The University of Manchester
Brooks World Poverty Institute

MANCHESTER 1824

Flying Blind? Constructing Evidence-based Poverty Reduction Policies in PRSP Adopting Countries

Richard Marshall[1]

Bernard Walters[2]

Frederick Nixson[3]

March 2009

BWPI Working Paper 85

[1] University of Manchester
[2] University of Manchester
[3] University of Manchester

Richard.Marshall@postgrad.manchester.ac.uk
Bernard.walters@manchester.ac.uk
Fred.nixson@manchester.ac.uk

Brooks World Poverty Institute
ISBN : 978-1-906518-84-4

Creating and sharing knowledge to help end poverty

www.manchester.ac.uk/bwpi

Abstract

The World Bank, alongside other development institutions and leading donors, has increasingly emphasised the importance of evidenced-based policymaking in addressing entrenched poverty. Additionally, widespread adoption of Poverty Reduction Strategy Papers (PRSPs) has necessitated establishing poverty baselines and updating them on a regular basis, to enable effective impact evaluation. To facilitate this, considerable resources have been devoted to providing sound longitudinal poverty data via the implementation of the World Bank's flagship Living Standards Measurement Survey (LSMS) programme. While a lively debate has emerged over the conceptual basis of poverty measurement within LSMSs, rather less has been said about the wider impact of basic changes in survey design and measurement approaches. We argue that the latter have been as important in shaping controversies over the comparability of national poverty aggregates and the value of LSMS data for policy analysis. Supported by a review of World Bank Poverty Assessments carried out since 2004, and a case study of the Mongolian experience, we find that variations in survey methods and subsequent major revisions are far from rare, and have generally been accompanied by inadequate disclosure. This paper shows that where revisions have been made, the lack of transparency undermines the application of standard analytical techniques and prevents the replication of key results. We conclude that insufficient efforts are being made to establish consistent and verifiable poverty data; and, moreover, that the lack of openness and rigour has permitted the fitting of results to predicted policy objectives and weakened the level of external scrutiny. An implication of this is that effective policymaking is made more difficult and national ownership is compromised. This supports the case for alternative, more eclectic approaches to the monitoring and evaluation of poverty reduction policies.

Keywords: Poverty, Poverty reduction policies, Monitoring and evaluation, Mongolia

Richard Marshall is a postgraduate in the School of Social Sciences, University of Manchester.

Bernard Walters is a Senior Lecturer in Economics, School of Social Sciences, University of Manchester.

Frederick Nixson is a Professor of Development Economics, University of Manchester.

1. Introduction

Poverty reduction has come to dominate the global development agenda and in turn the policy priorities of national decision makers in less developed countries (LDCs). Since 2000 the primary framework through which economic and social policies have been crafted in the developing world has been driven by the Poverty Reduction Strategy (PRS) process, which seeks to provide Poverty Reduction Strategy Paper (PRSP)-adopting countries with a strategic, multi-sectoral and, crucially, a poverty-focused and nationally-owned approach. Yet, initially at least, this shift in emphasis was undermined by the lack of a credible evidential record. Given that effective policymaking requires relevant and reliable data, the World Bank and other leading donors, have sought to develop an internationally benchmarked system of poverty surveys via the Living Standards Measurement Survey (LSMSs) programme. In addition to providing an international template of best practice,[1] this initiative has also directly supported a series of national surveys, involving the extensive modification (in some cases wholesale replacement) of pre-existing systems of household income and expenditure surveys. The global roll-out of LSMSs has also explicitly aimed to ensure longitudinal comparability, the promotion of evidenced-based policymaking within PRSP countries, and higher levels of transparency through the disclosure of core data and estimation techniques.

However, this paper identifies emerging evidence that a significant proportion of LSMSs have not been comparable, with large changes in poverty incidence being retrospectively reported, principally by the World Bank through its Poverty Assessment (PA) series. Moreover, these changes to established poverty records have often been made in a non-transparent manner, offering little opportunity for the conclusions to be subject to independent scrutiny. In several cases, we also find indications of a selective approach to the re-statement of poverty aggregates, where the previous results fail to accord with the orthodox expectations. We note that this inevitably lays the Bank and national authorities open to charges of fitting findings to the desired policy outcomes.

In such circumstances, we argue, first, that the use of longitudinal LSMS data to track impacts and trends via established analytical tools is flawed and by extension, the rational formulation of poverty reduction policies is compromised. Second, that the integrity of policymaking and national ownership are themselves damaged where the methods used in *ex post* revisions are not transparent, and where national constituencies cannot access the base data.

It is important to note that the core arguments advanced here are distinct from those within the growing poverty estimation literature, which are critical of the absolute basis of calorific poverty lines (see for example Palmer-Jones, 2007), nor are they directly associated with the ongoing contributions over measurement error (such as Lanjouw and Lanjouw, 2001). Instead, we make an empirical case, which emphasises the complexities and difficulties of

[1] In domains including the derivation of poverty thresholds, survey design and sampling approach.

anti-poverty policymaking in LDC environments with poor data and successive non-transparent adjustments to national poverty records.

The paper first examines the size of the problem and the general impact on the policy-making process, by reviewing the extent of retrospective changes to poverty aggregates made by the World Bank in recent years across a wide range of countries; we also highlight situations where the data on which these revisions were made are unavailable or opaque. The way in which these comparability problems compromise rational policy choice and national ownership is further developed by means of a case study of the Mongolian experience. This example is chosen purposively. The national authorities and major donors have devoted considerable attention to poverty during the country's transition and have conducted three LSMSs in 1995, 1998 and 2002. In addition, Mongolia provides a good example of how a re-statement of the results of an earlier survey has radically changed Mongolia's poverty reduction 'story'.

The paper is structured as follows. Section 2 identifies those countries which have attracted World Bank criticisms of their poverty records and those where material re-statements have been made since 2004. It also examines how transparent the process of revision has been. Section 3 switches perspective and presents findings from our Mongolian case study. It reviews the conclusions about poverty that have been derived from the three LSMSs and attempts to resolve the controversy surrounding the comparability of the 1998 and 2002 results. This is supplemented in section 4, where we attempt to measure the trajectory of poverty using standard analytical tools, chiefly a poverty-growth-inequality decomposition, over the periods 1995-1998 and 1998-2002. Section 5 concludes and makes an argument for a more eclectic approach to gauging the impact of policy on poverty outcomes.

2. How common are retrospective revisions to published poverty records, and what is their general impact?

The science of poverty measurement remains a work in progress. In addition to ongoing developments in the estimation of poverty lines, improvements in survey design and analysis occur continuously. Both governments and donor agencies, which support and carry out poverty surveys, have a responsibility to incorporate such advances in order to improve the statistical basis on which policy decisions are made. However, the extent of the criticisms articulated in World Bank Poverty Assessments (PAs) and the discontinuities forced by the external restatement of past findings cannot be attributed simply to continuous improvements in method. Moreover, revisions are almost always justified in terms of ensuring comparability rather than sophistication. A review of all Poverty Assessments published since 2004 demonstrates over 40 percent contained material criticisms[2] of past data and around 20 percent made major revisions to the previously published record.

[2] 'Material' is meant in the sense that Bank analysts questioned the reliability of the established record.

Table 1: World Bank Poverty Assessments objections and revisions to poverty data (2004–2007)

Region/ country (37)	*Year*	*Criticism of past record*	*Headcount re-stated?*	*Notes*
Sub-Saharan Africa (6)				
Benin	2004			Not applicable – no past survey
Burkina Faso	2005	✓	✓	Major restatement, increases the level of reduction
Ethiopia	2005	✓		Comparability concerns but no restatement
Guinea Bissau	2006			
Malawi	2006	✓	✓	Restatement of 1998 data , draft report only
Uganda	2005			
East Asia & Pacific (5)				
Cambodia	2006	✓	✓	Past non-comparability. Major restatement
Indonesia	2007			
Mongolia	2006	✓	✓	Major restatement, increases the level of reduction
Lao PDR	2004			
Vietnam	2007			
Lat America & Carib (8)				
Bolivia	2005			
Brazil	2006	✓		No restatement, but major concerns
Dominican Republic	2005	✓		Disputed poverty lines – final report not yet available
Ecuador	2004			
El Salvador	2005			
Honduras	2006	✓		Highly critical, although no re-statement
Mexico	2004	✓	✓	Re-statement/adjustment, but also presents alternatives
Peru	2005			
Nepal	2006			
Sri Lanka	2006			
Europe & Cent Asia (13)				
Armenia	2004	✓	✓	Restates past but still presents alternatives
Belarus	2005			
Kazakhstan	2004			
Kosovo	2005			Not applicable – no past survey
Kyrgyz Republic	2005	✓	✓	Major objections, restatement made
Latvia	2006			
Macedonia FYR	2005			
Moldova	2004			
Poland	2004			
Russia	2004			
Slovak Republic	2005			Non-applicable – not a standard LSMS
Tajikistan	2004			
Turkey	2005	✓		Critical of past and alternatives, but no re-statement
Mid East & Nth Africa (3)				
Jordan	2004	✓		Major change in methods & critical, but no restatement
Morocco	2004	✓		Locally driven adjustment
West Bank & Gaza	2004	✓		Critical of alternatives and past

Grand total of 37, adjusted total of 34 countries – i.e. excluding those with no past survey
Source: World Bank Poverty Assessments library – see http://go.worldbank.org/2SDF7W6AR0

The full results of our investigations are provided in Tables 1, 2 and 3, alongside more detailed discussion. However, at the outset it is worth setting out the considerations that govern the impact such widespread revisions pose for the core policy functions of feedback and analysis. These can be expressed in the form of four key questions:

- Are the changes generated by these *ex post* reviews of a size to render unreliable the policy inferences drawn from earlier studies?
- As an adjunct to this, is it still possible for national policymakers to discern the general trajectory of poverty and its proximate causes?
- Are the retrospective adjustments to poverty estimates presented in a sufficiently transparent way to allow understanding and debate, which enhance in-country scrutiny and national ownership?
- Why are revisions made in some cases and not in others, and thus what processes or procedures signal the need for the re-examination of past LSMSs?

If the answers to these questions suggest that the value of LSMS-type surveys are frequently compromised, then a further crucial question is raised:

- Would the construction of alternative metrics, reflecting changes in poverty incidence and severity, be more likely to provide results that would be more reliable over time, and therefore, more policy useful?

The first stage in answering these questions is to establish in more detail the size of the problem, by identifying the frequency and extent of revisions to previously published estimates of poverty. Table 1 reports the results of an examination of the 37 countries for which recent PAs (published between 2004 and 2007) have been undertaken by the World Bank's staff. PAs are undertaken on a rotating basis for all members receiving support, and remain the main vehicle through which the Bank audits, and if necessary re-states, poverty data. Our review identified those where serious criticisms were made, and where re-estimation and re-statement of the past poverty aggregates were considered necessary.

The period provides a geographically balanced sample, and Table 1 groups the countries by region. Of the 37 for which assessments were completed in this period, three countries were excluded because there was no past survey on which to situate any changes, leaving 34 which in principle had past records capable of being criticised and/or revised. Column 3 indicates those countries where previous LSMSs were strongly criticised, and column 4 indicates those for which the World Bank undertook a restatement exercise. Column 5 makes a number of further observations, which clarify the reported re-statements – some emphasised the inadequacy of the data; in others the method of calculating poverty thresholds was criticised and re-worked.

Summary results from this initial exercise are given in Table 2. Overall, we found that serious criticisms and re-statements of past poverty records are more common than might be expected, given the LSMS programme's comparability objectives. Of the 37 published in the period, 15 Poverty Assessments strongly criticised the earlier reported figures, representing 44 percent of the adjusted total. In addition, the poverty headcount was revised in seven cases, representing 21 percent of the adjusted total. These figures provide *prima facie* evidence of a comparability problem affecting national LSMSs and show that, in spite of considerable resource inputs, major discontinuities have been generated in national longitudinal poverty records. The impact on policymakers, as framed by the questions above, requires further investigation.

Table 2: Summary proportions of Poverty Assessments (PAs) identifying serious weaknesses and poverty re-statements published between 2004 and 2007

	Of all PAs	Of comparable PAs
Proportion where past survey results were challenged	41%	44%
Proportion where the poverty headcount was re-stated	19%	21%

However, a policy analysis problem is not generated by the fact of a re-statement taking place – the crucial issue is whether a consistent account of poverty, its trajectory, and the ability to identify the dynamics can be maintained. To examine these issues, we reviewed the affected seven countries in some detail, to clarify the size and direction of the restatements relative to the previously published figures, the given rationale, and the accompanying level of transparency. The results are reported in Table 3.

The seven countries are drawn from four continents, and six had adopted a PRSP. Column 3 reports the effect of the changes on the reported trend in poverty. It is possible to draw a distinction between two groups: two cases where minor changes were made to reflect revisions in measurement and estimation methods (Mexico and the Kyrgyz Republic), and a larger group of five where the changes were large and less well supported. In all of the latter cases, the changes were very substantial, and in three, large enough to reverse the implied trend of poverty. Moreover, this is linked in each case with PRSP-adoption, i.e. the revision serves to vindicate the introduction of new (and IFI-sponsored) policies. In only the single case of Malawi is there a substantial change in the opposite direction (World Bank 2004, 2005 2006a, 2006b, and 2007; World Bank and the Republic of Malawi 2006).

These outcomes follow from both the direction and size of the restatements relative to the most recent LSMS estimates; whatever its present level, falling poverty can be inferred if

previous poverty levels were sufficiently understated. The case of Mongolia is instructive: the 1998 figure of 35.6 percent of the population reported as poor was retrospectively revised (in 2006) to 43.1 percent, an increase of over 21 percent on the base. This implies that the most recently reported figure (for 2002) of 36.1 percent represents a substantial fall, rather than the previously reported small increase. Mongolia, like the others with re-statements that reversed the previously reported poverty trajectory, adopted a PRSP during the period in question. It is important to note that Mongolia is not an outlier – the impact on others in this group is of a similar magnitude and significance (World Bank, 2006b).

Column 6 of Table 3 provides a brief description of the degree of disclosure supporting the restatements made by the World Bank. This is in terms of both the method used to re-estimate the previous poverty aggregates and the degree of access to the data to which the new methods were applied. Clearly, replication and testing of the results would require both. Again, the results are disturbing.

Only in the cases where minor changes were made (the Kyrgyz Republic and Mexico) was there substantial disclosure, and only in the case of Mexico (the one non-PRSP country) were the base data and the methods fully disclosed. However, in these cases the revisions were of insufficient size to have any real policy or scrutiny implications (World Bank 2004 and 2007).

In contrast, for the majority of countries with very large revisions, there is strikingly little disclosure of either method or data. For example, in the case of Cambodia, although the 1994 study is restated (relative to 2004), the 1997 LSMS results are not even referred to. Drawing inferences about the trend of poverty over the whole period would, therefore, seem to be difficult. Yet the Cambodian PA confidently asserts ‘that poverty has fallen since the beginning of the transition‘, and strongly associates this trend with PRSP adoption (World Bank, 2006a). In spite of the inevitably contested nature of the revised data, the Bank’s confidence is also evident in the other cases. The most controversial is that of Burkina Faso, where the World Bank asserts that it has resolved the ’Burkinabe paradox‘ of stable growth alongside stagnant poverty (World Bank, 2005). The Mongolian PA is as bold in its claims, listing the substantial decline in the headcount since 1998 as an ’established fact‘ (World Bank, 2006b, p. iv).

Finally, Table 3 provides, in column 7, a brief summary of the rationale provided in the PAs for the restatement exercises. The amount of space and depth of discussion vary, yet it is apparent that in those cases where major revisions have been made, the examination of the issues, again, is more limited. Moreover, in only three of the seven is incontrovertible evidence of a fundamental break in estimation and/or data collection methods provided (Armenia, Kyrgyz Republic and Mexico). Once again, these are those with only minor revisions to the previously published data.

Before moving on to more detailed discussion of the Mongolian experience, it is worth reflecting on the questions raised at the beginning of this section, which we argued governed the policy impact of *ex post* revisions to the established poverty record. First,

Table 3: Case summaries of countries where poverty restatement exercises were undertaken by World Bank PAs (2004–2007)

Country	Year restatement published	Effect on trend in poverty (new change in poverty versus old)	Years (periods affected)	Direction of change and new headcount (base and % change)	Level of disclosure	Rationale (evidence of a fundamental discontinuity of approach)
Armenia	2004 on website (2002 on document)	Major impact: replaces an increase with a reduction in poverty levels (level of change not calculable).	1994 (1994 to 1996)	Increases 1994 headcount to 54.7% (from 31% rural; 25% urban – change varies).	Substantial explanation is given: text provides annex detailing approach and summary data. Mechanics of adjustments are less clear however.	Movement (in 1996) to an absolute poverty line. Also lists concerns regarding inflation treatment. A fundamental break in series therefore.
Burkina Faso	2005	Major impact: replaces an increase with a reduction in poverty levels (-15% versus +2%).	1998 (1998 to 2003)	Increases to 1998 headcount to 54.3% (from 45.3% – a (change of +20.5%).	Long on justification, short on explanation. Details of method given, but little data provided (even fails to disclose the value of the former headcount).	Changes in survey design, including sophistication, but no fundamental discontinuity apparent. Aims to address the so-called 'Burkinabe paradox' (stability and growth versus high static poverty levels), and subsequently shows sizeable reductions are taking place.
Cambodia	2006	Major impact: increases reduction in poverty levels (-34% versus -10%).	1994 (1994 to 2004)	Increases 1994 headcount to 47% (from 39% – a change of + 20.5%).	Weak explanation. Little data is provided. The original 1994 headcount is not given, and thus direct comparisons are not made. The 1997 value is not re-stated, and this is not commented on. Also acknowledges some sensitivity of results, depending on treatment of excluded areas. This could cut the 1994 value to 41% (versus their preferred 47%).	The PA argues that major changes in coverage and size of the basket make series non-comparable. But the case is not clear-cut; a restricted coverage estimate is not presented. Only restates 1994 value, ignores the 1997 survey (as was done on 1994 basis). Perplexingly, the text notes that the re-statement was only possible after the 1994 sample frame being 'found' (had previously been misplaced).
Malawi	2006 (draft, not yet published on website)	Major impact: reduces the reduction in poverty levels (-4% versus -20%).	1998 (1998 to 2005)	Reduces 1998 headcount to 54.1% (from 65.3% – a change of -14.4%).	Strikingly little discussion included: no reference made within main text. Limited details noted within one text box. Explains an annex will be provided – but none is given – also refers readers to the NSO website, yet this too provides little detailed explanation.	Fundamental non-comparability claimed, due to change in recall periods used. The former estimate is also correlate-based, not given by re-analysis/ re-working of the data. An absolute break in comparability seems difficult to sustain. Also not clear why re-working of data was not at least attempted.
Kyrgyz Republic	2005	Minor impact: due to re-basing of survey series.	2003 (2003 to 2005)	Increases 2003 headcount to 49.9% (from 49.2% – a change of +1.4%).	Thorough disclosure of method and limited supporting data. But this is largely unnecessary as there is no controversy. Approach involved running two survey methodologies concurrently. The change is in any case small.	Approach involved major changes to survey design and data collection (conceptually and widening of basket). Approach was to run both methods concurrently and publish two results – clearly a best-practice approach.
Mexico	2004	Minor impact: restatement of a national re-estimation.	2002 (2000 to 2002)	Reduces 2002 headcount to 51.3% (from an alternative restatement of 51.5% – a change of 0.4%.	Considerable disclosure – the PA is effectively a re-statement of a prior re-statement by the national authorities. Given the minor impact on levels, the change seems of marginal benefit.	Bank staff re-estimate poverty based on a new methodology – effectively auditing the approach of national authorities.
Kyrgyz Republic	2005	Minor impact: due to re-basing of survey series.	2003 (2003 to 2005)	Increases 2003 headcount to 49.9% (from 49.2% – a change of +1.4%).	Thorough disclosure of method and limited supporting data. But this is largely unnecessary as there is no controversy. Approach involved running two survey methodologies concurrently. The change is in any case small.	Approach involved major changes to survey design and data collection (conceptually and widening of basket). Approach was to run both methods concurrently and publish two results – clearly a best-practice approach.

we can safely conclude that the restatement of LSMS data is far from uncommon and has often been substantial enough to invalidate entirely past policy inferences. Indeed, in five of the seven cases where the World Bank's PAs made changes, the longitudinal poverty record was effectively rewritten several years after the initial surveys. In addition, and contrary to expectations, it is apparent that the level of disclosure varies negatively with the magnitude of the revisions made. Thus, we can also conclude that restatements generally render the task of analysing and tracking the dynamics of poverty, and crucially that of gauging the pro-poor impact of policy choices, extremely difficult. The implications for the quality of non-governmental scrutiny are likely to be still more severe.

The fourth question posed regarding the rationale through which restatement exercises are undertaken is impossible to answer with any certainty. Nevertheless, it does seem that PAs, which operate as a form of quality check cum-audit, are highly selective, with revisions concentrating on those countries with poverty histories which fail to match the outcomes expected by the World Bank. It seems reasonable to hypothesise that only those cases in which expectations are disappointed are subject to further scrutiny and, if merited, re-evaluation. In turn, this suggests that records that match expectations escape detailed examination, potentially masking the true comparability problems of LSMS data.

3: A case study: Making sense of poverty data in Mongolia

This and the following section examine the case of Mongolia, one of the seven countries for which re-statements of the key poverty aggregates were made in the World Bank PAs reported in Table 3. Mongolia was chosen as a case study for two reasons. First, the country has three high quality LSMSs, carried out in 1995, 1998 and 2002, under the sole or joint supervision of the World Bank and the United Nations Development Programme (UNDP).[3] In principle, these surveys offer an authoritative record of the evolution of poverty over that period. Second, Mongolia is a low-income country, with one of Asia's highest poverty rates, which has also pursued a very distinct shock-therapy transition strategy. It therefore provides an ideal vehicle for examining how changes in the evidence on which policymaking is based influences subsequent policy choices. In this section, we investigate the published poverty record and reveal serious inconsistencies within and between the surveys. Section 4 considers how effectively non-governmental, civil society organisations might be able to engage in policy debate on the basis of the data which is in the public domain.

Table 4 reports the results of the three surveys, with the percentage difference from the previous survey shown in brackets. It can be seen from column 1 that there was virtually

[3] The 1995 survey was supported by the World Bank, the 1998 by the UNDP, and the 2002 by the two organisations jointly.

no change in the headcount figure between 1995 and 2002, with all of the surveys giving a result close to 36 percent of the population below the poverty line. It is worth noting that, as the population increased significantly over this period, the total number of households in poverty, in fact, grew substantially.

Table 4: Results of Mongolian LSMSs 1995-2002

Survey year measure	Poverty headcount	Poverty gap	Severity (squared poverty gap)	Gini coefficient
1995	36.3	10.9	4.8	0.31
1998	35.6 [-1.9%]	11.7 [+7.3]	5.6 [+16.7]	0.35 [+12.9%]
2002	36.1 [+1.4%]	11.0 [- -6.0%]	4.7 [- -16.1%]	0.33 [-5.7%]

Note: Change on previous survey in parenthesis – positive change (+) represents deterioration.
Sources: LSMS 1995, 1998 and 2002 and authors' calculations.

The evolution of the other dimensions of poverty is more complex. The poverty gap and severity of poverty measures all show significant increases between the 1995 and 1998 surveys, followed by falls by 2002. The Gini coefficient (consumption) follows a similar pattern, although with a less substantial fall between the 1998 and 2002 surveys.

It was always acknowledged that there were differences in coverage and method between the surveys (Schelzig, 2001), but there appeared to be broad acceptance of the general result that there had been little change in the level of poverty, as measured by the headcount, over this period. However, in 2006 the World Bank published a radical revision of the poverty headcount data for 1998, which raised the headcount proportion substantially. This, in turn, suggested a large fall in poverty over the 1998-2002 period and, coincidentally, provided a vindication of World Bank policy prescriptions. The re-estimated figure for 1998 was 43.1 percent of the population below the poverty line, an increase of over 21 percent from the previously reported figure. In turn, this implied that by 2002 approximately 7 percent of the population had moved out of poverty, which represents a 16 percent fall in proportionate terms (World Bank, 2006) – an extremely large fall over a four-year period.

The reasons why the World Bank felt that re-estimation was necessary are given in 2006 Mongolia Poverty Assessment (World Bank, p. 4, 2006b). In brief, they note: that the 1998 survey was not sufficiently representative in terms of geographical coverage; that consumption, both in terms of items and the recall period, was poorly mapped; and that these problems were exacerbated by the application of inconsistent poverty lines. However, the precise calculations on which the changes were made were not published and the dataset has not been released. Little is disclosed other than the final headcount and the proportionate changes in mean consumption levels, together with limited

disaggregated data. The Assessment provides no more than six pages of discussion (pp.6-12) of the restatement procedure, within a document of over 100 pages. Neither the re-pricing basis used, nor the summary distributions, nor the new nominal poverty line for 1998, are reported. Only basic outline details are given of the aggregation calculations and no disclosure is made as to whether the threshold was fully re-estimated (World Bank, 2006b). Reaching an evidenced-based judgment on the reliability of World Bank's re-estimation procedure has, therefore, proved extremely difficult. [4]

In the absence of such information, it is impossible to replicate the results precisely, or to judge their plausibility. Yet, it is clear that the re-estimation procedure is not transparent, and that this makes impossible any authoritative questioning of the validity of the restated poverty aggregates. In addition, acceptance of such results has a strong policy message: if the period saw such a significant fall in recorded poverty, the policy package arguably had been successful. However, without an ability to replicate and, potentially, disagree with these results, local ownership of such policies is hardly possible. This stands in ironic contrast to the comment within the principal recommendations that: 'data should be made available to the public to encourage open dialogue and constructive policy debate on poverty issues' (World Bank, 2006b, p.xiii).

In the absence of data, a set of alternative, admittedly second-best, options have been followed, in order to investigate the validity of the reported changes. First, the plausibility of the reasoning advanced for the changes to the 1998 figures is examined. Second, the consistency of the implied changes in poverty over the subsequent period with other, generally accepted data is considered. Finally, an approximate re-construction of the results based on published data using the World Bank's Povcal software is attempted.[5] However, it is worth emphasising that such a major restatement is in itself a serious failure, and suggests there are major survey design and management problems. Indeed, an explicit objective of the survey series was the establishment of a sound longitudinal record (Schelzig, 2001).

At the heart of the Bank's case is the assertion that the preferred and more accurate 2002 poverty threshold is not consistent with the former surveys. A higher nominal (and real) threshold is specified for 2002. When these are projected backwards to 1998, a larger proportion of the population falls below the threshold; poverty is judged to have been higher in 1998 and, by comparison with the 2002 figures, to have fallen sharply.

[4] A direct approach to the lead authors has, as yet, not led to any answers to these questions.
[5] This routine fits Lorenz Curves to distributional data and after specification of key parameters (poverty lines and means), predicts the standard poverty aggregates. See Chen, Datt and Ravallion (1991) for a discussion of Povcal and a user guide; and Datt and Ravallion (1992), for an application.

The magnitude of the increase in the poverty threshold can be inferred from a comparison of the re-based values of the three poverty lines. The value of the poverty line in real terms should remain constant over time. Table 5, column 1 lists nominal poverty lines reported in the three surveys. Column 2 rebases the 1998 poverty line to 1995 and 2002, respectively, using the Mongolian consumer price index (CPI).[6] It is immediately clear that the 1995 and 1998 nominal figures are consistent with a constant real poverty line, assuming the applicability of the CPI. By contrast, the nominal poverty line specified in the 2002 survey is 30 percent higher than that implied by re-basing the 1998 line.

There are, of course, several conceptual and practical reasons why nominal poverty lines might not map onto each other (Lanjouw and Lanjouw, 2001). Major reasons include both the method of compilation and the appropriateness of using a price index reflecting an average consumption bundle rather than that of the poor.

Table 5: Comparison of real and nominal poverty thresholds

Year	Nominal poverty line	Value of 1998 line re-based to other periods	Difference with the 1998 line
1995	7,240	14,624	-0.3%
1998	14,674	14,674	–
2002	24,674	19,082	+30.0%

Note: All in Mongolian Togrogs.
Sources: LSMS surveys 1995 (reported in World Bank, 1996), 1998 and 2002 and authors' calculations.

Moreover, the quality of Mongolian inflation data is itself the subject of controversy. World Bank (2006b) offers four different reasons in support of adopting the higher poverty threshold.

- First, that the 1998 survey made use of several regional poverty lines, as opposed to a single national line in 2002;
- Second, that the size of the consumption basket on which the threshold was based and the survey was taken, expanded dramatically between the two periods;
- Third, that the recording period in 2002 was longer and adopted a different approach;

[6] There is a discrepancy between the CPI as reported in the Mongolian Statistical Handbook and that reported in the IMF International Financial Statistics. We have used the mid-point of the two in the calculations reported.

- Fourth, that the coverage varied, with the 1998 surveys sampling nine representative aimags (Mongolian provinces), including Ulan Bator, while all 22 aimags and Ulan Bator were covered in 2002.

While each of these represents real differences between the surveys, the size and direction of the adjustment that they imply to maintain comparability are far from straightforward.

The first contention, about the consequences of moving to a single national line, is perhaps the weakest. Different prices in different regions can be accommodated by having different poverty thresholds (the 1998 procedure) or different regional price indices. In principle, both methods should generate similar outcomes. For this to justify a higher national poverty line represents a strong claim about the method for adjusting for regional price differences. However, the procedure by which this was done, and the supporting data, are not provided.

The second is a stronger objection, as changes in the consumption basket are a well-known source of measurement error, biasing both the nominal poverty threshold and reported consumption upwards. However, the literature suggests the net effect on poverty is not clearcut and may run counter to that suggested by the Bank, since the impact on mean consumption is generally higher than that on the poverty line (Lanjouw and Lanjouw, 2001). The recording period and sample coverage are also significant issues. Yet their impact on poverty aggregates will vary according to, respectively, the nature and balance of the areas included (in terms of their relative consumption levels), and changes in the length of the recall period, the method and the time of year the survey was undertaken. The longer recall period would tend to decrease recorded consumption and therefore increase poverty (Deaton, 2004). By contrast, the larger number of items in the basket might raise recorded consumption. Once again, the net impact on the poor is unclear and likely to be complex. Finally, the assumption of the 2006 assessment that the nine *aimags* that overlap the two surveys can be projected onto the whole country when re-estimating the 1998 figures is certainly questionable, even if the data are insufficient to allow the degree of possible bias to be investigated. It is notable that the Bank discloses that mean consumption for 2002, calculated on the basis of the 1998 truncated consumption set, would yield a figure approximately 15 per cent lower than the 2002 recorded value. In other words, that the net effect of the changes in survey method on average would raise reported consumption by approximately 15 percent.

Although there are undoubtedly differences in coverage, it is also the case that the conceptual bases of the surveys are identical. These include the use of a two-part (food and non-food) consumption bundle to define the threshold, a Food Equivalent Intake approach, and calculation by equivalent individuals derived from household-based recording. On their own, the differences between the two surveys do not *a priori* imply

that the 1998 figure is understated and, therefore, that poverty has fallen dramatically between the two surveys.

However, the chief objection to accepting the re-statement is that the implied trajectory of poverty over the subsequent four years has a poor fit with other economic and social data. The 2006 assessment reports that economic growth rose by a cumulative 10 percent in real terms (World Bank, 2006b), consumption increased by a cumulative 30 percent on average, and inequality dramatically declined from a Gini coefficient of 0.38 to 0.31. The argument is, therefore, that a combination of robust economic growth, accompanied by rapidly declining inequality, generated the sharp fall in poverty; in short, over the 1998-2002 period, economic growth was, at least in the weak sense, pro-poor.

Table 6 reports a range of statistics for the growth of GDP, GNP and GDP and GNP per capita. These figures are difficult to reconcile with World Bank 2006. GDP per capita from end 1997 to end 2002 grew by only a cumulative 6 percent, according to the published record. In fact, GDP per capita fell for two consecutive years in 2000 and 2001. GNP per capita provides a more encouraging story, yet even here GNP per capita fell in 2000 and only grew substantially in the final year, during which the survey was taking place.

Table 6: Growth rates of income and income per capita

	GDP growth rate	GDP/capita growth rate	GNP growth rate	GNP/capita growth rate
1998	3.5	2.2	4.7	3.3
1999	3.2	2.0	2.8	1.5
2000	1.1	-0.4	0.9	-0.5
2001	1.0	-0.4	4.5	3.0
2002	4.0	2.6	6.5	5.1

Sources: IMF World Economic Outlook 2008; Asian Development Bank country database.

Any large increase in consumption over this period, given the limited growth of overall GDP growth, would imply a substantial change in the structure of demand. In fact, private consumption expenditure rose very sharply as a proportion of GDP: from 60.5 percent at the end of 1997 to 77.2 percent at the end of 2002. This largely reflects a dramatic reversal in the balance of trade, with net exports falling from a positive 6 per cent of GDP at the end of 1997 to minus 21.3 per cent at the end of 2002. However, the rate of growth of consumption per head based on the national accounts is known to be a poor indicator of the behaviour of the survey-based estimates of mean consumption (Deaton, 2004). In fact, mean consumption estimated from the national accounts typically grows substantially more quickly than survey-based estimates, so that even the increases in consumption per head recorded in the national accounts cannot easily be

reconciled with the survey-based claims for such a substantial increase in mean consumption claimed in the restatement. However, in addition, the commodity structure of imports during this period shows little evidence of increases in basic consumption items which would enter the shopping baskets of the poor; for example, food items do not dramatically increase their weight (Mongolian NSO, 1996, 1998, 29003, 2005).

Table 7: Growth of output by sector: Annual percentage change

	1997	1998	1999	2000	2001	2002
Agriculture	4.3	6.6	4.4	-15.9	-18.3	-12.4
Industry	-2.6	3.7	1.6	0.3	15.5	3.8
Services	7.5	1.1	3.1	15.3	6.1	11.6

Source: Asian Development Bank, Country database.

The sectoral breakdown is also revealing and is shown in Table 7. These figures reflect clearly the impact of the *Dzud,*[7] or extreme winter conditions, which killed over 11 million animals in the winters of 2000 and 2001. As a large proportion of the poorest Mongolians are nomadic pastoralists, and agriculture provided between 49 and 45 per

cent of the employed labour force during this period, these figures also seem extremely difficult to reconcile with reported dramatic increases in the consumption baskets of the poor. In addition, the claim of a strikingly large reduction in income inequality is hard to reconcile with both the aftermath of the *Dzud* and the impact of a range of privatisation measures that fed into large increases in the inequality of wealth, for example, the privatisation of residential property at no cost to sitting tenants (Nixson and Walters, 2006). In fact, World Bank (2006, p.8) notes '[t]he finding of reduced inequality is contrary to the general perception…'.

A more precise method of reconciling the macroeconomic data with the behaviour of poverty and inequality would be to use the World Bank's Povstat[8] software to attempt formally to reconcile the figures. Indeed, World Bank (2006b) claims that the application of the routine supports their re-statement exercise. However, in order to apply this program a range of distributional data are required, which are not reported in either the 2006 re-statement or in the reports of the Mongolian National Statistical Office (except for 1998). An alternative would be to construct the Lorenz curves synthetically, but this would offer no analytical advantages.

As an alternative, we apply the Bank's Povcal program to the data that are in the public domain. The program calculates, on the basis of decile or quintile distributional data, the

[7] A *Dzud* is the Mongolian word for an exceptionally harsh winter.
[8] See http://povlibrary.worldbank.org/library/view/12934 (accessed 12 March 2009).

implied poverty levels once the poverty threshold and mean income are specified (Datt et al., 2003). In principle, this allows the approximate reconstruction of the re-statement exercise.

However, an immediate and major problem emerges when trying to identify accurate published distributional data. The two major international databases of distributional data are Povcalnet, supported by the World Bank, and the World Income Inequality Database version 2.0c, May 2008 (WIID), supported by the UN WIDER. Both of these are purportedly based on the same sources (the original surveys), yet they provided dramatically different distributional data. The alternative sets of data are shown in Table 8 below.

There are two related problems. First, and foremost, the data are inconsistent with each other, not merely for the disputed 1998 survey but for each of the reported surveys. In addition, the level of aggregation is different in the two sources: for 1998 and 2002, WIID only reports quintile data. However, these quintile data are still inconsistent with the decile data reported in Povcalnet. The differences are most marked in relation to the 1998 survey, where the two sources report radically different distributions. These are illustrated in the Lorenz curves given in Figure 1 below. The estimated Lorenz curve based on the WIID database shows significantly higher levels of inequality than that based on Povcalnet. As they are based on the same surveys one, or both, have, presumably, been adjusted. However, there is no information provided in either database about whether or how any adjustments have been made, or their extent or direction.

Figure 1: Lorenz Curves for 1998 sourced from Povcalnet (PCN) and the World Income Inequality Database

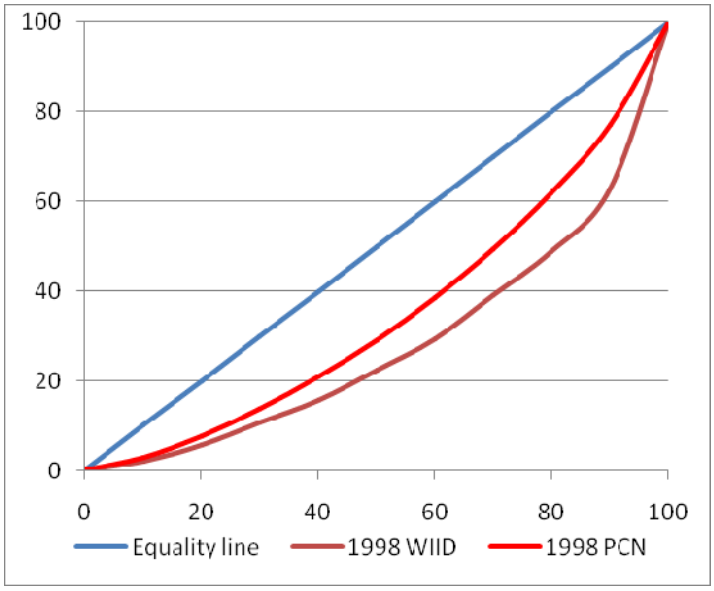


Table 8: WIID (W) and Povcalnet (P) reported distribution data for LSMS surveys

	Q1	Q2	Q3	Q4	Q5	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
1995W						3.06	4.31	5.45	6.58	7.74	9.01	10.54	12.55	15.80	24.96
1995P						2.98	4.29	5.42	6.55	7.79	9.15	10.73	12.71	15.57	24.81
1998W	5.60	10.00	13.80	19.40	51.20	2.10									37.00
1998P						3.04	4.68	5.99	7.14	8.26	9.44	10.78	12.52	15.23	22.92
2002W	7.47	12.16	16.79	23.10	40.48	2.99									24.64
2002P						3.02	4.36	5.49	6.61	7.83	9.16	10.72	12.66	15.49	24.66

Sources: WIID database at http://www.wider.unu.edu/research/Database/en_GB/database (accessed 12 March 2009), and Povcalnet at http://iresearch.worldbank.org/PovcalNet/povcalNet.html (accessed 12 March 2009).

A further inconsistency is evident when the distributional data are used to generate Gini coefficients. On the basis of the reported distributions, Povcal fits Lorenz curves and implied Gini coefficients. Table 9 shows the Gini coefficients listed on the databases; those calculated on the distributional data published by both WIID and Povcalnet; and those published in the LSMS documents.

Table 9: Different Gini coefficients based on the same survey data

	1995	1998	2002
Gini WIID listed	33.2	44.0	32.8
Gini WIID calculated	33.19(GQ);33.31(β)	44.81(GQ);44.42(β)	
Gini Povcalnet listed	33.2	30.27	32.84
Gini Povcalnet calculated	33.14(GQ);33.25(β)	30.27(GQ);30.37(β)	32.75(GQ);32.85(β)
Survey reported Gini	31.00	35.00	32.9

Note: GQ is the Generalised quadratic form of the Lorenz curve; β is the Beta form of the Lorenz curve.

It is evident that there are major inconsistencies, both between the two sources and between these sources and the originally published reports. The most egregious differences are in respect of the WIID 1998 data, where the reported and calculated Gini coefficients are strikingly different from either the Povcalnet database or the original survey reports. For this reason, in the subsequent simulations we discarded the WIID data and used the Povcalnet data instead.

However, further inconsistencies emerge when Povcal is used to simulate the poverty headcount proportions, given the published Povcalnet distributional data augmented with the reported poverty thresholds. The simulated poverty headcount figures generated are, in two out of the three cases, inconsistent with the alternatives reported in LSMS documentation. In addition, however, estimates generated from the (sparse) aggregate distributional data reported in summary tables within the LSMS reports are also inconsistent with both the Povcal estimates and the Gini coefficients and headcount proportions reported in the LSMS reports. The comparative results are shown in Table 10.

The first rows reports the Povcal estimates of the Gini coefficient, based on the different published distributional data, while the second row gives those reported in the LSMS document. For 1995 and 1998 surveys, the differences are extremely large. This means that the effective distributional data in the Povcalnet database and the LSMS data must be different. In the case of 1998, in particular, these differences are extremely large. Row 4 reports the headcount estimates generated by Povcal under the different distributional assumptions, but applying the same mean income and poverty threshold

figures. Once again, there are large differences; the difference in the estimated headcount is over 10 percent, with a 13.1 percent difference in the Gini. In fact, for 1998, the quintile data taken from the LSMS report provide an unambiguously better match, but even these are far from perfect – underpredicting the headcount by some 6.3 per cent. By contrast, there is a very close match between the Povcal estimates for 2002 based on the published deciles and the reported headcount and Gini coefficient.

Table 10: Povcal consistency testing results

Year	1995	1995	1998	1998	2002
Source:	PCN (World Bank)	LSMS Documentation	PCN (World Bank)	LSMS Documentation	LSMS
Type of data:	Deciles	Summary table		Summary table	Deciles
Povcal estimated Ginis	33.1	31.1	30.4	33.9	32.9
LSMS reported Gini	31	31	35	35	32.9
Difference	-2.1	-0.1	4.6	1.1	0
Povcal estimated headcounts	36.5	34	32	35.9	36
LSMS reported headcount	36.3	36.3	35.6	35.6	36.1
Differences	0.2	2.3	3.6	0.3	0.1

These outcomes are very far from satisfactory, raise further doubts about the reliability of the published distributional data, and question the integrity of the Povcalnet (and WIID) databases. Despite this, applying Povcal does allow a plausible reconstruction of the procedures underlying the 2006 re-statement. If it is accepted that there are two separate poverty lines, for 1998 and a higher one for 2002, and it is further accepted that the consumption mean has risen by 15 percent, then using Povcal applied to the Povcalnet published data allows two simulation experiments to be conducted: first, re-estimating the 1998 figures based on a backward-adjusted poverty line from 2002; and second, re-estimating the 2002 figures using a forward adjusted 1998 poverty line. The

results of these simulations and comparisons against the Bank's re-stated headcount are provided in Table 11 below.

Table 11: Simulated consistent headcounts for 1998 and 2002

Survey year	Mean (or estimated mean)	Poverty line (or re-priced line)	Published & estimated headcounts	World Bank re-stated headcount
	Revised Scenario 1: Using the higher (2002 basis) poverty line (all at 2002 prices):			
1998	32495	24743	43.3	43.1
2002	36747	24743	36.1 [-16.6]	36.1 [-16.2%]
	Revised Scenario 2: Using the lower (1998 basis) poverty line (all at 1998 prices):			
1998	21795	14674	35.6	n/a
2002	24647	14674	29.5 [-18.2%]	n/a

Note: Figures in brackets show the percentage change from the 1998 survey result.

These results show a strong correspondence with the re-statement exercise. The scenario based on their preferred higher poverty line (using the 2002 methodology) shows a strikingly similar result – a reduction of 16.6 percent versus the reported figure of 16.2 percent. The alternative scenario, using the lower line, suggests a still higher reduction of 18.2 percent.

It seems clear that the re-statement results are driven by the change in the poverty threshold and the claimed increase in the mean consumption basket, both of which follow from the changes in survey methods, the details of which remain opaque. However, it is worth noting that the modelling approach exercise is exceptionally crude. For want of access to data, Povcal has been parameterised by quintile data that are of questionable quality. In addition, the 15 percent increase in the consumption basket is not likely to have been uniform in both distributions, nor is this effect across each distribution guaranteed to be mapped by the shape of the two fitted Lorenz Curves.

In addition to these caveats, the major problem with accepting the legitimacy of the exercise is that these results remain at odds with external data and the events between 1998 and 2002. It must be underlined that this period saw only a 6 percent increase in per capita incomes, and included the *Dzuds* of 1999 and 2000, which decimated the livestock sector and in turn the incomes of the poor. Admittedly, consumption based on the national accounts grew at a substantial rate, but this is unlikely to have benefited the

poor, who have little or no opportunity to dis-save or build up debts and, in any case, as noted above, national accounts consumption typically grows significantly more quickly than survey-based measures (Deaton, 2004). Like the Bank's initial estimates, this spectacular level of poverty reduction still requires much explanation.

4. Understanding Mongolia's poverty reduction story

In this section, we place ourselves in the position of national actors – be they policymakers or external analysts – in attempting to track the dynamics of poverty in the face of the evidential difficulties identified above. Our approach uses standard analytical techniques, to gain additional insights into the evolution of poverty in Mongolia between 1995 and 1998, and between 1998 and 2002. Following Datt and Ravallion (1992) we seek to determine the relative contributions of growth and distributional change, and therefore the extent to which growth has been 'pro-poor' in each of these periods. This section seeks, therefore, to demonstrate the consequences of incomplete and non-transparent poverty data for the effective monitoring and evaluation of policy choices.

The techniques we use form a standard element in the analytical toolkit made available to PRSP countries (Datt et al., 2003) and include a poverty-growth-inequality decomposition and the plotting of Growth Incidence Curves (GICs). These techniques form an important part of the evidential base for economic and social policymaking in LDCs, and in the following we show that these devices are severely compromised where data are unreliable and non-comparable.

The comparable poverty record which we re-established in the previous section provides the starting point for the decomposition exercise. Our purpose is to attribute changes in poverty to growth versus changes in distribution between two survey dates. The formal approach we adopt is that given by Kakwani (1997) as applied in McCulloch and Baulch (1999).

We achieve this, again, by using Povcal to simulate different outcomes, by holding inequality constant while allowing growth in mean consumption to change, and then vice versa. We then difference the predicted values from the actual outcomes to calculate the partial contributions to poverty change. A core assumption, therefore, is that poverty lines remain constant in real terms.

Within our modelling, we adopt a baseline position consistent with the published 1995 and 1998 results, and our reworking of the 2002 data (based on the lower 1998-basis poverty line). Due to the poor fit of the distributional data described above, we also fix the start and end points to the published poverty proportions. The results are given in Table 12 below:

Table 12: Poverty decomposition results

Period	1995-1998 Change in headcount/ [% change on base]	1998-2002 Change in headcount/ [% change on base]
Growth effect	-1.9 [-5.2%]	-6.7 [-18.8%]
Inequality effect	+1.2 [+3.3%]	+0.2 [+0.6]
Total change in headcount	- 0.7 [-1.9%]	-6.5 [-18.2%]

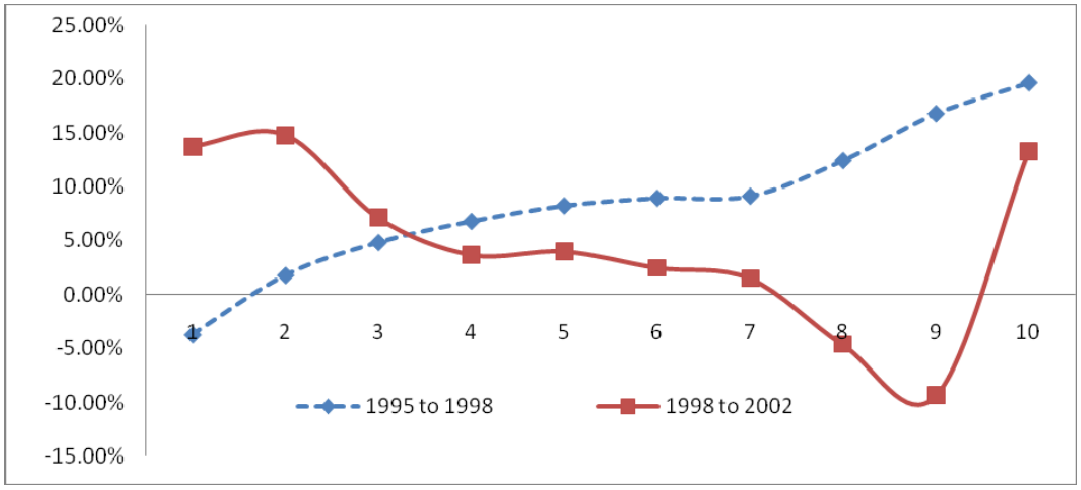
These results show strong poverty reduction, but also show very different dynamics were at work in the two periods. The poverty impact of the relatively weak levels of growth between 1995 and 1998 is offset by rising inequality, yet between 1998 and 2002 growth (in consumption) is seen to pass through directly to very substantial reductions in the headcount index. Thus, the turnaround in the second period is remarkable in two senses: growth in consumption appears to increase very dramatically, while simultaneously, it becomes more 'pro-poor'.

This pattern of strong improvement after 1998 is by and large replicated in the other analyses we undertook. The pro-poor indices (Marshall et al., 2008) show increases in the responsiveness of poverty levels to growth (given by the poverty elasticity of growth), and a striking improvement in the poverty impact of actual against distributionally-neutral growth. Although rather crude (due to the fact that we are forced to use decile-based data) the two Growth Incidence Curves, which depict changes in consumption across the distribution (provided below in Figure 2), also exhibit dramatically different poverty reduction experiences. In the first period, consumption growth favours the centre and the rich, with real terms reductions among the poorest groups, whereas in the second, the pattern of gains is reversed, with the eighth and ninth deciles seeing considerable falls in consumption.

Although the patterns of change depicted in the statistics and GICs are not entirely contradictory, the level of change does raise further questions about the credibility of the data and particularly the re-stated 1998 results. Indeed, the distributional patterns, given by the GICs, are near mirror images.

More significantly, the pattern of gains and losses depicted for the second period is unprecedented for transitional economies and runs counter to established theoretical perspectives. This period shows a very strong rebound in the poorest deciles and very

Figure 2: Estimated decile-based Growth Incidence Curves



weak growth in the centre of the distribution, during a period which saw ongoing economic liberalisation leading to rising inequality in asset distribution (Nixson and Walters, 2006) and further rationalisation of the state provision and welfare safety nets.

Furthermore, and as noted, the pattern contradicts other non-policy driven events within Mongolia – specifically, the social and economic dislocation associated with the *Dzuds* of 1999 and 2000. Other studies have shown that these disasters had a considerable impact on consumption levels and their incidence was greatest in the rural heartland, which, according to the 2002 Poverty Assessment, is where the poor are over-represented.

It also needs to be borne in mind that our analyses are based (again in the absence of fuller disclosure) on a rather crude replication method. Two particular issues that would alter the decomposition, and hence the foregoing results, are: how we modelled the distributions; and second, the re-pricing bases used. To attempt to get an estimate of the impact of these issues we undertook sensitivity analysis: the use of different base specifications (varying the Lorenz curve formulations); changing the spacing of the distributional data (using quintile versus decile-based sources); and making different re-pricing assumptions.

The sensitivity analyses were carried out alongside our initial tests of the credibility of the reported data for each of the surveys through the iterated use of Povcal. The first test, that of examining the impact of using alternative Lorenz curve specifications, the generalised quadratic (GQ) versus the Beta specification, had only a minimal impact on the predicted headcounts, and therefore no subsequent impact on our results. However, our second test, that of using finer distributional data for 1995 (deciles in place of

quintiles) reveals a much stronger variation.[9] As Table 7 shows, the change in simulated outcomes in this first period, although constrained by the low overall change, is substantial in terms of the balance between growth and distributional effects. The decile data attributes all of the change to the growth effect. It is worth noting that our choice of quintile data, due to better fit with the published results, was marginal.

Table 13: Impact of using quintile versus deciles data for 1998 on the poverty decomposition

Poverty change 1995 to 1998:	Selected (quintile-based) 95-98	Decile alternative
Due to growth	-1.9	-0.6
Due to change in inequality	+1.2	-0.1
Total	-0.7	-0.7

Thirdly, we tested the sensitivity of our results to the use of different re-pricing bases. Two distinct issues were addressed here: the re-pricing poverty thresholds with an average CPI-type inflation metric (given that the effective change in prices faced by the poor is unlikely to be the same as that incurred by the mean consumer); and, second, the impact of a discontinuity between the CPI series published by the IMF and the Mongolian Statistical Agency.[10] Re-pricing assumptions are important to our method, and changes will, in principle, have a substantial effect within the decomposition exercise, since a higher discount factor will reduce the growth effect in both periods. More significantly, there would be a direct impact on projected poverty levels in the second period, as our re-working of the data to correct for the 1998-2002 discontinuity relies on re-pricing the two nominal poverty thresholds.

Unfortunately, our options for demonstrating the first substantive re-pricing issue are rather limited, as there is no published index of prices facing the poor. Yet it seems likely that, in the second period at least, the effective price increase facing the poor because of the *Dzud* will be more substantial, given the domestic supply constraints and reductions in the availability of in-kind consumption. Sadly, we were unable to estimate this accurately.

[9] Deciles are available for 1995 and 2002, but were not used for the former year, as quintile data provided a better fit with the published aggregates.
[10] In the modelling work we re-price using the midpoint of these series

We were, however, able to re-estimate the impact on the comparability adjustment and the poverty change decomposition, using the IMF and Mongolian CPI series. The IMF series runs dramatically ahead of the national index in the first period, and somewhat below it in the second. The estimated results are given in Table 14 below.

Table 14: Impact of alterative IMF/ Mongolian statistics re-pricing indices

Poverty change:	Based on combined index	Based on IMF index	Based on NSO index
(a) 1995 – 1998 period (fixed start point)			
Due to growth	-1.9	+0.7	-4.4
Due to change in inequality	+1.2	-1.4	+3.7
Total	-0.7	-0.7	-0.7
(b) 1998 – 2002 period (fixed start and end points)			
Due to growth	-6.7	-7.0	-6.3
Due to change in inequality	+0.2	+0.1	+0.2
Total	-6.5	-6.9	-6.1

Clearly, the 1995 to 1998 period sees the most significant impact on the balance between the growth and inequality effects (this results from the much larger discrepancy between the two series in this period), but the overall change is still constrained. The change in the latter period is also worth highlighting, because of the direct effect on the projected headcount, which is driven by the rescaled comparability adjustment. The effect of using the lower price growth given by the IMF series is to increase the overall level of poverty reduction, whereas the higher price growth in the national series moderates the fall.

In closing, two key themes stand out from this section, which has sought to apply basic analytical techniques to the restated poverty data for Mongolia, and hence gain insights into the challenges faced by national analysts. The first is that the flaws in the evidential base are inevitably replicated in subsequent analyses. Indeed, the 'insights' gained from the analyses reveal further contradictions (versus the published record), and cast particular doubt on the reliability of the re-restatement exercise. Second, we have shown that these follow-on analyses are severely constrained by the lack of disclosure, both generally and specifically in relation to the Bank's restatement. Our (and, by implication, other external analysts') attempts to model the dynamics at work are hugely sensitive to a plethora of assumptions and data choices. As a result, pro-poor analyses can be formulated to support very different accounts of Mongolia's poverty reduction record.

5. Conclusions

This paper seeks to contribute to the wider debate over the value and role of minimum welfare-calibrated poverty lines within the policy monitoring and evaluation process. In contrast to the established literature, which has tended to examine either the conceptual basis of calorific poverty thresholds or the technical difficulties of through time comparisons, we focus on the operational difficulties faced in low-income countries (LICs). In this context, where weak data and inadequate capacity are defining features of the policymaking environment, we argue that measurement failures, *ad hoc* changes in survey design, and inadequate disclosure are as important in limiting the ability of national governments to make sound policy responses. Furthermore, that these pressures have also hampered the exercise of effective domestic scrutiny.

The paper shows that, in spite of the best efforts of national authorities and their international sponsors (principally the World Bank), the gathering of authoritative longitudinal poverty data remains, for many LICs, a challenging and incomplete task. Moreover, in contrast to its laudable objectives, the LSMS initiative has not secured a scientific basis for the tracking of policy choices to poverty outcomes. An appraisal of World Bank Poverty Assessments undertaken since 2004 shows that substantive criticism and, in some cases, the re-estimation and restatement of the main poverty aggregates, are common if not commonplace. These exercises often involve radical revisions and are rarely fully transparent. Indeed, there appears, superficially at least, to be a negative correlation between the controversy generated by re-statements and the level of disclosure.

It is also troubling that major revisions undertaken by the World Bank tend to take place in those cases where the data record runs counter to orthodox expectations. While a degree of self-selection might be expected from an *ad hoc* process, the combination of this with limited and weak disclosure opens the Bank to charges of fitting the data record to desired policy outcomes. This is a major failing, but it also implies that there might be many more cases, in which the published outcomes were in line with expectation, where serious errors remain uncovered.

Our review of the Mongolian case brings home the impossibility of effective policymaking and the exercise of external scrutiny within an environment of weak data and an opaque external review process. The World Bank's *ex post* revision to the 1998 data radically re-writes Mongolia's poverty story, transforming a period of immiserating growth (Bhagwati, 1958) into one of spectacular improvement. That such a turnaround can be 'announced' to the authorities, without any real disclosure of method or data, not merely lacks credibility, but destroys the essence of national ownership, which, it is asserted, lies at the heart of the PRSP process.

We also sought to replicate and validate Mongolia's re-stated poverty statistics, but with little success. We found that primary and secondary data sources are often contradictory and offer inadequate detail to enable even the most basic examination. Furthermore, we show, through the application of a poverty-growth-inequality decomposition, that the 1998 revision, and its associated lack of transparency, render standard analytical techniques near redundant. In such circumstances, analysts charged with supporting policymakers would be forced (as we were) to make arbitrary assumptions and manipulate methods in order to generate plausible results.

Although these problems are exacerbated in Mongolia's case by overtly strict data security laws, our initial review of global experience shows that Mongolia is not alone in facing these challenges. In the years between 2004 and 2007 we identified five cases where very substantial revisions were made to the established record and where the level of disclosure was judged inadequate. It seems likely that poverty analysts and decision makers in these countries are genuinely 'flying blind'.

In closing, we reach two major conclusions. Firstly, while we consider that revisions are inevitable and that poverty measurement remains an evolving science, effective monitoring and evaluation require that consistent signals are gained from longitudinal data. The primary mechanism through which these tensions are likely to be resolved is the provision of adequate levels of transparency, requiring openness regarding methods and direct access to supporting data. At present, and in spite of rhetoric to the contrary, even the World Bank's own review process, as given by its Poverty Assessment series, fails to meet these basic standards. Ensuring transparency here would be a good place to start.

Second, we conclude that in the data environment of most PRSP countries, the use of complex surveys, and the use of cutting-edge estimation techniques, can give the impression of scientific certainty where none exists. The importance of judgment should not be denied in this arena, and we argue that much would be achieved by the adoption of a more eclectic approach to poverty monitoring, combining the measurement of income and consumption poverty with non-monetary measures of deprivation, such as health and educational indicators. An approach which makes use of poverty correlates, chiefly easily measured household characteristics, both in validating other data and in tracking the impact of policy, also has much to commend it.

As a final note, we do however acknowledge, that many of these difficulties are intrinsic to the difficult operational process of successive survey taking, and the inevitably politicised nature of poverty data. If the issues we identify are to be tackled effectively, an awareness of these pressures also needs to be acknowledged. This inevitably requires the management of expectations over the absolute accuracy that can be gained from surveys undertaken in the challenging environments of most LICs.

Bibliography

Bhagwati, J. (1958). 'Immiserizing growth: A geometrical note'. *Review of Economic Studies* 25 (June), 201-205.

Chen S., Datt, G. and Ravallion, M. (1991). *Povcal: A Program for Calculating Poverty Measures from Grouped Data*. Mimeo. Washington DC: Policy Research Department, World Bank. .

Datt, G., Ramadas, K., ,van der Mensbrugghe, D., Walker, T. and Wodon Q. (2003). 'Predicting the effects of aggregate growth on poverty'. In Bourguignon, F. and Pereria da Silva, L. A. (Eds.), *The Impact of Economic Policies on Poverty and Income Dstribution: Evaluation Techniques and Tools.* New York, USA: Oxford University Press.

Datt, G. and Ravallion (1992). 'Growth and redistribution components of changes in poverty measures: A decomposition with application to Brazil and India in the 1980s'. *Journal of Development Economics* 38, 275-295.

Deaton, A. (2004). *Measuring Poverty in a Growing World (or Measuring Growth in Poor World*. NBER Working Paper 9822 [Online resource available at www.nber.org/papers/w9822.pdf] Accessed: 5 March 2009.

Kakwani, N. (1997). 'On measuring growth and inequity components of changes in poverty, with application to Thailand'. *Discussion paper no. 97/ 17*. Sydney, Australia: University of New South Wales (UNSW), School of Economics.

IMF (2008), *IMF World Economic Outlook 2008*. Washington DC: IMF.

Lanjouw, J. O. and Lanjouw, P. (2001). 'How to compare apples and oranges: Poverty measurement when the definition of consumption is different'. *Review of Income and Wealth* 47(1), 25-42.

Marshall, R., Nixson, N. and Walters, B. (2008). 'Tracking poverty and economic policy outcomes in Mongolia'. *Brooks World Poverty Institute Working Paper* 27. [Online resource available at http://www.bwpi.manchester.ac.uk/resources/Working-Papers/bwpi-wp-2708.pdf] Accessed: 5 March 2009.

McCulloch, N. and Baulch, B. (1999). 'Assessing the poverty bias of growth Methodology and an application to Andhra Pradesh Uttar Pradesh'. *IDS Working Paper 98*. Brighton, UK: IDS

National Statistical Office of Mongolia (1996). *Living Standards Measurement Survey 1995*. Ulaanbaatar, Mongolia: NSO.

National Statistical Office of Mongolia (1999). *Living Standards Measurement Survey 1998*. Ulaanbaatar, Mongolia: NSO.

National Statistical Office of Mongolia (2003). *Statistical Year Book*. Ulaanbaatar, Mongolia: NSO.

National Statistical office of Mongolia (2005). *Household Income and Expenditure Survey/ Living Standards Measurement Survey*. Ulaanbaatar, Mongolia: NSO.

Nixson, F. and Walters, B. (2006). 'Privatization, income distribution, and poverty: The Mongolian experience'. *World Development* 34(9), 1557-1579.

Palmer-Jones, R. (2007). 'Poverty measurement in India and Bangladesh: A great Indian rope trick?' *IDPM/ Development Economics Seminar Series* (September), University of Manchester, [Online resource available at http://www.uea.ac.uk/~d280/] Accessed: September 2008.

Schelzig, K. (2001). 'Escaping poverty: Behind the numbers'. *Public Administration and Development* 21, 259-269.

World Bank (1996). *Mongolia Poverty Assessment 1996*. Washington, DC: World Bank.

World Bank (2004). *Poverty in Mexico: An Assessment of Conditions, Trends and Government Strategy*. Washington, DC: World Bank.

World Bank (2005). *Burkina Faso Poverty Assessment 2006: Reducing Poverty through Sustained Equitable Growth*. Washington, DC: World Bank.

World Bank (2006a). *Cambodia: Halving Poverty by 2015 – Poverty Assessment 2006*. Washington DC: World Bank.

World Bank (2006b). *Mongolia Poverty Assessment 2006*. Washington, DC: World Bank.

World Bank (2007). *Kyrgyz Republic: Growth, Employment and Poverty – Poverty Assessment 2007 Volume I*. Washington, DC: World Bank.

World Bank and Republic of Malawi (2006). *Poverty and Vulnerability Assessment: Investing in our Future*. Lilongwe, Malawi: World Bank/ ROM.

The University of Manchester
Brooks World Poverty Institute

MANCHESTER 1824

Executive Director
Professor Tony Addison

Research Director
Professor Michael Woolcock

Associate Director
Professor David Hulme

The Brooks World Poverty Institute (BWPI) creates and shares knowledge to help end global poverty.

BWPI is multidisciplinary, researching poverty in both the rich and poor worlds.

Our aim is to better understand why people are poor, what keeps them trapped in poverty and how they can be helped - drawing upon the very best international practice in research and policy making.

The Brooks World Poverty Institute is chaired by Nobel Laureate, Professor Joseph E. Stiglitz.

Contact:

Brooks World Poverty Institute
The University of Manchester
Humanities Bridgeford Street Building
Oxford Road
Manchester
M13 9PL
United Kingdom

Email: bwpi@manchester.ac.uk

www.manchester.ac.uk/bwpi

www.manchester.ac.uk/bwpi