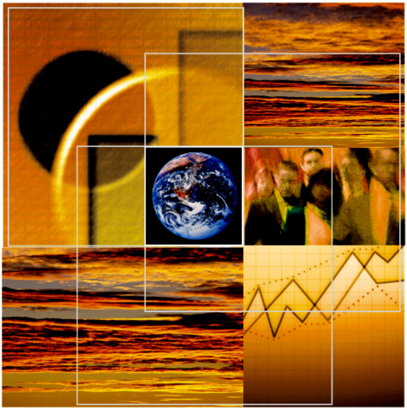
Employment Strategy Papers

International
Labour
Office
Geneva

Global poverty estimates and the millennium goals: Towards a unified framework

By Massoud Karshenas



Employment
Analysis
Unit

2004 / 5

Employment
Strategy
Department

Employment Strategy Papers

Global poverty estimates and the millennium goals: Towards a unified framework

By Massoud Karshenas

Department of Economics, SOAS, University of London and
Institute of Social Studies, The Hague

Employment Analysis Unit
Employment Strategy Department

2004/5

Acknowledgements

I am grateful to Angus Deaton, Ben Fine, Charles Gore, Nomaan Majid, Farhad Mehran and Graham Pyatt for comments on an earlier draft. I am responsible for all the remaining errors and omissions.

Preface

This paper discusses the compatibility of different global poverty estimates under a unified framework, and examines the compatibility of various international poverty lines used in the literature under different purchasing power parity exchange rate estimates.

The paper also addresses the issue of compatibility of survey means and national accounts data. It is argued that the non-compliance hypothesis that is usually invoked to assume away the incongruence between the two types of data is not supported by the empirical evidence.

The paper puts forward an alternative approach to deal with the inconsistency between survey and national accounts data, which consists of calibrating the survey means using the national accounts data as external calibrating information. Estimates of regional and global poverty are made on the basis of this new approach, and the results are compared to the World Bank estimates.

Duncan Campbell
Director a.i.
Employment Strategy Department

Contents

Figures and tables

1. Introduction

Counting the poor at the international level has become the subject of a growing literature with a growing number of alternative poverty estimates. One of the main sources of divergence between the different proposed estimates is the discrepancy between the national accounts and survey mean consumption or income and the way this has been reconciled by various authors. Another important source of discrepancy has been the differences in purchasing power parity exchange rates used in different estimates of international poverty. This latter source of discrepancy has received less attention in the ongoing controversy on international poverty comparisons, but as we shall show in this paper it can be of equal if not more importance in creating incompatibility between the various estimates. In fact, the differences in purchasing power parity exchange rates are so important that they render no only the poverty estimates but also the international poverty lines used in different studies incompatible.

In this paper we compare the different existing estimates within a unified framework which utilizes the information in both the national accounts and survey means. The method consists of calibrating survey means using national accounts averages as the external calibrating variables. Before attempting the calibration of the survey data, we address the issue of the lack of compatibility between the different purchasing power parity exchange rates. We also use the calibrated survey data to estimate income poverty at the country, regional, and global levels and compare the results with other available estimates and with the millennium goals on income poverty.

In the next section we start by a discussion of the various sources of data on the purchasing power parity exchange rates used in different poverty estimates. We compare the different purchasing power parity exchange rates and discuss their implications for the international poverty lines and poverty comparisons in general. In sections 3 and 4 we put forward the idea of calibrating survey means as a method of reconciling the national accounts and survey based estimates. The calibrated survey means are also used to discuss the bias in the available survey based and national accounts based poverty estimates. In section 5 we address some of the possible problems arising from the correlation between survey mean error and the error in the scale or shape of the distribution. The World Bank estimates of international poverty are based on the idea that the errors in survey means are by and large non-compliance errors, which are neutralized by errors in income distribution (see, e.g., Ravallion 2003). We test this hypothesis and show that the empirical evidence does not support the hypothesis of cross-country survey mean deviations being due to non-compliance of the rich. In section 6 we discuss the poverty estimates based on the calibrated survey means, and compare the aggregate regional and global poverty estimates with those of the World Bank.

2. Purchasing power parity exchange rates and poverty lines

The $1 a day poverty line originally proposed by the World Bank has become the standard poverty line for international poverty comparisons. The millennium goals on income poverty are also monitored in relation to the World Bank's standard $1 a day poverty line in 1985 PPP exchange rates. However, as the estimates of the PPP exchange rates have been revised over the years, the definition of the $1 a day poverty line has also changed. The original $1 a day poverty line proposed by the World Bank was derived on the basis of official poverty lines in a number of low income countries, measured in 1985 PPP exchange rates based on the Penn World Tables version 5.6 (hereafter PWT5.6). A detailed discussion of the derivation methods and the data used in calculating the $1 a day poverty line is provided in Ravallion *et al* (1991). This has become the standard international poverty line used by various researchers.

In later years new estimates of PPP consumption rates were produced by the World Bank for the year 1993, and a new poverty line of $1.08 a day in 1993 prices was introduced. The global poverty estimates by Chen and Ravallion (2000), are based on the new PPP exchange rates and the $1.08 a day poverty line. Since the terminal year of the data series in PWT5.6 dataset is 1992, it has been difficult to evaluate the relationship of the World Bank's new $1.08 a day poverty line in 1993 prices with the standard $1 a day international poverty line based on PWT5.6's 1985 PPP rates. However, with the availability of the new Penn World Tables version 6.1 (PWT6.1), which extends the PWT data series to the year 2000, it is now possible to shed some light on the relationship between the new and the old international poverty lines.

Lack of attention to the fundamental differences between these two poverty lines has led to some confusion in the literature. For example, Bhalla (2002 and 2003) goes to a great length to show that the international price inflation factor between 1985 and 1993 is 1.3 and not the 1.08 apparently assumed by the World Bank. He therefore opts for the $1.3 a day poverty line in 1993 prices rather than the World Bank's estimate of $1.08 a day. As we shall see below, however, the difference between the two poverty lines is not solely due to the inflation factor. The differences largely result from the fundamental changes in PPP exchange rates as well as the change in the method of calculation of the poverty lines. Similarly, Sala-i-Martin (2002) makes a valiant effort to show that his national accounts based aggregate global poverty estimates are not all that different from the World Bank estimates, by introducing various modifications to his first round estimates. However, as we shall see below, his $1 a day poverty line in 1985 prices is not consistent with the poverty line used in the World Bank estimates. If at all comparable with the Chen and Ravallion (2000) estimates, because of the differences in poverty lines alone Sala-i-Martin should be getting much higher poverty estimates than the former, and not lower estimates as the case is. A discussion of the international poverty line estimates under different PPP exchange rates will help shed light on these issues.

Currently there are three sources of PPP exchange rate data under which global poverty measures have been estimated. First is the PWT5.6 on the basis of which the original $1 a day international poverty line was estimated by Ravallion *et al* (1991). Second is the World Bank 1993 consumption PPPs which formed the basis for the $1.08 a day poverty line estimated in Chen and Ravalllion (2000). Third is the PWT6.1 which extends the estimates of the earlier series of Penn World Tables and can act as a link between the former two series.[1] The histogram of the ratio of the World Bank to PWT6.1 consumption PPPs for the sample of developing countries included in global poverty measurement is

1 The World Bank PPP data and its household survey based deciles distribution data are available on the World Bank Web Site http://www.worldbank.org/reseearch/povmonitor/index.htm. Penn World Tables 5.6 and 6.1 versions are accessible on http://pwt.econ.upenn.edu/aboutpwt.html.

shown in *Figure 1*.[2] As can be seen, there are large differences between the two series. Close to fifty per cent of the World Bank estimates diverge by at least 20 per cent on either side from the PWT6.1 PPP estimates in 1993. In more than 15 per cent of the countries the World Bank estimates are higher than the PWT6.1 figures by 40 per cent or more.

Figure 1. Histogram of the ratio of the World Bank to PWT6.1 consumption PPP exchange rates, 1993

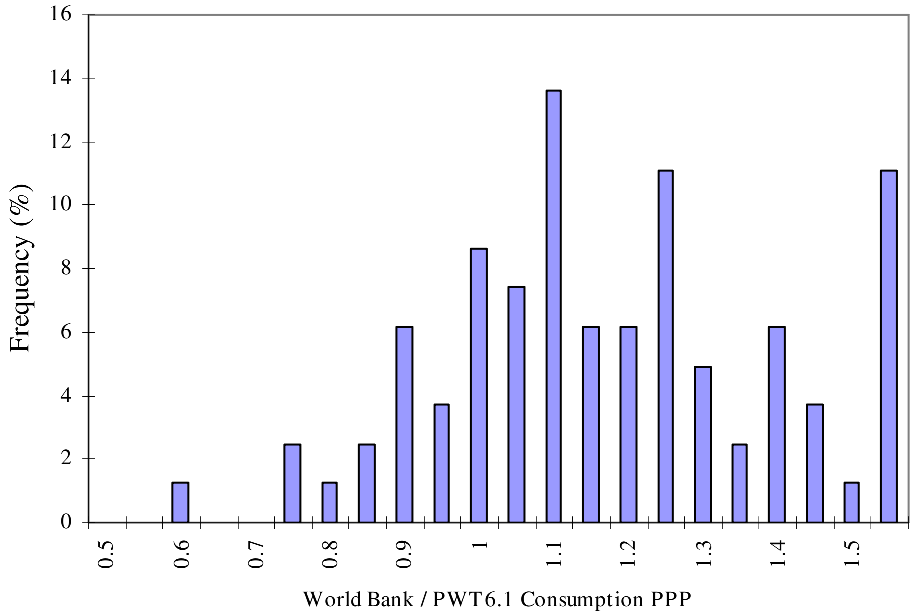
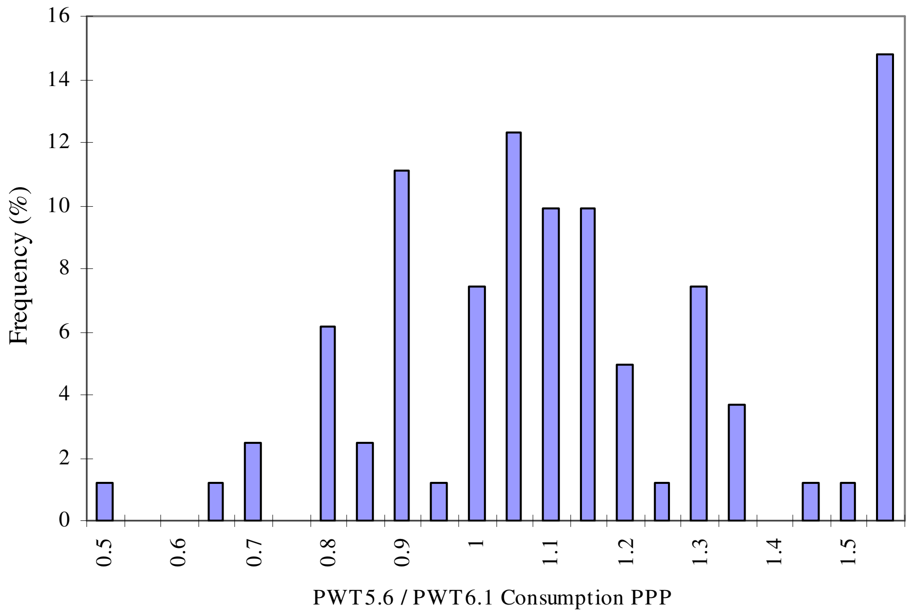


Figure 2. Histogram of the ratio of PWT5.6 to PWT6.1 consumption PPP exchange rates, 1985



2 The countries included in the histogram are the same as in Chen and Ravallion (2000) and the countries for which distribution data is provided on the World Bank poverty monitoring web site, excluding the Eastern European and Central Asian countries.

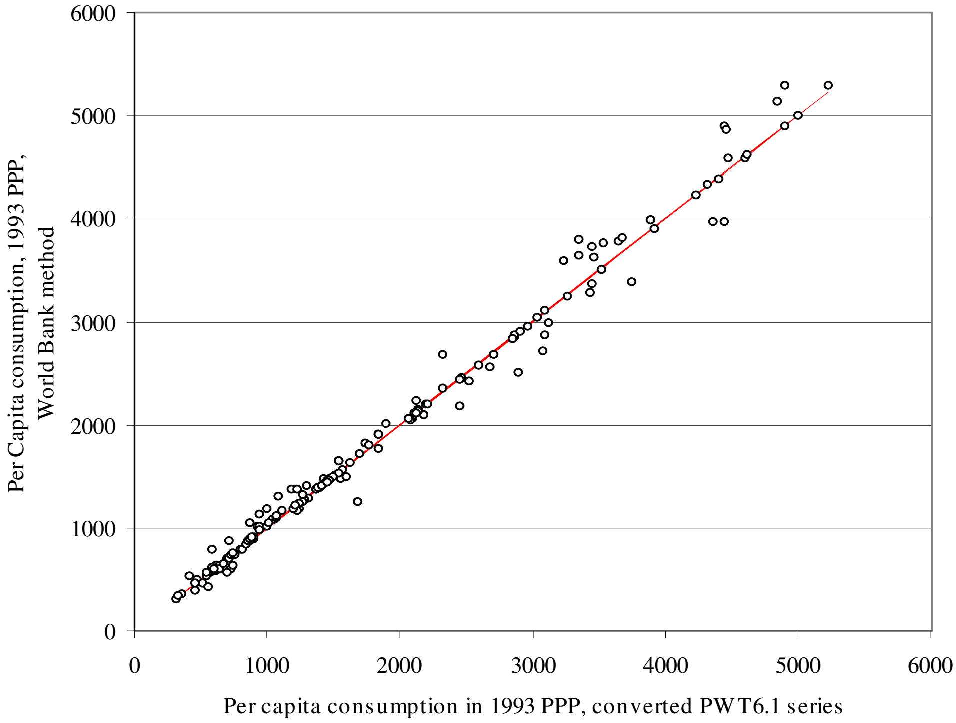
Similar discrepancies exist between the PWT5.6 estimates of consumption PPP exchange rates and its updated version PWT6.1, as new data have become available over time and the number of benchmark countries has increased from 62 to 115.[3] The histogram of the ratio of the PWT5.6 to PWT6.1 series on consumption PPP for the year 1985 is shown in *Figure 2*, which indicates relatively large and wide spread divergence between the two estimates. Since the PWT5.6 series terminate in 1992, we shall use the PWT6.1 version to link the former PPP estimates with the World Bank's 1993 estimates. We thus have three distinct series of consumption PPP estimates, which can give rise to different international poverty estimates. Of course the preferred choice of PPP exchange rates for international poverty comparisons has to be the PWT6.1 estimates for a number of reasons. Firstly, the PWT6.1 data contains all the information on relative prices for benchmark countries used in World Bank's 1993 data. Secondly, thought the World Bank publishes the 1993 PPP data on its poverty monitoring web site, there is no documentation attached to the data as to the methods of estimation used. A brief discussion of the methodological differences between the two is provided in Aten and Heston (2003), which indicates significant methodological differences. However, given that the survey data is provided by the World Bank in its own 1993 PPP rates, and that the $1 a day poverty line was originally defined in terms of the PWT5.6 1985 PPP rates, in order to proceed we need to link these three data series.

Before proceeding to the discussion of poverty lines under different PPP estimates, we should address the issue of inter-temporal relationship between the World Bank and PWT data. This relationship is used in linking the World Bank and the PWT5.6 PPP estimates which will be used in updating the international poverty lines as well as in comparing the PWT6.1 national accounts averages and survey means in the next section. The World Bank estimates of consumption PPP exchange rates refer to the single year 1993. To convert the survey means in various years into the 1993 PPP values, the World Bank uses the domestic currency consumer price indices first to convert the survey means into 1993 domestic currency values, and then converts them into internationally comparable values by using its 1993 PPP rates (see, Chen and Ravallion, 2000). To compare the inter-temporal properties of the World Bank estimation method with the Penn World Tables series, we follow the same procedure as the World Bank to convert first per capita household consumption expenditure in domestic currency in various survey years into 1993 domestic values and then into 1993 PPP values, using the domestic consumer price deflators and the World Bank PPP rates respectively in the two stages. Current household consumption expenditure and the consumer price deflators are based on the World Bank WDI 2001 databank. This is then compared with per capita household consumption expenditure from PWT6.1 for each country and each survey year in 1993 PPP values, which are converted to the World Bank PPP values using the conversion factors depicted in *Figure 1*. The scatter plot of the two series is shown in *Figure 3*. As can be seen, excluding one or two outlying observations, the scatter plot closely follows the 45 degree line, with a particularly close fit for per capita consumption levels below $2000 per annum which is the consumption range where the $1 a day poverty is most relevant. Provided that the World Bank uses the same consumer price deflator for updating the 1985 individual country poverty lines as the consumption price deflators in WDI national accounts, this inter-temporal consistency between the two series will allow us to calibrate the $1 a day poverty line to different datasets and different years.

3 See, Heston *et al* (2002), Appendix61 for further discussion.

The wide variations in the PPP exchange rates in different datasets, particularly in the case of low-income countries, as more and better data become available, makes it essential to have a consistent and transparent method of updating international poverty lines. How can the standard $1 a day poverty line be translated in a consistent manner in moving from one set of PPP exchange rate estimates to another in the three data sets available? Of course it would be impossible to devise poverty lines which give rise to exactly the same global poverty measure under different sets of PPP estimates in individual countries and regions and at the global level. A consistent and transparent set of rules for varying the poverty line as we move from one data set to another is however essential.

Figure 3 Per capita consumption expenditure;
The World Bank method vs the PWT6.1 series



The $1 a day international poverty line was originally supported by the observation that a number of low income countries seemed to have official poverty lines close to $31 a month in 1985 PPP exchange rates (see, Ravallion *et al,* 1991). This rule of thumb, however, breaks down in moving from PWT5.6 PPP estimates to other datasets, as the same cluster of countries now show divergent poverty lines. For example, moving from PWT5.6 to PWT6.1, the same official poverty line in Bangladesh moves to much lower than $30 levels while in Indonesia the opposite is the case, and so on. In updating the poverty line to the 1993 base, the World Bank adopts a more formal approach in choosing the $1 a day poverty line. As Chen and Ravallion (2000) point out, since the relationship between the poverty line and per capita consumption in the lowest income countries seems to be flat, they choose the median of the ten lowest poverty lines as the international poverty line, which is the $108 a day poverty line used by the World Bank. There are of course numerous other ways of defining the international poverty line on the basis of Ravallion and Chen's data on official poverty lines, e.g., various forms of curve fitting and other ways of averaging the tail of the poverty line / consumption curve. But the Chen and

Ravallion method is as good as any other method, and given the lack of sensitivity of the median to outliers it may be preferred to other averaging methods. In any event, since we are using the World Bank estimates as benchmarks, trying to reduce the differences resulting from the choice of the poverty line, we should follow the same procedure here.

Table 1. International poverty line under different PPP exchange rates

	PWT5.6 1985 PPP	PWT61 1985 PPP	World Bank 1993 PPP
1. Countries Ranked by Poverty Line in 1985 PWT5.6 PPPs			
$ per month	25.19	29.51	31.09
$ per day	0.83	0.97	1.02
2. Countries Ranked by Poverty Line in 1985 PWT6.1 PPPs			
$ per month	30.80	28.72	31.62
$ per day	1.01	0.94	1.04
3. Countries Ranked by Poverty Line in 1993 World Bank PPPs			
$ per month	25.19	28.72	31.09
$ per day	0.83	0.94	1.02
4. Countries Ranked by Poverty Lines separately for each case			
$ per month	25.19	28.79	31.09
$ per day	0.83	0.94	1.02
5. Countries Ranked by Poverty Line in 1993 World Bank PPPs (adjusted)			
$ per month	26.53	30.25	32.74
$ per day	**0.87**	**0.99**	**1.08**

Notes: International poverty line is measured as the median of the ten lowest poverty lines for the relevant year and PPP estimates.
Row 4 is the same as row 3, with estimates adjusted upwards to match the $1.08 poverty line defined by the World Bank.
Sources: As discussed in the text.

Even with the apparently simple median rule one is unlikely to get unique results, as the ranking of countries can change in different datasets. *Table 1* shows the median of the ten lowest poverty lines for each dataset according to different country rankings. The individual country data are based on 1985 PWT5.6 PPP values used in Ravallion *et al*, (1991). The first row of data report the median poverty line for the three data sets when the countries are ranked according to the 1985 values based on PWT5.6, the second row is based on ranks of PWT6.1, and the third row is based on the ranking according to World Bank 1993 values. Row 4 reports the median when poverty lines according to each dataset are ranked separately. The results for the PWT5.6 and the World Bank 1993 rankings in rows 1 and 3 are similar. What stands out in these two rows is that the new median rule adopted by the World Bank substantially reduces the value of the international poverty line measured in PWT5.6 PPPs. Furthermore, the 83 cents a day poverty line in 1985 PWT5.6 values, when adjusted for inflation on the basis of the international price index from the PWT data, is translated into exactly the $1.08 a day poverty line in 1993 prices. The median of the poverty lines in 1993 World Bank PPP shown in the last column of Table 1 is however consistently less than $1.08. This could be due to the differences between the price deflators used by the World Bank and the international deflators used here, or it can be the case that the World Bank's $1.08 figure is inflation adjusted of the median of PWT5.6 in 1985 prices. In row 5 of Table 1 the median poverty lines of row 4 are adjusted upwards such that the World Bank median attains the $1.08 reported by Chen and Ravallion (2000).

In all the alternative rankings in Table 1 the median poverty lines measured at 1985 PWT6.1 PPPs are very close to the $1 a day standard. We shall therefore adopt the $1 a day poverty line using he PWT6.1 data in this paper. This is not to say that the $1 a day poverty line is a more 'accurate' measure of international poverty line than other values that one can estimate on the basis of the PPP values in PWT6.1. However, it has the advantages of being consistent both with the World Bank's median rule, as well as with the popularly accepted $1 a day norm.

In considering poverty measures estimated on the basis of the $1 a day international poverty line, it should be kept in mind that only five countries have poverty lines which are below this norm. The rest of the countries all have poverty lines above the $1 a day in 1985 PPPs. The $1 a day poverty line therefore may be more appropriate for measuring extreme absolute poverty in the poorest countries. For the range of incomes of countries normally included in measuring global poverty trends, the $2 a day poverty line is arguably a more appropriate indicator of extreme poverty on a global basis. We shall examine global poverty trends for both poverty lines, focusing on the headcount measure of poverty.

3. National accounts and survey means: Substitutes or complements?

Two main datasets are used in estimating global poverty in this study. We use the household survey data provided by the World Bank for the mean and distribution of consumption expenditure / income, and the latest version of Penn World Tables (PWT6.1) for obtaining the calibrating national accounts variables. The World Bank data provides summary statistics in the form of the decile distribution of consumption or income and the survey mean based on household expenditure surveys, which excluding the Eastern European and Central Asian countries, constitutes some 64 countries and 156 observations.[4] The survey means, measured in 1993 PPP exchange rates, combined with the decile distribution data provided by the World Bank, and assuming $1.08 a day poverty line, reproduce the poverty measures published by the World Bank.[5] Contrary to what is sometimes asserted, therefore, the combination of the survey means and decile distribution data contains the full information used in the World Bank poverty estimates. Any discrepancies between our poverty estimates and those produced by the World Bank

4 According to Ravallion (2003), there are now 400 surveys representing 100 countries available, but unfortunately this data is not in public domain.

5 To be consistent with the World Bank estimates we use the POVCAL program for measuring poverty provided by the World Bank.

will be therefore mainly due to methodological differences and not the basic information regarding the distribution of income.

In order to discuss the relationship between the national accounts and survey means, we first convert the World Bank's survey means into domestic currency using World Bank's 1993 PPP exchange rates and reconvert them back to international values using PWT6.1 PPP exchange rates. Both the national accounts and survey means are then converted to 1985 PPP values using the international price indices of PWT6.1. This ensures that the divergence between the two series is not due to variations in the PPP conversion factors. Since the behaviour of survey means in consumption and income surveys are different we shall discuss them separately, starting with household consumption means.

The inconsistency of the household survey means and the national accounts averages has been widely discussed in the literature.[6] Pyatt (2003) notes that this has been a longstanding problem which has been neglected in World Bank poverty estimates. Deaton (2002, and 2003) discusses the possible sources of divergence between the two series, and concludes that despite the weaknesses of the household survey data, they should be used together with their accompanied decile distribution of consumption to measure global poverty, and not the national accounts per capita consumption. Deaton argues that household surveys are designed to measure individual welfare, but national accounts data on macroeconomic aggregates are not designed for this purpose. Another argument put forward by Ravallion (2003) is that since the survey underestimate of true consumption mean is likely to be largely due to under-reporting and non-response of the higher income groups, the existing biased survey mean would produce relatively more accurate poverty estimates. Under these circumstances, the use of national accounts mean consumption, even if it can be assumed to be accurate estimators of the true survey mean, would lead to an underestimation of poverty. These plausible arguments, to which we shall return below, but they do not go far enough to address the problems of variations of survey means across the countries. We may be able to do better if we do not discard the information contained in national accounts off hand.

Apart from non-compliance there are various other idiosyncratic phenomena arising from the differences in sample design, questionnaire design, recall periods, the nature and treatment of non-response, etc. that can affect the survey mean in a significant way. For example, according to Deaton (2002, p.14), 'when the Indian NSS experimentally changed the recall period for food from 30 to 7 days, the estimated poverty rate was cut by a half'.[7] To close our eyes to the relatively large sampling variations arising from these other sources, and hoping that the main source of the survey mean bias is the non-compliance of the rich which can be conveniently ignored, is not very wise. Large deviations of the survey means from their true values, would make poverty estimates across countries and over time non-comparable. And comparability is essential for global poverty measurement.

For these reasons the survey means need to be calibrated by external information in order to render them comparable across countries and over time. Rather than discarding the national accounts information, we can use them as calibrating external information to adjust the survey means. For this purpose, following Karshenas (2003), we divide the deviations of survey mean from national accounts averages into two components; the

6 See e.g., Pyatt (2003), Ravallion (2000, 2001, 2003), Deaton (2000, 2002, 2003), Bhalla 2000, Karshenas (2001, 2003), Sundaram and Tendulkar 2002, and Kulshreshtha and Kar 2002

7 See also Bhattacharya (2002).

systematic and the stochastic parts. The systematic part arises from the definitional differences between the two series. For example the national accounts consumption estimates contain elements such as imputed rent of dwellings and consumption by non-profit organizations, which are not part of consumption measured by the household expenditure surveys. On the other hand, there are other elements such as production and consumption through informal activities which can be only picked up at the micro-level by household expenditure surveys and may be missed in macro-level national accounts data. This can have a systematic effect on the deviations between the two means as the weight of such activities normally declines in the economy as the economy develops (see, Deaton 2002 and 2003).

The stochastic part on the other hand consists of all sampling errors in survey mean which are not due to definitional differences between the two means. In contrast to the non-compliance argument, which implicitly assumes a negative sampling error in survey mean, we assume the stochastic part to be a random variable with mean zero, which contains measurement errors that can act on the measured survey mean in a negative or positive way, with possibly non-constant variances. The reason for this is that in using the national accounts variables as calibrating factors we do not necessarily assume that these variables are a measure of the true mean consumption or income. They are basically regarded as external variables which are used to filter out the noise in survey means.

4. The devil is in the tail

We start with a non-parametric functional form for the calibrating equation of the form:

$$Y_i = f(X_i) + \omega_i \qquad i = 1, n$$

where Y is the survey mean and X is the national accounts mean and each of the n observations for a particular country and date in the data is treated as an independent observation. The function $f(X_i)$ represents the systematic relationship between the two means and ω_i captures the stochastic variations in the survey mean.[8] We estimate $f(X_i)$ non-parametrically, but a desirable property that we may wish to impose on the function is that its first derivative should be positive. Both per capita household consumption and per capita GDP were tried as possible calibrating variables and as the consumption variable had the closest fit to the data it was chosen for both the consumption and income surveys. *Figure 4* shows the scatter plot of the data for the consumption surveys in the World Bank dataset in 1985 PPPs, and the fitted nonparametric regression line, which is a locally weighted regression with bandwidth 0.6 (hereafter lowess regression line). The bandwidth of 0.6 was the smallest bandwidth where the regression line appeared smooth, and for the most part monotonically increasing. The last property, namely a positive slope does not appear to hold for the very poor countries under any bandwidth.

8 Measurement errors in national accounts means introduce added complications which need to be dealt with, by possibly some kind of instrumentation, in future research.

Figure 4. Survey vs national accounts consumption means

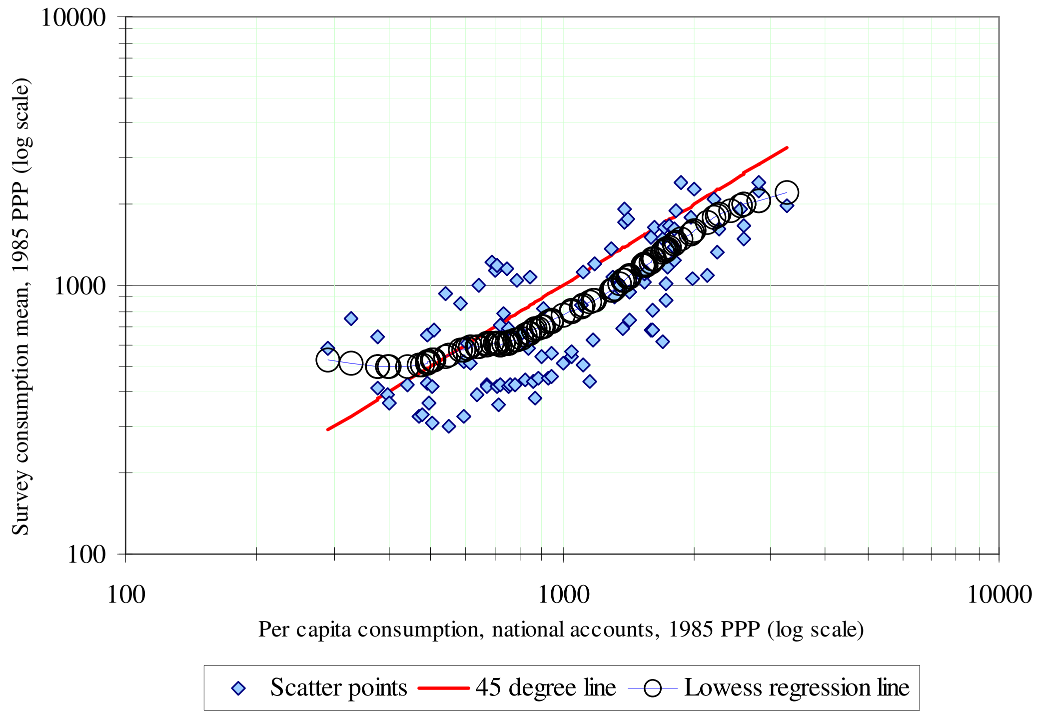


Figure 5. Survey vs national accounts consumption means in low income countries
(below $500 per captia consumption)

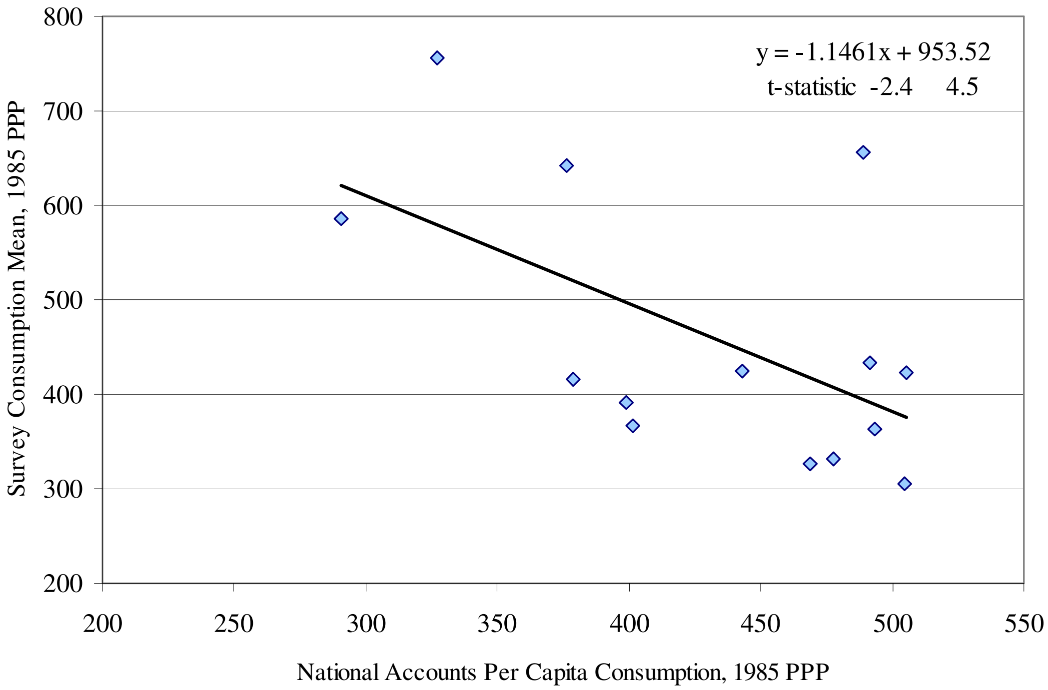


Figure 6. Survey vs national accounts consumption means with fitted regression lines

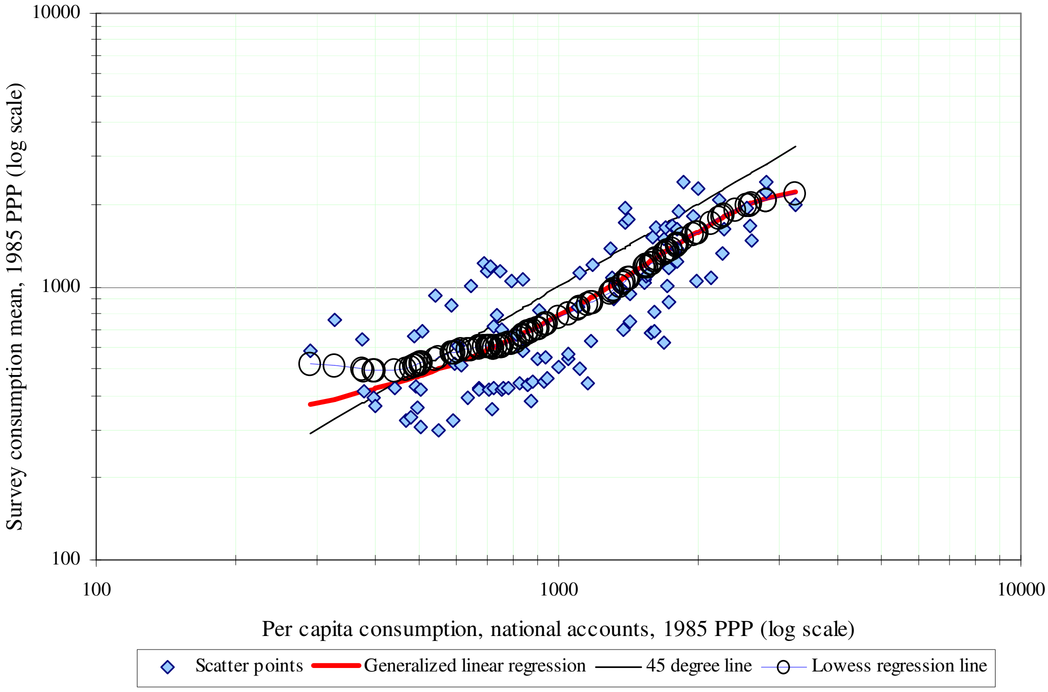
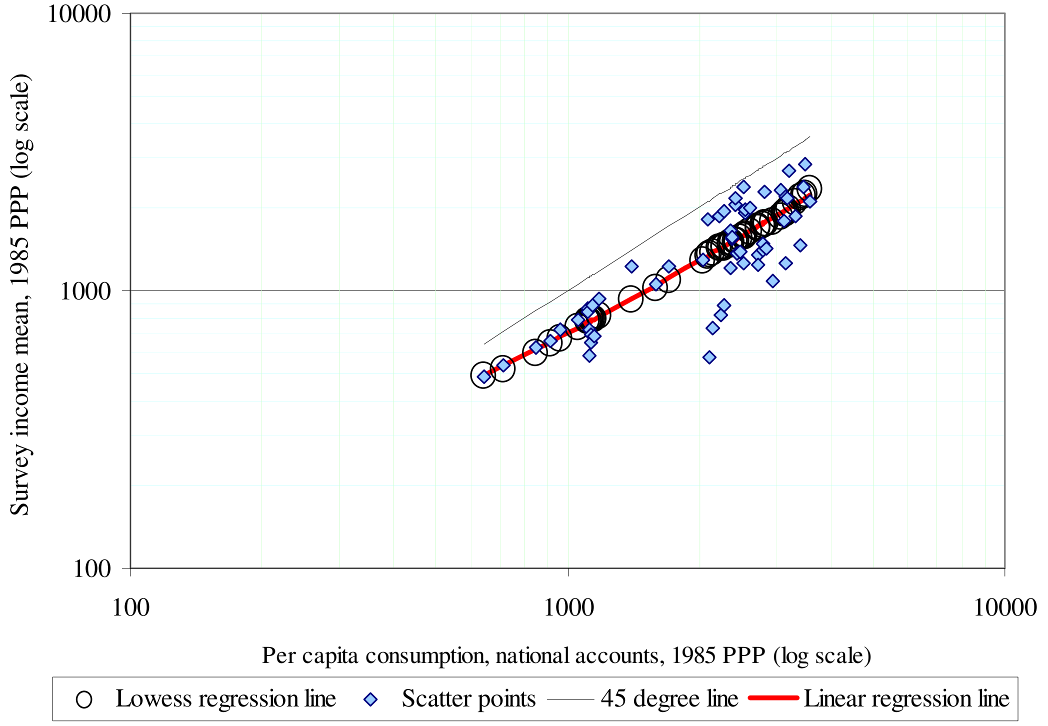


Figure 7. Survey mean income vs national accounts consumption mean



The lower tail of the lowess regression line is dominated by a few observations from the poorest countries where the data quality is not very high. In fact a blow-up of the relationship between the two means, shown in *Figure 5*, indicates that there is a statistically significant and negative relationship between survey mean and national accounts per capita consumption. There is strong reason to believe that the relationship between the survey mean and the national accounts consumption mean is likely to be flatter amongst low income countries than higher income groups. All the elements that create a wedge between the two means, such as imputed rents, consumption of non-profit organizations, etc., are likely to grow faster than household consumption at the early stages of development. The negative relationship indicated by the survey data for the countries below $500 annual per capita consumption, however, is not plausible and is likely to be the result of weaknesses in the data. For this reason, we searched for a parametric functional form, which is flexible enough to fit the lowess regression, but is less amenable to influence of observations in the lower tail of income groups. The following generalized linear model was estimated which closely fitted the lowess regression line as well:

$$\log(\hat{u}) = 5.559 + 0.0013\,X - 1.97\text{e-}07\,X^2$$

Standard Error	0.205	0.00023	6.09e-08

Where $\hat{u}$ is the estimated expected value of the survey mean, and X is the national accounts mean consumption. *Figure 6* shows the fitted generalized linear model together with the lowess regression. As can be seen, the generalized linear model fits the lowess curve almost perfectly as far as the per capita consumption level of $700 from above, but below that threshold is less influenced by the observations in the lower tail than the lowess regression. The generalized linear curve is flatter at lower income ranges, as required on a priori grounds, but follows the lowess curve at middle income ranges parallel to the 45 degree line, and has a lower slope again at per capita consumption ranges of over $2000. We use the predictions from the generalized linear curve as calibrated survey means.

The same procedure was followed for the income surveys. The scatter plot of the survey mean against per capita consumption from national accounts is shown in *Figure 7*, along with the lowess curve with bandwidth 0.6. The sample of countries with income surveys happens to fall in the middle income ranges, and perhaps for this reason the lowess curve has a linear shape. The lowess curve is very close to a simple linear regression of survey means on per capita consumption from national accounts. The linear regression line is shown by the solid line in Figure 7, with a slope coefficient of 0.585 (standard error 0.064) and intercept 114.44 (standard error 148.47). The fitted curves in Figure 7 have a lower slope than the 45-degree line. We use the predictions from the lowess regression as calibrated survey means in the case of income surveys for measuring individual country poverty, and the linear regression line for predicting values, where necessary, for the aggregate regional and global poverty estimates.

Figures 6 and 7 can be helpful in comparing the different global poverty measures put forward by different authors. Bhalla (2002) and Sala-i-Martin (2002) for example base their estimates on national accounts averages, which is tantamount to assuming the true survey means are located on the 45 degree line, or in other words it is as if they are calibrating the survey means such that they are located on the 45 degree line. As can be seen from Figure 6, apart from misrepresenting the non-linearity in the lower income range, this will also predict a higher survey mean for the higher income group of countries. Bhalla (2002) makes an adjustment to the national accounts means to make them on average the same as the survey means, which has the effect of locating his averages on a line parallel but below the 45 degree line. This will exacerbate the underestimation of mean consumption in the very poor countries, but brings his estimates closer to the lowess curve for the middle income countries. However moving along the 45 degree line would in general predict a much faster rate of poverty reduction as a result of GDP growth for the lower income countries than is warranted on the basis of the information conveyed by the surveys. These two alternative estimates of global poverty are biased in both the level and trends of global poverty because they ignore the information contained in household survey means, as the World Bank estimates suffer because they discard the information contained in the national accounts. Of course, as discussed in the previous section, the differences in the adopted poverty lines also make these alternative estimates non-compatible with each other and with World Bank's estimates.

5. The non-compliance hypothesis and the calibration of survey means

The calibrated survey means can be used along with the decile distribution data from the household surveys to estimate $1 and $2 a day poverty measures. This procedure has been subject to a number of criticisms that need to be addressed before we can proceed. The most important criticism is based on the non-compliance argument. If the error in survey mean is by and large due to the non-compliance of the rich, the survey mean will be biased, but the correction of this bias can lead to underestimation of poverty by unduly increasing the income of the poor. This is a plausible argument if it can be shown that the variation in the cross-country mean survey deviation from national accounts averages is due to under reporting of the rich. However, as seen above, the relationship between the survey mean and the national accounts averages is more complex than a simple negative bias is the survey mean, and indeed in many countries the survey means are higher than national accounts averages.

Few attempts have been made in the literature to empirically test the non-compliance hypothesis and examine its implications for international poverty comparisons. Deaton (2003) models selectivity bias resulting from non-compliance by using truncated probabilities due to non-response for higher income groups as a function of income. Assuming a log normal distribution and constant elasticity of non-compliance with respect to income, he shows that under these assumptions, and if truncation begins at very low income levels, non-compliance keeps the variance of log income constant, and that the ratio of observed mean to true mean varies with the variance of log income but is not correlated with mean income. With slight relaxation of the form of the non-compliance function, Deaton shows that the observed variance of log income is no longer equal to the variance of the true distribution, but mean deviation remains independent of average income. He subjects his results regarding the relationship between mean deviation and variance of the distribution to empirical tests using the World Bank survey data, but finds no significant correlation between the log ratio of survey mean to national accounts averages and variance of log incomes or the gini coefficient. However, since in the more general case discussed by Deaton the observed variance of log incomes or the gini coefficients also deviate from their true values, this empirical test remains inconclusive.

Mistiaen and Ravallion (2003) attempt to estimate the impact of non-compliance on income distribution in the United States. Under fairly general conditions they show that with non-compliance increasing as a function of income, survey mean and distribution as well as poverty estimates can be biased. In the case of the US they show that non-compliance affects the upper end of the distribution much more than the lower income groups, but the error in estimated poverty seems marginal. It is difficult to see the implication of their finding for the sample countries we are dealing with in a quantitative sense, because they do not report the decile distribution of income in the pre- and post adjusted data in their results. The results that they do report, namely the change in average income of the different income cohorts resulting from non-compliance adjustment, are not inconsistent with the mean calibration performed here which allocates a very small share of total adjustment to the low income cohorts. Furthermore, in cross-country comparisons much larger errors might overshadow the non-compliance error, and hence their results do not necessarily support the idea that survey mean calibration is unwarranted. Nevertheless, their findings regarding the relationship between non-compliance error and observed mean deviation in the case of the US data is important and highlights the need for more rigorous statistical verification in the case of our sample countries before we can proceed.

Ravallion (2003), uses the non-compliance of the rich hypothesis in defending the World Bank's practice of using unadjusted survey means in poverty measurement, by maintaining that, 'more plausibly, the underestimation of mean income from a survey tends to come hand-in-hand with an underestimation of the extent of inequality'(ibid, p.12). More specifically, drawing on Mistiaen and Ravallion (2003), Ravallion argues that, 'Recent research has studied how poverty and inequality measures from survey data can best be corrected for the tendency of richer households to not want to participate in such surveys… Results from the US suggest that without such corrections, surveys tend to appreciably underestimate both the mean and the extent of income inequality, but that these two effects are roughly offsetting for measures of poverty' (Ravallion 2003, p.14). In other words, under the non-compliance hypothesis the errors in the survey mean are likely to be neutralized by changes in the distribution of income.

To examine this argument, two questions need to be addressed. First, what are the likely implications of the non-compliance hypothesis for poverty measurement? And second, to what extent is it reasonable to believe that the discrepancy between survey and national accounts means are caused by the non-compliance of the rich? Starting from the hypothesis that the non-compliance of the rich is responsible for the mean deviation in the observed survey data, we may be able to devise more conclusive tests under more general conditions. This is after all the hypothesis which seems to be put forward by Ravallion (2003) in his critique of survey mean adjustment. We refer to this as the strict non-compliance hypothesis, which maintains that the poor have full compliance and non-compliance is due to non-response and/or under reporting of income by the non-poor. Before attempting to test this hypothesis, however, we need to examine the implications of the strict non-compliance hypothesis for poverty measurement based on survey information.

Non-compliance can be modelled in two ways, depending on whether it arises from the non-response or the underreporting of income by the rich. The first one gives rise to sample selection bias as modelled by Deaton (2003) and Mistiaen and Ravallion (2003). The second one results in biased mean and variance in the observed sample, but the bias is not due to sample selectivity. The implications of the two sources of non-compliance for poverty measurement can be very different. We shall consider both cases, starting with sample selectivity case.

The non-response case

We assume the population is composed of two groups, the respondents and the non-respondents. We denote the unobservable true population distribution and density functions for income or consumption by $F(x)$ and $f(x)$ with population mean μ. To be consistent with Deaton (2003) and with the available distributional indicators in the World Bank's data set – namely, decile distributions and gini coefficients – we measure all variables in logs. The density functions of the two groups of respondents and non-respondents are denoted by h(x) and $h_{nr}(x)$ with population means μ_r and μ_{nr} respectively. We assume a fairly general probability of compliance function p(x), such that $0 < p(x) \leq 1$, and $p`(x) \leq 0$, i.e. p(x) is non-increasing in x. More specifically under the strict non-response hypothesis, we assume p(x) = 1 for x < z, and $p`(x) < 0$ for $x \geq z$, where z is the poverty line. It is clear that under these assumptions, $\mu_r < \mu < \mu_{nr}$ as long as there is non-compliance in the population.

The density function from which observations are drawn, or the true density conditional on compliance, can be therefore written as:

$$h(x) = \frac{p(x)f(x)}{\int_0^\infty p(u)dF} = \frac{p(x)}{\theta} f(x)$$

Where θ is the expected value of the probability of response, or the proportion of respondents in the population, which is between 0 and 1. Under the assumption of full compliance of the poor, $p(x) = 1$ for $x < z$, and hence the proportion of the poor amongst the respondents, P_r is:

$$P_r = \int_0^z dH = \frac{1}{\theta}\int_0^z p(x)dF = \frac{1}{\theta}\int_0^z dF = \frac{1}{\theta}P \qquad \text{Eq. 1}$$

where P is headcount poverty measure in the population as a whole. Under the assumption of strict compliance of the poor, therefore, the number of the poor amongst the respondents is equal to the number of the poor in total population, as in a population of size N:

$$\text{The number of the poor amongst respondents} = \theta N \int_0^z dH = \theta N \frac{1}{\theta}\int_0^z dF = N\int_0^z dF$$

Which is equal to the number of the poor in the total population. This would be, however, of interest to the measurement of headcount poverty only if the entire population is contained in the original survey. In sample surveys where the sample proportions are multiplied by an inflation or expansion factor to estimate the number of the poor, headcount poverty measured as sample ratio is of interest. Under the non-response hypothesis, however, headcount poverty measured as sample proportion is biased, even if we assume full compliance of the poor. Considering a sample of size n, with indicator variable y_i taking the value of 1 for a respondent with income or consumption below the poverty line, and zero otherwise, the estimated headcount poverty will be $\hat{p} = \sum_{i=1}^{n} y_i / n$.

Since the sample is drawn from amongst the respondents only, $E(y_i) = P_r$, which is the poverty rate in the respondent population. Denoting headcount poverty in the entire population by P, we will have:

$$E(\hat{p}) = \sum_{i=1,n} \frac{E(y_i)}{n} = \sum_{i=1,n} \frac{P_r}{n} = P_r = \frac{1}{\theta} P$$

Which implies a bias of $(1-\theta)/\theta$ times the true headcount poverty ratio. The bias is over 10 per cent of true measure of poverty for a 10 per cent overall non-compliance rate, rising rapidly to close to 70 per cent for non-response rate of 40 per cent, and 100 per cent and over for non-response rates of 50 per cent and more. Under the assumption of strict compliance of the poor, the bias will only vanish if there is full compliance by the rich as well. Under more general non-compliance assumptions however, the bias will depend on the average compliance of the poor relative to the rich. Paradoxically, the closer the compliance of the poor to the average compliance of the total population, θ, the lower will be the bias in poverty measurement. In fact assuming the same rate of compliance for all the poor as θ, would lead to disappearance of the bias, but this goes against the assumption of a monotonically declining p(x) for the non-poor.[9]

The variation in the response rates in sample surveys in different countries is thus another source of difficulty in cross-country comparisons of poverty. Of course in a well-designed and well-implemented survey, and assuming the strict compliance of the poor, the surveyor may be in a position to calculate an accurate estimator for θ whereby this bias may be corrected. For example, in a representative sample the average response ratio $\hat{\theta}$ is an unbiased estimator for the population response rate, which can be used to correct the bias. Assuming that this is done by the World Bank surveyors, under the assumption of full compliance of the poor the headcount poverty measure will be correct even with the assumption of non-response by a fraction of the non-poor.

Since we do not have any information regarding the income or consumption of non-respondents, however, the survey mean will continue to be biased. To show this, consider a sample of size n, x_1, x_2, …x_n, with sample mean $\bar{x} = \sum_{i=1,n} x_i / n$. Note that since the sample is drawn from the population of respondents $E(x_i) = \mu_r$. The bias of the sample mean will be:

$$B = E(\bar{x}) - \mu = \mu_r - \mu = \mu_r - (\theta\mu_r + (1-\theta)\mu_{nr}) = (1-\theta)(\mu_r - \mu_{nr})$$

Since under the assumption of a decreasing response probability p(x), μ_r will be always less than μ_{nr}, there is a negative bias in sample mean, even when poverty as measured by sample proportion of the poor is correct. This is the case hypothesized by Ravallion (2003), where the World Bank headcount poverty measures are argued to be correct even if the survey mean is underestimated. It is important to keep in mind the assumptions required for these results to hold. Firstly, non-response amongst the poor should be negligibly small, and secondly, the required adjustment to sample proportions are made by estimating θ with a reasonable degree of accuracy for the different countries. Under these

9 I am grateful to Graham Pyatt and Angus Deaton for pointing this out to me in reading a previous version of this paper.

conditions, any attempt to estimate poverty by using the correct population mean and the distribution of income from the sample will lead to underestimation of poverty as pointed out by Ravallion (2003). If we use the national accounts means for this purpose, the bias will be even larger, as the national accounts means are likely to overestimate the true household income and its growth over time (Deaton 2003).

We next need to investigate the implications of the strict non-compliance hypothesis in terms of testable propositions that can be tested on the basis of the available cross-country data. If we differentiate the bias in the above equation with respect to z, the poverty line, we will have:

$$\frac{dB}{dz} = \frac{d(\mu_r - \mu)}{dz} = \frac{d\mu_r}{dz} > 0$$

Since, under the assumption of full response of the poor and strictly declining response probability for the non-poor, the higher the poverty line the higher will be the proportion of respondents in the population and the closer will be the mean income of the respondents to population mean. That is, B approaches zero from below, or the bias falls in absolute value as the poverty line increases. In other words, given the mean and distribution of income, under the strict non-response hypothesis the mean deviation between respondents and total population declines as poverty increases. In the international poverty comparison the poverty line is fixed at \$1 a day, and it is the mean and distribution of income that vary across the countries. In that context, the above result implies that conditional on the shape or scale of the distribution, there is an inverse relationship between headcount poverty and the absolute bias in the observed mean $|B|$. Given that under the strict non-response hypothesis the World Bank poverty estimates are expected to be close to the true measures of headcount poverty – with the proviso of correction for mean non-response as discussed above – we should be able to use the World Bank estimates to empirically test this hypothesis.

The under-reporting case

Before turning to empirical tests, however, we need to consider the effect of the second version of strict non-compliance error, that is, underreporting of incomes. Defining s and x as the observed and true log of income, with density functions h(s) and f(x), we specify a general non-compliance function corresponding to the underreporting hypothesis as: $s = g(x)$ for $x \geq z$ with $g(0) = 0$, $g'(x) > 0$ and $g''(x) \leq 0$, where z is the poverty line as before. For values of $x < z$ we define $s = x$.[10] Under these conditions the observed distribution will produce the same poverty as the true distribution, as Probability (s < z) = Probability (x < z). This fully specifies the under reporting version of the strict non-compliance hypothesis under fairly general conditions. We can use this specification to measure the true and the observed income means. Since s is a monotonically increasing function of x the observed income mean μ_s can be derived as:

$$\mu_s = \int_0^z x dF + \int_z^\infty g(x) dF$$

The true mean income μ_x can be derived as:

10 Since the g(.) function is defined with reference to the poverty line z, another requirement for this function is that the under-reporting should not make a non-poor person poor (see, e.g., Pyatt 2003).

$$\mu_x = \int_0^\infty xdF = \int_0^z xdF + \int_z^\infty xdF$$

The mean difference, or log mean ratio in this case, can be written as:

$$D = \mu_s - \mu_x = \int_0^z xdF + \int_z^\infty g(x)dF - \int_0^z xdF - \int_z^\infty xdF$$ That is:

$$D = \mu_s - \mu_x = \int_z^\infty [g(x) - x]dF < 0$$

Survey mean is less than the true mean as under the non-compliance hypothesis $g(x) < x$ for all values of x above the poverty line. As noted above, however, using the observed survey mean, s, combined with the observed distribution h(s) would still produce the correct poverty estimates. If we differentiate the mean deviation variable, D, with respect to the poverty line we will get:

$$\frac{dD}{dz} = -[g(z) - z)f(z) > 0$$

That is, mean deviation increases, i.e., the observed survey mean gets closer to the true mean, as the poverty line increases. These are similar to the results obtained under the non-response assumption, and hence the same empirical tests would cover both versions of strict non-compliance hypothesis.

Empirical tests

In order to test the non-compliance hypothesis empirically, following Deaton (2003), we take the log ratio of the survey to national accounts mean as a proxy for mean deviation $\mu_s - \mu_x$, for the observations where the survey mean is lower than the national accounts consumption mean. To test the strict non-compliance hypothesis, we investigate the relationship between the mean deviation variable and World Bank measures of headcount poverty, conditional on the scale or shape of the distributions as measured by variables such as the gini coefficient or the variance of log income. Under the hypothesis of non-compliance, the World Bank poverty measures can be equal to the true headcount poverty despite the errors in mean and variance of surveys, with the proviso of mean response adjustment made in the non-response case as discussed above. However, under the noncompliance hypothesis the available scale indicators such as the gini coefficient or the variance of log income from the surveys are different from the shape indicators of the true distribution.[11] Though in testing the noncompliance hypothesis we are not interested in the relationship between the scale indicators and mean deviation per se, the inclusion of such indicators in the regression would prevent possible bias due to omission of variations of income distribution across countries. However, given that the bias due to the omission of distribution indicators would favor the noncompliance hypothesis, such omission turns out to be immaterial to the results of the test.

We regressed the log of the ratio of survey mean to national accounts mean on the World Bank headcount poverty estimates for both the \$1 a day and \$2 a day poverty lines, with or without distribution variables such as the gini coefficient and the variance of log income. As in Deaton (2003), in none of the regressions the coefficients of the distribution variables were significant. The key coefficient for testing the non-compliance hypothesis, namely that of the poverty variable, was also insignificant in all cases and had

11 This is discussed by Deaton (2003), and Mistianen and Ravallion (2003) in the non-response case, where it has been shown that the direction of the bias is not determinate a priori. In the under-reporting case it can be shown that under the strict non-compliance hypothesis, the expected value of sample variance is always less than population variance.

the wrong sign. The scatter plots of the log mean ratio for the consumption surveys, against headcount poverty for the \$1 and \$2 a day poverty lines, together with the fitted regression lines are shown in *Figures 8 and 9.* As can be seen, if anything, contrary to the predictions of the noncompliance hypothesis the absolute value of mean deviations seem mildly to increase with poverty. Considering that these tests do not even include the observations where the survey mean is higher than the national accounts mean, in which case the noncompliance hypothesis is a fortiori untenable, they clearly do not lend any support to the noncompliance hypothesis.

Figure 8. Log Mean Ratio against \$1 a day headcount poverty

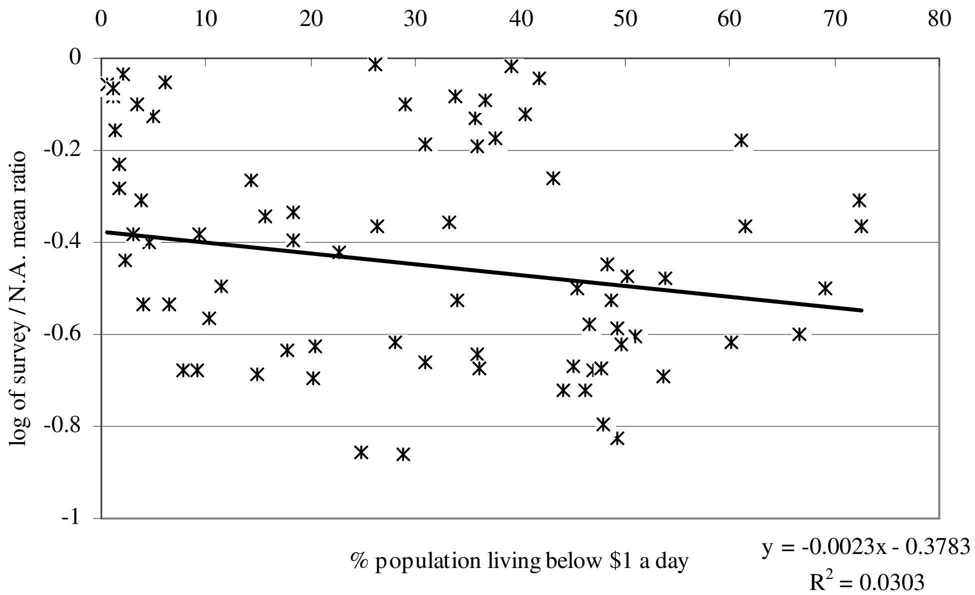
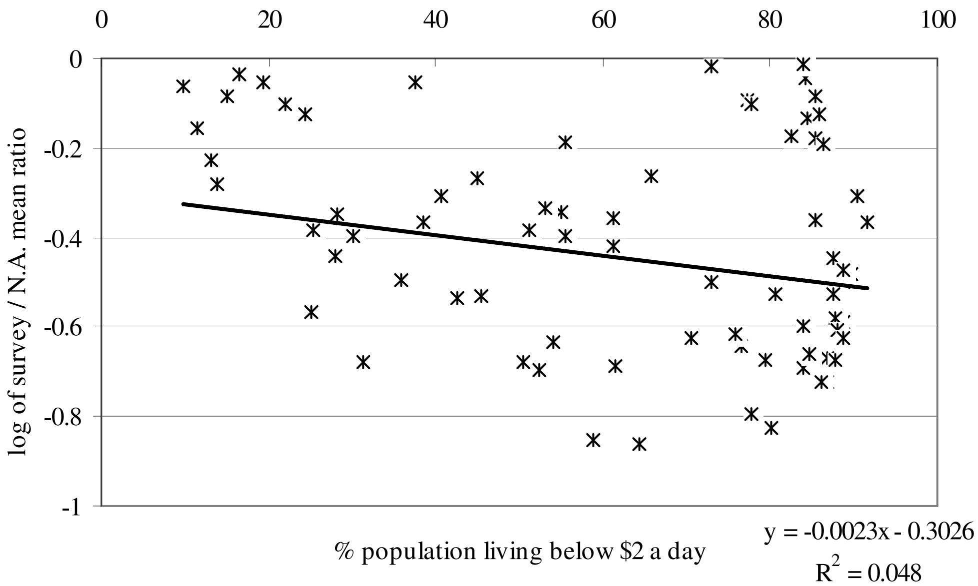


Figure 9. Log Mean Ratio against \$2 a day headcount poverty



One may relax the assumption of strict non-compliance, by assuming that for example some of lower income brackets of the non-poor also follow full compliance, without changing the basic structure of the above test. These results do not of course mean that noncompliance is not an important source of error in survey data. What the above tests indicate is that in cross-country comparisons there may exist other more important sources of error that overshadow the noncompliance error. One may therefore argue that the lack of a significant relationship between poverty and the mean deviations between the surveys and the national accounts data, lends support to the practice of using the calibrated survey means combined with the distribution indicators from the surveys in poverty measurement followed here. This practice, however, has been questioned in that it corrects for the errors in survey means but pretends the distribution information from the same surveys are correct (see, e.g., Ravallion 2003 and Deaton 2003). As correctly pointed out by Ravallion (2003), how can we assume that the shape of the distribution is correct but its mean is error ridden? In fact both the observed location and shape indicators in survey data are likely to be error ridden. The question is, how significant these errors are, and what can be done about them?

In an ideal world one would have liked to know both the true shape and mean of the distribution of household income. Such an ideal, however, is not attainable. Even if our surveys cover the whole of the population and non-compliance rates fall to zero so that sampling errors are non-existent, non-sampling errors resulting from differences in questionnaires would be present. If the questionnaire are made uniform and with the same recall periods, error will still persist because of the fallibility of human memory in recalling past expenditures. Measurement errors will always exist, but the question that we need to address is how important the errors are and how significantly they affect the poverty estimates. The errors in survey means appear too large to be ignored. The coefficient of variation of the log ratio of survey to national accounts mean is over 1.60 for the consumption surveys. This is eight times higher than the coefficient of variation of the Gini coefficients of sample countries of only 0.20, and the coefficient of variation for the share of the tope 20% income groups which is 0.19.

The variations in the mean and distribution of course do not affect poverty in a uniform way. To form an idea of the relative significance the two in our sample countries, we have compared the poverty estimates based on the calibrated means with a number of alternatives. First, keeping distribution constant, we measure $1 a day headcount poverty using both calibrated and non-calibrated survey means in 1985 PWT6.1 PPP values. Second, we keep the means constant at the level of calibrated survey means, and vary the gini coefficient by ± one standard deviation of the gini coefficients in the sample (which is about ±9.5). All the measurements for the purpose of this exercise are done by assuming a lognormal distribution. The results indicate that using poverty measures with the calibrated survey means and original survey distribution as the baseline, the root-mean-square-deviation of poverty for the non-calibrated survey mean is 13.1, which is 2 to 3 times higher than the root-mean-square-deviation of poverty measures due to the change in distribution (5.2 for the lower gini coefficient assumption and4.4 for the raised gini coefficient alternative). Considering that the ± 9.5 adjustment to the gini coefficient implies a substantial ±20 per cent change in the gini coefficient for most sample countries, these results signify the fact that variations in survey means are too important to be ignored.

The high variation in survey means would imply that the penalty for not correcting for the error in survey means can be very high. For example a glance at Figure 6 will show that there are a large number of countries at per capita consumption (national accounts) range of $1000 per annum whose standards of living according to the survey averages are lower than countries with per capita consumption (national accounts) of below $300 per annum! As we have argued above, to assume that such large deviations will be ironed out by variations in income distribution in the surveys, as assumed by the non-compliance argument, is unrealistic.

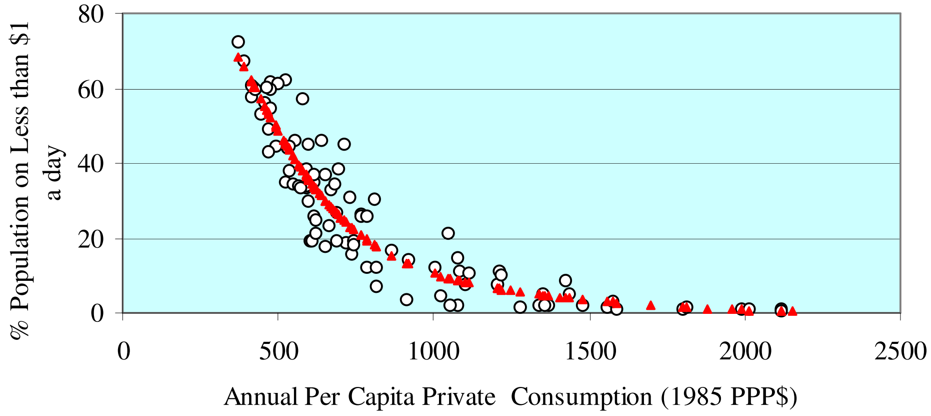
Given the lack of significant observable relationship between survey mean errors and distribution indices, it is not clear how best to adjust the decile data without further research. Given the relative stability of the decile distributions the best strategy may be to leave them as they are. As we shall see below, variations in mean consumption appear to have a more significant impact on poverty than distributional changes in the income ranges of below $1000 in 1985 PPP where the $1 a day poverty line is really relevant. Furthermore, considering that in our sample countries on average over 70 per cent of expenditure or income belongs to the top 40 per cent of income groups, much of the adjustment in survey mean, keeping the decile distribution intact, will be allocated to the richer households.

6. Regional and global poverty estimates

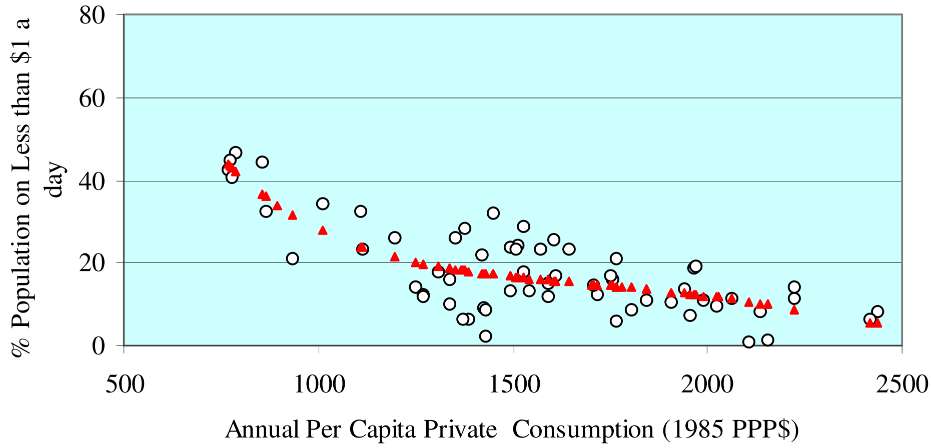
Using the calibrated survey means in 1985 PPP values and the decile distribution from the surveys we estimate headcount poverty for individual countries for the $1 a day and $2 a day poverty lines. Apart from the use of the calibrated survey means, our estimates differ from the World Bank estimates in a number of ways. First, we use the PWT6.1 PPP values and the corresponding $1 a day poverty line in 1985 prices, compared to the World Bank's use of its own 1993 PPP exchange rates and $1.08 poverty line in 1993 prices. There is also a difference in the treatment of income surveys. In the case of income surveys, the practice by the World Bank, as described in Chen and Ravallion (2000), is to adjust survey means by using the household consumption / income ratio from the national accounts. This procedure is questionable for a number of reasons. Household income surveys usually report survey means which are not very different from expenditure means, and often below expenditure means. Deflating income further means by national accounts consumption ratios can unduly inflate poverty estimates. The assumed relationship between the national accounts savings rates and household savings rates is also questionable, particularly in the case of poor households. Furthermore, to combine the adjusted survey means, assumed to reflect mean consumption, with decile distributions which supposedly refer to income distribution is problematic. As in the income ranges below the $1 and $2 a day poverty lines household savings are likely to be negligible, in the case of income surveys we use the calibrated survey means along with their associated distributions to estimate poverty.

The estimated headcount poverty measures are plotted against average per capita household consumption in *Figures 10 and 11* for the $1 and $2 a day poverty lines respectively. The Latin American countries are on the whole higher income countries in the sample, with relatively higher Gini coefficients relative to other countries, and they constitute the bulk of the observations in the sample countries with income surveys. We have, therefore, shown the observations belonging to the lower income African and Asian countries in panel (a) and those of Latin American countries in panel (b) in Figures 10 and 11. Four countries in Southern Africa, namely Zimbabwe, South Africa, Botswana and Lesotho, are also moved to panel (b), as they show similar income distribution traits as the Latin American group. The combined sample is shown in Panel (c) of Figures 10 and 11. Each Figure also shows the poverty curve, which is a logistic curve fitted to the data.

Figure 10. $1 a day headcount poverty estimates and poverty curves
(a) Asian and African Countries



(b) Latin American Countries



(c) Combined Sample

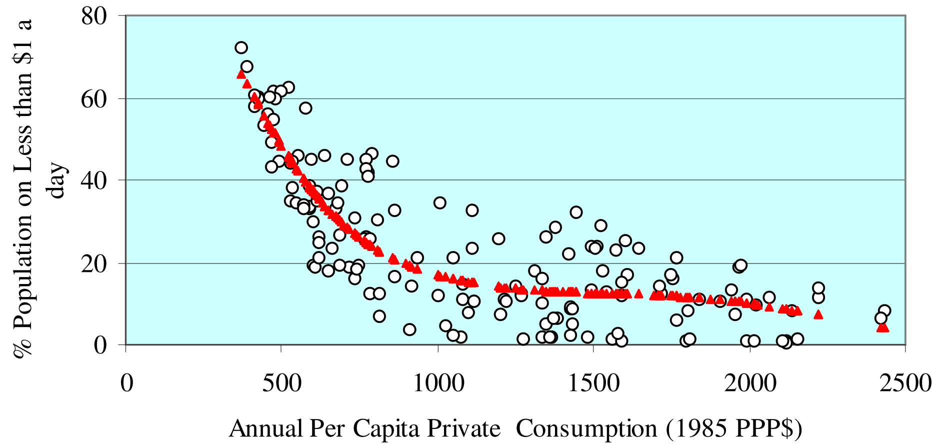
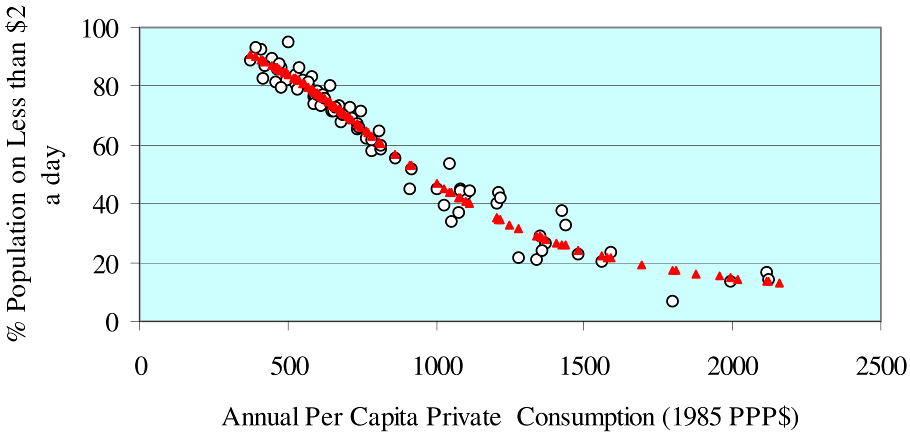
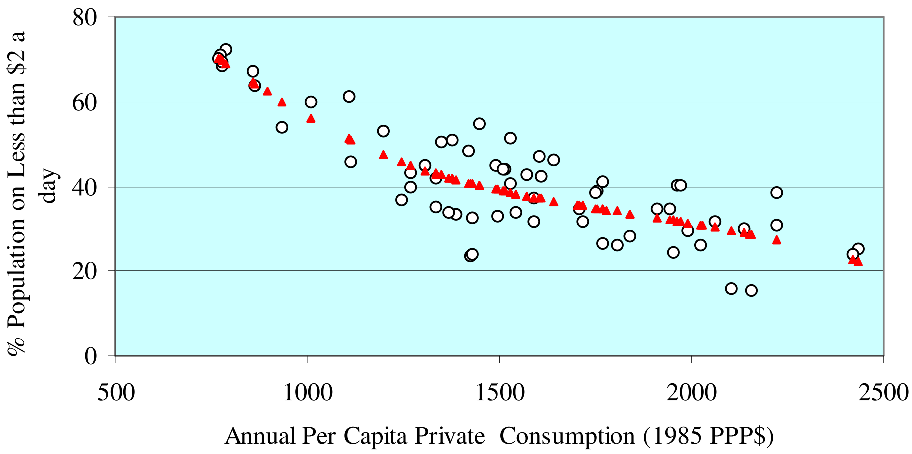


Figure 11. $2 a day headcount poverty estimates and poverty curves

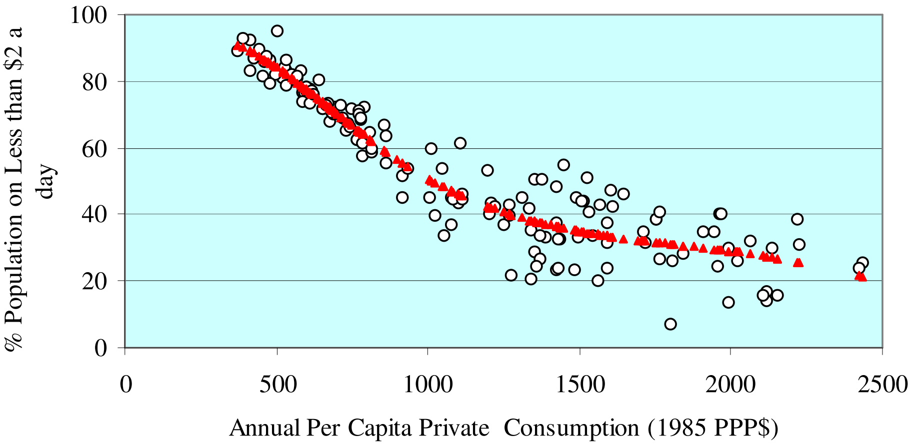
(a) Asian and African Countries



(b) Latin American Countries



(c) Combined Sample



As can be seen from Panel (a) of Figures 10 and 11, in consumption ranges of below $1000 a year, there seems to be a close association between per capita consumption and poverty. In countries at middle and higher income ranges in Panel (b), this association is less pronounced and variations in income distribution overshadow the effect of mean variations on poverty. In general the further to the right of poverty line lies the mean of the sample, the more pronounced will be the variations in income distribution in explaining variations in poverty across the countries. Since we are using the same data on income distribution and the same poverty line as the World Bank, and since our calibration method preserves the mean of the sample, the plot of World Bank poverty estimates against the survey means will show exactly the same phenomenon. The individual country results, and the relationship between the two poverty estimates with the national accounts per capita consumption will be however very different between our estimates and those of the World Bank. And so will be the regional and global poverty estimates.

In measuring regional and global poverty estimates we adopt a methodology that is as far as possible consistent with Chen and Ravallion (2000). We measure aggregate regional poverty for the same years as Chen and Ravallion, namely, 1987, 1990, 1993, 1996, and 1998. For countries that have only one survey, the World Bank estimates assume the same income distribution for all the years, but use the national accounts data to extrapolate survey mean. As we have observed above, this practice is not consistent with the observed relationship between survey means and national accounts averages at low-income levels. For this type of country we use the calibrated survey mean conditional on per capita national accounts consumption, on the basis of regression curves discussed in Section 4.

For countries that have more than one survey, we use the calibrated survey means combined with distribution data in the year closest to the year for which measurements are made. In countries where more than two household surveys exist, we check for the possible errors in decile distribution in particular years. For example, Brazil 1997 data shows a gini coefficient which is much lower than 1996, and is below the historical trends in that country. This anomaly cannot be due to the non-compliance of the rich, as there is also a large increase in the ratio of survey mean to national accounts mean in that year. Under these types of circumstances, rather than deleting the observation, we use the calibrated survey mean in conjunction with what appears a more plausible distribution, which in the case of 1997 Brazil is the 1996 distribution data. This is made possible by the fact that the calibrated survey means in the case of individual observations are estimated from the fitted regression lines and are minimally influenced by the distribution errors in individual observations. The World Bank practice in the case of such anomalies is to take the average of poverty estimates for the years that lie between two surveys with anomalous distributions (See, Chen and Ravallion 2000).

The aggregate regional and global poverty estimates for the $1 a day and $2 a day poverty lines are shown in Tables 2 and 3. Our estimates are listed in panel a of Tables 2 and 3, with Chen and Ravallion's estimates in panel b of the respective tables for comparison. In addition to headcount poverty rates, the tables also report the absolute number of the poor in each region, which is derived by applying the poverty rates to the total population of each region. The aggregate global poverty rates for the $1 a day poverty line are relatively close to the World Bank estimates. For example between 1990 and 1998, global poverty falls from over 28 per cent to over 23 per cent in both cases. The aggregate figures, however, conceal important differences between our estimates and the World Bank estimates at the regional level. In comparing the aggregate figures it should be also taken into account that our estimates exclude the Eastern European and Central Asian countries.

Table 2a. Population living below $1 a day in 1985 PPP

	% of pop. living below $1 a day					Number of poor (millions)				
Region	1987	1990	1993	1996	1998	1987	1990	1993	1996	1998
East Asia	33.3	31.4	29.6	21.3	20.6	505.3	501.0	492.4	367.6	364.3
(excluding China)	21.3	17.1	15.5	11.3	10.1	92.6	78.9	75.7	57.9	53.2
Latin America &Caribbean	20.3	20.6	18.1	15.3	15.2	84.4	90.3	83.6	74.1	76.2
MENA	2.6	2.6	2.3	1.9	1.8	5.7	6.2	5.9	5.1	5.2
South Asia	30.2	23.9	23.9	18.1	22.7	317.9	268.3	283.8	228.3	296.3
Sub-Saharan Africa	45.6	45.7	49.9	49.2	48.2	212.6	232.7	274.5	293.1	302.7
Total	31.1	28.4	27.8	22.1	23.3	1125.8	1098.4	1140.2	968.3	1044.7
(excluding China)	27.6	23.9	24.0	20.5	22.4	713.2	676.3	723.5	658.6	733.6

Notes: Totals exclude Eastern Europe and Central Asia

Table 2b. World Bank's estimates of population living below $1.08 per day at 1993 PPP

	% of pop. living below $1 a day					Number of poor (millions)				
Region	1987	1990	1993	1996	1998	1987	1990	1993	1996	1998
East Asia	26.6	27.6	25.2	14.9	15.3	417.5	452.5	431.9	265.1	278.3
(excluding China)	23.9	18.5	15.9	10.0	11.3	114.1	92.0	83.5	55.1	65.2
Latin America &Caribbean	15.3	16.8	15.3	15.6	15.6	63.7	73.8	70.8	76.0	78.2
MENA	4.3	2.4	1.9	1.8	2.0	9.3	5.7	5.0	5.0	5.6
South Asia	44.9	44.0	42.4	42.3	40.0	474.4	495.1	505.1	531.7	522.0
Sub-Saharan Africa	46.6	47.7	49.7	48.5	46.3	217.2	242.3	273.3	289.0	290.9
Total	28.3	28.4	28.2	24.5	24.0	1183.2	1276.4	1304.3	1190.6	1198.9
(excluding China)	28.5	23.9	27.7	27.0	26.2	879.8	915.9	955.9	980.5	985.7

Notes: World Bank's Total includes Eastern Europe and Central Asia. Source: Chen and Ravallion (2000).

Table 3a. Population living below $2 a day in 1985 PPP

	% of population living below $2 a day					Number of poor (millions)				
Population, total	1987	1990	1993	1996	1998	1987	1990	1993	1996	1998
East Asia	73.5	71.0	62.2	53.5	52.1	1115.6	1134.4	1035.6	926.0	922.4
(excluding China)	58.1	51.6	44.4	35.8	33.5	252.5	238.1	216.2	183.4	176.8
Latin America &Caribbean	40.9	41.1	38.1	34.5	34.3	169.5	180.0	176.2	167.4	171.6
MENA	17.3	16.4	15.6	14.1	13.9	37.5	39.1	40.0	38.4	39.4
South Asia	73.8	69.3	65.7	59.2	60.5	775.4	775.9	780.8	745.8	790.0
Sub-Saharan Africa	75.2	74.9	75.9	75.9	75.1	351.0	380.8	417.8	451.6	471.3
Total	68.2	65.7	60.6	54.5	54.2	2449.1	2510.2	2450.4	2329.2	2394.7
Excluding China	62.4	59.0	56.2	51.4	51.5	1585.9	1613.8	1631.0	1586.6	1649.1

Notes: Totals exclude Eastern Europe and Central Asia

Table 3b. World Bank's estimates of population living below $2.15 per day at 1993 PPP

	% of population living below $2 a day					Number of poor (millions)				
Population, total	1987	1990	1993	1996	1998	1987	1990	1993	1996	1998
East Asia	67.0	66.1	60.5	48.6	49.1	1052.3	1084.4	1035.9	863.9	892.2
(excluding China)	62.9	57.3	51.6	42.8	45.0	299.9	284.9	271.6	236.3	260.1
Latin America &Caribbean	35.5	38.1	35.1	37.0	36.4	147.6	167.2	162.2	179.8	182.9
MENA	30.0	24.8	24.1	22.2	21.9	65.1	58.7	61.8	60.6	62.4
South Asia	86.3	86.8	85.4	85.0	84.0	911.0	976.0	1017.8	1069.5	1095.9
Sub-Saharan Africa	76.5	76.4	77.8	76.9	75.6	356.6	388.2	427.8	457.7	474.8
Total	61.0	61.7	60.1	56.1	56.0	2549.0	2718.4	2784.8	2724.1	2801.0
Excluding China	58.2	58.8	58.6	57.8	57.6	1796.6	1918.8	2020.5	2096.5	2168.9

Notes: World Bank's Total includes Eastern Europe and Central Asia. Source: Chen and Ravallion (2000).

An important difference between the two sets of estimates at the regional level is the much lower poverty rates in South Asia in our estimates compared to the World Bank estimates. This is to some extent compensated at the aggregate level by relatively higher poverty rates in our estimates for East Asia. According to World Bank estimates poverty rates in South Asia are not very different from sub-Saharan Africa for the $1 poverty line, and they are as much as ten percentage points higher in the case of the $2 poverty line. This is the direct result of the inconsistency of the survey means in low income countries, which give rise to the flat tail in the lowess regression curve as observed in Section 4 above. As a consequence, countries in sub-Saharan Africa with per capita GDP of just over $300 in 1985 PPP, have the same survey means as South Asian countries with double per capita GDPs. This is the main reason for the higher poverty estimates by the World Bank in South Asia relative to sub-Saharan Africa.

As far as the attainment of the millennium goals are concerned, poverty trends are more critical than the estimated levels. Aggregate poverty trends are of course also affected by the regional poverty levels. For example if poverty is more concentrated in the faster growing regions then the aggregate trends will produce a more optimistic picture. One difference between our results and the World Bank estimates regarding poverty trends which stands out, is that in our estimates of poverty rates in sub-Saharan Africa show about 2.5 percentage points increase between 1990 and 1998, with about 70 million increase in the total number of people living below the $1 poverty line between these two dates. World Bank estimates show a slight decline in poverty rates, and some 58 million increase in absolute numbers. Latin America and Caribbean region also seems to have a sharper decline in poverty between 1990 and 1998 in our results as compared to the World Bank.

7. Concluding remarks

The proliferation of different global poverty estimates is indicative of both unresolved methodological issues and lack of transparency in presentation and manipulation of data. This paper has been an attempt towards establishing a more unified approach. Without a unifying framework, the current state of affairs would only add to the confusion through increasing proliferation of incompatible poverty estimates. Progress on methodological side has been slow as the problem of incompatibility of survey means with national accounts data has come to light late in the day. Sometimes ex post rationalizations have been invoked to assume away the discrepancies brought to light only recently about age-old practices. The lack of compatibility between the different purchasing power parity exchange rate estimates has added to the confusion. International poverty lines have been adjusted at the whim of one or a few researchers, etc. To plod one’s way through the resulting confusion assumes more the character of detective work than straightforward research.

Much of this confusion is, however, unnecessary. World Bank living standard surveys are a rich source of data for researchers on various aspects of development. As far as the poverty measurement work is concerned, it would be best if the individual household data were made available at domestic currency units on the World Bank’s poverty monitoring web site. If this is not practicable then the provision of data on survey mean and average income of twenty or more income cohorts in domestic currency would go a long way towards ending the current confusion. This would substantially enhance the transparency and clarity of the data, and it would help end the need for second guessing the conversion factors and the methods used by the World Bank to convert the survey data into international prices under different base years. Timeliness of the availability of data is another important issue. We would have been years ahead in research on methods of reconciliation of the survey means with national accounts, had the survey mean data been available from the outset, like other World Bank data, e.g., in WDI.

Similarly the confusion in relation to the question of international poverty lines is not due to intrinsic difficulties of this concept, which are plenty and important on their own account, but the result of the change of the base year and the purchasing power parity exchange rate estimates by the World Bank and the difficulty to link these to the older Penn World Table estimates. With the availability of the PWT6.1 PPP data, it would be important to establish an appropriate poverty line for the new PPP estimates. The $1 a day poverty line in 1985 PWT6.1 PPP rates suggested in this paper needs to be investigated further and an appropriate poverty line has to be endorsed by an international forum to prevent further confusion and duplication of work with incompatible poverty lines. The World Bank methodology for estimating PPP exchange rates is different from the Penn World Tables, and in order to prevent two parallel global poverty estimates from emerging it may be appropriate to adopt the PWT6.1 PPP measures as the international standard estimates for poverty measurement. Over time as the PPP estimates are revised, it may be necessary to adjust the international poverty line accordingly. However, this has to be done in an international forum according to explicit and clear rules, rather than at the discretion of particular researchers, in order to prevent confusion and duplication of work.

The standardization of living standard surveys suggested by Deaton (2003), to minimize errors resulting from differences in survey design, can go a long way in making the household surveys comparable across countries. However, it is unlikely that survey means will be fully comparable, and the need for calibrating survey means by using external national accounts based information will always be present. At present, given the low quality of data in the poorest countries, it appears that calibration of the survey mean is needed not only to filter out the random errors but also to adjust the systematic component of survey mean in the case of these countries. However, with the availability of more surveys for the low-income countries, more precise calibration can be attempted. As we have been trying to show in this paper, ignoring the mean error in the surveys and hoping that these will be neutralized by differences in scale variables is not realistic.

Bibliography

Aten B. and A. Heston (2003), 'Use of Penn World Tables for International Comparisons of Poverty Levels: Potentials and Limitations', *Global Poverty Workshop*, Initiative for Policy Dialogue, Columbia University, March 31-April 1, 2003.

Bhalla, S. (2002), *Imagine there is no country: poverty, inequality, and growth in the era of globalization*, Institute for International Economics, Washington D.C.

Bhattacharya, N. (2002), 'Results of a Pilot Survey on Suitability of Different Reference Periods for Measuring Household Consumption', Report by NSSO Expert Group on Non-sampling Errors, Paper presented at the *Workshop on Poverty Monitoring and Evaluation*, New Delhi, January 11-12, 2002.

Chen S. and M. Ravallion (2000), 'How did the world's poorest fare in the 1990s?', mimeo, World Bank: Washington D.C.

Deaton, A. (2003), 'Measuring Poverty in a Growing World; (or Measuring Growth in a Poor World)', *NBER Working Paper Series*, Working Paper 9822, National Bureau of Economic Research, Cambridge MA.

Deaton, A. (2002), 'Data for monitoring the poverty MDG', mimeo, Research Programme in Development Studies, Princeton University.

Deaton A. (2000), 'Counting the world's poor: problems and possible solutions', mimeo, Research Program in Development Studies, Princeton University.

Heston A., R. Summers and B. Aten, Penn World Table Version 6.1, Center for International Comparisons at the University of Pennsylvania (CICUP), October 2002

Karshenas, M. (2001), 'Measurement and Nature of Absolute Poverty in Least Developed Countries', Economic Working Paper Series, no. 129, SOAS, University of London.

Karshenas, M. (2003), 'Global Poverty: National Accounts Based versus Survey Based Estimates', *Development and Change*, vol.34, no.4 September, pp.684-712.

Kulshreshtha A.C. and A. Kar (2002), 'Estimates of Food Consumption Expenditure from Household Surveys and National Accounts', Paper presented at the Workshop on Poverty Monitoring and Evaluation, New Delhi, January 11-12, 2002.

Mistiaen, J.A. and M. Ravallion (2003), 'Survey compliance and the distribution of income', The World Bank; Washington D.C.

Pyatt, G. (2003), 'Development and the Distribution of Living Standards: A Critique of the Evolving Data Base' Review of Income and Wealth, Vol.49, No.3, Page 333-358.

Ravallion M. (2003), 'The Debate on Globalization, Poverty and Inequality: Why Measurement Matters', World Bank, Washington D.C.

Ravallion M. (2000), 'Do National Accounts Provide Unbiased Estimates of Survey-based Measures of Living Standards?', mimeo, World Bank, Washington D.C.

Ravallion (2001), 'Measuring aggregate Welfare in Developing Countries: How well do National Accounts and Surveys Agree?', mimeo, World Bank, Washington D.C.

Ravallion, M., G. Datt and D. van de Walle (1991) 'Quantifying Absolute Poverty in the Developing World', *Review of Income and Wealth*, no.37, pp.345-361.

Sala-I-Martin, X. (2002), 'The disturbing 'rise' of global income inequality'. NBER Working paper No. 8904, Cambridge, M.A.

Sundaram, K. and S.D. Tendulkar (2002), 'Recent Debates on Data Base for Measurement of Poverty in India: Some Fresh Evidence', Paper presented at the *Workshop on Poverty Monitoring and Evaluation*, New Delhi, January 11-12, 2002.

World Bank (2001), *www.worldbank.org/research/povmonitor/index.htm,* World Bank: Washington D.C.