



The Initiative for Policy Dialogue

Report on Global Poverty Workshop

New York, 31 March – 1 April 2003

Sudhir Anand and Joseph Stiglitz opened the Workshop with general introductions and a statement of its goals. Anand discussed the motivation for the Workshop, and went through the agenda and background papers. Stiglitz reminded participants why global poverty is a matter for international concern. Not only is it a human tragedy for those in poverty, but the consequences of deprivation can spread across borders. Poverty can contribute to social and political instability and is potentially a breeding ground for terrorism. The Millennium Development Goal for poverty reduction represents the international commitment to a concerted and coordinated approach to the problem.

This Workshop has been called in response to the uncertainty in our estimates of global poverty. Even the question of whether absolute poverty is increasing or decreasing is contested. According to World Bank estimates it is going down – although it is increasing outside China, and rates of change vary widely by region. But there are alternative estimates showing different trends.

There are three broad questions. First, how should a poverty line be set? Second, how are comparisons made over time and across countries? Third, how should we measure it empirically?

Session 1: From Inequality to Poverty

Branko Milanovic presented his findings on global inequality. By using income and expenditure data directly from household surveys he finds that global interpersonal inequality (denoted *concept three inequality*) increases between 1988 and 1993 and then falls slightly in 1998. On the other hand, population-weighted inequality between countries (*concept two inequality*), when measured using national accounts (NA) estimates of GDP per capita, falls steadily over the period. Concept two inequality typically comprises two-thirds of concept three inequality so the question arises: why do the two move in different directions between 1988 and 1993?

The methodology for Milanovic's concept three estimates differs from NA-based concept two estimates in four respects. First, concept three estimates count within-country inequality while concept two estimates by definition do not. Second, he splits some large countries into two. For instance, he separates rural and urban China, and rural and urban India. Third, the mean per capita expenditure (or income) in a country is taken directly from surveys rather than the NA. Fourth, since he is restricted to those countries with household surveys of sufficient quality, his sample has a different coverage – 90 countries rather than the 140 with NA data.

Which of these differences explain the differing trends? He finds that the inclusion of within-country inequality in his data makes a difference to the level of world inequality but not to the trend. That is to say, estimating concept two inequality using his concept three dataset, with incomes and expenditure taken directly from household surveys, does not change the trend. Data coverage and splitting up (two) large countries make very little difference. So what drives the difference in trend between 1988 and 1993 is the use of mean expenditure (income) from surveys versus GDP per capita from the NA. Between-country inequality appears to have increased when measured using mean expenditure (income) from surveys, and to have decreased when measured using GDP per capita from the NA.

Howard Nye pointed out that NA estimates of household consumption expenditure are different from NA estimates of GDP. If one is going to use the National Accounts for estimating global inequality, then the former would seem to be more appropriate. This is the method used by Bhalla for estimating global poverty, while Sala-i-Martin uses GDP for both inequality and poverty.

Robert Johnston presented a discussion of the history of poverty lines, noting that a variety of methods have been used since poverty became a subject of research in the UK in the mid-nineteenth century.

Sakiko Fukuda-Parr presented the human development approach to deprivation, a perspective concerned with multiple dimensions of well-being and poverty rather than just command over resources (whether measured through incomes or expenditure). She pointed out that a number of countries have suffered increasing mortality rates and decreasing school enrolment rates in recent years, in addition to declining incomes. Fifteen countries saw declines in their Human Development Index (HDI) between 1990 and 2001. Some countries in South Asia and Sub-Saharan Africa have seen declining immunization rates. The most precipitous declines in life expectancy in Sub-Saharan Africa are due to HIV/AIDS, but this is not the only cause of the declines.

On the other hand, she noted that great progress is possible even in stock variables or variables that are typically slow-moving. Sri Lanka increased its life expectancy by 12 years over a period of just seven years between 1946 and 1953, while Botswana's net primary school enrolment rate approximately doubled in the 15 years after independence in 1970.

Angus Deaton observed that life expectancy, like other indicators based on mortality rates, is not a stock variable like literacy. It can therefore decline rapidly such as during an epidemic, and then bounce back to its former level. Such sudden declines should not be especially surprising. Sudhir Anand agreed but noted that the declines in life expectancy in Sub-Saharan Africa and parts of the former Soviet Union have been sustained over some years and in some countries are recovering only very slowly, while in others they have not recovered at all.

Augustin Fosu in his discussant's comments observed that inequality within countries matters even if ultimately we care only about poverty, because the shape and dispersion of the distribution of expenditure/income affects the elasticity of poverty reduction with respect to per capita expenditure/income growth. Moreover, we should be concerned that growth may raise inequality in the short term, and bypass the poor. He also noted that Milanovic's inequality estimates lacked robustness tests.

Session 2: Problems in National and Global Poverty Measurement

Albert Berry made a presentation on inequality and poverty in Latin America (LA), where surveys typically measure household income but not household expenditure. Very high levels of income inequality in this region imply that poverty could be substantially reduced by modest transfers from the rich to the poor. Measured Gini coefficients in the region are typically between 0.4 and 0.6, not counting capital income. Including capital income, which is usually not measured, would probably raise an observed Gini of 0.5 to about 0.55.

His subjective estimate of a 95% confidence interval due to measurement error would be about +/- 4 to 5 Gini points. This level of uncertainty makes it very difficult to estimate trends, and therefore to judge the effects of policies on inequality. Such policy analysis would require one to be able to reliably measure a change of 2 or 3 Gini points over a period of a few years, which is currently not possible. The result is that one cannot reliably estimate the direction of change of inequality over the last 25 years in most Latin American countries. Some shorter-term changes are clearer, but we do not know, for instance, the overall effect on inequality of the policy changes in Chile since the 1970s.

Much of income inequality in Latin America is due to the very large income shares of the top 1% or top 5%. Changes in their income shares have a large impact on inequality but are hard to measure, partly owing to disproportionate undersampling of – and underreporting of incomes by – the rich. We are likely to know more about what is happening in the lower parts of the distribution, including those below various poverty lines. Nonetheless, there are also problems of measurement in the lower parts of the distribution, viz.:

- Mismeasurement in rural areas is very serious. Many countries in Latin America did not have rural surveys at all until recently. Brazil is an exception,

but Colombia and Argentina, for example, still have no reliable rural surveys. And rural incomes are difficult to measure even when there are household surveys. Small-scale agricultural incomes are particularly difficult to measure.

- Public goods provided to individuals are typically not measured. Since countries in LA are relatively high-income countries with some very poor people, the value of public goods can be a high proportion of the incomes of the poor.
- Price differences between rural and urban areas, and differences in access to services, are poorly measured if at all.
- Instability of income is not taken into account. Since surveys in the region are mainly income rather than consumption surveys, variability of incomes adds a great deal of noise when what we are typically concerned about is inequality of consumption expenditures.

Until we have substantial improvement in these areas we cannot have much confidence in poverty or inequality figures from Latin America.

Priorities for research are to determine the effects on poverty of market reforms, human capital formation (education), and policy towards small- and medium-sized enterprises (SMEs).

The two most serious attempts at analyzing the effects of market reforms have derived different conclusions. One cross-country time-series analysis by ECLAC finds that reforms have been mildly positive for growth and mildly negative for inequality – on average leading to an increase in inequality of about 0.5 Gini points. On the other hand, a deeper country-by-country analysis, which allows for lagged effects of reforms, suggests a large negative impact on inequality of 3 to 5 Gini points. Given this disagreement, and the cautionary notes above regarding the reliability of data, all we can say with confidence is that market reforms almost certainly did not reduce inequality. Whether they made it worse, and by how much, is not possible to say.

Not only is there measurement error in the data, but often the available data are not used well. Different data can appear to be, or actually be, contradictory, and reconciling them into plausible estimates requires careful country-by-country analysis that few researchers have undertaken. We need more integrated studies at the country level. We need people spending a year or two looking at all available distribution data within a country, from household surveys, firm surveys, wage data, etc. Cross-checking the various possible sources is important, and if this is not done serious mistakes may be made.

This kind of research allows the analysis of mechanisms that underlie changes in inequality and poverty, and can help to avoid errors. For instance, Colombia introduced a policy of giving vouchers to the poor for secondary schooling, but found that secondary school enrolment rates did not improve. It was subsequently discovered that the enrolment data contained biases. Enrolment figures reported by states affect government transfers to them, leading to an incentive for state officials to

overreport enrolment rates. Alternative methodologies were then used that were able to determine that secondary enrolment rates were increasing.

Stiglitz noted that economic reforms often increase income volatility. Berry agreed, but said that we cannot tell how much of an effect this has on inequality.

Deaton argued that the poor quality of LA surveys cannot be generalized to other developing countries. The LA surveys are of poor quality, but in addition they suffer from being income rather than expenditure surveys. Income surveys face particular problems. For instance, households with zero reported income may be quite well off with relatively high consumption expenditures.

Stiglitz further noted that growth in LA was much worse in the 1980s and 1990s than in the 1950s to 1970s, and suggested that those growth spurts that did occur in the later period are likely to have been unsustainable. They were partly driven by privatization of assets and resulting inflows of external finance. In order to allow for this, growth figures should be based on net national income rather than on GDP.

Vijayendra Rao presented a paper arguing that the poor in three small villages in India systematically pay higher unit prices for food than do the rich, taking account of (fine) variations in quality. Liquidity constraints, such as the refusal by shop owners to extend credit to poor households, compel the poor to buy goods in small quantities. Non-linear pricing results in small quantities costing more per unit than the typically larger quantities bought by richer households, with prices up to two or three times higher per unit. The estimated Gini within these villages is 0.257, but taking account of price differences faced by the poor this rises to 0.316.

One cannot generalize from this small sample to all cases, but the fact that most surveys fail to take account of such price differences suggests that they may understate inequality and poverty.

Deaton commented that the survey used in Rao's paper does not properly price own consumption of production. Furthermore, NSS data on unit values suggest that nationally the problem of differential prices does not arise to the extent found by Rao. Rao maintained that NSS misses the finer grades of quality variation required to establish this.

Ivo Havinga made a presentation on National Accounts Systems, focusing on estimates of household final consumption expenditure (HFCE). He pointed out that the system used had changed over time, with major changes occurring in 1968 and 1993. The change in 1993 led to estimated GDPs increasing by an average of two to three percent.

NA estimates of disposable income comprise a number of elements, estimated in different ways. There is currently considerable variation in calculation methods across countries. The UN is encouraging convergence in methods, and in particular is encouraging countries to use household surveys to estimate HFCE in their NA. When NA data use household surveys there will no longer be the possibility of the inconsistency that now occurs.

Survey and NA estimates of HFCE differ in concept. For instance, the latter includes imputed charges for financial intermediation, consumption by non-profit organizations, and the imputed value of owner-occupied housing. In principle, surveys could measure the latter, but often they do not.

There was discussion of the possibility of imputing the value of social benefits in kind through surveys – hospital stays, school education, etc. Havinga argued that it is possible, and that it is supposed to be included in the NA. Deaton questioned the possibility of correctly pricing such services. Ravallion commented that food transfers are included in NA estimates of HFCE.

It was also noted that the first Millennium Development Goal (global poverty reduction by half by the year 2015) refers only to private consumption and does not include transfers in kind. Stiglitz observed that this can have perverse results: a change from publicly-provided education to private education can reduce measured poverty with no change in real consumption.

Surjit Bhalla argued that the most reasonable assumption is that the poor do not benefit from government expenditures. Deaton agreed that this is true in India but added that it may vary across countries.

On the gap between survey and NA estimates of household consumption, Deaton reported that the financial services and insurance charges included in NA HFCE but not in surveys can play a large role. Not only does their inclusion lead to NA estimates being higher, but as countries become richer the share of their financial sectors in the economy grows. This conceptual difference between NA and survey consumption may explain part of the observed widening of the (proportional) gap between the two in some countries. Since these services are not likely to be used by the poor, this argues against using NA HFCE growth rates to estimate changes in poverty.

Deaton also reported that in India there has been substantial double-counting of some commodities in the NA estimates of HFCE. HFCE is calculated commodity-by-commodity as the residual of total production (net of imports and exports) minus government consumption and intermediate (business) consumption. But for some commodities, notably edible oil, the intermediate consumption component has been underestimated. Some oil used by restaurants is therefore counted twice, once as edible oil from national production attributed directly to households, and once as part

of the price of the restaurant meal. Such double-counting may explain some of the divergence between survey and NA household consumption that is not due to differences in concept. It is another reason not to use NA HFCE growth rates for poverty estimation.

Suresh Tendulkar stressed that the residual nature of HFCE entails that mistakes made anywhere else in the NA – whether in aggregate production, intermediate consumption, or government consumption – will end up in NA estimates of HFCE. Thus there is little reason to believe NA estimates of commodity consumption over survey estimates. The latter are at least direct observations, despite suffering from their own problems.

Deaton noted that the main problems with surveys are under-representation of the rich and underreporting of income or consumption by the rich who do choose to respond.

Martin Ravallion reported that in China there is virtually no difference between survey consumption and NA HFCE once relevant corrections are made. In Eastern Europe, he calculates that survey consumption growth has no correlation with NA HFCE growth. Globally excluding Eastern Europe he estimates survey consumption growth to be about 85% of NA HFCE growth, and that this figure is statistically indistinguishable from 1. Bhalla claims he finds this same ratio to be 27%, and statistically indistinguishable from zero. This disagreement was not resolved.

Anand stated that one important outcome of this meeting should be that the people working on National Accounts at the UN liaise and collaborate with those at the World Bank working on these issues.

Teleconference with Nick Stern of the World Bank

Nick Stern expressed reservations about the attention given to the international \$1-a-day poverty measures. First, what happens within each country is more important than international comparisons, and should not be obscured by global figures. Second, the whole distribution of consumption matters, and not only that of people below the poverty line. The MDG on poverty reduction should be interpreted as a more general concern with those near the bottom end of the distribution. Finally, the \$1-a-day poverty line should not detract attention from non-income (/non-consumption) dimensions of poverty. On the other hand, the international poverty line has been successful at focusing international attention on poverty.

There are several problems with the estimates of global poverty. First, the PPP exchange rate database is an embarrassment. Estimates are nearly 10 years out of date, cover too few countries, and are generally unsatisfactory. The next International Comparison Project (ICP) should improve matters, but estimates from it will not appear until 2005. So two questions need to be resolved by the research community:

How do we best make use of what we have now? And how do we make the next set of estimates as good as possible?

On the question of using survey estimates of consumption directly versus using survey distributions scaled up to NA HFCE means, Stern prefers to use surveys directly.

Stern described the steps the World Bank is taking to improve transparency. More and more information will appear on its websites, and the Bank will engage in consultations with expert groups over the next year. Once these processes have developed to the stage that experts outside the Bank are able to rigorously check the Bank's work and produce their own alternative estimates, the Bank will engage in open public consultation.

Session 3: PPPs for Measuring Global Poverty Over Time

Bettina Aten and Alan Heston presented the Penn World Tables (PWT) approach to calculating PPP exchange rates. In calculating a national aggregate price level, goods are weighted by the amount of money spent on them. In this sense the weights are plutocratic. This entails that the goods bought by the rich receive higher weighting than those bought by the poor, so the PPP exchange rate may not accurately reflect the prices faced by the poor for the goods they buy. Alternative weights could in principle be used that would better reflect prices paid by the poor. Food prices would figure prominently in such a price index. The necessary data for doing this do exist in some cases.

They have recently started considering the possibility of different groups paying different prices for the same goods, but as yet this is not allowed for in PWT estimates. The EU now records differential prices for the same goods. Aten reported finding that in the late 1980s a poverty consumption bundle in New York or Los Angeles cost nearly 50% more than in some less metropolitan areas, reinforcing Rao's point that different people may face widely varying prices.

The estimation of PPPs at PWT proceeds in two steps. They start with price data from 'benchmark' countries, which took part in the price surveys of ICP. These data are from the years 1990 to 1995. Prices in some non-benchmark countries are taken from external sources, such as the Employment Conditions Abroad organization in London and the International Civil Service Commission. When none of these sources is available they use benchmark data from previous years. Countries are classified in the documentation by whether or not they are 'benchmark' (i.e. whether or not they took part in the ICP). Price and expenditure data in different countries are then aggregated to construct an international price vector using the Geary-Khamis (GK) method to obtain price levels in each country for Consumption (C), Government expenditures (G), Investment (I), and Net Foreign Balance (NFB). Prices are

projected backwards to construct price series from 1950 to 2000, with a 'consistentization' process used to reconcile these estimates with estimates from previous ICPs.

Countries are graded in PWT documentation by the reliability of their data and whether they are benchmark countries. Confidence intervals for price levels have not been provided. (The 1991 Summers and Heston paper describing PWT Mark 5 has some estimates of confidence intervals for non-benchmark countries. The 95% confidence interval for real GDP estimates ranges upto 60% up or down for countries with GDP per capita one tenth or less of that of the US.)

Reddy found using PWT raw data that PPPs for food in poor countries are 30-40% higher than general PPPs, still worse for staple foods like cereals and grains. This implies that poverty is higher than measured since the poor face higher prices than suggested by the aggregate PPP exchange rate.

Stiglitz asked whether the difference between food prices and aggregate prices changes over time. Deaton stated that the relative price of food falls secularly with development. Heston cited one study of China that found a convergence of prices across regions, although this does not imply anything about changes in the price of food relative to other goods.

Deaton also commented that small changes in the PPP rate used can make a huge difference to measured poverty.

On the relationship between PPP exchange rates and market foreign exchange (FX) rates, Aten reported that some evidence suggests that trade openness moves the two closer together. Stiglitz observed that, on the other hand, capital market liberalization leads to FX rates being determined more by capital flows – much of it short term – than by real price differences. This would tend to reduce the correlation between FX and PPP rates.

Yuri Dikhanov presented on the World Bank method of estimating PPP rates and on his (and Michael Ward's) estimates of global inequality. The primary differences between the World Bank's and PWT's estimation methods for PPP rates are that the Bank uses the EKS rather than the GK method, and imputes prices to non-benchmark countries using a method different from PWT. The difference between GK and EKS estimates of real GDP can be up to 50% for some countries in some years.

There is no reason to expect consistency between national price deflators and changes in PPP rates over time. Growth rates calculated using the two can therefore differ. Moreover, countries sometimes change their method for calculating national price deflators, leading to further non-comparability. The US, for instance, has recently changed its method of calculating its consumer price index (CPI).

For their estimates of global inequality, Dikhanov and Ward use World Bank consumption PPP exchange rates with NA estimates of mean HFCE, and for inequality within countries they rely on data from the UNU/WIDER World Income Inequality Database (WIID). Thus mean per capita consumption expenditure is taken from the NA rather than from the surveys which provide their estimates of inequality within countries. They find that inequality increases over time according to both the Theil and Gini indices. On eyeballing the shifting world income distribution, Deaton expressed scepticism of the finding that inequality had increased, but this was resolved by the observation that the increase in inequality was driven by increases in the share of the richest 10%, which were difficult to spot on their income distribution graphic.

Anand noted that the Dikhanov-Ward estimates showed concept 2 (population-weighted between-country) inequality to be increasing, unlike *all* other studies presenting estimates based on NA data (e.g. those by Milanovic, Schultz, Sala-i-Martin, Bhalla). This, he claimed, was surprising.

Session 4: Poverty Trends in China and India

Carl Riskin presented on poverty levels and trends in China. Official poverty figures report that rural poverty (at a very low line, about 66¢ per day in 1985 terms) fell from about 33% to 3% over 1978 to 2000. (Independent estimates of poverty are typically somewhat higher than official estimates but still show a rapid decline.) Despite this dramatic decline, the government is still very concerned about rural poverty and continues to consider it a serious problem.

Riskin described the following problems with official estimates of poverty (based on work by Park and Wong):

- They are restricted to rural China. Only recently have official attempts been made to calculate urban poverty, using different poverty lines. It remains very difficult to estimate conditions for rural-urban migrants, of whom there are about 40 million each year.
- The rural poverty line has not kept up with inflation. One calculation says this leads to the 2000 poverty line being 13% too low.
- The food component of the consumption bundle defining the poverty line excludes ‘non-essentials’, so that foodgrains comprise 88% of the defined basket while poor households actually spend closer to 70% on foodgrains.
- The non-food share of the consumption bundle is correspondingly lower than what the poor actually buy.
- The official poverty line is treated as an income rather than an expenditure line, which matters particularly because China has a high savings rate.

Using the PPP\$1-a-day expenditure line yields much higher poverty rates. This is due partly to the line being higher, but partly also to its being applied to the expenditure rather than the income distribution.

It is important to realize that China has never participated in an ICP. The best existing estimates of national prices come from 1986 list prices.

Reddy reported that there are four respectable estimates of Chinese prices, which range by a factor of two. We can therefore have no confidence in any estimates. Dikhanov disagreed with this conclusion.

Ravallion pointed out that changes over time are calculated using a rural CPI, so while the estimate of the PPP exchange rate makes a difference to estimated levels of poverty, it does not make a difference to the estimated rate of change of poverty.

Ravallion and Riskin described the process of rural poverty reduction in China. During the period 1979-2000 after the start of the reforms, rural poverty declined in two waves. It declined substantially in the first five years of this period, and then again in the period 1993-1996. These declines were due primarily to changes in the terms of trade between rural and urban areas of China, which in turn were due to a change in government policy on grain prices. But in the late 1990s, the rural-urban terms of trade shifted away from favouring rural areas. Poverty levels stagnated over those years but they did not increase as one would expect if the terms of trade had been all that mattered.

Abhijit Banerjee commented that it is extremely important that so much of world poverty reduction is due to these simple (price) changes in Chinese government policy, and that this is much less well-known than it should be.

Riskin reported that in the past there was very little urban poverty because the government favoured urban compared with rural areas. Urban poverty has now grown due to the laying-off of millions of workers in the former state enterprises and rural-urban migration in recent times.

Suresh Tendulkar presented two controversies in the measurement of poverty in India. The first relates to the question of comparability of the different rounds of the National Sample Survey (NSS). The second is concerned with the growing divergence between consumption expenditure as measured by the NA and as measured by the NSS.

On the first issue, a change of recall period for the NSS for the 1999/2000 round created the possibility of non-comparability with previous rounds. With independent samples a 7-day recall period leads to 30% higher measured consumption than a 30-

day recall period, while asking the same households their consumption over both the last 7 days and the last 30 days yields virtually identical reported (average) consumption. It appears that respondents adjust their answers to ensure consistency. Examining the Consumer Expenditure Survey (CES) and the Employment and Unemployment Survey (EUS) he finds that the 7-day response is being reported as lower than it would have been had they not also been asked their 30-day consumption, and that the 30-day consumption is being recalled accurately. It seems that when respondents are asked *only* for their 7-day consumption they include items consumed longer ago than 7 days back. This is denoted as *telescoping error*.

He asked what made the NSS and NA estimates of consumption expenditure growth so different? The concept of consumption was different, counting different elements; and HFCE in the NA was indirectly observed – being estimated as a residual – whereas NSS measured consumption directly (albeit with error). There have also been discontinuities in the methods that have been used to calculate the NA.

It should be noted that many different types of survey are used for the estimation of components of the NA. Aggregate national production requires surveys for estimation – for instance, small-scale agricultural production that is not in official records, and intermediate consumption by businesses. These surveys are typically of lower quality and more out-of-date than the NSS, so that in estimating household consumption expenditure Tendulkar argued there is no reason to believe that NA is better than NSS.

Deaton commented that a further problem in the NA is deciding how to price goods that are recorded as quantities. He also expressed disagreement with the finding by Abhijit Sen that there has been no serious divergence between NA HFCE and NSS consumption. He finds a clear divergence even when everything possible is controlled for. This may be due to the fact that as consumers get richer they spend a higher proportion of their incomes on processed goods such as restaurant meals, some of which is double-counted. This increasing double-counting would lead to increasing divergence in the two estimates of consumption.

Bhalla argued that rapidly increasing wage rates of very poor ploughmen in India appear to contradict claims of increasing inequality, and undermine NSS estimates. Tendulkar responded that ploughmen were at the same time working fewer days per year so the increased wage rate did not necessarily translate into increased income or consumption. Deaton commented that the wages of ploughmen do not necessarily reflect other rural wages.

Stiglitz observed that in the US all the income growth between 1973 and 1993 was in the upper part of the distribution, and incomes at the bottom were declining. If there were underreporting of upper incomes then one would expect the kind of divergence we see in India.

Banerjee said that in India it is true that the richest have seen huge increases in their incomes, but even so their underreporting cannot account for all or even most of the difference between NA and NSS.

Stiglitz asked about changes in agricultural productivity in India. Ravallion replied that it has not grown much, which is more consistent with the lower poverty reduction as measured through NSS than the higher poverty reduction when NA means are used to scale up distributions. Ravallion argued that when all biases had been removed that it was possible to remove, it looked as though the rate of poverty reduction in the 1990s was no greater than that in the 1980s.

Angus Deaton presented on inequality, poverty and growth, and the divergence between NA and survey consumption worldwide. He finds no systematic relationship between growth and changes in inequality, implying that growth is good for the poor *on average*. But there remains disagreement over how much poverty reduction there has been – for example, between the World Bank and Bhalla – because of the different data that have been used to construct the respective estimates.

Surveys are probably more reliable than NA, but if this is the case then aggregate GDP growth rates recorded in NA are probably wrong. So if surveys show there has been less poverty reduction than do the NA, and we believe them, then we ought to conclude that the NA overstate GDP growth.

Deaton has collated data from 498 surveys worldwide from 1979 to 2000 using a variety of sources: Shaohua Chen, his own estimates for India, and published sources. His sample has varying coverage of the world population, going up to about 60% in the late 1990s. In these surveys, the ratio between survey consumption and NA consumption overall is 0.86, with a standard deviation of 0.314. The ratio varies substantially by region, and has fallen over time, particularly in the poorer countries. For India, the growth in his estimates of survey consumption is 1.1% per annum slower than that in the NA consumption estimates. In the UK, the difference is 0.5% per annum; and in the US 1.3% per annum, although the ratio there of Consumer Panel Survey (CPS) *income* (as opposed to expenditure) to NA *personal income* is not going down.

We certainly cannot simply assume that either surveys or NA are correct. So why is there a difference? First, there is non-response in surveys by the rich: when the response rate decreases with income, the observed mean will be biased downwards and observed inequality will also be biased downwards. Moreover, the assumption of a straightforwardly decreasing response rate with income is probably too simple. Response rate is likely to depend on other factors too, such as whether the potential respondent lives in rural or urban areas.

Non-response is surely part of the reason for the difference between surveys and the NA, but it does not explain why income surveys seem to suffer less divergence from

the NA than do expenditure surveys. It may be that income surveys are less onerous than expenditure surveys for respondents, so that rich people are not as reluctant to respond to them.

Imputed goods such as own production are probably better measured by surveys than by the NA, as are illegal activities.

Part of the difference arises from different definitions of consumption. HFCE in the NA includes financial intermediation services and consumption by non-profit organizations. The former increased in India from 0% to 2.5% of household consumption over 10 years. The fact that the concepts are different makes it conceptually illegitimate to scale surveys to NA total consumption: it is quite possible for measured inequality according to the survey definition of consumption to remain the same, overall consumption according to the NA definition to increase, and the poor to get poorer.

Session 5: The first Millennium Development Goal, Growth and Poverty Reduction

Jeffrey Sachs presented on the first Millennium Development Goal (MDG) – global poverty reduction by half by the year 2015. He noted that dimensions of deprivation other than income or expenditure poverty also suffered serious measurement problems. For instance, there were no reliable data on numbers of malaria deaths.

He then argued that the World Bank and the IMF pay too little attention to the ultimate goals of policies and too much to intermediate outcomes. Their policy aims typically refer to outcomes such as credit expansion rather than to poverty reduction or hunger alleviation. Furthermore, the Bank and Fund are too inclined to tell countries to adjust expectations downwards. One Bank mission told their client-country government that it could not expect to meet the MDG for poverty reduction and should therefore adopt a more modest target. Instead, they should have made a serious effort to find out what is preventing such a goal from being achievable, and helped the country to calculate how much assistance would be required to achieve it.

Most of the developing world is on track to achieving most of the goals, primarily because big countries like India, China, Mexico, and Brazil are doing relatively well. But there are still at least a billion people living in countries that are nowhere near on track. Many countries in Sub-Saharan Africa, the Andes, parts of Central America, some of FSU, some of South Asia, are not on track.

In those countries that are really struggling, the obstacle is geography more than policy. Certainly bad macroeconomic policies can ruin a country, but a number of poorly-governed countries that are large or coastal are performing relatively well, and

small inland economies find it very hard even with good policies. Large countries are more able to benefit from efficient division of labour and to provide adequate supplies of it, making them attractive to foreign investment. Inland countries struggle with transport costs, keeping them out of world markets. ‘Success stories’ like Uganda are not the successes they are sometimes portrayed to be. Exporting primary commodities is not a route to growth.

For some problems the binding constraint is simply money. The poorest countries cannot combat malaria by good policies alone: they need cash.

Countries at the bottom of the pile cannot sit back and rely on free-market forces to develop their economies. The marginal product of capital in countries with no capital is not infinity but zero, *pace* standard growth theory (i.e., the so-called Inada conditions). Investment in infrastructure such as roads is particularly important, and often a prerequisite for other investments (private and public) to yield returns. Capital does not simply flow to capital-poor areas in search of returns.

Banerjee argued that the PPP\$1-a-day line is in some ways effective, but that were poverty to go down in the aggregate it would not obviate the need to worry about poverty. Even if India and China drive the global poverty numbers down, there remain real problems in certain countries, as well as in certain areas of India and China. He agreed with Sachs that it is important not to ignore these pockets of poverty.

Furthermore, when we are discussing the case for or against ‘globalization’, or policies generally, it is a mistake to point to the success of China and India as a validation of the global economic system. From the point of view of policies, they comprise only two observations. Moreover, they are exceptional in some respects and therefore cannot easily be adopted as models for other countries. Their size makes them particularly attractive to international investment. Smaller countries, and those without established industries, find it very difficult to compete. China and India are not only exceptional, they are also part of the problem: they pull investment away from other countries. But if we focus only on global poverty numbers we will fail to recognize this problem, because we will overlook smaller countries.

We should care not simply about falling aggregate poverty, but seek to understand why it is falling. No one really knows, for instance, why rural poverty in India is falling, since agricultural productivity has not increased much. It may partly be due to remittances from migration, but we have little data on migration (the Census measures some movement but does not properly account for circular movement). If the explanation is indeed migration and remittances, then the reduction may be unsustainable. If it is productivity growth, then it is likely to continue.

Inequality matters independently of poverty because of its effect on political stability. In India, those at the very top of the income distribution are seeing their incomes rising much faster than everyone else and this is a matter for concern.

On the MDGs, the international statistical and research community would be wiser to focus on a subset of the Goals, so that we can measure and monitor these properly – rather than trying to measure them all and failing to measure any of them accurately.

T. N. Srinivasan disagreed with Sachs on the fundamental importance of geography as a determinant of growth, pointing out that geography does not change while growth rates do. Even dire health indicators do not condemn a country to stagnation. In the middle decades of the twentieth century, countries like India had abysmal life expectancies, yet over time they have managed to develop. Conversely, China had good human capital by the mid-1960s yet did not start growing strongly until the policy reforms at the end of the 1970s. Exporting primary commodities also cannot explain poor growth performance.

Stiglitz wondered about the importance of having very precise estimates of poverty, because a view on which policies will help to reduce or increase poverty does not require such accurate figures. For example, NAFTA is going to bring down the price that poor farmers in Mexico receive for their corn, and will thus contribute to increasing poverty. We know this without having precise estimates of poverty. And as discussed previously, the change of the rural-urban terms of trade in China has an easily predictable effect on poverty without its being measured accurately.

Moreover, focusing on the number or proportion of people below the \$1-a-day poverty line may lead us away from other poverty measures, such as the poverty gap. We might find that numbers below the poverty line fall while those that remain poor fall into deeper poverty.

Pogge commented that several aspects of the MDG for poverty reduction remain unclear, including: whether it is income or consumption poverty; whether the goal is to halve the proportion or the number of people in poverty; what the agreed method is for calculating poverty numbers (for instance, the use of surveys directly or scaling up survey distributions to HFCE in the NA); whether it is aggregative or country-by-country poverty; and the 1990 baseline relative to which success will be assessed.

Deaton commented that the problem is not simply that we cannot get accurate \$1-a-day poverty figures, but that we have little (or poor quality) data on living standards of poor people generally. On the question of aggregative versus country-by-country poverty counts, he argued that it is indeed appropriate to weight the importance of poverty reduction in countries according to their population. To give Benin as much weight as India in our concern for poverty reduction is to say that poor people in Benin have a greater claim on resources than equally poor people in India. The unit of concern must remain the individual.

Sachs responded by stating that politically it is important to think on a country-by-country basis.

Session 6: Monitoring Global Poverty: The World Bank Approach

Martin Ravallion presented the World Bank's method for measuring global poverty, opening with the observation that measuring poverty is indeed very difficult. The audience for global poverty figures is the international community: these figures are not of concern to country-level analysts who do the work of trying to reduce national poverty. Country-level analyses use poverty lines appropriate to what "poverty" means in each country. Increasingly country work also looks at multiple poverty lines and uses dominance tests to get a deeper picture than that achieved by using one particular poverty line (e.g., the equivalent of PPP\$1-a-day in national currency).

Two methods that have been tried in setting a poverty line are costing the bundle of goods required to satisfy basic needs, and finding the lowest level of income in a distribution at which a sufficient number of calories are consumed.

When making international comparisons, a problem with the first is that one cannot use the same bundle of goods for different countries since relative prices (which determine consumption bundles), and what counts as staple goods, vary widely. The problem with the second approach is that the minimum level of income at which people consume the required number of calories varies even when the prices faced are the same. In Indonesia, for example, it was found that people in rural areas consumed the required number of calories at lower income levels than people in urban areas despite facing the same prices.

Paul Segal suggested that this may be because people in urban areas require more, or different and more expensive, goods in order to achieve a given quality of life. For instance, getting work or social respect in the city may require more expensive clothes than in rural areas, even if clothes cost the same in rural and urban areas. This would imply that achieving the same level of well-being costs more in urban areas, making it appropriate that the poverty line be higher.

Ravallion replied that this is probably not the reason for the difference: rather, it is simply that people in cities choose to eat more expensive calories, even if that entails eating 'too few' calories.

Given such problems with alternative methods for choosing a poverty line, the Bank in 1990 based its international poverty line on the national poverty lines of the poorest countries translated into PPP\$. Many of these were around the PPP\$1-a-day mark so they chose this as the original international line. For the later estimates, they took the

median poverty line of the ten poorest countries in 1993 PPP\$, getting a figure of \$1.08-a-day. It is important to note that no calculation of inflation rates went into the switch from \$1-a-day at 1985 prices to \$1.08-a-day at 1993 prices. If one used US inflation then the poverty line for 1993 would be considerably higher. The consistency between the two lines lies in the fact that they both reflect the domestic poverty lines of the poorest countries in the world in the year in question.

The Bank uses more than 400 surveys of acceptable quality, covering about 100 countries. Conditions of acceptability include that the surveys are nationally representative, and impute the value of own consumption. All but 20 of the 100 countries have more than one survey. All of the Bank's poverty estimates are based on primary data: they do not rely on estimates made by anyone else. And for their latest revision, the Bank used its own PPP exchange rates because the PWT PPP rates were not available in time. For estimation of poverty in a country between (or just outside) survey years, the Bank applies NA consumption growth rates to the survey expenditures of households.

Despite the Bank's best efforts, problems with surveys remain. Some are income surveys while others are consumption surveys; the surveys use different recall periods and valuation methods; and so on. The Bank is making efforts to improve data quality, with mixed success. Sub-Saharan Africa, the Middle East, and North Africa have much greater coverage now than in 1990. In previous estimates they adjusted income surveys in order to make them consistent with expenditure surveys. But incomes are typically found to be both higher and more unequally distributed than consumption, and the Bank have now found that these two factors approximately cancel each other out so they no longer make any adjustment. This change in method explains why, for instance, the new estimates of poverty levels in Latin America are somewhat different from the old estimates.

The Bank finds that the rate of global poverty reduction has been 0.7% a year since 1990, which is on track to halving the 1990 rate by 2015. Improvements in between-country income inequality contributed only about 1% to overall poverty reduction between 1987 and 1998. The reduction in global poverty is largely driven by China's rapid growth: excluding China there is no change in the proportion of people in the world living in poverty.

In countries with rising mean incomes, as measured by surveys, the rate of poverty reduction depends to a large extent on inequality: those with falling inequality saw much higher rates of poverty reduction than those with rising inequality. And the higher is initial inequality, the lower is the rate of poverty reduction at a given rate of growth.

In spring 2004 the Bank will produce three-yearly estimates of poverty from 1984 to 1999. They will provide researchers with access, through POVCALNET, to primary distributions from the Bank's Global Poverty Monitoring Data website. It will allow the use of alternative PPP rates, country groupings, etc. At the moment, about 20% of

the data are legally confidential and this is a constraint on making them publicly available.

The upcoming 2004 ICP is very promising. The new PPP exchange rates will be better than the current ones. Sub-national PPP rates may also be available. PPPs aimed specifically at consumption of the poor ('poverty PPPs') are clearly desirable, but are possible to produce for only about 30 countries.

Sanjay Reddy presented a critique of the World Bank method, based on his work with Pogge. In response to Ravallion's claim that global poverty figures have no bearing on policy at the national level, and in particular cannot be used to justify such policies, he cited James Wolfensohn as saying that the fall in global poverty indicates that globally policies have been good. He urged that economists in the World Bank who argue that this inference cannot be made should inform their President of this fact. He also observed that some development agencies, e.g. the UK Department for International Development (DfID), use World Bank poverty figures to allocate funds among countries. This implies that difficulties in estimation procedures really do matter to development policies and funding.

Reddy specified three minimal requirements for any poverty line: it should have an appropriate meaning; it should be the same everywhere (in some appropriately abstract sense); and it should be the same over time. The current World Bank approach fails to satisfy all three criteria.

First, it lacks meaning. It does not correspond reliably to domestic poverty lines in poor countries and has no direct interpretation in terms of the needs of human beings. It therefore does not reflect what people, including policy makers, think of as poverty, i.e., serious deprivation.

Second, it lacks consistency across countries. \$1-a-day at PPP exchange rates in fact yields different purchasing power for poor people in different countries. In developing countries typically, the domestic price of food is considerably higher than the overall PPP price level, the amount varying across countries. In India, for instance, food prices are 33% higher than the 'general consumption' PPP. This implies that poor people face higher prices than suggested by the PPP exchange rate. The ICP contains the data required to make these food price calculations, so we could better estimate the level of purchasing power needed to avoid poverty in different countries.

Third, the meaning of the international poverty line within countries varies over time. Measuring poverty in year t at 1985 \$PPP, and inflating/deflating by the domestic CPI, gives a different estimate from using 1993 \$PPP and inflation/deflation by the domestic CPI. In particular, changes in the PPP poverty line over time do not reflect changes in prices faced by the poor in a given country.

The proposed alternative is to specify poverty in terms of very basic and uncontroversial capabilities such as adequate nourishment. In each country, a set of commodities possessing the required characteristics could be chosen and then priced. It would require collection of data on their prices within each country, but this is feasible for the next round of the ICP.

On the issue of ‘poverty PPPs’, they would be a great improvement over the current method, but they entail a certain circularity because one needs to identify the poor before one can determine which goods they consume. Once the goods consumed by the poor have been priced, then the proposed alternative method that is conceptually preferable can be implemented.

Srinivasan argued that the Reddy-Pogge proposal is unworkable. In addition to the problems highlighted by Ravallion, it is not possible to specify a universally applicable calorie requirement. Different people in different countries and circumstances have different calorie needs. When poverty lines within countries have been specified by reference to food energy intake, or a consumption basket, then these lines have been updated over time according to domestic CPIs, and not by recalculating the (food- or calorie-based) line by the original method as it should be.

Turning to the use of household surveys he pointed out that some are designed to be consistent but not unbiased, and that what we want for current purposes are surveys that are designed to be unbiased. Surveys that are adequate for one purpose may be inappropriate for another. It is also often unclear in the published literature which of four possible types of distribution is being used: households by household expenditure, households by per capita household expenditure, persons by household expenditure, or persons by per capita household expenditure. There is no obvious logic to say which is more appropriate, although researchers should certainly specify the distribution that they are using. Ravallion responded that the Bank always uses fractiles of individuals ranked by per capita household expenditure.

Srinivasan also stated that he could see no policy at the global level that is being proposed or defended by reference to the global poverty numbers, despite the rhetoric of certain individuals.

Deaton argued that updating poverty lines using domestic price indices is not much worse than updating GDP by GDP deflators. What causes real confusion is when estimates of PPP exchange rates are updated with similar frequency to poverty estimates. This can give the misleading impression that each new poverty estimate corresponds to a new PPP estimate, which is not (and should not be) the case. It would therefore be preferable, as argued in his paper (“How to monitor poverty for the Millennium Development Goals”), to update PPPs only once every ten years or so, while updating poverty estimates with new survey data much more often. Reddy and Pogge’s claim that the Bank encourages switching between PPP estimates over

time is incorrect because the Bank explicitly states in its publications that different PPPs are not comparable. But frequent PPP updates do foster confusion and the Bank could be clearer.

He agreed with Reddy and Pogge's argument that PPPs could be improved, and in particular that we could base them on food or general consumption of the poor.

Session 7: Monitoring Global Poverty: Alternative Perspectives

Surjit Bhalla presented on growth and its effects on poverty. His estimates differ substantially from those of the World Bank only after 1988, indicating much faster poverty reduction. The reason for the divergence is that he uses NA mean consumption (scaling up survey distributions) while the Bank's estimates are based on survey mean consumption directly.

Bhalla uses an adjusted HFCE from the NA rather than mean consumption from surveys. He notes that since inequality is underestimated in surveys (due to underreporting and non-response by the rich), simply scaling up within-country distributions from surveys to NA HFCE would overestimate consumption of the poor and hence underestimate poverty. To compensate for this downward bias he adjusts NA HFCE downwards by 15% and scales up survey consumption to this figure.

Focusing on India, he finds that NA estimates of household consumption of each commodity are systematically higher than NSS estimates. Consumption of food, for instance, has gone up substantially according to the NA. Since there is a limit to how much food the rich can consume, some of it must have gone to the poor. Such increased consumption is inconsistent with the slow rate of poverty reduction found through reliance on NSS alone.

Whether one uses NA or NSS, one needs to be consistent when estimating the effect of growth on poverty. One cannot use NSS estimates for poverty and NA estimates for growth and thence argue that consumption by the poor has not kept pace with aggregate growth, and that growth has not been pro-poor.

To find whether or not growth was pro-poor, Bhalla takes the proportion of people who were poor in each country at the beginning of the period (e.g., 28% in India in 1987) and compares the income growth of that bottom quantile over time with aggregate income growth. Combining such groups across countries, he finds that their incomes grew at 1.8 times aggregate income growth and concludes that globally growth was indeed pro-poor.

Deaton repeated the argument that NA overstates consumption, e.g., of edible oil, sugar and other foods, and that this overstatement is likely to have increased over

time. This implies that the increase in food consumption reported in the NA may be an illusion. Bhalla replied that many different types of commodities, and not just food, show substantial differences between NA and NSS, and they cannot all be subject to double-counting in the way Deaton described food to be.

Tendulkar cited two detailed studies, published last year, comparing NA and NSS and concluding that NSS estimates are more reliable. Bhalla responded that he had seen the studies and disagreed with their conclusions.

Howard Nye presented on problems common to recent estimates of global poverty. He argued that there are four such problems. First, inappropriate methods and concepts are used for assessing individual advantage or disadvantage. Second, the criteria used for comparing levels of poverty across space and over time are inappropriate. Third, inappropriate generalizations or interpolations are made from limited data – for instance, the kernel estimation method used by Sala-i-Martin to impute distributions from quintile shares. Fourth, estimates that use NA for mean income and household surveys for income inequality are inconsistent.

It is also harder than it should be to find documentation on the World Bank's methodology. It took Nye, Pogge and Reddy a lot of time to find out basic facts about World Bank procedures from their website and staff. Information should be much more easily available. Deaton seconded this comment.

Ravallion presented again on the methodology of poverty measurement. He noted that the difference between NA and survey estimates of consumption vary widely across regions. For instance, in Eastern Europe and Central Asia there is no correlation between the two. In sub-Saharan Africa, on the other hand, the two are fairly close and are not diverging. The World Bank approach is to go with surveys. They are more direct, unlike the residual estimation method used in NA to obtain HFCE, and they are conceptually more appropriate. For instance, the consumption of financial services and consumption by non-profit organizations, which are included in NA HFCE, should not be a part of household consumption. And while surveys do underestimate mean consumption they also underestimate inequality, the two tending to act in opposite directions on poverty estimates.

Moreover, as described in Mistiaen and Ravallion's paper ("Survey compliance and the distribution of income"), one can estimate the degree of bias due to disproportionate under-response by the rich. By comparing mean consumption with response rate across regions of a country one can estimate the relationship between consumption level and the likelihood of responding. With this information we can correct for differential response rate by consumption level, and estimate the bias in both mean consumption and inequality. Using this model on US income distribution data from surveys, they find that inequality is substantially underestimated – by 4 to 5 Gini points – but that the downward bias in the estimated mean entails that measured poverty is not substantially biased. One clear conclusion is that underreporting due to

survey non-response is not distribution-neutral, so that scaling up survey distributions to the correct mean (income or consumption) will measure poverty incorrectly.

Deaton presented again on poverty measurement, commenting that decades of debate have shown that measuring poverty on the basis of a nutrition-based consumption basket is not workable. This means we have to use PPPs one way or another. However we use them, there will be index number problems which are well known to economists. The question is how to minimize the impact of these problems, not how to eliminate them. Developing a 'poverty PPP' exchange rate that is based on goods consumed by the poor would be a good step forward.

He has argued in his paper that PPPs should be updated only once every ten years or so in order to reduce problems of comparability over time. The upcoming ICP 2004 sounds very promising: it should lead to much better PPPs and should be supported.

It should also be a priority to improve surveys globally and to find out more about non-compliance and its effects.

To recap, the problems with NA estimates of HFCE are that in some countries they double-count some commodities; they include financial services, which increase as a proportion of GDP as countries get richer; they include consumption by non-profit organizations; and they are indirect residual estimates and therefore they compound errors elsewhere in the NA. In India, when we take out financial services the difference in annual growth rates over the last 20 years between NA HFCE and NSS consumption is only 0.9% per year, which is small enough to be due to increased double-counting, and is therefore consistent with NSS estimates being right and NA estimates being wrong. Nonetheless, we need to understand the problems with NSS and other surveys better than we do now.

In conclusion, Milanovic commented that all researchers should make greater efforts to document their calculations and methodology comprehensively. Stiglitz noted that poverty figures and their implications are clearly not being well understood by the press and by policy makers. Anand concluded that explaining the difference and the growing divergence between NA and surveys in many countries is an important research imperative. He also commented that careful thought and substantive consultation must go into the design of the 2004 ICP and the estimation of PPPs, including poverty PPPs.