased significantly since 1992 and rose to an average of eight data points per year between 1992 and 1999. During this time, the number of poverty estimates also increased significantly in both the LAC and SSA regions.

The years 2000 to 2005 brought about another spike in the availability of poverty data; the number of countries with available poverty data almost doubled from 37 countries in 2000 to 60 countries in 2010 and remained fairly constant since then with 62 countries producing poverty estimates in 2011. The ECA region was again a key driver of this second wave of improvements in data availability. The number of poverty estimates in the ECA region increased from 9 in 2000 to 26 in 2004 while the number of poverty estimates in other regions fluctuated, but overall showed improvements. The average of all data points in a given 10 year period illustrates significant improvements in Sub-Saharan Africa with increases in the availability of data points by almost 50 percent from 5.1 in the 1990s to 7.5 in the 2000s.

It is worth noting that the decline in the availability of data points for 2012 and 2013 is not indicative of a decline in actual surveys carried out. Processing of data and estimation of poverty is time-consuming, and in many developing countries poverty estimates become available in the WDI one to two years after surveys

were actually carried out. The reduction in poverty estimates in the last two years (2012 and 2013) does not necessarily reflect a reduction in surveys, but it simply displays the lag between data collection and the publication of poverty estimates and we therefore ignore years after 2011 in this note.[7]

Figure 1: Number of Poverty Data since 1976
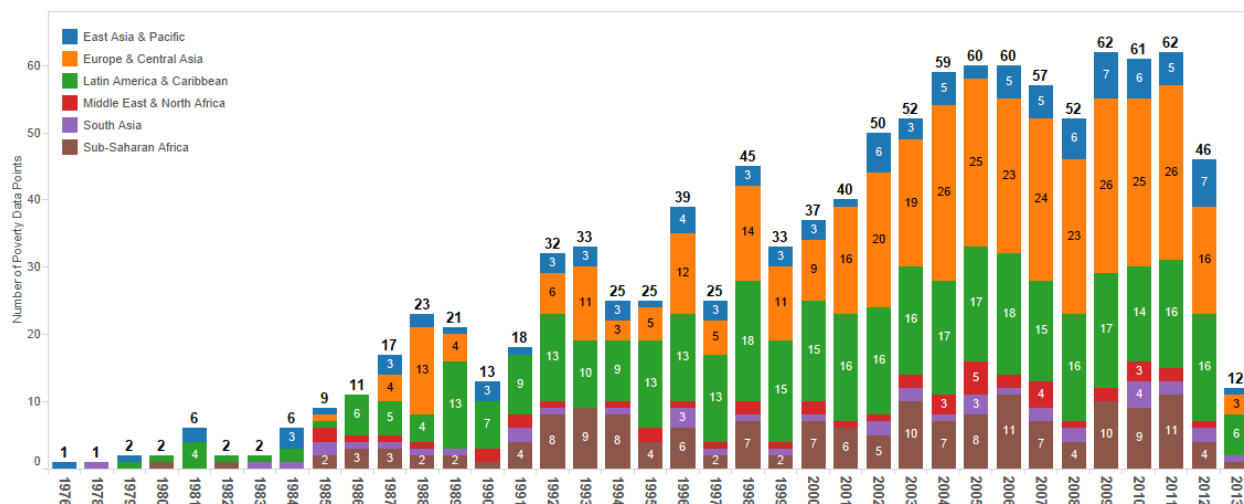
Counting the number of poverty estimate per year is a useful exercise to assess data availability, but to assess a country's ability to monitor poverty and shared prosperity, the frequency of poverty data matters. For example, to grasp whether poverty in a country is rising or declining, at least two data points within a 'reasonable' time interval are required. Setting a standard on the ideal interval is however not an easy task. Though valuable for decision making, surveys typically impose significant demands on financial and human resources. The General Data Dissemination System (GDDS)[8], developed as an attempt to improve the quality of statistics, recommends that poverty statistics be updated in at least 3 to 5 year intervals (http://dsbb.imf.org/pages/GDDS/TableB.aspx). While not necessarily endorsing the GDDS, we find its periodicity recommendation a useful benchmark and position much of our discussion on data availability around it.

Consequently, we examine the frequency of poverty data by inspecting the availability of poverty data across countries during ten year intervals. If a country collected poverty data every five years, it would have two data points in every ten year time period. If a country collected poverty data every three years, it would have three to four data points every ten years. Therefore, if a country satisfies GDDS's recommendation of updating poverty data every 3 to 5 years, it should have two to four poverty estimates in a ten year time period.

---

[7] This observation illustrates another important issue of global poverty monitoring – lack of timeliness of poverty data. Although this is a very serious issue for global and country poverty monitoring, it is beyond the scope of this paper, and we will focus on the data availability issue in this paper.

[8] The GDDS is a structured process through which International Monetary Fund (IMF) member countries commit voluntarily to improving the quality of the data compiled and disseminated by their statistical systems in accordance with a set of recommended standards (e.g., see IMF 2013). The World Bank worked closely with the IMF to articulate the socio-demographic component of the GDDS guidelines, including poverty data.

This metric of data availability, similar to our earlier discussion on the number of poverty data points per year, also points to clear improvements over time. As Figure 2 shows, the total number of countries with zero or one data point declined from 78 during 1990-99 to 57 during 2002-11, implying that over a 12 year period 21 additional countries were collecting two or more poverty data points every ten years. The improvement is striking when comparing the periods of 1997-2006 and 2002-11 – i.e., looking back ten years at the end of 2006 versus at the end of 2011 – when 16 more countries were collecting two or more data points.

Figure 2: Number of Countries with Less Than Two Data Points
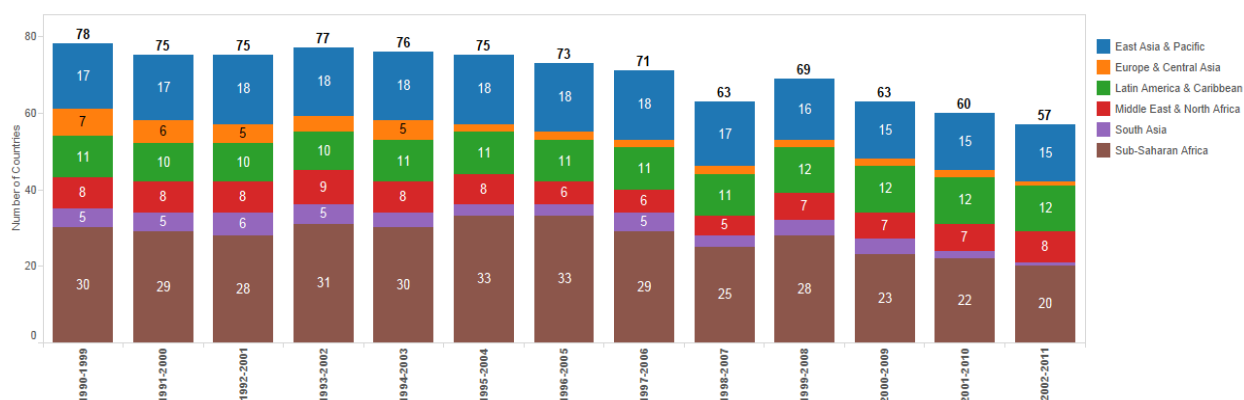
There appear to be two broad phases in improvements in data availability. The first phase was between 1990-99 and 1997-2006 in which the number of countries with zero data points declined from 50 to 31 while the total number of countries with zero or one data points only declined slightly from 78 to 71 (Figure 3). Thus, the main shift took place due to countries moving from no poverty data point to one data point. Subsequently, the second phase was between 1997-2006 and 2002-11 with countries shifting to more frequent collection of poverty data. The number of countries with one data point declined from 40 to 28 while the number of countries with two or more data points increased. Overall, the number of countries with zero or one data point declined to 57 for the time period of 2002-11.

To monitor poverty adequately and to satisfy the GDDS's minimum requirement (of one data point every five years), countries should have at least two data points in ten years with surveys no more than five years apart. However, even one data point every five years might not be enough to monitor progress towards poverty reduction. Compared to labor statistics, poverty is often thought to be less volatile and does not need frequent monitoring. However, this is not necessarily reflected in empirical evidence. For example, according to PovcalNet (as of October 2014), the poverty headcount rate in rural areas in India declined 10 percentage points in just two years between 2009/10 and 2011/12. Similarly, the poverty headcount rate in urban areas in India declined seven percentage points during the same period. As an earlier example, Paraguay experienced nearly 10 percentage points increase in the poverty headcount rate measured at $4 a day in just one year between 2001 and 2002. In the middle of Argentina's debt crisis, the poverty headcount rate in urban areas of Argentina also rose nearly 10 percentage points within one year (again 2001-02). Since both Argentina and Paraguay have annual data on poverty, it was possible to monitor and analyze the hardship people faced during these difficult times between 2001 and 2002.

Therefore, it would be preferable to have even more frequent data on poverty and inequality than at a five year interval. Since collecting household survey data is costly and time-consuming, annual data collection can be burdensome for many National Statistics Offices (NSOs) and data quality could suffer as a result. We are aware that recent innovations in survey data collection and imputation techniques allow us to fill the data gaps with less cost and time. However, the reliability of such techniques also depends on the availability and frequency of traditional household consumption surveys.

Considering the necessity to monitor the progress of poverty reduction reliably, keeping in mind data quality, and looking at data from 2002-11, the latest 10 year interval for which we can infer reliable statistics on data availability, we classify countries into five categories:

i) Countries with **extreme data deprivation** in monitoring poverty are those without any poverty data point in the ten year reference period. These countries would require the strongest efforts to reach the benchmark of two data points in any ten year interval.

ii) Countries with **moderate data deprivation** are those with one data point in any ten year interval.

iii) Countries that are **vulnerable to data deprivation** are those with two data points in 2002-11, but at an interval of more than five years apart. These countries are at risk of falling short of having adequate data for poverty monitoring (as they will not have two poverty data points in *any* ten year interval).

iv) Countries with two data points in ten years with the surveys 5 years apart are deemed to satisfy **a minimum requirement for data needs.**

v) Countries with 3 or more data points in any 10 year interval are considered to be **satisfactory for data needs**.

Note that while the above classification is prepared by frequency of poverty data, we have also attempted to incorporate the notion of 'regularity' in defining countries that are vulnerable to data deprivation. Regularity of data collection is, however, a somewhat amorphous concept to measure when looking at a 10 year time period. For example, even if poverty data of one country are available in 2005 and 2010, it is difficult to infer that this country collects poverty data *every* five years. The possibility, that this country collected data in 2005 and 2010 but will not collect poverty data in 2015, remains.[9] To see whether a country collects data every five years, we could increase the time period, for example to a 20 year period. Since for most countries, poverty data are not available prior to 1990, such measures are difficult to obtain and we remain at a 10 year time period to measure data frequency and regularity.

Global trends

Using the classification suggested above and looking at the 2002-11 time period, we see that 63 of 155 countries have satisfactory data availability with three or more data points and 15 countries meet the minimum requirement for data needs with two data points with 5 year intervals (Figure 3). Clearly, over time the number of countries with satisfactory data availability has increased. At the same time, countries with extreme data deprivation have declined (from 50 in 1990-99 to 29 in 2002-11).

---

[9] Jordan, e.g., conducted its flagship household survey in two year intervals between 2006 and 2010 followed by a 4 year interval.

Challenges, however, remain. About half of the countries – 77 of 155 – are deprived of adequate data, 57 of them quite acutely. While 20 countries are vulnerable to data deprivation, another 57 have either moderate (28) or extreme data deprivation (29).

Figure 3: Availability of Poverty Data
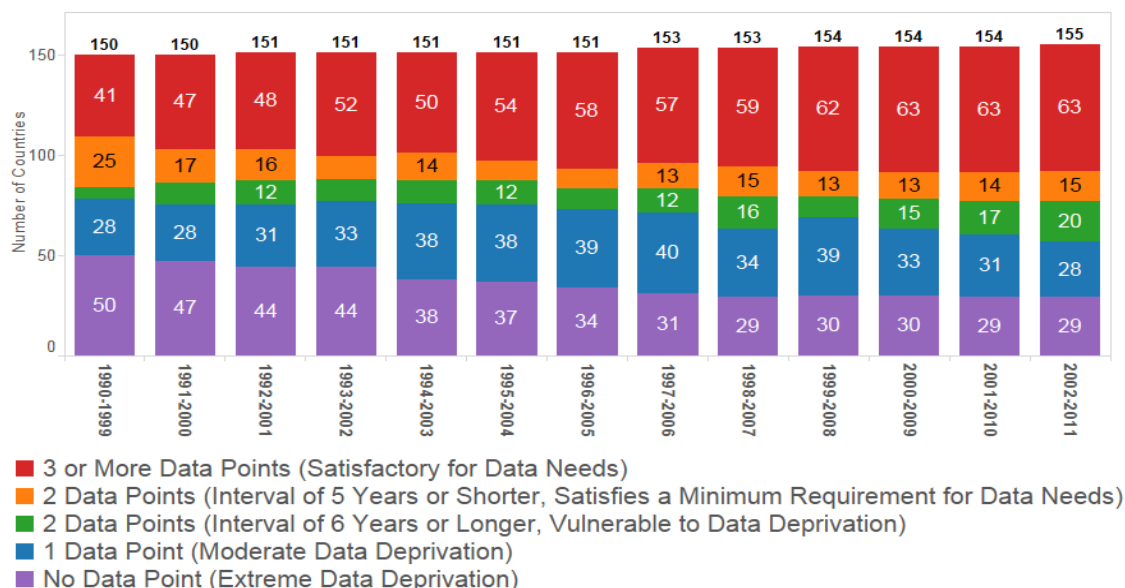
While lack of survey data precludes country level poverty monitoring efforts, it also hinders global monitoring efforts by reducing the precision of global estimates. Global (and regional) '$1.25 dollar a day' poverty is estimated every few years by aggregating internationally comparable country level poverty estimates to the global (and regional) level.[10] However, household consumption surveys for countries do not always coincide with the reference year for which global poverty is being estimated. Between the various survey years for countries and the reference year, economic growth may have taken place and cost of living may have changed. Survey data for all countries are lined up to a given reference year following adjustments based on national accounts data on real consumption growth. This is a reason the reference year is called the "lining up year." The greater the interval between the survey year and the lining up year, the more imprecise the estimate of the lining up year and thus the global/regional poverty estimates become. Figure 4 a shows that global poverty estimates since 2005 are typically calculated from survey data from the same year for about a third of all countries (45 to 54); a majority of countries need to be lined up. When weighted by population the situation tends to improve: for the 2010 global poverty estimates, 67 percent of the developing country population was represented by the household surveys carried out that year. A year later it lowered to 44 percent (Figure 4b), a large fluctuation driven mainly by a populous country like India. From the perspective of global poverty monitoring the more populous countries carry more weight. In general, with more frequent household consumption surveys global poverty estimates would become more reliable.

---

[10] For details see Chen and Ravallion (2010), Chen, Ravallion, Sangraula (2008), or Jolliffe *et al* (2014). Also, estimates are available in the PovcalNet website (http://iresearch.worldbank.org/PovcalNet/).

Figure 4: Interval between Household Survey Year and Lining-up Year for Global Poverty Estimate in PovcalNet

a. **Number of Countries**
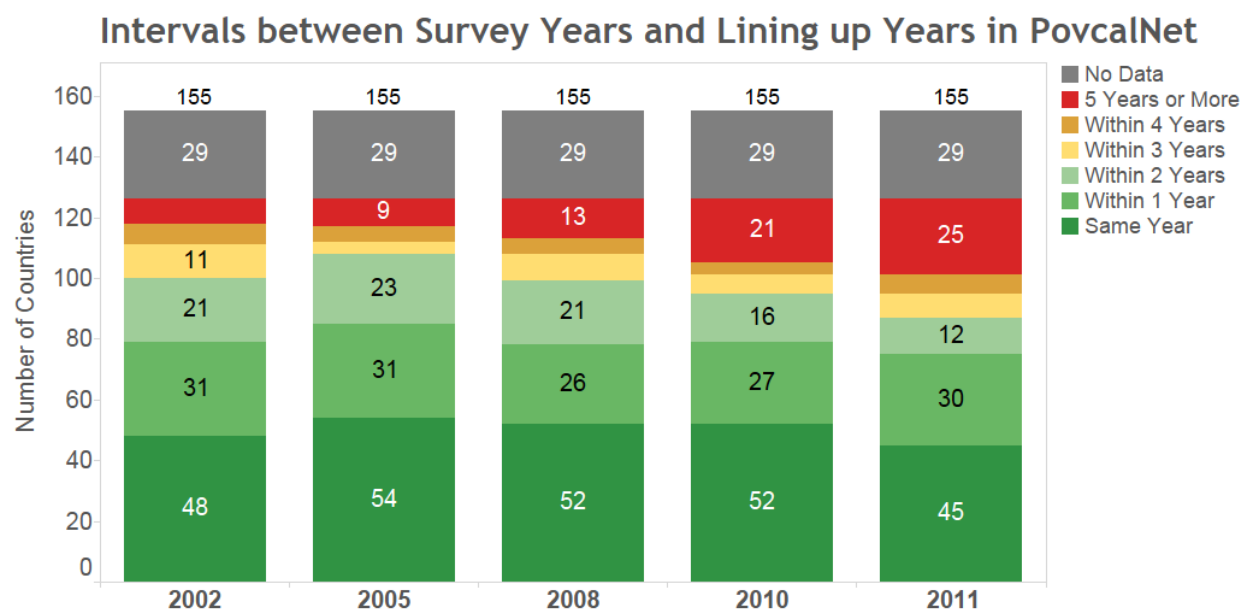
b. **Population weighted (share of population)**
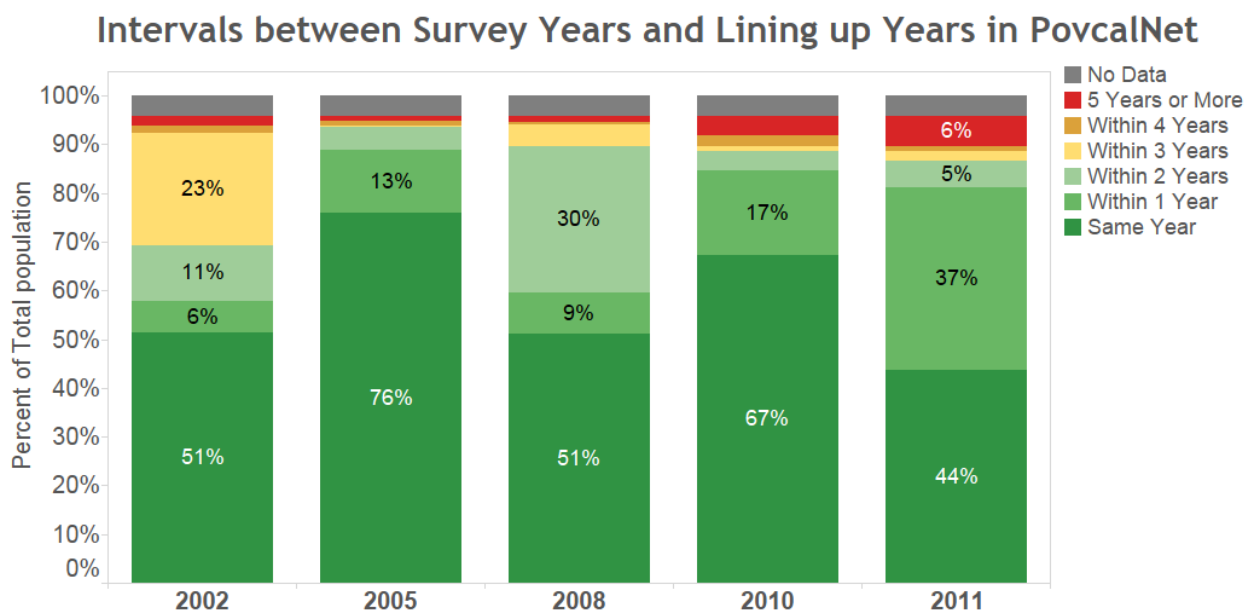
Regional trends

In the ECA region, almost all countries have satisfactory levels of data availability. Only one country does not have two data points in the ten year interval (Table 1). The LAC and EAP (East Asia and Pacific) regions have countries on both ends of the spectrum. Almost half of all countries in these regions meet

satisfactory or minimum requirement levels of poverty data, but almost half have an extreme data deprivation. This results from the WDI not including poverty data for most small island countries in the Caribbean Sea and the Pacific Ocean.

In Sub-Saharan Africa, only 5 of 48 countries in the WDI database have satisfactory data availability and another 9 meet the minimum requirement for data needs. On the flip side, 34 countries – 70 percent – do not have adequate data. More notably, 20 countries have moderate and extreme data deprivation and it is therefore impossible to monitor poverty or to estimate the shared prosperity indicator for these countries (40 percent of the countries in the SSA region). Moreover, in another 14 countries data are more than five years apart.

Challenges remain in the Middle East and North Africa region (MNA) with more than 60 percent showing inadequate data; two countries in fact have extreme data deprivation as they have no poverty data point in the ten year period. However, in some countries in the MNA region, such as Iran or Algeria, household survey data are produced but not shared with the World Bank and it is therefore important in the MNA region to improve data access to increase data availability.

In South Asia only one country has a moderate data deprivation while all other 7 countries have at least two data points in the ten year period. Another two countries are vulnerable to data deprivation as they have two data points in the ten year period which are more than 5 years apart. Considering that the South Asia region (SAR) has the second largest poor population in the world, it is essential to improve the frequency and regularity of poverty data.

Table 1: Availability of Poverty Data by Region between 2002 and 2011

Category	East Asia & Pacific	Europe & Central Asia	Latin America & Caribbean	Middle East & North Africa	South Asia	Sub-Saharan Africa	Grand Total
No Data Point	12 48%	1 3%	12 39%	2 15%		2 4%	29 19%
1 Data Point	3 12%			6 46%	1 13%	18 38%	28 18%
2 Data Points (Interval of 5 Years or Shorter)			1 3%	2 15%	3 38%	9 19%	15 10%
2 Data Points (Interval of 6 Years or Longer)	3 12%	1 3%			2 25%	14 29%	20 13%
3 or More Data Points	7 28%	28 93%	18 58%	3 23%	2 25%	5 10%	63 41%
Grand Total	25 100%	30 100%	31 100%	13 100%	8 100%	48 100%	155 100%

4.   Defining Indicators of Data Availability

To monitor poverty and shared prosperity, ideally data availability should be satisfactory. Based on the two principles of minimum requirement for data needs and satisfactory levels of data needs for poverty, we suggest two indicators of data availability:
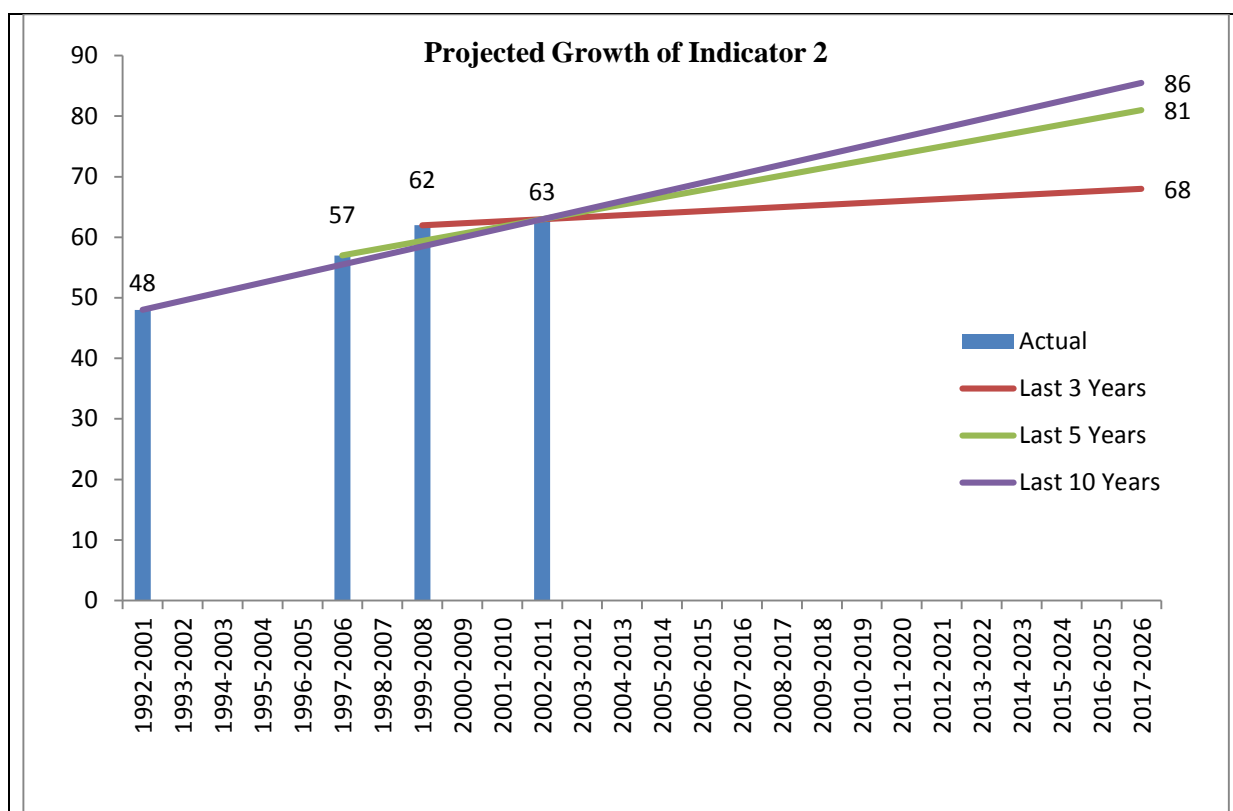
(i)   Indicator 1: *The number of countries with two or more data points in the last ten years*; and
(ii)   Indicator 2: *The number of countries with three or more data points in the last ten years*.

As mentioned earlier, the GDDS recommends every country to have poverty data every three to five years. The first indicator satisfies the minimum requirement of the GDDS's recommended frequency of poverty data every five years. The second indicator approximately measures whether a country collects poverty data every three years. Empirical evidence shows that poverty rates can change even within relatively short time intervals. Experiences from the global financial crisis in 2009 suggest that the lack of frequent poverty data can pose challenges to policy makers. For example, in Bangladesh, where poverty data are collected every five years, the timing of the household consumption survey did not coincide with the economic shock in 2009, and as a result, policy makers had no information on the effect of the global financial crisis on poverty in the country and could not formulate policies to lessen the negative impact of the crisis based on strong empirical evidence. Thus, more than two poverty data points in ten years is desirable.

5.   Setting Targets and a Goal

To monitor progress regarding data availability *vis-a-vis* the two indicators defined above, targets have to be set. To develop targets we go through the process of projecting the current trends of improvement/deterioration of the two indicators forward into 2030. As the discussion in section 3 shows, while both indicators show significant improvements, the pace of improvement differs across different time periods. Therefore, we calculated the annual average growth rate of these indicators for three different time periods: the last three years, the last five years, and the last ten years. Figure 5 summarizes the results.

Figure 5: The pace of growth in Indicator 1 and 2

Projected Growth of Indicator 1

Projected Growth of Indicator 2
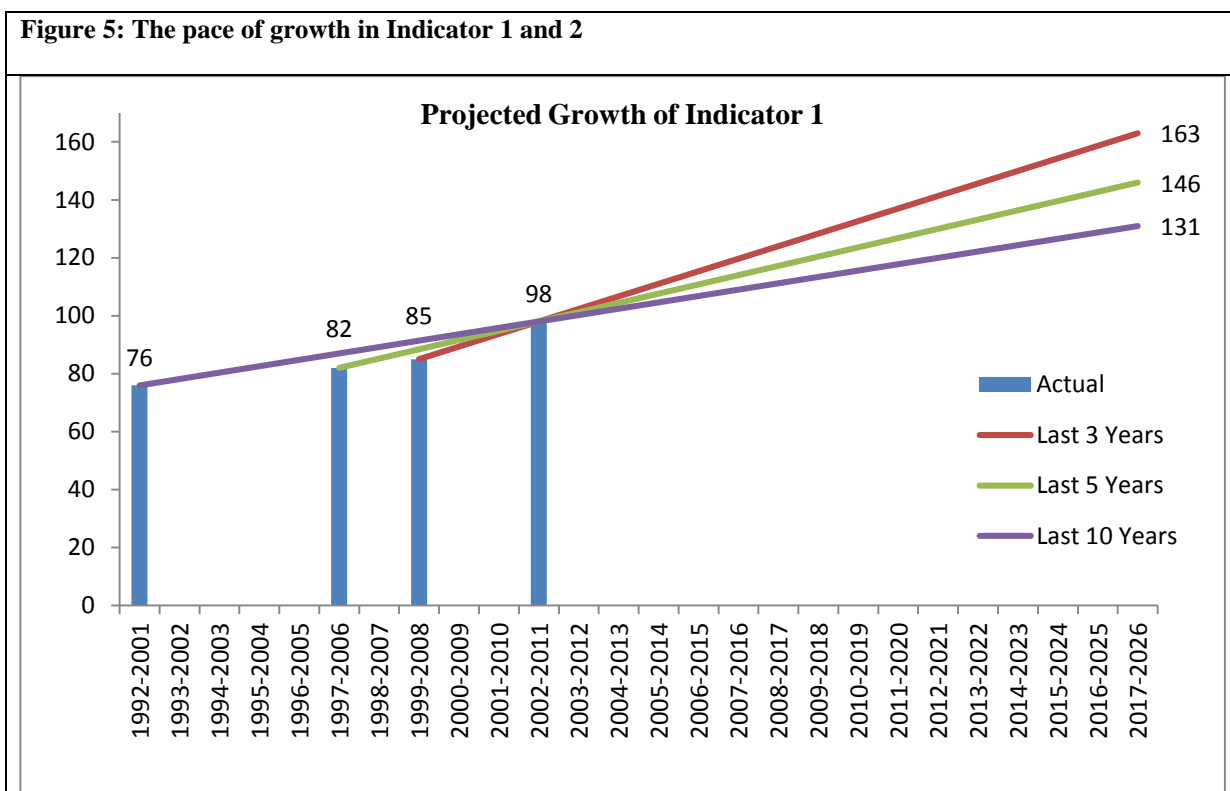
Although all projections assume the pace of growth in an indicator to be linear for the future as well, in reality the pace of growth can be very different. In fact, there are many reasons for the pace of growth to be non-linear. For example, recent improvements in the technology for data collection will certainly reduce the burden of data collection by NSOs. As shown earlier, there is a clear lag of nearly two years between data collection and dissemination of poverty data. The use of Computer Assisted Personal Interview (CAPI) has the potential to cut this lag significantly. Currently, very few countries use CAPI for data collection, and as more countries adopt CAPI, the pace of growth in both indicators of data availability could increase. Needless to say, the current delays between data collection and final dissemination are often due to weak capacity or bureaucratic inertia, and the success of CAPI would depend on making progress on those limitations.

There is also a possibility that the pace of growth declines. Some countries that can easily increase the frequency of poverty data, e.g., ECA countries, might have increased it already. The remaining countries may not have increased data collection simply because for them it is not easy to do so. For example, many countries under extreme data deprivation are small island countries that face much higher survey implementation costs per capita. Improving the frequency of poverty data in these countries is much more challenging than in other countries. Similarly, many fragile and conflict prone countries face real challenges in collecting data.

The complexity of projecting the pace of growth in both indicators can be seen from how significantly the future trends differ based on what reference period is chosen. Also, the relationship between the reference period and the pace of growth is not the same for the two indicators. Therefore, all projections should not be seen as predictions but rather seen as a means to set aspirational targets.

13

The annual growth rate of indicator 1 is fastest if the growth is measured by changes in the last three years. According to figure 5, the number of countries with two or more poverty data points has increased by four every year in the last three years (or more precisely, from 85 between 1999-2008 to 98 between 2002-11).

The annual growth rate of indicator 2 is the fastest if the growth is measured by changes in the last ten years. The number of countries with three or more poverty data points has increased by 1.5 annually for the last ten years. However, if growth is measured by changes in the last three years, it remained virtually unchanged.

Figure 5 shows projections of Indicators 1 and 2 into 2030 based on different growth scenarios. If both indicators grew at the pace of the last three years, all countries would have two or more poverty data points in ten year intervals, while 68 countries would have three or more poverty data points.[11] If both indicators grew at the pace of the last five years, 146 countries would have two or more poverty data points in ten year intervals by 2030, while 81 countries would have three or more poverty data points. Finally, if both indicators grew at the pace of the last ten years, in 2030 around 131 countries would have two or more poverty data points while 86 countries would have three or more poverty data points.

Based on these projections, we propose the following targets:

- Target 1: *End "Data Deprivation" by 2030. All countries will have two or more poverty data points in the last ten years by 2030.[12]*
- Target 2: *86 countries will have three or more poverty data points in ten years by 2030.*

Both targets are set using the best case scenarios; the pace of growth in the last three years is used for setting the target for Indicator 1 while the pace of growth in the last ten years is used for setting the target for Indicator 2. The target for Indicator 1 implies that alongside setting the goal of ending extreme poverty by 2030, the World Bank also would set the goal of ending data deprivation by 2030. It is not an exaggeration to assert that to ensure the end of extreme poverty, the end of data deprivation is critical. It might be attractive to set a more aggressive target 2. For example, a target could be set to push all the 35 countries that currently collect two data points in ten years (between 2002-11) to collect three of more data points. While perhaps appealing intuitively, setting such a target would have to make a clear break with historical trends.

The targets are indeed ambitious. One way to capture the challenge is to see how many extra surveys would be needed to meet the targets. To meet target 1 an additional 86 household surveys would have been needed in the last ten years: 58 surveys for the 29 countries with no data in the ten year window (2002-2011) and 28 more surveys for the 28 countries with one data point. To meet target 2 an additional 23 household surveys would have been needed in order to raise the number of countries with three or more data points from 63 to 86. This means that to meet both targets at least 109 additional surveys would have been needed

---

[11] The projection for indicator 1 in 2030 is 163, which is higher than a total number of countries in our dataset. Therefore, we say "all countries are projected to have two or more poverty data points in ten years".

[12] It would be imperative to set targets to ensure that data are reasonably current. An examination of the all of the countries with two data points between 2002 and 2011 shows that the latest survey was conducted after 2009, i.e., in the last 5 years or even more recently. For this reason we do not include explicitly include how recent the last survey was in the targets.

during 2002-2011, which is an average of 11 more surveys per year. Does this mean the total number of household surveys for every year needs to increase by 11 surveys on average?

Note that, going forward, the 109 additional surveys are a minimum requirement and are not necessarily sufficient to meet the two goals for the following reasons. First, this number assumes that all countries would repeat the same number of surveys carried out in the 2002-2011 period. If any country were to reduce surveys, other countries would need to fill the gap, which means an increase in the number of surveys needed to meet both the goals. Second, this number does not include any additional surveys for countries that already carried out more than three surveys in the 2002-2011 period.

We now interpret this analysis at an annual base assuming the number of household surveys in 2011 is a steady state.[13] In 2011, a total of 62 poverty data points were available and meeting the two targets appears to imply an average of about 73 data points per year, an increase of 11 data points per year or around 18 percent. However, just monitoring of the total number of household surveys per year is not enough; rather, it is necessary to see where the data are coming from. As mentioned above, surveys coming from countries that have already three or more surveys in the ten years would not count toward meeting the targets. Table 2 shows that 51 of the 62 data points include countries that already conduct three or more surveys per year; only 11 are from countries with less than three data points. Since the additional 11 data points would need to come from countries with less than three data points in the 2002-2011 period, the total number of surveys from those countries need to be doubled.[14]

Table 2: Breakdown of 62 countries which had poverty data points in 2011(by the frequency of surveys carried out in past ten years)

Number of Data Points between 2002-2011	Number of Countries	Cumulative Percentage
1	1	1.6
2	10	17.7
3	4	24.2
4	2	27.4
5	4	33.9
6	1	35.5
7	2	38.7
8	6	48.4
9	7	59.7
10	25	100.0
Total	62	

Furthermore, to meet target 1, 86 additional surveys need to be carried out over ten years in the countries with one or no survey, an average of 9 additional surveys per year. According to Table 2, only one out of

---

[13] This assumption is not so strong because the number of surveys per year is quite stable around 60 since 2002 (see Figure 1.)

[14] The numbers are similar when other ten year time intervals are considered (e.g., 2009, 2010, etc.).

57 such countries implemented a survey in 2011. Thus, the number of surveys in these countries need to increase drastically.

Finally, we propose to add a goal concerning improving the quality of poverty data. The World Bank's recent Policy Research Report (Jolliffe *et al* 2014) stressed the importance of improving the quality of data for monitoring the WBG's twin goals. Maintaining comparability of data over time and adopting and executing a poverty measurement methodology that meets the global standard are key markers of data quality. Since 2010, the Bank's Global Poverty Working Group (GPWG) has been gathering available poverty data across the world, and at the same time vetting the quality of the associated poverty statistics before sending them to the WDI. The number of poverty estimates that had to be dropped from the WDI – because the quality of the estimates was not good enough or the estimates were not comparable to those in the previous rounds – was quite high. Indeed, for WDI 2011 (the first WDI update after the GPWG was created) the GPWG could add 412 poverty estimates but also needed to drop 66 poverty estimates due to weak quality; about 86 percent of all estimates were deemed to be of good enough quality for the WDI database. Quality issues surrounding poverty estimates remain a concern, and given how important and expensive the collection of household survey data is, it is vital to ensure quality.

In this context, we propose another target for improving the quality of poverty estimates:

- Target 3: *All new poverty estimates are of good quality and thus can be included in the WDI.*

This target would imply that all, i.e., 100 percent, poverty estimates that are generated by countries should be of sufficiently high quality to be included in the WDI database. This again would be a challenging target, especially since target 1 calls for more surveys in countries which currently collect little or no poverty data, and with bold initiatives to expand survey coverage data quality could suffer unless special attention is paid.

6. Practical Considerations in Achieving the Data Availability Targets

Target 1: *All countries will have two or more poverty data points in the last ten years by 2030.*

As of October 2014, 57 countries have a pronounced inadequacy or deprivation of poverty data as they have less than two data points in the ten years interval of 2002-11. Out of these 57 countries with less than two data points, 29 countries do not have any data in the last ten years and 28 countries have only one data point. For those countries with data deprivation, a medium to long term commitment to improve data availability has to be established. The issue with these 57 countries seems twofold, 1) data simply do not exist, or 2) data cannot be accessed by the WBG.

Most of the 29 countries[15] with extreme data deprivation are small island states in the Caribbean Sea or the Pacific Ocean. Since sampling errors do not change much with population size, the per capita cost of reaching similar levels of sampling errors tend to be higher for smaller countries compared with larger countries. As a result, the financial burden of regularly carrying out sample household surveys to collect consumption data is high for most small island states in the Caribbean Sea and the Pacific Ocean. It is likely

---

[15] These countries are: St. Kitts and Nevis; Vanuatu; Tuvalu; Turkmenistan; Trinidad and Tobago; Tonga; Suriname; St. Vincent and the Grenadines; St. Lucia; Somalia; Samoa; Palau; Myanmar; Micronesia, Fed. Sts.; Marshall Islands; Libya; Korea, Rep.; Korea, Dem. Rep.; Kiribati; Haiti; Grenada; Guyana; Eritrea; Dominica; Cuba; Belize; Antigua and Barbuda; American Samoa; and Algeria.

that increasing both technical and financial support from international development agencies will be necessary to aid survey implementation.

Despite the high financial burden (in per capita terms), we would like to argue that it is vital for all countries – regardless of their size and impact on global poverty rates – to maintain regular monitoring systems of poverty and shared prosperity. Many of the small island states face frequent weather shocks, which are likely to be magnified by climate change, and these countries are also vulnerable to global economic shocks.

Several countries with no poverty data availability, such as Algeria, do in fact carry out household consumption surveys and maintain poverty estimates. However, they do not share data and estimates with the World Bank. For those countries, creating closer relationships and collaboration with the National Statistical Offices (NSO) would be essential.

For the 28 countries with a moderate data deprivation in monitoring poverty[16], it is vital to benefit from the countries' already existing capacity and structures (of conducting household consumption surveys) by providing additional technical and financial support. Most countries in this group are low income countries in Sub-Saharan Africa where the technical capacities of NSOs tend to be low and domestic funds for household surveys are limited. It would therefore be essential to build statistical capacity through a longer term collaboration of NSOs and the World Bank.

Reducing vulnerability to data deprivation requires that household consumption surveys are carried out not only frequently but also regularly with the same intervals (of five years). Supporting countries in this effort is demanding as countries with irregular poverty estimates are often countries with low statistical capacity or countries that are affected by external shocks (e.g., conflict) that prevent them from regular data collection.

Lastly, it is instructive to discuss the timeframe of achieving the targets. Poverty data is available with a lag due to the long data processing time: collecting, cleaning, and preparing data for analysis tends to be a prolonged process, at times, spanning multiple years from the start of fieldwork to the time when the data are ready for analysis.[17] For this reason this paper, written toward the end of 2014, uses the period 2002-2011 as the reference period for setting targets for 2030. If the same trend persists, the reference period to judge whether or not the targets have been met in 2030 would be 2017-2026 (and so the projections in figure 4 use this as the end period). This suggests a 12 year window to achieve the targets of ending data deprivation, a rather tight operational period. It would be expected that due to the technological progress being made, timeliness of poverty data can be improved and an active effort should be directed to this end. If data eventually become timelier it would add some extra time to the operating period for the targets. However, as matters stand at present, the timeframe to achieve the targets is quite small.

---

[16] These countries are: Zimbabwe; Seychelles; Yemen, Rep.; Timor-Leste; Tanzania; Swaziland; Sudan; South Sudan; Solomon Islands; São Tomé and Principe; Papua New Guinea; Morocco; Mauritius; Maldives; Liberia; Lebanon; Kenya; Iraq; Iran, Islamic Rep.; Ghana; Gabon; Equatorial Guinea; Djibouti; Congo, Dem. Rep.; Comoros; Cameroon; Burundi; and Angola.
[17] For example, of the 61 data points in 2011 (Figure 1), only one data point became available in 2011, 15 in 2012, 21 in 2013, and 24 in 2014. It took three years for all of the data points in 2011 to be finalized and available in the WDI.

Target 2: *86 countries will have three or more poverty data points in ten years by 2030.*

To achieve the target for indicator 2, the main focus will have to be placed on increasing data points for countries that already have two data points. For the period of 2002-11, 17 countries have two data points regularly while 18 countries have data irregularly (and are considered vulnerable to data deprivation). Incidentally, of these 35 countries, 18 are low income countries and the remaining 17 are middle income countries.

In order to reach the target of 86 countries with three or more data points, at least 23 countries that currently have two data points or less (in a ten year interval), will have to increase data frequency to at least 3 data points. Various existing efforts of increasing the number of data points from two or less to three or more can be capitalized on to achieve this target. For example, as already discussed, innovations in data collection methods and modes have changed the way in which data can be collected and analyzed. Traditional, face to face household survey data collection is costly, time-consuming and complex. Recent innovations in Information and Telecommunication Systems (ITS) and Computer Assisted Personal Imputation (CAPI) allows NSOs to switch to more affordable interview modes and techniques without impeding the quality of data.

Target 3: *All new poverty estimates are of good quality and thus can be included in the WDI.*

The target for indicator 3 measures whether new poverty estimates fulfill the quality criteria set forth by the Global Poverty Working Group (GPWG) and can thus be included in the WDI database. In some instances, poverty estimates are not included in the WDI after careful review of the methodology and comparability across years by the GPWG. This is often due to changes in the methodology, the lack of transparency of metadata information, or the lack of capacity and technical know-how on poverty measurement of NSOs.

One of the major quality issues is the lack of comparability of poverty estimates over time. Paradoxically, often the challenge arises from good intentions. The methodology of poverty measurement needs to reflect the consumption patterns of the people, but since consumption patterns may change over time, the methodology for poverty measurement also needs to be revisited once in a while. If the existing methodology does not reflect the current consumption patterns of consumption, it needs updating. A dilemma is that updating the methodology will likely make new poverty estimates incomparable to previous ones. However, this sort of challenge can be resolved through planning ahead. One way to maintain the comparability of poverty estimates is to prepare a subsample for which both data and the methodology of poverty measurement are the same as before. Using the subsample it is possible to produce poverty estimates that are comparable to the previous ones. Furthermore, using the rest of sample, poverty can be estimated based on the new methodology.

Lack of capacity, on the other hand, is also a serious source of quality concern and requires attention. To be successful, capacity building initiatives need to be hands-on and practical and build on existing local abilities, and very importantly, engender a local commitment toward sustainable improvements in data. For this, capacity building should not be narrowly focused or be one-time stand-alone initiatives, but rather be part of a longer term comprehensive engagement. It may also require governments to commit staff and financial resources to statistical efforts on a more permanent basis (rather than staffing tasks as they come along in an *ad hoc* manner).

*Importance of better planning and coordination on data collections*

It is important to entertain the possibility that the poverty data gaps discussed in this note may not result from the lack of household survey data *per se*, but result from the lack of planning or coordination in survey implementation. According to Demombynes and Sandefur (2014), who utilized the International Household Survey Network (IHSN) database, the SAR region on average has two household surveys every year. However, according to the WDI data used to construct the statistics in this note, each country in the SAR region on average only has slightly more than one poverty estimate for every five years. Similarly, according to Demombynes and Sandefur (2014), the SSA region has 1.5 surveys every year on average but the database of poverty estimates used in this note indicates that countries in Sub-Saharan Africa on average have roughly around one poverty estimate every five years.

Our own examination of the IHSN database and the World Bank's Microdata Library in conjunction with the WDI database reinforces that there is a disconnect between available surveys and poverty data. Figure 6 shows that low income countries on average had 16.6 household surveys between 2002 to 2011, slightly lower than the average number of surveys for lower-middle income and upper middle income countries (with about 19 surveys between 2002-11). However, low income countries had far fewer poverty estimates available, as demonstrated by their very high survey to poverty data point ratio (10.8) compared with lower middle income and upper middle income countries (7.5 and 5.6 per year).

These differences stem from the fact that the IHSN includes all surveys conducted in countries, while we focus on surveys that include consumption data used for poverty estimation. Thus, while many countries may have an abundance of surveys, most of the surveys are not suited for or tailored to poverty monitoring. While this does not mean that every survey carried out should include a consumption module, carrying out surveys in a more coordinated fashion would be effective in reducing key data gaps.

Figure 6: Availability of Household Surveys vis-à-vis Poverty Estimates

(a) By Region (between 2002-2011)
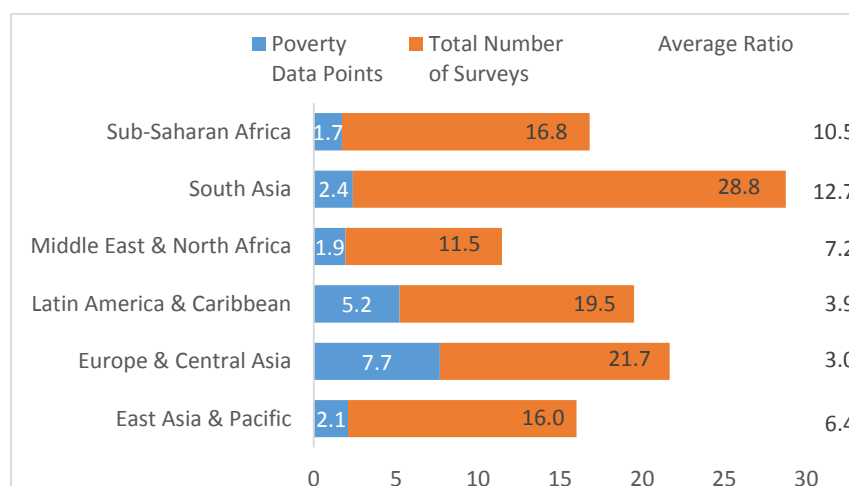
(b) Country Income Classification (between 2002-2011)
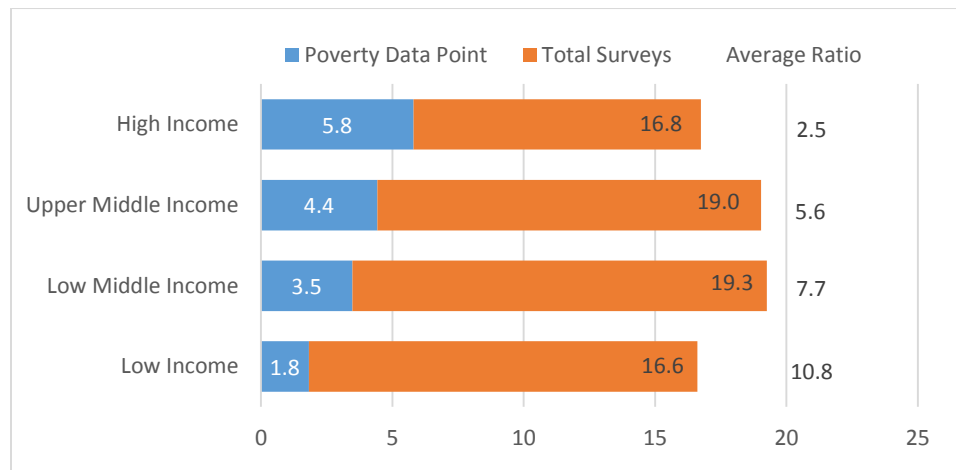
This observation leads us to suggest that attempts to improve the coordination or planning across different types of surveys would increase the availability of poverty data (and other data as well). To facilitate the coordination across different survey initiatives, PARIS 21 and development communities encourage countries to prepare National Strategies of Developing Statistics (NSDS). The NSDS typically include plans to undertake household surveys within five years or so, and if the NSDS are implemented as intended, the frequency of poverty data can be increased without increasing survey implementation costs. It will be very useful to evaluate the performance of NSDS in reducing the gaps between the number of household surveys in the IHSN and in the database used for developing this note.[18]

7.   Concluding Remarks

This paper analyzes the availability of poverty data using World Development Indicators, which include both national poverty data measured at country-specific poverty lines and international poverty data measured at the $1.25 or $2 a day poverty line. The WDI database in principle includes comparable poverty estimates that are vetted by the World Bank's poverty measurement specialists. Since efforts in global poverty monitoring such as the Millennium Development Goals and the World Bank Group goals use this database, the WDI database is appropriate for analyzing the availability of poverty data.

This paper defines the status of data availability in the following way:

- if a country does not have two or more poverty estimates in ten years, the country is categorized as "data deprived";
- if a country has two poverty estimates in ten years but the interval of surveys is more than five years apart, then the country is categorized as "vulnerable to data deprivation"; and

---

[18] Writing in the context of Sub Saharan Africa, Devarajan (2013) places some of the blame of the failure in the NSDS on the donors: "Many donors, including the World Bank, undertake statistical activities without ensuring that they are consistent with the NSDS. Why? Because donors may want the data for their own purpose—to publish a report, for example. Their incentive is to collect the data as quickly as possible, whereas building statistical capacity takes time. Even though the country owns the data, donors often behave as if they do. They often publish the data without any recognition of the source of the data". In a different context, Sandefur and Glassman (2014) point out how a coordination failure between donor and statistical agencies can lead to flawed data being produced (with there being a frequent disconnect between statistics produced from household surveys and from administrative data).

- if a country has three or more poverty estimates in ten years, the country is categorized as "satisfactory".

Based on this classification using the latest data (2002-11), we find that 57 of 155 countries face data deprivation, and only 62 countries have satisfactory frequency of poverty data. Furthermore, there are 20 countries that are vulnerable to becoming data deprived.

Based on projections using the past data, this paper proposes the following three targets:

- Target 1: *All countries will have two or more poverty data points in the last ten years by 2030.*
- Target 2: *86 countries will have three or more poverty data points in ten years by 2030.*
- Target 3: *All new poverty estimates are of good quality and thus can be included in the WDI.*

With the World Bank Group's recently proposed twin goals, acceptance of target 1 would imply that the world would mobilize efforts to end extreme poverty by 2030 and to end data deprivation by 2030 as well. While SDG indicators are not yet finalized, in the publicly available draft documents, the target of ending extreme poverty features prominently as well. This dual goal of ending extreme poverty and ending data deprivation is critical to ensure that the declaration of "ending extreme poverty" can actually happen. The analysis based on projections suggests that the targets are ambitious but feasible. The proposed targets also rather fittingly align with the call of the United Nations for a 'data revolution' – by exploiting advances in knowledge and technology, utilizing resources for capacity development, and improving coordination of efforts among key actors – to mobilize sustainable development (IEAG 2014). [19]

Finally, this paper discusses some practical considerations in achieving these targets by carefully looking into countries that face data deprivation or are vulnerable to facing data deprivation. Also, a comparison between the IHSN and WDI databases indicates that a potentially key problem is a lack of coordination between household surveys rather than lack of surveys. The role of NSDS and other coordination mechanisms will be key to improving data availability in a cost-effective manner.

The scope of the present paper has several limitations. The paper does not examine the question of data access. This is a complicated issue as there are several notions of access, which are often difficult to neatly classify. For example, some countries may provide full public access to their microdata, while others may choose to make publicly available a subset of variables (e.g., it may exclude location variables) or a subsample of the survey. Other countries may provide 'negotiated' access to their surveys on the basis of a clear understanding of what the data will be used for and some countries may provide no access to their data at all. While lack of data access can hamper poverty monitoring efforts in the same way as lack of data due to consumption surveys not being conducted, solving it would require a different approach and should be examined in greater detail.

The present paper does not delve into granular issues of implementation either. For example, a crucial practical consideration is to cost an initiative to end poverty data deprivation. While this paper does not attempt that, future work ought to carry out a careful analysis of costs based on different plausible scenarios. The database underlying this paper is well suited to that exercise as it pinpoints country-specific poverty gaps. The costs of surveys vary across countries, as do the funding sources (e.g., governments themselves,

---

[19] For more information see http://www.undatarevolution.org.

multilateral donors, bilateral donors, private sector actors). Once costs have been estimated, at country levels as well as at the global level, an effective engagement strategy to end data deprivation can be developed. The key driver of change would be the countries themselves, with the donors playing an enabling role by providing support ranging from technical assistance to funding.

References

Chen, S., and M. Ravallion, 2010. "The Developing World is Poorer than We Thought, But No Less Successful in the Fight Against Poverty." *Quarterly Journal of Economics*, 125 (4): 1577-1625.

Chen, S., M. Ravallion, and P. Sangruala. 2008. "Dollar a Day Revisited." World Bank Policy Working Paper 4620.

Demombynes, G., and J. Sandefur. 2014. "Costing a Data Revolution." Working Draft Paper, Copenhagen Consensus Center.

Devarajan, S. 2013. "Africa's Statistical Tragedy." *Review of Income and Wealth*. Series 59, Special Issue.

IEAG (Independent Expert Advisory Group on a Data Revolution for Sustainable Development). 2014. *A World That Counts: Mobilizing the Data Revolution for Sustainable Development.* IEAG Secretariat. (http://www.undatarevolution.org/report/)

IMF. 2013. *The General Data Dissemination System: Guide for participants and users*. Washington D.C.: International Monetary Fund.

Jolliffe, D., P. Lanjouw, S. Chen, A. Kraay, C. Meyer, M. Negre., E. Prydz, R. Vakis, and K. Wethli. 2014. *A Measured Approach to Ending Poverty and Boosting Shared Prosperity: Concepts, Data, and the Twin Goals*. Policy Research Report. Washington, DC: World Bank, 2014.

PARIS21. 2004. "National Strategy for the Development of Statistics (NSDS) Essential." Paris.

Sandefur, J., and A. Glassman. 2014. "The Political Economy of Bad Data: Evidence from African Survey and Administrative Statistics," *Journal of Development Studies*, forthcoming.