



**What was the poverty headcount
in 2004 and how does it compare to
recent estimates by van der Berg et al?**

by
Charles Meth

About the Author

Charles Meth is a Senior Research Fellow in the School of Development Studies at the University of KwaZulu-Natal. He is also a Research Associate of the Southern African Labour and Development Research Unit at the University of Cape Town, and an Academic Visitor in the Department of Social Policy and Social Work at the University of Oxford.

Recommended citation

Meth, C. (2006) What was the poverty headcount in 2004 and how does it compare to recent estimates by van der Berg et al?. Southern Africa Labour and Development Research Unit Working Paper Number 06/01. Cape Town: SALDRU, University of Cape Town.

ISBN: 0-7992-2311-5

© Southern Africa Labour and Development Research Unit, UCT, 2006

Working Papers can be downloaded in Adobe Acrobat format from www.saldru.uct.ac.za.
Printed copies of Working Papers are available for R15.00 each plus vat and postage charges.

Orders may be directed to:

The Administrative Officer, SALDRU, University of Cape Town, Private Bag, Rondebosch, 7701, **Tel:** (021) 650 5696, **Fax:** (021) 650 5697, **Email:** brenda.adams@uct.ac.za

Abstract

The publication by Servaas van der Berg and his colleagues at the University of Stellenbosch last year of a set of estimates of poverty headcounts and gaps, pointed to substantial success in the battle against poverty in South Africa in the period 2000–2004, an improvement attributable mainly to massive increases in social grants. The present paper takes issue with the van der Berg *et al* headcount of 15.4 million in 2004, arguing that a more likely figure is in the region of 18–20 million. This difference is far too large to be ignored by policymakers. Using household survey data from Statistics South Africa (as opposed to the All Media and Products Survey data used by van der Berg *et al*), the paper develops a method for dealing with under-reporting by specifying limits to the size of the corrections that can be applied to income estimates without straying beyond the bounds of the plausible. The paper questions the uncritical assignment of all disability grant payouts to poverty alleviation. Although it concentrates on estimating the headcount and poverty gap in 2004, the paper does offer a tentative comparison of its results with those made by other researchers for the year 2000. The paper ends with recommendations for a series of changes that need to be made to the General Household and Labour Force Surveys conducted by Statistics South Africa, if these surveys are to fulfil their potential role as suppliers of information about poverty.

Acknowledgements

The paper survived its first public outing relatively unscathed. Version 1 was presented to the SALDRU (South African Labour and Development Research Unit) seminar at the University of Cape Town on 18th January 2006. Participants made many useful suggestions — attempts were made to incorporate as many of them as possible. Thanks are due to Justine Burns, Reza Daniels, David Lam, Miguel Lacerda, Murray Leibbrandt, Jo Lorenzten, Julian May, Martin Wittenberg and Ingrid Woolard, all of whom have either been coerced into reading and commenting on various versions of the paper, checking its workings and results, or have been harangued into discussion of its method and findings. Special thanks are due to Anna McCord, whose diligence helped me to eradicate a grievous error from the paper, and whose careful reading of it has resulted in many changes. Daniela Casale and Miguel Lacerda gave invaluable assistance with STATA programming. As discussant, David Lam offered useful comments on Version 4 of the paper at a workshop held in Pretoria on 17th February 2006 (all of the comments have been incorporated into it). The workshop was organised by the HSRC, and its purpose was to compare the findings and methodologies of the present paper with those of the van der Berg *et al* (2005) paper with which it takes issue. Miriam Altman of the HSRC has earned my gratitude for making possible the first presentation of the findings in roughly their present form, to a wide audience of academics and senior government officials, several of whose members (notably John Kruger of the National Treasury) made useful comments. Amended after the HSRC gathering, Version 5 of the paper received an airing at the seminar of the School of Development Studies at the University of KwaZulu-Natal, Durban, on 2nd March 2006. I am grateful to the participants, in particular to Dorrit Posel and Julian May, who subjected the paper (and its author) to a thorough grilling. Version 7 of the paper was presented to senior staff of Statistics South Africa on 16th March 2005. My thanks go to the Statistician-General Mr Pali Lehohla and to Dr Liz Gavin, who arranged the presentation, as well as to participants in the seminar, who offered useful criticisms. Finally, warm thanks are due to two fine people, my son Samuel Meth, who waited just long enough to be born to allow me to complete a draft for submission to another fine person, Murray Leibbrandt, who kindly performed the unenviable task of reading the manuscript prior to its publication as a working paper. Although it seldom seems to, it should go without saying that despite the numerous eyes that have been cast over it, all errors, omissions and misrepresentations in the paper are my own.

Table of contents

1. Introduction	1
2. Declining poverty: The van der Berg et al interpretation of reality	5
3. Extracting poverty estimates from the LFS & GHS:	
A first glance	9
Do disability grants alleviate poverty?	16
Poverty estimates in FGT format	19
Disaggregating the headcounts	21
A tentative comparison with the Hooegeveen and Özler findings for 2000	25
4. Extracting poverty estimates from the surveys: The details	31
Imputation of expenditure categories from income estimates	35
Adding in social grants and removing implausible zeroes	45
5. Estimation of actual from nominal under-reporting errors	51
6. Distributional implications of the poverty estimates	63
7. Other methodological and conceptual problems	68
8. Conclusions and recommendations	71
Recommendations	76
The GHS	77
The LFS	78
General	81
References	83

Index of Tables

Table 1	Poverty in South Africa, 1993, 2000 and 2004	6
Table 2	Poverty headcounts (1000s) in the bottom five expenditures categories, (R0-2500 per month) South Africa 2003–2004 ($z = R250$ in 2000 prices)	10
Table 3	FGT Poverty measures (ratios and gaps) in the bottom five expenditure categories (R0-2500) per month South Africa 2004 ($z = R250$ in 2000 prices)	20
Table 4	Poverty headcount (1000s) – Sept 2004 LFS ($z = R250$ in 2000 prices)	23
Table 5	FGT Poverty Measures (ratios and gaps) in South Africa, 1995–2004	27
Table 6	Headcounts in Meth (2006) and Hoogeveen and Özler (2004)	30
Table 7	Total monthly household income from all employment by monthly household expenditure (Labour Force Survey September 2004)	38
Table 8	The impact of the purging process on household distribution and incomes (Labour Force Survey September 2004)	40
Table 9	Child support grants in the 2004 GHS and September 2004 LFS (unpurged data)	49
Table 10	Converting nominal into effective error levels	58
Table 11	Impacts of varying mean income assumptions on effective error levels	60
Table 12	Cumulative distributions of consumption & population – various errors	64
Table 13	Household shares of expenditure, 2000	67
Table 14	Distributions (1000s) of individuals 16 years and over in three surveys	70

1. Introduction

In the field of political economy in South Africa, there are not many more sensitive subjects than that of the extent to which government policies aimed at alleviating poverty have succeeded.¹ There is broad agreement among academics that income poverty rose in the period 1995–2000. Government counters this with the claim that the finding fails to take account of the ‘social wage’ (social spending). To date, nobody has succeeded in demonstrating that the ‘social wage’ offset increases in income poverty prior to 2001.² Because of the absence of authoritative data, what has happened in the years after 2000–2001 is less clear. This has added to the contentiousness of the competing claims that are made. Possibly mindful of this, Lesetja Kganyago, Director-General of the National Treasury, is reported recently as proclaiming that

At this moment, there is just hot air and noise about whether we are making progress in reducing poverty (*Business Report*, Monday 28 November 2005).³

It would be surprising, however, if he were not tempted to exclude from this generalisation the most recent findings on the matter by van der Berg *et al* (2005). They report a significant fall in poverty headcounts and poverty gaps in the period 2000–2004. If true, their results would silence critics of government’s anti-poverty policy. Unfortunately for the poor, it seems unlikely that the optimistic picture painted by van der Berg and his colleagues is correct.

With observations for three separate years, 1993, 2000 and 2004, they are able to create something of a history of the trajectory of poverty. The

¹ As may be seen in Table 1 below, even van der Berg and his colleagues have poverty rising between 1993 and 2000, as measured by both the poverty gap ratio and the headcount ratio.

² See Meth (2005a) for a review of the debate and for a set of estimates of poverty, taking the social wage into account.

³ The context was the announcement of the assembly by government of a “task team made of local and international experts, to develop a poverty benchmark for the country” in an article by Andile Ntingi, headed “Poverty measure proposed: SA’s own system to evaluate government’s success”.

present paper harbours no such ambitions — its main purpose is to present a competing set of estimates of the poverty headcount and poverty gap for the year 2004 to those offered by the van der Berg *et al* (2005) paper.

It is sometimes asserted, particularly in discussions about national accounting magnitudes, that trends are as important, if not more so, than absolute levels. The poverty estimates under consideration here are most decidedly not an instance where such a claim could validly be made. In the first place, such reductions in poverty as may have occurred in the past few years are almost certainly the consequence of an unprecedented increase in the reach of government's social grant system, the expansion of which government has repeatedly stated that it wishes to limit.⁴ If future growth of the system is restricted to current commitments (mainly the extension of the reach of child support grants), the trend observed over the period 2000–2004 will not be repeated. In the second place, it matters very much to those affected whether there are 15 million people below the poverty line, or 20 million of them. Rather obviously, with a given rate of poverty reduction, the chances of meeting the poverty-halving goal to which government has committed itself, are considerably diminished if there turn out to be many more people to be raised out of poverty than is thought to be the case. In short, the poverty headcount (the absolute number of poor people), for all its arbitrariness, is extremely important. It follows from the foregoing that challenging the validity of the estimate for a single year (2004) is a legitimate, and (I would argue) necessary enterprise. Since the severity of poverty is also a matter of extreme importance, the same conclusion must hold for the comparison of competing estimates of the magnitude of the poverty gap ratio for a single year (2004).⁵

Having read a disavowal of any ambition to construct a time series of poverty estimates, readers may be excused for being a little baffled when, in Section 3 of the present paper, they come across (in Tables 5 and 6) an example of what appears to be precisely that which has been forsworn, namely, a time series for the years 1995, 2000 and 2004. Constructed by

⁴ The *Ten Year Review*, for example, saw the key social challenges as “[a] massive Public Works Programme, improved access to social security measures, with better vehicles to improve service delivery while reducing the number of citizens dependent on grants” (PCAS, 2003, p.114).

⁵ As social indicators, poverty gaps have the advantage of being able to convert directly into the size of the transfer required to raise to any desired level, the consumption of all those currently consuming at some level below it.

placing my estimates for 2004 with those made by Hoogeveen and Özler (2004) for the years 1995 and 2000, the resulting ‘series’ lacks the integrity of the van der Berg *et al* estimates. The Table 5 and 6 figures, constructed as they are using different methods from disparate sources, cannot be treated as though they formed a coherent series. The story that they appear to tell, if it is a story at all, must be regarded as entirely conjectural.

Returning to the 2004 estimates which are the central interest of the paper, it is hoped that by presenting them, interest may be kindled in the major household surveys conducted by Statistics South Africa as possible sources of information on poverty. Whatever its shortcomings, the van der Berg *et al* paper has the virtue of having re-invigorated a debate that, of late, had become increasingly stale and bad-tempered.

Having begun by expressing the limits of what is to be attempted in this paper, it will probably be useful at this point to say what else the present paper is not going to attempt to do. Most importantly, the paper does not attempt to examine the headcounts for any other year in any detail.⁶ The decision not to look into the past obviously means that little can be said here about improvements in the lot of the poor that may have taken place. Given the very large amount of money that has been thrown at the poverty problem in recent years, much of which appears to have reached its target, I would be very surprised if the headcount had not fallen, or at least not increased. Despite its importance, the hypothesis that poverty headcounts have declined is not tested here. The fact that finding suitable data for an earlier year is a job that has still to be tackled should not, however, be allowed to stand in the way of the publication of the results in the present paper. If the method by which they were obtained (and the data set from which they come) stands up to the scrutiny of the social scientific community, then the numbers that the method generates are of self-evident importance. Anyone wishing to reject my results has merely to demonstrate that (a) the method used to generate them is faulty, and/or (b) the data on which they rest are unreliable.⁷

⁶ Results for 2003 are presented. They are not intended, however, to be used as though they formed part of a time series.

⁷ To facilitate this, the spreadsheets used to generate the results, which have the STATA ‘Do File’ commands embedded in them, plus the relevant STATA Do Files, will be made available by the author on request at either meth@ukzn.a.za or chasmeth@telkomsa.net, once it is felt that the results are stable.

Furthermore, although brief references to aspects of the method that van der Berg *et al* use are made, there is no systematic attempt in the paper to grapple either with their primary sources or with the many data sources they call on to bolster their findings. They offer, for example, an optimistic assessment of possible job growth among the poor. Other than to remark that the labour market data may be too fragile to support this claim, it will not be contested.

Earlier versions of the paper ignored their treatment of the social grant system. More recent versions devote some attention to the Child Support Grant (CSG) and the Disability Grant (DG). By my reckoning, the CSG was reaching about 35 per cent of the children under the age of 15 years in households where expenditure was reported⁸ to be less than R800 per month in July 2004.⁹ Since almost all of these households would meet the eligibility criteria for the CSG (as will many in the next expenditure categories), it is clear that there is still quite a bit to be achieved by this component of the social grant system. Welcome as the grants must be among poor households, the distribution of CSGs that emerges in the 2004 GHS suggests that the grant may also be being paid in respect of as many as half-a-million children who do not qualify for it. In other words, there is tentative evidence that the CSG suffers from the usual shortcoming of means-tested grants — it excludes many who qualify, and includes many who do not.

⁸ These results are based on the ‘unpurged’ figures (purging is explained in detailed below). Since it is not the intention to engage in the debate over social grants here, no greater precision will be attempted, nor is any attempt warranted. There may be about 10–11 million under-fifteens eligible for the CSG. In July 2004, the GHS found 4.6 million CSG recipients. Because the numbers of people in the bottom expenditure categories fall, the use of the purged figure will raise coverage ratios. Also, the proportions of households covered are greater than the proportions of individuals covered.

⁹ The total number of CSGs picked up by the 2004 GHS matches quite closely the number reported by the administrative data of the Department of Social Development (the SocPen database).

2. Declining poverty: The van der Berg *et al* interpretation of reality

Selected findings from the van der Berg *et al* paper are reproduced below in Table 1. Use is made of two poverty lines, one set at R3 371 per capita per annum, the other at R3 000 per capita, or R250 and R281 per month respectively (all in constant 2000 prices). Although they are suitably cautious about the results in their conclusions, commenting that the trend in poverty rates over the first period is “not so clear” (2005, p.22), their work discloses a slight worsening of poverty between 1993 and 2000 — each of the poverty measures P_0 , P_1 and P_2 increases a little. Corresponding to the small rise in P_0 is a fairly substantial rise in poverty headcount numbers, to which they refer (2005, p.22). During the succeeding period, 2000–2004, they find, however, that poverty decreased in severity (headcount ratios and poverty gap ratios both fell). Poverty headcounts and poverty gaps are obviously smaller, the lower the poverty line. Looking more closely at the results in the table, we note that the drop in the headcount between 2000–2004, using the higher of the two poverty lines, was apparently about 0.7 million, while the ever-so-slightly-lower line is said to be associated with a fall in the headcount of 3.1 million. If true, this would suggest a massive clustering of people just below the R3 000 poverty line, requiring (and receiving) only a tiny income increment to lift them out of poverty.

That a mere three-quarters of a million or so people apparently find their way out of poverty when the poverty line is raised to R281 per month seems odd. Oddities require explanation. Since we do not have access to their data sets and calculations, the only thing that can be done here is to scrutinise the arithmetic in the table. Step one is to check the total population figures yielded by adding their estimates of the numbers of the poor and non-poor for each of the two poverty lines, for each of the three years.

Table 1 Poverty in South Africa, 1993, 2000 and 2004

	1993	2000	2004
Using poverty line of R3 371 per capita per year			
P ₀ (Headcount ratio)	0.441	0.446	0.366
P ₁ (Poverty gap ratio)	0.224	0.230	0.169
P ₂ (Squared poverty gap/Poverty severity ratio)	0.144	0.146	0.100
Number of poor (million)	16.9	18.7	18.0
Number of non-poor (million)	23.0	26.0	28.4
Sum of poor and non-poor (million)	39.9	44.7	46.4
Using poverty line of R3 000 per capita per year			
P ₀ (Headcount ratio)	0.406	0.413	0.332
P ₁ (Poverty gap ratio)	0.200	0.205	0.146
P ₂ (Squared poverty gap/Poverty severity ratio)	0.126	0.127	0.085
Number of poor (million)	16.2	18.5	15.4
Number of non-poor (million)	23.7	26.2	31
Sum of poor and non-poor (million)	39.9	44.7	46.4
Source: van der Berg <i>et al</i> (2005, Table 2, p.17)			

Checking the results(i) Implied total populations (Headcount * 1/P₀)

Using poverty line of R3 371 per capita per year	38.3	41.9	49.2
Using poverty line of R3 000 per capita per year	39.9	44.8	46.4

(ii) Implied P₀ if 'Number of poor' is correct

Using poverty line of R3 371 per capita per year	0.424	0.418	0.388
Using poverty line of R3 000 per capita per year	0.406	0.414	0.332

Corrected headcounts at R3 371 poverty line if P₀ & total population are correct

	17.6	19.9	17.0
--	------	------	------

The figures so obtained are close to those most commonly used.¹⁰ Next, we multiply the numbers of poor by the reciprocals of headcount ratios to obtain total populations (operation (i) in the table). For the R3 000 per capita line, the population totals agree with the sums of poor and non-poor. For the R3 371 poverty line they do not. Another way of checking the results (operation (ii) in the table) is to assume that the poverty counts and total populations are correct and then to estimate the implied poverty headcount ratios. Once again, one sees agreement of the relevant figures at the R3 000 poverty line, and disagreement at the R3 371 line. Reference to their Figure 17 (2005, p.39) confirms that their Table 2 estimates at the R3 371 poverty line are incorrect. It is possible that the poverty headcount ratios are correctly reproduced in their work. That would imply poverty headcounts at the R3 371 poverty line of 17.6, 19.9 and 17.0 million in 1993, 2000 and 2004 respectively. Figures of that order of magnitude make a lot more sense than those in the table — the decline of 2.9 million between 2000 and 2004 being much closer to that reported at the nearby R3 000 poverty line.

Mistakes such as this, which all of us make from time to time, are possibly the result of some (nightmare) transcription error. We do not dwell on this, but rather derive an alternative set of estimates of the poverty headcount for 2004. Constructed from the official (Statistics South Africa) household survey figures, these figures are argued to have a pedigree that is at least as credible, if not more so, than those produced by van der Berg *et al* (2005).

There is no national survey dedicated to the measurement of poverty in South Africa. To fill this gap, researchers make use of the income and/or expenditure data gathered in numerous household surveys, none of which is explicitly designed for the task. Chief among them are the Income and Expenditure Surveys (IES) (conducted every five years),¹¹ the Labour Force

¹⁰ Discovering the size of the population is no simple matter. For the early days, reliance must be placed on estimates which had to 'reconstruct' the country to take account of the populations of the so-called 'TBVC states'. One of the recent estimates of the size of the 1993 population, including the TBVC populations, is 37 802 000 (Statistics South Africa 2001, p.1.1).

¹¹ In the past, income and expenditure surveys relied on the 'recall' method for gathering data (respondents were required to recall how much has been spent on a wide variety of goods and services). The 2000 IES used the same sample as the September 2000 Labour Force Survey. In 1995, the October Household Survey (OHS) and IES were paired. The recall method has been dropped in favour of the

Surveys (LFS) (twice yearly) and the annual General Household Survey. Use is also made of the population censuses. By comparing estimates of the total values of household income and/or expenditure with the figures gathered independently for the national accounts, it has been demonstrated (repeatedly, and in many countries) that household surveys (and population censuses) under-report both. It should not be assumed, however, that national accounts are themselves beyond reproach (Deaton, 2003), a circumstance that must cause all poverty estimates based on figures raised to national accounts levels to be viewed with suspicion. The force of Deaton's criticism is fully acknowledged by van der Berg and his colleagues. Nevertheless, when all is said and done, the position they adopt is one in which they

... trust national accounts data for aggregate household income, [and] trust survey data¹² for the distribution of such income between households (2005, p.11).

To support this position, they draw on such subsidiary evidence as is available. The poverty headcounts yielded by the method I propose in the present paper are much higher than the van der Berg *et al* figures. There are three possibilities: (i) the van der Berg *et al* figures could be (roughly) correct, or (ii) my figures could be (roughly) correct, or (iii) we could both be wrong. If either (ii) or (iii) is true, then it follows that their approach is incorrect. Were that so, there would be no need to examine the attempts they make to back up their findings with corroborating evidence dredged from alternative sources. Since I proceed on the assumption that the method I propose here, and the results it yields, are reliable, no attention will be paid to the evidence they adduce in favour of their claims that poverty has fallen.

Reduced to its barest essentials, their approach apparently consists in the scaling of distributional means (by race) derived from the surveys until they equal per capita income means (by race) derived from the national accounts

'diary' method for collecting the data (participants keep diaries of income and expenditure). There is no link to any other household survey. This means that the IESs can no longer be used for poverty studies in the way that they have been in the past.

¹² The survey data that they trust is that gathered by the annual All Media and Product Surveys (AMPS). They do, however, check their results on other, publicly available data.

(van der Berg *et al*, 2005, pp.28–30). Despite the appeal to international authority¹³ (in the form of a reference to the use of this technique elsewhere), it is not obvious why the magnitude of adjustments resulting from such a procedure should bear any resemblance to the size of actual under-reporting errors by income or expenditure class (which, all agree, are likely to plague any survey instrument).

There is no indication in the van der Berg *et al* paper of how zero income households are treated. We know that they are aware that there are many households that contain no wage earners (2005, p.21). The only other reference to the problem in their paper occurs in a brief discussion of the fact that other poverty researchers have shown that it is significant (2005, p.9). But enough now of their approach — let us turn to recent Labour Force and General Household Surveys, from which evidence on both under-reporting and the plausibility of zero income reports may be extracted.

3. Extracting poverty estimates from the LFS & GHS: A first glance

Table 2 presents the results of an attempt to use the LFSs for September 2003 and 2004, and the GHSs for 2003 and 2004 as sources of information on poverty in South Africa. Headcounts for the lower of the two poverty lines used by van der Berg *et al* only ($z = \text{R}250$ per capita per month in 2000 prices) are presented. This paints the most ‘optimistic’ picture of poverty in the country.¹⁴ The 2003 results are presented only because they act as a benchmark (of sorts) against which

¹³ The use of national accounts in poverty estimation is currently the subject of debate. A paper from the United Nations Statistics Division (Havinga, Kamanou and Vu, undated, but which must be relatively recent because it refers to a 2004 event), states that “... national accounts estimates of household final consumption and income, as used in the current procedures to estimate poverty, are not appropriate” (p.1). Reza Daniels of UCT (*pers. comm.* 18th January 2006), respondent at a conference presentation of the van der Berg *et al* results, brought the paper to my attention.

¹⁴ If desired, the R281 poverty line can be used on the LFS and GHS data. There is, however, no point in attempting a comparison of the results with the corresponding van der Berg *et al* figures until such time as the confusion about headcounts at the R281 line is eliminated.

to test the consistency of the 2004 results. It is not intended that the results in Table 2 be read as though they constituted an historical record.

Rather than using correction factors derived from relationships between survey and national accounts estimates of consumption and/or income, I have used the survey data to purge the estimates of the numbers of households (and individuals), by expenditure category, whose reported income is greater than the upper bound of the expenditure category in question ($Y > E$). Households in expenditure categories where income exceeds expenditure have been relocated to higher expenditure categories. Purging is thus a process of imputation — expenditure categories are imputed from income data.

Table 2 Poverty headcounts (1000s) in the bottom five expenditures categories, (R0–2500 per month) South Africa 2003–2004 ($z = R250$ in 2000 prices)

Data purging stage	2004		2003	
	LFS	GHS	LFS	GHS
1. Relocate if $Y > E_{\max}$				
Under-reporting error = 0%	23 518	23 174	23 615	23 782
Under-reporting error = 10%	22 665	22 336	22 840	22 995
Under-reporting error = 20%	21 814	21 553	22 045	22 178
Under-reporting error = 50%	19 527	19 080	19 798	19 389
Under-reporting error = 100%	15 580	15 478	16 390	16 277
		Δ Head-		Δ Head-
2. Relocate if $Y > E_{\max}$, include	2004	count	2003	count
CSGs and OAPs	LFS	(Panel 2	LFS	(Panel
		– 1)		2 – 1)
Under-reporting error = 0%	20 031	-3 486	21 897	-1 717
Under-reporting error = 10%	19 000	-3 665	20 913	-1 927
Under-reporting error = 20%	18 068	-3 746	19 970	-2 075
Under-reporting error = 50%	15 366	-3 891	17 406	-2 392
Under-reporting error = 100%	11 817	-3 763	13 806	-2 584

3. Simulation 1: Award Disability Grants to households containing recipients of the OAP		Δ Head-count (Panel 3 – 1)	Δ Head-count (Panel 3 – 1)
Under-reporting error = 0%	18 813	-4 705	20 673
Under-reporting error = 10%	17 843	-4 822	19 707
Under-reporting error = 20%	16 976	-4 838	18 795
Under-reporting error = 50%	14 418	-4 839	16 307
Under-reporting error = 100%	11 068	-4 512	12 887

4. Simulation 2: Award Disability Grants to households containing recipients of the CSG		Δ Head-count (Panel 4 – 1)	Δ Head-count (Panel 4 – 1)
Under-reporting error = 0%	19 089	-4 429	21 297
Under-reporting error = 10%	18 048	-4 617	20 296
Under-reporting error = 20%	17 061	-4 753	19 365
Under-reporting error = 50%	14 459	-4 798	16 796
Under-reporting error = 100%	10 959	-4 621	13 211

Note: Unless otherwise stated, the results in all tables in the present paper have been estimated from the data in the relevant CD ROMs supplied by Statistics South Africa.

Households receiving child support grants and/or state old age pensions have been treated separately from households which receive neither, in order to obtain comprehensive conditions in different households. Apart from disability grants (DGs), the other social grants on which information is collected, in both the LFSs and GHSs, have been ignored. Implausible zero income reporters have also been identified and removed. Rich though both the LFSs and GHSs are in information relating to economic wellbeing (or illbeing), both have weaknesses. The LFS does not collect individual information on social grants (something it could very easily have done), while the GHS does not collect information on migrants. Since the shortcomings of the LFS relating to social grant information can be overcome (albeit awkwardly), while the GHS' lack of migrant data cannot, the former constitute the basis on which the poverty estimates in the present paper come to rest. The operation described as 'purging' may be performed on both data sets. Because of the existence of data on migrant remittances in the LFSs, however, imputation may be taken a lot further with the data from those surveys than it can with the data from the GHSs.

The results presented in Table 2 (and subsequent tables) are for the folk in households where total monthly expenditure in 2004 was less than R2 500 per month. It will be argued that the lowest estimate that could be regarded as a plausible contender for the title of poverty headcount exceeds the van der Berg *et al* figure by more than three million. The first panel of Table 2 shows the numbers of households that fall below the poverty line before receipt of any social grants. The only income accruing to these households is that earned by household members and that remitted by migrants attached to the households. The table thus shows the effects of what may be described as a simple purging operation in which all households with an earned and/or remittance income greater than the upper bound of the expenditure category in which they are located (i.e., $Y > E_{\max}$), are relocated to the expenditure category appropriate to their reported income level.¹⁵

In the second panel, which makes use of the LFS data only, both the child support grant and the state old age pension are taken into account in estimating income levels. Panels 3 and 4 contain the results of simulations designed to examine the impact on poverty of allocating the total amount paid out in the form of disability grants to the income received by households. In both cases, it is assumed that the total value of the benefit goes towards poverty alleviation. Although it seems to be taken for granted by some researchers that it is appropriate to do so, it is not obvious that this is appropriate. This assumption will be critically examined below.

To perform the simulations that incorporate the disability grants, a way had to be found to navigate around limitations of the model which make it difficult, without a great deal of effort, to assign them to the households in which they accrue.¹⁶ To bypass the obstacle, disability grants were given either to households already receiving old age pensions (Simulation 1), or to households already receiving the child support grant (Simulation 2).

¹⁵ The model also allows for the estimation of poverty levels from the raw (unpurged) data. This, in effect, is what I have done in the past, applying what I thought at the time were sufficiently large correction factors for under-reporting of income. The more rigorous approach in the present paper shows that faith in the validity of that approach was misplaced.

¹⁶ The LFS does not (did not in the past) gather information on individual social grant receipts. This is (was) gathered at a household level. The problem here lies not so much with the data source as it does with the capacity of the model to assimilate the data.

Because the numbers of recipients of each grant type differ quite considerably, the value of the disability grant had to be scaled to ensure that the correct total value of benefits was paid out.¹⁷

In all four panels in the table, estimates of the headcounts with nominal under-reporting errors of between 0 and 100 per cent are presented. A nominal under-reporting error of 50 per cent, for example, means that if a household reported that they were in household expenditure category R0–399 per month, their actual maximum potential consumption level would be $R399 + R399 * 50/100$, which equals R598.5. An individual in a two-adult household in this expenditure category, instead of consuming roughly R200 per month, would thus consume about R300 per month of goods and services. Ignoring dissaving (borrowing, drawing down previous savings) the sources of the income necessary to make consumption possible are earned incomes (of in-house workers and migrants) and social grants. The value of the latter can be estimated with some precision, so the full burden of the error adjustment falls on earned incomes. Estimating the actual or effective under-reporting errors corresponding to any nominal assumed under-reporting rate is not a simple matter. Techniques for doing so are considered further (see the discussion to Tables 10 and 11). Suffice it to say at this point that the ratio between the two types of error is influenced by the ratio of the number of individuals in the expenditure category to the number of workers in it, and by the assumed category mean income (expenditure). To jump ahead in our story, it is argued that it is unlikely that the nominal under-reporting error would exceed 10 per cent, and extremely unlikely that it could exceed 20 per cent.¹⁸

¹⁷ To explain this method of allocating disability grants, an example is necessary. In the bottom expenditure category, according to the purged 2004 GHS results, 97 093 disability grants, 883 089 child support grants, and 235 819 old age pensions were paid out. The value of the disability grant in 2000 prices was R597.6. If it is assumed that disability grants are paid into households that receive child support grants, then the distributed value of disability grant income that matches this is R65.7 ($597.6 * 97\ 093 / 883\ 089$). If the disability grants go into households receiving pensions, the value is R246.0 ($597.6 * 97\ 093 / 235\ 819$). This is a thoroughly unsatisfactory way of dealing with the problem. Unfortunately, the model admitted of no other solution that was not impossibly time-consuming.

¹⁸ If that is so, the reader is entitled to ask why results at error levels higher than 20 per cent are presented in Table 2. The answer is apparent in the table — to obtain headcount estimates as low as those estimated by van der Berg *et al*, the data used to construct Table 2 suggest nominal under-reporting errors of somewhere in the region of 50 per cent.

If the proposition that the nominal under-reporting error cannot plausibly exceed 20 per cent is accepted (it is defended at length in Section 5 of the paper), then the poverty headcount excluding disability grants lies between 18 and 20 million. Assuming that the full value of all disability grants paid out goes towards poverty alleviation, then the poverty headcount lies between 17 and 19 million in both simulations (give or take a few thousand). Comparing the zero error level results in the first panel with those in Panel 2, it would appear that between them, the child support grant and the old age pension lift some 3.5 million people (or 7.5 per cent of the population) above the poverty line. Adding in the disability grants (in Panels 3 or 4) raises a further one million or more people out of poverty (as measured by the R250 per capita poverty line in 2000 prices), pushing to about ten per cent the proportion of the population elevated above the poverty line by the social grant system.

As one would expect, poverty headcounts fall with rising assumed nominal under-reporting error levels, converging on the van der Berg *et al* estimates soon after the 50 per cent level in Panel 1 (and, as we shall see below, at about 38 per cent in Panels 3 and 4). As will be demonstrated in the discussion of the method, however, there is little likelihood of such a large error being present in the results after social grants and remittances have been taken into account. Neither are there sufficient income earners in the lower household expenditure categories to make such a thing possible, nor are the mean incomes (expenditure levels) that have to be assumed in order to generate such a result, plausible.

Also as one would expect, the numbers of people raised out of poverty by the social grant system change as the assumed under-reporting error level rises, increasing until the error level reaches 50 per cent. Larger assumed errors than this see the difference fall, or even, as is the case in Panel 3, become negative. Clearly, the numbers are influenced by the other assumptions about the distribution of grants. Most markedly in Panel 2, where the difference between the largest and smallest estimates of the numbers lifted over the poverty line is about 400 000, the differences (either positive or negative) remind us that the manner used to estimate the impact of the disability grants is unsatisfactory. Greater precision requires the resolution of the conceptual and empirical problems associated with disability grants. Even so, the fact that independent household surveys (conducted a few months apart, in July and September) yield broadly

similar results before social grants are taken into account, and suitably different results (changes in headcounts) after social grants and remittances have been admitted into the calculus, does inspire some confidence in the surveys.¹⁹

Having expressed some confidence in the Table 2 results, it must also be said that the temptation to conclude from a comparison of the 2003 and 2004 results that the poverty headcount dropped by almost two million, has to be resisted. Although the reported year-on-year changes between 2003 and 2004 would undoubtedly be statistically significant, they are so large as to make one suspect that either the 2004 LFS is finding too few poor, or the 2003 surveys are finding too many. Until such time as there is a thorough check of what the potential poverty reducing effects of such changes as took place over the period might be,²⁰ it is as well not to attempt to draw any conclusions from a comparison of the 2003 and 2004 results.

Although the present paper relies on the September 2004 LFS as the source of its poverty estimates, presenting results from the other three surveys (the two GHSs and the September 2003 LFS) is not without point. Placing them side-by-side allows us to raise questions about the quality of the surveys. Common sense suggests that very large surveys conducted not far apart in time should generate similar population totals and very similar distributions of individuals among households. In addition, it does not seem unreasonable to postulate that there should not be large differences in the extent to which households under-report income and expenditure. Part of the process of determining the suitability of the household surveys for use as poverty measuring instruments (and it will be argued that the GHS should fulfil this role) is thus a thorough comparative analysis of the differences between the four surveys used to generate the results in Table 2.

¹⁹ I am mindful here of the injunction against the use of the survey results as time series data, when they are separated by so short a period as a year. This notwithstanding, it would have been disconcerting to see smaller headcounts emerging from the 2003 bottom panel figures than those yielded by the 2004 survey.

²⁰ The number in employment rose by 200 000, while the number of officially unemployed fell by 300 000. As a result, the official participation rate dropped from 54.8 to 53.8 per cent. The number of discouraged rose over the same period by 170 000, while the expanded participation rate fell marginally. Disability grant recipients increased in number by about 300 000, while the number of CSGs increased by almost 1.7 million.

Do disability grants alleviate poverty?

Estimating the poverty alleviating effects of disability grants is no simple matter. As McLaren, Solarsh and Saloojee (2004, p.164) point out, the difficulties of using quantitative data to understand the problems of the disabled are manifold. The 2001 population census collected information on disabilities, endeavouring as it did so to ascertain whether or not the disability in question prevented “full participation in life activities (such as education, work, social life)”. Results for the country as a whole appear in Report No. 03-02-04 (Statistics South Africa 2001). Most people suffering from the disabilities listed, even those affecting the intellect, would still be capable of leading productive lives if suitable employment could be found for them. The report is silent, as it must be, on this particular issue. Thus although global totals for each of the disabilities are given, they do not advance our understanding of the extent to which the needs of the disabled are met.

Similar questions to those in the 2001 population census are asked in the GHSs. The surveys have the advantage of being able to link disability status to as many relevant variables (e.g., employment status, receipt of disability grant, household income, etc.) as is thought to be necessary. It is thus possible to gain some insight into the conditions in which the disabled live. The 2004 GHS finds over one million DG recipients, 95 000 in the bottom expenditure category of the unpurged data, and over 450 000 in the next category (R400–799 per household per month). As ‘generous’ as the old age pension in 2004 (R740 per month), the more than R11 billion that disability grants cost dwarfed the child support grant’s R8.8 billion price tag. A check on the reliability of the GHS figures may be performed by referring to the SocPen database figures produced by the national Department of Social Development. These show that in April 2004, the disability grant attracted 1.27 million successful claims. Including the amount paid out as disability grants in the poverty headcount calculus has, as we have seen above, a not insubstantial effect on the numbers in poverty. Given, however, that we have switched to the LFS as the primary source of data for the poverty estimates made in the present paper, and given that the LFS does not collect information on disabilities, it is not clear what use can be made of the information extracted from the GHS.

Two features of the disability grant make it difficult to measure its poverty alleviating effects. Firstly, it is difficult to assess the ability of the grant to meet the needs of those for whom it is intended. There is some knowledge

about the distribution of disabilities by degree of severity in South Africa (CASE, 1999). Much less is known, however, about the needs of the disabled.²¹ Poverty lines are difficult to determine — poverty lines for disabled people even more so. Because it is inappropriate to apply an arbitrary poverty line of R250 per month to the ‘income’ of disabled persons, it is not possible to determine the extent to which the grant of R740 per month actually met the minimum needs of the disabled in 2004.²²

Secondly, it is difficult to determine the extent of the (inevitable) benefit dilution that takes place when categorical grants of this nature are made available in very poor communities. Suppose that all of the disabled were wholly incapable of supporting themselves. Suppose further that the grant did not cover the cost of doing so at a level equivalent to that implied by the national poverty line. In such a case, the disability grants would have no poverty alleviating effects at all. Imagine, by contrast, a situation in which nobody’s disability is serious enough to prevent them from earning their keep. Imagine, however, that for whatever reason, there were very few jobs available for the disabled. If, in such circumstances, those responsible for granting the certificates of disability were moved more by humanitarian than by technical motives, each disabled person could qualify for the grant. In this case, the disability grant would be a poverty alleviation measure, pure and simple. The true state of affairs in South Africa (and elsewhere) lies somewhere between these poles — where, it is impossible to say with any precision.²³

By its very nature, the concept of ‘disability’ carries with it unavoidable ambiguities. Try as they will, policymakers cannot tailor the legislation and the regulations to fit every circumstance. The legislative intention is to provide social assistance for people suffering from a wide variety of “physical, sensory, mental and intellectual” disabilities²⁴ that render them

²¹ The Taylor Committee recommended the conducting of “needs-assessment” of the disabled by the Disability Sector and “economists” (DoSD, 2002, p.105).

²² Disability ranges from relatively minor handicaps that “prevent full participation” to conditions requiring elaborate and extensive care. Meeting special needs could, in many cases, cost a lot more than the R740 paid as a grant in 2004.

²³ A recent survey by Nattrass (2006) considers these and many of the other problems associated with the disability grant. Not unreasonably, she makes no attempt to address the question of the extent to which the grant alleviates poverty as opposed to meeting the special needs of the disabled.

²⁴ These are the categories used in the Taylor Report (DoSD, 2002, p.104).

“unfit” to earn the means to meet their needs.²⁵ Disabilities range from minor handicaps to a total inability to perform anything other than involuntary acts. In some instances, the grant functions as a means of providing replacement income for those incapable of earning an income (the incapability could be a permanent or temporary). Viewed in this light, the benefit could be described as an incapacity grant.²⁶ Individuals suffering from certain illnesses, such as TB and HIV/AIDS, can, for example, be expected to become fit for work with the appropriate medical treatment. When the health of the individual concerned returns to a (medically determined) state where they would be capable once more of earning their keep, they are no longer eligible for the grant. In a tight labour market, they would simply be re-absorbed into employment. A market saturated with a supply of low-skilled and unskilled labour, such as one finds in South Africa, will, however, create the hypothetical conditions outlined above. Clearly, in these circumstances, benefit claimants have a strong incentive to manipulate the system in ways that enable them to retain the grant.²⁷

²⁵ “A person is [subject to certain conditions] eligible for a disability grant, if he or she (a) has attained the prescribed age; and (b) is, owing to a physical or mental disability, unfit to obtain by virtue of any service, employment or profession the means needed to enable him or her to provide for his or her maintenance.” Social Assistance Act, No. 13 of 2004.

²⁶ The Taylor Committee argued that “[e]ligibility should not be based on the person’s incapacity to work, as often their lack of work is due to the poor economic climate and prejudice in the work place, as opposed to their physical or mental inability to perform the job” (DoSD, 2002, p.104).

²⁷ Anecdotal evidence suggests, for example, that some TB sufferers stop taking medication part of the way through the course of treatment, in order to continue to qualify for a benefit that is generous in relation to the mean incomes of people at the bottom end of the labour market. (One shudders to think of the implications of such acts for the emergence of drug-resistant strains of the disease. One shudders, as well, to contemplate the desperation of people prepared to treat their bodies in this way). Other weaknesses of the system (concerning, for example, the role of the medical profession in making the judgements necessary for people to qualify as incapacitated or disabled) are suspected in some quarters to lead to widespread abuse. Natrass (2006, pp.3–5) describes the changes in the regulatory framework that preceded the flood of successful grant applications, as well as the subsequent tightening of the regulations in an attempt to undo the introduction of “‘social’ factors . . . into what the legislation envisaged should be a purely ‘medical’ decision” (p.5). The appendix in Natrass (2006, p.23), a table showing differences in disability policy application in the different provinces, illustrates the dangers of granting the provincial sphere of government too much leeway in formulating

If the grant is being ‘abused’ (used as a substitute for unemployment benefits, to be shared by all and sundry in a household), the unintended consequence of its payment is poverty reduction. By how much, nobody can say. Unless the extent of such ‘abuse’, and the extent to which the needs of the disabled are met by the grant, can be established (two exceedingly difficult tasks, the first more than the second), a question mark must hang over the validity of poverty headcounts that uncritically include the full value of the disability grants as part of the poverty alleviation package. Nonetheless, it is clear that the disability grant does reduce poverty. For purposes of comparison, therefore, it is necessary to make allowance for this. Given the limitations both of the model and of the data it uses, the only gesture in this direction is the all-or-nothing approach used to construct the poverty estimates in Table 2. That gesture is repeated in the Foster-Greer-Thorbecke (FGT) type-estimates, to which we now turn.

Poverty estimates in FGT format

Everyone who dabbles in poverty analysis is aware of the limitations of simple headcounts.²⁸ Earlier versions of the present paper ignored these problems, taking it for granted that since the purpose of the paper was to assess the van der Berg *et al* headcount estimates for 2004, there was no need to have recourse to the more sensitive measures available. Convinced by a discussant that there was merit in presenting these more conventional measures, the standard FGT-type estimates have been calculated. Table 3 gives estimates of the headcount ratio (P_0) and poverty gap ratio (P_1) that correspond to the van der Berg *et al* figures reproduced in Table 1 above.²⁹ Estimates of P_2 (the squared poverty gap ratio) have not been made.³⁰

In the case of the poverty gap ratio estimates, estimating the FGT measures entailed a change from the approach I had adopted in setting up the model.

welfare policy.

²⁸ See, for example, Woolard and Leibbrandt, 2001, p.57.

²⁹ The table contains estimates for 2004 only. This is because the statistical news release that contained the revised LFS results for the period 2000–2005 (SR P0210 of 26 September 2005) did not contain a revised total population estimate for 2003.

³⁰ Although the rationale behind the estimation of this measure is simple (giving more weight to the poverty of the very poor), I have never been able to make any kind of intuitive sense of the figure, a predicament I apparently share with others (see Woolard and Leibbrandt, 2001, p.57).

Table 3 FGT Poverty measures (ratios and gaps) in the bottom five expenditure categories (R0–2500 per month) South Africa 2004 ($z = R250$ in 2000 prices)

	P_0 headcount ratio	P_1 Poverty gap ratio
2. Relocate if $Y > E$, and include CSGs and OAPs		
Under-reporting error = 0%	0.430	0.195
Under-reporting error = 10%	0.408	0.179
Under-reporting error = 20%	0.388	0.164
Under-reporting error = 50%	0.330	0.129
Under-reporting error = 100%	0.254	0.088
3. Simulation 1: Award Disability Grants to households containing recipients of the OAP		
Under-reporting error = 0%	0.404	0.183
Under-reporting error = 10%	0.383	0.168
Under-reporting error = 20%	0.365	0.155
Under-reporting error = 50%	0.310	0.121
Under-reporting error = 100%	0.238	0.083
4. Simulation 2: Award Disability Grants to households containing recipients of the CSG		
Under-reporting error = 0%	0.410	0.176
Under-reporting error = 10%	0.388	0.162
Under-reporting error = 20%	0.367	0.148
Under-reporting error = 50%	0.311	0.116
Under-reporting error = 100%	0.235	0.078

Note: The numbering of the panels in this table corresponds with that in Table 2. Data are from the September 2004 LFS

For my own part, I prefer the bluntness of a figure which says that to raise all those in poverty above the poverty line, Rx billion will have to be transferred each year — that is a figure policymakers can understand. Constructed originally to churn out these ‘in-your-face’ figures, the model says that the Panel 2 poverty gap with zero under-reporting error is R27 billion in 2000 prices (or roughly the sum of the OAP and the CSG in 2004,

all over again). At the 10 per cent error level it is R25 billion, while a 20 per cent under-reporting error would reduce the poverty gap to about R23 billion. Expressed in 2004 prices, the latter would have amounted to about R28 billion.

Since the object of the present paper is to challenge the van der Berg *et al* estimates for 2004, the Table 3 FGT-type estimates make it possible to extend the comparison between the figures from the two papers to one more variable, the poverty gap ratio, using comparable measures. The comparison confirms the Table 2 conclusion that to obtain the van der Berg *et al* 2004 P_0 and P_1 results, a nominal under-reporting error of somewhere between 20 and 50 per cent would be required. This is shown by the tendency towards convergence with the van der Berg *et al* results as error levels rise.

Comparing the figures in Table 3 with the van der Berg *et al* results in Table 1, we note that the value of P_0 at the R250 per month poverty line (R3 000 per annum) for the year 2000 (0.413), is similar to the zero-error estimates in Panels 3 and 4 of Table 3. The P_1 values at the zero error level differ by a fair amount. The van der Berg *et al* P_1 for 2000 was 0.205; that for 2004 was 0.146. If the claim in Section 5 that the under-reporting error level is unlikely to exceed ten per cent is sustained, then the 2004 P_1 value, assuming all of the disability grant goes to poverty alleviation, would be between 0.16 and 0.17. Expressed in terms that policymakers can understand, the difference between my estimates of the poverty gap in 2004 and those made by van der Berg *et al* amounts to somewhere between R2 billion and R6 billion per annum, a sizable claim on the fiscus.

Disaggregating the headcounts

Table 4 opens up the results of Panels 2, 3 and 4 in Table 2, to allow closer scrutiny of the extent of poverty in each of the bottom five expenditure categories (expenditure in the range R0–2 499 per month). Information on four basic household types (compositions) is presented in each of the five panels into which Table 4 is divided. In the first panel, the aggregated results for all households are given. In the second panel, none of the individuals concerned lives in a household in which either a child support grant or state old age pension is received. The third panel counts individuals in households which receive at least one child support grant, but into which no old age pensions flow. The fourth panel is concerned with households receiving at least one old age pension, but no child support grants, while the fifth panel picks up the individuals in households that receive at least one

child support grant and one state old age pension. Assumed error levels for the results presented in Table 4 are zero in each expenditure category.³¹

The total headcount of 20 031 000 in the first panel of Table 4 thus matches the first estimate in the lower panel in Table 2. Constructed only from ‘purged’ data (i.e., data in which all households where income exceeds the upper bound of the expenditure category have been relocated to the appropriate expenditure category), and in which social grants in the form of CSGs and OAPs (as well as migrant remittances) are taken into account, the results do not throw up many anomalies. Most obvious of the anomalies present are (a) the few thousand adult-only households that report receipt of child support grants and (b) the even fewer two-adult households in the R0–399 per month expenditure category that receive old age pensions. Non-eligible child support grant and non-eligible old age pension recipients are so few in number that they may safely be ignored.³²

Of the 5.8 million people who live in households in which total monthly household expenditure is less than R400 per month, more than five million are below the poverty line. In the next expenditure category, 8.3 million of the 10.7 million people it contains are in poverty. Although the bulk of poverty is to be found in households containing adults and children (that is, after all, where the bulk of the population is to be found) there are still fairly large numbers of poor in adult-only households. As one would expect, the poverty rate falls as household expenditure levels rise. There are, however, quite a few poor in the top two expenditure categories in the table, congregated entirely (once more, as one would expect) in multiple-adult or adult-and-child households.

The figures in Table 4 seem to tell us that despite the much-vaunted ‘generosity’ of South Africa’s social grant system, much more is required. In 2004, there were still about six million individuals living in poverty in households that receive no child support grants, where some substantial proportion is almost certainly eligible for such grants. They are visible in the second panel of the table.

³¹ Nominal error levels in each expenditure category are assumed to be equal. The model allows them, however, to take on different values.

³² The value of the state old age pension in 2004, expressed in 2000 prices, was about R598. This means that the individuals in single-adult or two-adult households in the bottom expenditure category in which at least one individual received the full pension, could not fall below the poverty line. A possible explanation of the observed numbers is receipt of less than the full pension. The greater likelihood is that the result is merely statistical noise.

This will, of course, have changed considerably between September 2004, the period to which the results in Table 4 refer, and the present, because of the drive to extend the reach of the child support grant. As may be seen by comparing the headcount ratios in Panels 3 and 4 of the table (a little mental arithmetic is required here), the old age pension has a much bigger impact on poverty than the child support grant. This is hardly to be wondered at, given their relative magnitudes — R740 for the pension and R170 for the child support grant (in 2004 prices). The mean number of child support grants was about 1.5 per household in 2004. Even if this rises to about two per household as the grant is extended to the under-15-year-olds, the CSG will, on average, have less than half the poverty alleviating impact of the OAP.

Compounding the woes of the poorest households is either the absence or the smallness of migrant remittances. If the second panel gives an indication of how many people lack the wherewithal to keep body and soul together (about seven million people), the third and fourth panels suggest that poverty is so deep, that even with access to social grants there are still more than ten million people below the poverty line. Almost all of them, as one would expect, are adult-and-child-type households. Relatively large household sizes ensure that poverty rates (numbers of individuals in poverty divided by the total number of individuals in any particular category) remain high, except in the highest expenditure category in the table. The lesson to be drawn from this is that even when the grants have been made available to all those currently targeted (these will almost all be child support grants), the poverty headcount will remain high.

The story told here repeats itself, but with smaller numbers, as the assumed under-reporting error level is increased. It is a tale of the deep intractability of poverty down at the lower end of the distribution. With a 50 per cent assumed nominal error, the headcount in the bottom expenditure category falls by only 380 000, a mere seven percentage point drop. The impact in the next expenditure category is somewhat larger (predictably, because the numbers of people just below the poverty line increases). A 50 per cent under-reporting error would cause the headcount in the second expenditure category (R400–799) to fall by 1.4 million, a 17 percentage point decrease.

Table 4 Poverty headcount (1000s) – Sept 2004 LFS ($z = R250$ in 2000 prices)

1. In all households						
	R0-399	R400-799	R800-1199	R1200-1799	R1800-2499	R0-2499
Single- adult households	284	0	0	0	0	284
Two-adult households	255	177	7	0	0	439
More than two-adult households	189	328	106	54	13	689
Adult(s) and child(ren)	4 424	7 779	3 696	1 945	774	18 618
Total numbers in poverty	5 153	8 284	3 809	1 999	787	20 031
Total no. of people in category	5 748	10 733	5 818	4 982	3 439	30 720
Poverty rates (%)	89.6	77.2	65.5	40.1	22.9	65.2
2. In households receiving neither CSGs nor OAPs						
	R0-399	R400-799	R800-1199	R1200-1799	R1800-2499	R0-2499
Single-adult households	282	0	0	0	0	282
Two-adult households	249	174	7	0	0	431
More than two-adult households	150	230	77	43	12	513
Adult(s) and child(ren)	1 413	2 301	1 191	765	333	6 003
Total numbers in poverty	2 094	2 705	1 275	808	345	7 228
Total no. of people in category	2 417	3 394	1 941	2 103	1 817	11 672
3. In households receiving CSGs but not OAPs						
	R0-399	R400-799	R800-1199	R1200-1799	R1800-2499	R0-2499
Single- adult households	3	0	0	0	0	3
Two-adult households	4	3	0	0	0	7
More than two-adult households	1	5	3	0	0	9
Adult(s) and child(ren)	2 454	3 011	1 599	825	371	8 259
Total numbers in poverty	2 461	3 019	1 602	825	371	8 278
Total no. of people in category	2 521	3 267	1 922	1 449	1 033	10 192
4. In households receiving OAPs but not CSGs						
	R0-399	R400-799	R800-1199	R1200-1799	R1800-2499	R0-2499
Single-adult households	0	0	0	0	0	0
Two-adult households	2	0	0	0	0	2
More than two-adult households	35	92	26	10	1	165
Adult(s) and child(ren)	295	1 340	354	103	15	2 108
Total numbers in poverty	333	1 432	380	114	17	2 275
Total no. of people in category	506	2 549	932	743	309	5 039

5. In households receiving CSGs and OAPs

	R0-399	R400-799	R800-1199	R1200-1799	R1800-2499	R0-2499
Single-adult households	0	0	0	0	0	0
Two-adult households	0	0	0	0	0	0
More than two-adult households	2	0	0	0	0	3
Adult(s) and child(ren)	262	1 127	552	253	55	2 248
Total numbers in poverty	264	1 127	552	253	55	2 251
Total no. of people in category	304	1 523	1 023	687	280	3 817

The conclusion that van der Berg *et al* draw from their reading of the evidence is that further poverty alleviation can only come from growth. This conclusion is challenged by those who argue that the grant system should be made more generous.

A tentative comparison with the Hooegeven and Özler findings for 2000

One of the criticisms made of earlier versions of the present paper centred on the absence of results for any other year than 2004. My argument that presenting the results for a single year was justified, given the contentious nature of the van der Berg *et al* claim about the level of poverty in that year, failed to mollify critics. As a nod in their direction, the historical account below, which uses the Hooegeven and Özler (2004) results for 2000 and my figures for 2004, is offered (not without trepidation).

The Hooegeven and Özler paper makes use of two poverty lines. These have their origins in a piece of work carried out by a team at Statistics South Africa (Babita *et al*, 2003) of which Özler, a World Bank staffer, was a member. The following passage from their paper explains the provenance of the poverty lines in Table 5 below:

We used the “cost-of-basic-needs” approach ... to draw normative poverty lines for our analysis... According to these calculations, a reasonable poverty line for South Africa must lie between 322 Rand (lower-bound poverty line) and 593 Rand (upper-bound poverty line) per capita per month in 2000 prices. In this paper, we report poverty using the lower-bound poverty line as well as the \$2/day poverty line,³³ which is equivalent

³³ Hooegeven and Özler treat their \$2/day line as an “extreme” poverty line.

to 174 Rand per capita per month. The \$2/day poverty line is close to the poverty line used by Deaton (1997), and also reasonably close to our food poverty line of 211 Rand (2004, p.9).

Following Woolard and Leibbrandt (2001) I have used the 1997 Deaton line (calling it the ‘Deaton dollar-a-day’ line) elsewhere (Meth, 2005a). Deaton (1997, p.157) suggested an amount of R105 per person per month in 1993, noting as he did that it was “... much lower than poverty lines commonly discussed in South Africa, but is approximately equal to \$1 (U.S.) per person per day ...” Inflating this poverty line to 2000 prices produces a figure of R178.47, a little less than R6 per day.

Panel 1 of Table 5 reproduces the Hooegeven and Özler results for the two poverty lines $z = R174$ and $z = R322$. Panel 2 reproduces my results with $z = R178.47$ and $z = R322$, for the three different social grant payout combinations used in Tables 2 and 3, with an assumed under-reporting level of zero. Panel 3 is similar, except that the assumed under-reporting error level is 20 per cent. To jog the memory, Panel 4 reproduces the van der Berg *et al* results from Table 1 above.

Comparing my zero under-reporting error results for the country with the Hooegeven and Özler results (Panel 1) at the lower poverty level ($z = R178.47$ or R174), we see a modest fall in P_0 and P_1 if disability grants are excluded (Panel 2a), and a substantial drop if they are included (Panel 2b or 2c). The higher of the two poverty lines yields a P_0 of 0.507 (Panel 2a) and 0.487 (Panel 2b) or 0.494 (Panel 2c), and a P_1 of between 0.257 and 0.239, all a lot lower than the corresponding Panel 1 figures. Moving to the 20 per cent error levels in Panel 3 gives rise to huge falls in both P_0 and P_1 .

Understanding how these results compare with the van der Berg *et al* figures — it may be recalled from Table 1 that they claim that the drop in the headcount between 2000 and 2004, using a poverty line of R250 per month per capita (in 2000 prices), amounted to 3.1 million (from 18.5 million in 2000 to 15.4 million in 2004) — is much simpler if they are converted to headcounts. To do so, some consideration needs to be given to the way in which the Hooegeven and Özler results are to be treated. There are (at least) two ways in which this may be done. One is to accept them at face value, i.e., to assume that the repairs done by these authors to the 1995 and 2000 Income and Expenditure Surveys, and to the October Household Survey for 1995 and the Labour Force Survey for September 2000 from all of which their results derive, are adequate, and that as a consequence, errors of under-reporting do not significantly affect the outcomes.

Table 5 FGT Poverty Measures (ratios and gaps) in South Africa, 1995-2004

	P ₀ Headcount ratio	P ₁ Poverty gap ratio	P ₀ Head- count ratio	P ₁ Poverty gap ratio
1. Hoogeveen and Özler (2004) results	z = R174		z = R322.00	
South Africa 1995	0.32	0.11	0.58	0.27
South Africa 2000	0.34	0.13	0.58	0.29
2. Meth (2006) results: zero error	z = R178.47		z = R322.00	
a. Include CSGs & OAPs – South Africa 2004	0.314	0.122	0.507	0.257
b. Include CSGs & OAPs. DGs to hhs containing OAP recipients – South Africa 2004	0.292	0.116	0.487	0.243
c. Include CSGs & OAPs. DGs to hhs containing CSG recipients – South Africa 2004	0.286	0.106	0.494	0.239
3. Meth (2006) results: 20 per cent error				
a. Include CSGs & OAPs – South Africa 2004	0.269	0.097	0.468	0.224
b. Include CSGs & OAPs. DGs to hhs containing OAP recipients – South Africa 2004	0.250	0.093	0.448	0.211
c. Include CSGs & OAPs. DGs to hhs containing CSG recipients – South Africa 2004	0.243	0.084	0.453	0.208
4. van der Berg et al (2005) results	z = R250.00			
South Africa 1993	0.406	0.200		
South Africa 2000	0.413	0.205		
South Africa 2004	0.332	0.146		
Data source for Meth (2006, i.e., the present paper) results is the September 2004 LFS.				

The other is to assume that the criticisms made in the review of the poverty literature in South Africa by van der Berg *et al* (2005, pp.6–10), along with

the claim that most poverty studies overstate the extent of poverty in the country, are valid. Having done that, one can make use of the information presented in the van der Berg *et al* (2005) and Hooegeveen and Özler (2004) papers to attempt to estimate the possible extent of error in the latter for the year 2000.

There are a number of steps to be taken. Firstly, from the chart in the Hooegeveen and Özler paper showing cumulative distribution functions (Figure 1, 2004, p.34) one can gauge that the headcount ratio for a poverty line of R250 per capita per month is probably in the region of 0.45.

Next, we estimate the van der Berg *et al* poverty headcounts for the poverty lines used by Hooegeveen and Özler ($z = R174$ and $z = R322$ per month). There are two ways in which this could be done. The first relies on Figure 17 in the van der Berg *et al* paper, which shows estimated trends in poverty using different poverty lines (2005, p.39). The second would make use of a pair of cumulative density curves which appear to have been incorrectly labelled (van der Berg *et al*, 2005, p.40). For our purposes here, a pair of rough estimates extracted from the figures in their Figure 17 provide a good enough guide. Their lowest poverty line has a value of R2 000 per annum, i.e., R166.67 per month. This poverty line is very close to the lower line used by Hooegeveen and Özler. About 29 per cent of the population in the year 2000 fell below the R166.67 per month line. Let us assume, for argument's sake, that at the slightly higher poverty line (R174), roughly 30 per cent of the population would have been in poverty.

The topmost poverty line that van der Berg *et al* use is set at R4 000 per annum in 2000 prices, i.e., about R333.33 per month. This is not far from the R322 used by Hooegeveen and Özler. At the R4 000 level, van der Berg *et al* find about 49 per cent of the people in poverty in the year 2000. Lowering the line a little (to R322) would reduce the headcount ratio slightly. Let us, however, leave the headcount ratio at an assumed 49 per cent — as will become clear below, doing so is conservative.

To convert the headcount ratios (P_0) so derived into headcounts, we multiply by the population estimate figure for September 2000 (44 043 000) obtained from the Labour Force Survey conducted in that month. At the R174 (R166.7) poverty line where Hooegeveen and Özler find 14 975 000 people, the van der Berg *et al* figure would be about 13 213 000, a 13.3 per cent difference. At the R250 poverty line, the van der Berg *et al* figure is

18.5 million (see Table 1 above). The corresponding Hoogeveen and Özler number looks as though it would be about 19 819 000, a difference of 7.1 per cent. At the R322 (R333) line, Hoogeveen and Özler have 25 545 000 people in poverty, while the van der Berg *et al* number is in the region of 21 581 000, a difference of 18.4 per cent.

Stand back for a moment from this speculative exercise and recall that what we are trying to achieve is an estimate of the change in the number of poor between 2000 and 2004, using the Hoogeveen and Özler figures for 2000, and my estimates for 2004. My 2004 headcounts are larger than the Hoogeveen and Özler figures for 2000, so the larger the number living in poverty in the year 2000, the bigger the decrease in the poverty headcount. When, therefore, we decrease the size of the original Hoogeveen and Özler numbers on the grounds that their van der Berg *et al* counterparts are lower, we reduce the estimate of the impact of anti-poverty policy. In other words, the larger the percentage difference between the Hoogeveen and Özler and the van der Berg *et al* estimates, the smaller the impact of anti-poverty policy.

Table 6 contains the information necessary for the making of a set of estimates of the changes in poverty headcounts between 2000 and 2004. In the first panel of the table there are six variants (two sets of three grant settings) on my results for each of the two poverty lines discussed above. One set uses a zero under-reporting assumption, while the other set assumes that expenditure is under-reported by 20 per cent. As before, grant setting #1 gives the CSG and OAP to households. Grant setting #2 distributes the disability grant to households containing recipients of the OAP, and grant setting #3 gives it to households containing recipients of the CSG.

The results in Panels 2 and 3 are obtained using the assumptions discussed in the argument above. The set in Panel 2 assumes that the Hoogeveen and Özler figures are correct (i.e., that they do not over-estimate poverty). The second set (in Panel 3) assumes that they overstate poverty at both the R174/R178.47 and the R322 poverty lines by (a conservative) ten per cent. It is a simple matter, if we assume zero error in the Hoogeveen and Özler results, to obtain a poverty reduction effect of similar magnitude to the van der Berg *et al* figure of 3.1 million (at the R250 poverty line). Using grant setting #2 and an under-reporting error that is greater than zero, but less than 20 per cent, looks as though it would do the trick at both poverty levels.

Table 6 Headcounts in Meth (2006) and Hoogeveen and Özler (2004)

	0% error	20% error
1. Poverty headcounts (1000s) for 2004 in Meth (2006)	Grant	Grant
	setting #1	setting #1
z =R178.47	14 618	12 519
z =R322	23 604	21 794
	Grant	Grant
	setting #2	setting #2
z =R178.47	13 569	11 639
z =R322	22 662	20 867
	Grant	Grant
	setting #3	setting #3
z =R178.47	13 293	11 317
z =R322	23 011	21 095
2. Differences, 2004 – 2000	0% error	20% error
Zero error in the Hoogeveen & Özler results, Grant setting #1		
z =R178.47	-356	-2 456
z =R322	-1 941	-3 751
Grant setting #2		
z =R178.47	-1 405	-3 336
z =R322	-2 883	-4 677
Grant setting #3		
z =R178.47	-1 681	-3 658
z =R322	-2 534	-4 450
3. Differences, 2004 – 2000	0% error	20% error
10% error in the Hoogeveen & Özler results, Grant setting #1		
z =R178.47	1 141	-958
z =R322	613	-1 196
Grant setting #2		
z =R178.47	92	-1 838
z =R322	-329	-2 123
Grant setting #3		
z =R178.47	-184	-2 160
z =R322	21	-1 895

Note: Data sources of the population estimates for converting the headcount ratios in Table 5 above into the headcounts in this table, are the September 2000 and 2004 LFSs, Table 1.1, p.1 in each.

Matters take on a different complexion when we assume that Hoogeveen and Özler overstate poverty by (as little as?) ten per cent (as opposed to the 13–18 per cent suggested above). For a zero under-reporting error in my figures, poverty rises in grant setting #1, and is more or less constant in grant settings #2 and #3. It is the disability grant, distributed among more than a million households, that converts an increase in the headcount into rough constancy. This suggests that the claims of those who insist that the numbers of the poor are increasing may not be so outrageous after all. Although allowing the under-reporting error to rise to 20 per cent does away with the increase in the number of poor, it also suggests that anti-poverty policy has not succeeded in reducing the headcount by much more than one-and-a-half million or so.

This conclusion obviously relies heavily for its validity on the validity of my estimates of the level of poverty (and the invalidity of those of Hoogeveen and Özler). About the latter, there is nothing more to be said here. We can, however, look in more detail at the way in which my findings have been extracted from the survey data, a task to which we now turn.

4. Extracting poverty estimates from the surveys: The details

This section of the paper delves into the process by which the results are extracted from the data sets. It commences with a brief historical account of the ways in which the method of treating the data used here, and the model into which the data feeds, have evolved. The various amputations that have had to be done to tailor the data to fit the model are then discussed. This is followed by a discussion of the process of imputation I have called ‘purging’. The section ends with a look at the steps that have been taken to harvest social grant information from a deficient source, the Labour Force Surveys (LFSs).

Although the household surveys from which the poverty headcounts presented in Tables 2, 3 and 4 are drawn were not designed specifically for such a purpose, I have long been persuaded that useful information on

poverty could be extracted from them.³⁴ The paper of mine (written with my colleague Rosa Dias) to which van der Berg *et al* (2005, p.8) refer (Meth and Dias, 2004), represented a very first attempt to begin doing so. In December of the same year I published a second set of results which reduced considerably the increase in poverty reported in Meth and Dias (2004). Further refinements to the approach were embodied in Meth (2005a).³⁵ These did not alter headcounts by much. The results offered in the present paper are the outcome of extensive improvements to the method of wringing poverty estimates out of the household survey data. Although not directly comparable (the most recent estimates in Meth (2005a) are for the year 2002), the earlier results do not differ greatly from those presented above. This, however, has less to do with science than with good fortune — the original method, as we shall see below, does not have much to recommend it. That said, it should also be noted that results in the present paper are far from being the last word on the topic — the approach is capable of being refined still further.

Designed originally to estimate the impact of the social wage on the poverty of folk in households where expenditure was between R0–399 and R400–799 per month, the calculating engine used to extract the results presented above, has, as noted above, had to be revised extensively in order to enable it to do so. Apart from a slew of improvements made after the first paper in this series was published, the big change made to the engine, for the purposes of the present paper, is the addition of the data for households in

³⁴ Statistics South Africa's Ros Hirschowitz insisted on numerous occasions, while I was a member of the Statistics Council, that it was not appropriate to use the household surveys as poverty measurement instruments. By common agreement, the use of consecutive surveys as indicators of change over a short period of time is not advisable. If, however, a competently administered survey were conducted over a lengthy period of time, say, a decade or so, it is hard to imagine what objection there could be to comparisons being made between conditions at the beginning and end of the period. Unfortunately, we do not have a long run of more-or-less identical household surveys — there have been numerous important changes since the first truly national October Household Survey was conducted in 1994. In addition, doubts as to the reliability of both the LFS and the GHS have been expressed. Neither of these conditions, however, absolves us of the duty, in the absence of other data, to extract every useful bit of information from the surveys that have been conducted. I suspect that we are still some way away from exhausting the available resources.

³⁵ It is a pity that van der Berg *et al* did not cite this later paper rather than its more primitive forebear. The blame, however, lies with us for accepting the offer of publication of what was really a first draft.

the expenditure categories R800–1199; R1200–1799 and R1800–2499 per month. With that, however, and with the demand that it cope with more than two social grants, it has reached the point where the price of getting it to do its job properly is an almost intolerable increase in the complexity of an already cumbersome device. This makes it appropriate to pause for a moment to reflect on the way in which the design of the model has unfolded.

It is common, when measuring income distributions (for that, ultimately, is what the calculating engine does) to sum income from all sources accruing to individuals and/or households, and then to use the results to rank them from richest to poorest. This paper uses expenditure categories to measure household wellbeing. Debates about the relative merits of income and expenditure aside, the reason for this choice is that with a minimum of supplementary information (household composition and mean household size), the dimensions of social distress (or wellbeing) could be conveyed in a manner that was as easy to understand, as it (apparently) was to extract from the data. In particular, the dimensions of such phenomena as workerless households, unemployment, and service delivery of all sorts become very interesting when viewed through the prism of household expenditure categories. This approach, with all its drawbacks, is the foundation upon which the model rests.

In essence, the approach consists in the estimation of the maximum potential consumption of each individual in the household survey. Individuals are then ranked by expenditure level to form a distribution. To accomplish this, households are first sorted by type (by type I mean household composition in terms of the number of adults and/or adults and children it contains) and by expenditure category. In, for example, a household in the R0–399 per month expenditure category that contains a single (adult) individual, the maximum level at which that individual could conceivably consume is R399 per month. Maximum possible consumption levels for a household in this category containing two individuals are obviously R199.50 each, if available income is equally shared, and so on.³⁶

³⁶ The assumption of equally-shared consumption within the household is problematic. In the absence of detailed intra-household distribution, though, there is little other one can do than to deploy it.

Households fall into one of four types or compositions (as we saw in Table 4): single adult; two adult; multiple adult, and adult and child.³⁷ The initial data sorting process takes in as many adults and children as are reported in the surveys — there are thus examples of households containing 15 children and three adults, or in one spectacularly isolated case, 14 adults and 11 children.³⁸ These outliers, which may or may not represent real households, are eliminated by truncating the adult and child households to those containing a maximum of six adults and 12 children. Doing this eliminates roughly five per cent or so of the individuals (at most) from the adult and child household population. Multiple adult households are truncated at ten individuals. The effect of this is negligibly small.

Truncating the data sets (done, it must be acknowledged, to ease the programming burden) is conservative, in the sense that it will cause poverty levels to be understated. Because most of the households that are removed contain large numbers of people, individual consumption levels within them are low. In the highest expenditure category (R1 800–2 499 per month), for example, the deflated value of the upper boundary is R2 018 per month. At the zero error level, any household in this category that contains more than eight individuals will fall below the poverty line. Implausibly large under-reporting errors are required to lift bigger multiple-person households out of poverty.

Within any expenditure category, a household can fall into one and only one of 82 cells (one each for single- and two-adult households, eight for multiple-adult households, and 72 for adult-and-child households). Once the households have been sorted into the correct household type, i.e., allocated to the appropriate cell, the individuals in those households need to be distributed within that cell. Income distribution studies frequently reveal

³⁷ The model is capable of converting children to adult equivalents (at any desired child-cost ratio). It is also capable of taking account of household economies of scale. Because van der Berg *et al* work with simple per capita poverty lines, these features of the model do not need to be used. Incidentally, the model is also capable of estimating the impact on consumption levels of the ‘bankable’ effects of social spending on individuals (the difficulty, of course, is estimating how large this effect is).

³⁸ The surveys point to the possible (likely?) existence of child-headed households, i.e., households in which there is no adult present. For the moment, the numbers are small enough to be ignored. As the AIDS epidemic takes its course, this may no longer be the case in the future.

lognormal distributions. The arithmetic involved in employing assumed distributions of this sort in 82 cells in each of the expenditure categories examined is unwieldy, to say the very least. To make life easier (without, as will be seen below, apparently sacrificing too much), simple linear distributions within each cell have been assumed. The means of such distributions are determined by the assumed value of the bottom boundary in each cell. For the expenditure categories R400–799; R800–1199; R1200–1799 and R1800–2499, the lower bound is fixed by definition. Means are thus approximately R600; R1000; R1500 and R2100 respectively.³⁹ For the bottom category, R0–399, a bottom bound of R200 has been assumed. This fixes the mean at R300 per month. These assumptions will bias the results in a conservative direction, i.e., they should cause poverty levels to be understated, a set of circumstances to which we will return in the discussion below on the relationship between nominal and actual or effective under-reporting errors (Section 5 of the paper).

Imputation of expenditure categories from income estimates

A moment's reflection on the purging process reveals that it is nothing other than a means of adjusting for expenditure under-reporting. The trick (an extremely difficult one) is to measure the extent of this adjustment. Before we get to that (in Section 5 of the paper), a thorough understanding is necessary of the way in which the imputations have been carried out.

If one plots the information on income obtained in the LFS (and GHS) against that on expenditure, patterns similar to those in Tables 7 and 8 emerge. Depicting the relationship between these two variables, as is done in Tables 7 and 8, makes the need to 'purge' the data immediately obvious. To do so (i.e., to impute expenditure levels from income levels), one simply pushes any household for which $Y > E$, into the expenditure category implied by Y . Table 7 presents the raw or unpurged distributions for two types of households, those containing at least one adult and one child, and

³⁹ An assumption that means lie below mid-points of linear distributions would be in line with the findings reported by Keswell and Poswell (2002). Working with the 1997, 1998 and 1999 OHSs, that gave categorical (income bracket) estimates, they devised a technique for dealing with the problem caused by the absence of actual income (point) estimates. Critical of approaches that use the mid-point of an income bracket, these authors (who prefer to assume a lognormal distribution of income) show that the use of the mid-point assumption overstates true means, leading to over-representation of better-off individuals (see their Appendix B, pp.42–43 and their Figure 1, p.50).

those containing a single adult only, gathered by the September 2004 LFS. The figures in the cells in Table 7 above the total line are column percentages — those immediately below it are numbers of households. Incomes are earnings from all employment of persons in the household. Migrant remittances are not included. On the stepped lines in Table 7 (and Table 8) income is equal to expenditure, and (obviously) above the lines, income is less than expenditure.

The mechanics of the purging process are simple — from the bottom expenditure category, for example, 11.3 per cent of households with a reported income in the range R400–799 are relocated into the second expenditure category. A further 5.5 per cent of households in the expenditure category R0–399, along with 8.2 per cent of the households in expenditure category R400–799, are shifted to the expenditure category R800–1199. The process is repeated until there are zeroes everywhere below the stepped line. As a method of imputation, this technique is undoubtedly crude — ignoring, as it does, all other household characteristics. More sophisticated imputation techniques exist.⁴⁰ Until an opportunity to acquire the skills required to apply them presents itself, purging, for all its crudeness, will be used; although imputation of an expenditure category from an income estimate is crude, it is not hopelessly so.

The results in Table 7 (and Table 8) appear not to be unique. Very similar patterns emerge from the 2004 GHS. They were present also in the 2003 GHS and the September 2003 LFS. On two of the three occasions on which the present paper has been given an airing before an academic audience, the imputation technique used in it has come under fire. Expressing surprise at not finding expenditure greater than income, critics claimed that the surveys results were too weak to bear the burden of the imputations. The relative stability of the relationships between income and expenditure in four separate surveys, however, gives the lie to such doubts. While it remains necessary to explain why a certain proportion of households understates expenditure, it is also the case that critics have to explain how two independent samples (successive LFSs use part of the same sample, as do GHSs) generate such similar patterns of under-reporting. To note that the surveys generate similar patterns is not to suggest, however, that such an outcome is either desirable or acceptable. It is important, therefore, for Statistics South Africa to develop the means to detect inconsistencies (in

⁴⁰ The sequential multiple regression technique used by Barnes *et al* (2005) on the 2001 South African population census data, is a case in point.

households where income and expenditure differ substantially) while households are being surveyed, and to rectify these during the interview.

Relocating the households in the cell immediately below the line in any expenditure category imparts a conservative bias to the poverty estimates (they will tend to be understated). In the expenditure category R400–799 per month, for example, a hypothetical household that consumed R395 per month out of earnings of R405 (admittedly, an extreme example) would meet the criteria for relocation. Assumed mean consumption would rise from R300 to R600 per month in the process. Even if there are not many cases like this, the overall effect of the purging process will still be conservative.

To carry out the purging process, it is necessary to separate households containing migrants from those not containing migrants and then to relocate all inappropriately located households in the manner discussed above. The details of the process are of no immediate concern to us here.⁴¹

Most striking of the results is the large number of households at the bottom of the distribution in which there are no earned incomes, and by implication, no workers. This is linked with another feature of the results, namely, the rapidity with which the percentage of households above the line falls as expenditure rises — by the time the fifth expenditure category (R1800–2499 per month) has been reached, it has dropped from almost 74 per cent to below 40 per cent. An important part of the explanation for this is the presence of so many zero-income households in the bottom expenditure categories. Among the zero-earned-income households, a significant minority receives migrant remittances; most of those that do not survive on social grants, chiefly the child support grant, the state old age pension, and the disability grant. To check the reliability of the zero-income results, the numbers of households reporting zero earned incomes were compared with the numbers of households containing working age people in which nobody was employed and which reported zero incomes. The difference between them (a rough estimate of the ‘implausible’ zeroes) appears in the last rows of the two panels in Table 7.

⁴¹ Those interested in the mechanics of the process can trace it in the relevant spreadsheets.

Table 7 Total monthly household income from all employment by monthly household expenditure (Labour Force Survey September 2004)

	Total household expenditure, R/month									Total	
	R0-399	R400-799	R800-1199	R1200-1799	R1800-2499	R2500-4999	R5000-9999	R10 000+	Don't know		Refuse
Zero earned income	59.3	51.0	35.1	25.2	16.5	10.8	16.0	19.2	43.8	84.4	39.8
R1-99	0.8	0.3	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.3
R100-199	4.4	1.4	0.8	0.4	0.2	0.0	0.0	0.5	0.9	0.0	1.5
R200-399	9.1	3.9	2.9	1.7	1.4	0.2	0.0	0.1	0.7	0.0	3.7
R400-799	11.3	14.5	7.9	4.9	1.4	1.0	0.6	0.0	3.7	0.8	8.7
R800-1199	5.5	8.2	8.3	5.1	1.0	0.4	0.0	0.1	5.9	0.0	5.6
R1200-1799	4.7	8.3	14.2	15.2	7.1	1.3	0.4	0.9	8.8	2.7	7.7
R1800-2499	2.3	4.7	9.9	14.5	11.5	2.4	0.9	0.7	5.7	0.2	5.7
R2500-4999	1.9	5.7	14.4	22.4	30.6	28.0	3.8	2.7	13.1	0.0	10.8
R5000-9999	0.5	1.8	4.8	8.1	23.1	42.0	35.6	6.1	10.9	10.4	9.3
R10000+	0.0	0.2	1.5	2.4	7.5	13.7	42.6	69.7	6.7	1.3	6.9
Total	1 451 000	2 315 000	1 147 000	618 000	471 000	546 000	468 000	220 000	102 000	57 000	7 390 000
% above line	73.7	71.1	55.1	52.6	39.0	44.3	57.4	99.9			
No. of hhs with no workers & zero income	815 523	1 082 664	350 994	128 871	48 481	18 551	7 668	2 226	14 099	3 958	2 473 530
No. of hhs reporting zero income	860 933	1 179 441	402 754	155 929	77 595	59 120	74 639	42 123	44 340	48 035	2 944 168
% difference	5.3	8.2	12.9	17.4	37.5	68.6	89.7	94.7	68.2	91.8	16.0

Table 7 Continued

	Total household expenditure, R/month										Total
	R0-399	R400-799	R800-1199	R1200-1799	R1800-2499	R2500-4999	R5000-9999	R10 000+	Don't know	Refuse	
Zero earned income	51.6	38.2	24.6	24.2	24.0	29.8	27.6	25.8	62.6	96.9	36.8
R1-99	0.5	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
R100-199	2.2	0.5	0.3	0.2	0.0	0.3	0.0	0.0	0.5	0.0	0.8
R200-399	8.5	2.3	1.9	1.4	0.5	0.1	0.0	0.0	0.7	0.0	3.1
R400-799	14.3	17.5	3.9	1.3	0.7	0.7	0.0	0.3	4.7	0.0	8.6
R800-1199	7.2	11.9	10.9	2.0	0.7	0.4	0.4	0.0	1.6	0.3	6.4
R1200-1799	7.2	11.4	20.2	16.3	2.6	1.2	0.8	0.9	4.7	0.0	8.9
R1800-2499	3.5	7.4	13.5	17.6	12.8	2.3	0.3	0.9	0.9	0.0	7.0
R2500-4999	4.0	8.6	17.8	26.0	35.7	22.2	3.7	2.7	1.2	1.4	12.5
R5000-9999	0.9	1.9	5.8	10.0	17.1	28.5	23.4	6.7	6.6	1.1	8.2
R10000+	0.0	0.3	1.2	1.1	5.9	14.5	43.7	62.7	16.5	0.8	7.6
Total	1 259 000	1 271 000	626 000	447 000	359 000	495 000	361 000	170 000	94 000	67 000	5 146 000
% above line	62.8	58.6	41.6	45.3	41.3	57.0	56.3	99.9			
No. of hhs with no workers & zero income	617 000	450 000	132 000	97 000	70 000	106 000	44 000	12 000	36 000	15 000	1 574 000
No. of hhs reporting zero income	649 000	485 000	154 000	109 000	86 000	148 000	100 000	44 000	59 000	65 000	1 895 000
% difference	5.0	7.3	14.4	10.6	19.6	28.5	56.1	73.0	38.8	78.4	17.0

Table 8 The impact of the purging process on household distribution and incomes (Labour Force Survey September 2004)

Total monthly household income, by category, from all employment, all households (1000s of households)									
	R0-399	R400-799	R800-1199	R1200-1799	R1800-2499	R2500-4999	R5000-9999	R10 000+	Total
Zero earned income	1 549	1 681	559	265	163	207	174	86	4 897
R1-399	353	168	59	20	10	4	1	2	615
R400-799	344	559	115	36	9	10	3	1	1 083
R800-1199	172	342	166	42	8	5	2	1	740
R1200-1799	159	337	290	167	43	14	6	4	1 031
R1800-2499	79	203	198	169	100	24	6	4	787
R2500-4999	77	242	276	255	273	263	32	11	1 442
R5000-9999	19	65	91	96	171	371	251	26	1 110
R10000+	2	10	24	19	57	147	357	260	896
No. before purging (1000s)	2 751	3 602	1 776	1 065	829	1 041	828	390	12 598
No. after purging (1000s)	1 901	2 751	1 410	1 313	979	1 645	1 282	1 003	12 597
% change	-31	-24	-21	23	18	58	55	157	0
% above line before purging	69	67	50	50	40	50	57	100	
% one category below line	82	76	67	65	73	86	100		

Table 8 Continued***Estimate implicit correction for under-reporting***

Assumed mean										
household	monthly income	R0-399	R400-799	R800-1199	R1200-1799	R1800-2499	R2500-4999	R5000-9999	R10 000+	Total
	300	570	555	185	85	52	63	52	26	1 653
	600	206	335	69	21	5	5	2	0	649
	1000	171	341	165	42	8	5	1	0	739
	1500	238	505	435	249	64	21	8	5	1 546
	2100	164	426	415	355	209	50	12	7	1 651
	3750	288	907	1 035	954	1 020	983	118	39	5 404
	7500	142	482	680	718	1 276	2 778	1 880	189	8 324
	25000	28	249	598	472	1 417	3 653	8 913	6 488	22 392
	Face value expenditure level	825	2 161	1 775	1 597	1 741	3 902	6 208	9 732	27 940
	Expenditure after purging –									
	(i)	570	1 096	932	1 575	1 698	5 331	8 148	22 085	42 386
	Expenditure after purging –									
	(ii)	570	1 650	1 409	1 969	2 056	6 166	9 611	25 028	48 460

Note: Estimates in this panel are R millions/month in current prices. 'Don't knows' and 'Refuses' have been omitted

Also striking, if we believe the income data provided, is the fact that a significant proportion of households apparently misreports expenditure levels. While there is little difficulty in reconciling the expenditure and income reports of a household which has the former above the line, and the latter in the category immediately below the line, when the two are further apart than this, the difficulties multiply. In the Table 8 figures, some 18 per cent of households in the bottom expenditure are more than one income category below the line, as are 24 per cent of households in expenditure category R400–799, 33 per cent of households in the category R800–1199, 35 per cent in the category R1200–1799 and 27 per cent in the category R1800–2499 per month.

Unless one can find some testable hypothesis that could explain why poor people should lie consistently about their incomes, their expenditure levels and their labour market status, the figures in Table 7 suggest that in South Africa, under-reporting of income and expenditure is least severe among the poorest, becoming steadily worse as expenditure levels rise. By the same token, of course, it is also necessary to explain why some proportion of the population gives inconsistent answers about income and consumption. Except in those cases where income is only one category above the consumption upper bound, appeals to savings to explain the differences (using the identity $Y \equiv C + S$) cannot succeed (implied savings ratios would be too high). Some respondents may have greater difficulty recalling consumption than income levels. This explanation sounds more than a little weak, however, when income exceeds expenditure by a substantial margin.

Further progress in explaining the reported errors is not possible with the information at our disposal. Let us, therefore, take leave of the problem, and consider instead the outcome of the purging or imputation process. The upshot of the reported errors (disagreements between income and expenditure estimates) is a migration of the whole distribution towards its better-off pole. As may be seen at the bottom of the upper panel in Table 8, the number of households in the bottom expenditure category falls by 31 per cent, and that in the category R400–799 by 24 per cent. The three top expenditure categories each show substantial increases (157 per cent at the very top, and 55 and 58 per cent respectively, in the next two down).

To examine the income/expenditure implications of this migration we consider the results in the bottom three rows of the table. The values in the row labelled 'Face value of expenditure' are obtained by multiplying the

unpurged number of households in each expenditure category by the assumed monthly mean expenditure level for that category; for example, in the R0–399 per month category, mean expenditure is assumed to be approximately R300 per month.⁴² Multiplying this by 2.75 million yields a sum of R825 million. The sum of these unadjusted totals is R27.9 billion.

The values of expenditure after the purging process has been completed, have been estimated in two ways. The first of them, labelled ‘Expenditure after purging (i)’, takes the income figures more or less as reported,⁴³ and applies these to the numbers of households remaining after purging. The R1.096 billion in expenditure category R400–799 consists of the sum of the R555 million and R335 million above the line, plus the R206 million from expenditure category R0–399. The adjustment errs on the side of generosity, because the zero-earned-income households are assumed to have an income of R300 per month, as are those with reported earned incomes of R1–300 per month. The sum of the values generated by this set of calculations is R42.4 billion, which is 52 per cent greater than R27.9 billion.

The row labelled ‘Expenditure after purging (ii)’ takes the number of households after purging and multiplies it by the category mean expenditure level. In the bottom expenditure category (R0–399), for example, total income (or expenditure — I am conflating the two at this point), amounts to R570 million. In the second expenditure category (R400–799), 2.75 million households multiplied by an assumed R600 per month equals R1.65 billion. The sum of the products in the nine expenditure categories of this even more generous approach is R48.5 billion, which is 73 per cent more than the value obtained in the row labelled ‘Face value of expenditure’. It is this set of estimates that finds its way into the model that estimates poverty in the

⁴² The mean income or expenditure level of households in the open category (R10 000 plus) is assumed to be R25 000 (in 2004 prices). Various techniques have been developed to help researchers guess at the mean of the open category. There is no way to discover the ‘true’ mean. Simkins (2004, Table 7, p.14) estimated that individual incomes at the 90th, 95th and 99th percentiles in 2001 were R62 896, R112 838 and R331 084 respectively. Given a mean household size of about 3.46 in the open expenditure category (using the purged estimates), the figure offered here ‘feels’ like a reasonable estimate for the richest five per cent or so of the province’s households.

⁴³ The information on incomes in Table 8 is gathered by question 4.15a in the LFS. It reads “What is ...’s salary/pay at his/her main job? Including overtime, allowances and bonus, before any tax or deductions.”

present paper. The proportional and absolute impact of this adjustment is determined by the relative densities within each expenditure category of zero- and low-income households. Since these are located overwhelmingly at the bottom end of the income distribution, it is there that the effect of the adjustment is greatest. This imparts a conservative bias to the results, i.e., it will tend to cause poverty to be understated. A consequence of the purging operation, a major upward revision of expenditure estimates, is that the total in the last row of Table 8 (R48.5 million) requires an adjustment of only 7.5 per cent to bring it into line with the national accounts estimate of compensation of employees (R624.5 billion for the year in current prices, or R52 billion monthly).⁴⁴

The actual process of imputation is more elaborate than is suggested by the figures in Tables 7 and 8.⁴⁵ In practice, it is necessary to dip into each expenditure category eight times. In the first place, there are two sources of earned income, one from workers within the household, the other in the form of migrant remittances. Two types of social grant, the CSG and the OAP, are also taken into account. That means distinguishing four ‘types’ of household: those that receive no grants; those that receive only the OAP; those that receive only the CSG, and those that receive both CSGs and OAPs. Each ‘type’ of household may also have either or both earned income (within the household) and migrant remittances. Migrant remittances are, however, not included in the income measure in Tables 7 and 8. Although remittances only find their way into a minority of households in the bottom expenditure categories, they make an important, nay a vital contribution to household wellbeing.⁴⁶ As noted above, the decision to base the primary estimates of the headcounts on the LFSs rather

⁴⁴ See *South African Reserve Bank Quarterly Bulletin*, March 2006, p.S-112.

⁴⁵ No attempt has been made to impute expenditure categories to the ‘Refuses’ and ‘Don’t knows’. Typically, they account for somewhat less than three per cent of all responses to the expenditure question in the household surveys. Judging by their relative percentages of implausible zero incomes, the latter look as though they belong somewhere near the middle of the distribution, while the former look as though they are drawn from the R10 000-plus expenditure category. Neither is likely to have much effect on the poverty estimates.

⁴⁶ The purged data suggest that among the 1.8 million households in expenditure category R0–399, 24 per cent have migrant workers attached to the household. Among the 2.6 million households in the category R400–799, the figure is 25 per cent. It is 19 per cent in the category R800–1199 among 1.4 million households, and seven per cent among the 1.3 million households in the category R1200–1799.

than on the GHSs was made on the grounds that both the September 2003 and September 2004 LFSs contain information on migrant remittances.⁴⁷

Adding in social grants and removing implausible zeroes

Purged of households for which income is greater than reported expenditure (the possibility that some of this income might be saved is ignored), the distributions in Tables 7 and 8 need to be adjusted to reflect (a), the social grants that accrue to households, and (b), implausible zeroes. Social grants have been accommodated by expanding the structure of the calculating engine in a manner, and to an extent, that would have delighted Heath Robinson.⁴⁸ Households (of all types or compositions) are separated into the four groups presented in Table 4, namely, those into which no social grants find their way; those receiving child support grants but no old age pensions, those receiving old age pensions but no child support grants, and those receiving child support grants and state old age pensions. Distributions of household by type (composition) are then estimated. The results for each category of household (no grants, CSG only, CSG and OAP, or OAP only) are then exported to separate (subsidiary) calculating engines for processing. Having been sorted into distributions, the results are re-exported to the main calculating engine. Summing the different distributions allows for a chart like Figure 1 (shown later in this section) to be constructed.

Incorporating information on child support grants and state old age pensions into the model entails a number of compromises. For the purposes of the present paper, it is necessary to be aware of the fact that the LFS and GHS asked questions about social grants in different ways. In the GHS, each individual in the household who gave a positive response to a question on whether or not they had made use of welfare services (questions 1.47 and 1.48 in the 2004 GHS), was directed to question 1.50 where they were required to provide information on which of the various grants they had

⁴⁷ Even though the General Household Surveys do not collect the necessary information on migrants, the close temporal proximity of the surveys in question, and the broadly similar results they throw up in Table 2 above, must surely suggest that the conclusions on poverty that would be drawn from the GHSs if only they collected migrant data would be similar to those drawn using the LFSs.

⁴⁸ Readers who are too young to understand the reference are invited to Google Heath Robinson. Anyone taking up the offer above to make available the spreadsheets that make up the model and the STATA Do Files associated with it, will also receive instruction in how to drive the thing (a service of which they will be in sore need).

received. Answers to this question made it possible for a record of total grants paid to be constructed, one which could then be compared with the administrative data collected by the national Department of Social Development (the so-called SocPen data base).⁴⁹ In the LFS, by contrast, the question sought to discover whether any member of the household received any of the social grants in a list (question 7.28 in the September 2004 LFS). The LFS was thus incapable of detecting the receipt of multiple grants (such as the child support grant, or the state old age pension), in any household. Some means has to be found for dealing with the LFS's deficiency. We begin by examining, in Table 9, the respective distributions of CSGs by expenditure category revealed by the 2004 GHS and September 2004 LFS.⁵⁰

That there were 12.1 million households in the GHS and 12.6 million in the LFS, conducted a mere two months later, is a little worrying. So, too, is the fact that there are differences in the distributions of the households among the different expenditure categories. In the bottom category, for example, where the GHS has 2.25 million households, the LFS has 2.75 million. Still, as was noted in the commentary on Table 2 above, after the first data purging stage, the differences between the poverty headcounts yielded by the two surveys were not all that large (about half-a-million or so).

Of more immediate interest here is the fact that the two surveys find roughly the same number of households reporting that at least one CSG had been received. Distributions of CSGs among households by expenditure category differ, but not dramatically. Crude household coverage rates (crude, because I have simply expressed households receiving at least one CSG as a percentage of households containing at least one child under the age of 15 years), in the expenditure categories with which we are

⁴⁹ There were 4 289 932 CSG payments on 1st April 2004 and 5 648 800 on 1st April 2005. The ratio of CSGs to caregivers (beneficiaries) on the latter date was 1.504, which is close to the average number of CSGs per household obtained from the 2004 GHS (1.557).

⁵⁰ The GHS continues to collect data on a wide range of welfare grants. The LFS collected the more limited information described above up until September 2004. All of the household questions (40 of them in Section 7 of the September 2004 questionnaire), which included the questions on social grants, have been deleted from the September 2005 questionnaire. This means that if a household survey is to be used in the future for measuring poverty, the task must fall, by default, onto the shoulders of the GHS.

concerned, are close enough to allow us to use the GHS results to help fill the gap caused by the shortcomings of the LFS. Some compromise, however, is unavoidable. If the GHS were the primary instrument being used to estimate headcounts, we could simply tell STATA to add the appropriate amount to the income of every individual concerned. Instead, the mean number of CSGs per household has been awarded to all households containing children that reported receiving a CSG.

This cavalier approach does harm to the truth by ignoring the fact that some proportion of the households receiving CSGs contains only one child (in the expenditure category R400–799, for example, 97 000 of the 615 000 households receiving the grant contained only one child). The existing calculating engine is poorly equipped to cope with a problem of this sort. It would be possible to devise a way around it, but the only way that suggests itself is exceedingly cumbersome. When consideration is given to the relatively small impact that the child support grant has on household poverty (for the typical household, full benefit dilution sees to it that each individual in the household benefits to the tune of a mere R40 or so), this compromise is deemed to be a small price to pay for the larger benefit of taking the grant into account.

Pensions have been treated in a similar way. Although they present similar problems as far as households containing only a single pensionable adult are concerned, the intensity of the problem is reduced somewhat by virtue of the fact that the average number of pensions per household among those households receiving at least one pension is much lower than the corresponding ratio for CSGs. Of course, both problems present themselves simultaneously in households receiving both pension and CSG. Fortunately (for the researcher, rather than the household), such households are a small minority.

As far as the (apparently) implausible zeroes are concerned, the usual technique for dealing with this problem is to use sequential regression software to impute expenditure or income categories to the implausible zero households. Lacking access at this point to such sophisticated devices, the problem is addressed by the arbitrary method of simply removing the households concerned from the distribution. Tables 7 and 8 give a rough indication of the impact of removing the ‘implausibles’. They overstate the case. Not only are the Table 7 and Table 8 figures unpurged (even in a simple manner), they also take no account of migrant incomes and social grants.

When proper purging is carried out, the distribution of households across the various expenditure categories changes quite a bit. The tally from the purged data set suggests that for the September 2004 LFS figures, some 180 000 households would be removed because their reports of zero income are implausible. These households probably contain somewhere near 700 000 people. Many of them are likely to be below the poverty line, so that the method of dealing with implausible zeroes, although arbitrary, is nonetheless conservative, in that it reduces the poverty headcount.

Once migrant remittances and social grants have been included in household incomes, and the necessary purging operations have been carried out, per capita consumption levels can be estimated. To estimate headcount ratios and poverty gap ratios, all that is now required is a poverty line (or lines) in constant prices of the relevant year. Poverty lines in van der Berg *et al* (2005) are given in prices of the year 2000. It is necessary to transform (deflate) the expenditure category boundaries, which are given in the current prices of 2003 and 2004, to reflect this.⁵¹ For this purpose, the all items Consumer Price Index (CPI) series published by the South African Reserve Bank has been used.⁵²

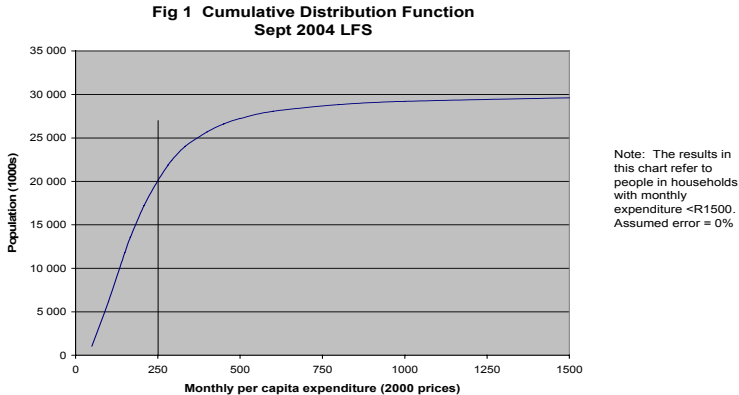
⁵¹ R400 in the current prices of 2004, for example, when expressed in the prices of the year 2000, amounts to about R323.

⁵² There is a case to be made for using the very low income CPI measured by Statistics South Africa. The problem is that the series does not seem to have been repaired after the 2003 CPI débâcle (the experience of having an outside commentator point out that the accommodation rental component of the CPI was causing inflation to be overstated).

Table 9 Child support grants in the 2004 GHS and September 2004 LFS (unpurged data)

	R0-399	R400-799	R800-1199	R1200-1799	R1800-2499	R2500-4999	R5000-9999	R10 000 plus	Don't know	Refuse	Total
2004 GHS											
Total no of hhs receiving csg (1000s)	598	1 104	552	293	127	90	13	3	21	3	2 800
% of all hhs receiving csg	21.4	39.4	19.7	10.5	4.5	3.2	0.5	0.1	0.8	0.1	100.0
Avg no of CSGs per hh	1.48	1.58	1.62	1.58	1.60	1.39	1.28	1.13	1.43	1.02	1.56
Nos of eligible hhs (1000s)	1 104	2 084	1 084	741	435	633	397	162	81	60	6 775
(contain child <15)											
Household coverage (%)	54.2	53.0	50.9	39.5	29.2	14.2	3.3	1.9	25.9	5.0	41.3
Total no of hhs in category (1000s)	2 255	3 487	1 724	1 265	818	1 216	784	269	184	149	12 143
% of all hhs	18.6	28.7	14.2	10.4	6.7	10.0	6.5	2.2	1.5	1.2	100.0
Sept 2004 LFS											
Total no of hhs receiving csg (1000s)	780	1 088	569	220	124	56	13	3	23	4	2 875
% of all hhs receiving csg	27.1	37.8	19.8	7.7	4.3	1.9	0.5	0.1	0.8	0.1	100.0
Nos of eligible hhs (1000s)	1 366	2 164	1 089	585	440	500	429	192	91	53	6 904
(contain child <15)											
Household coverage (%)	57.1	50.3	52.2	37.6	28.2	11.2	3.0	1.6	25.3	7.5	41.6
Total no of hhs in category (1000s)	2 750	3 603	1 775	1 065	829	1 041	828	390	197	124	12 597
% of all hhs	21.8	28.6	14.1	8.5	6.6	8.3	6.6	3.1	1.6	1.0	100.0

Given per capita consumption levels, it is possible to plot cumulative distribution functions. Figure 1 below does this for the zero under-reporting error assumption.



With a graphic representation of the function before us, it is possible to see, at a glance, the implications of some of the assumptions used to derive the poverty estimates. It may be recalled, for example, that it was necessary, from an operational point of view, to assume that the distribution of individuals within each of the 82 cells isolated in the data set is linear. As may be seen in Figure 1 below (which uses the zero under-reporting error data, and which excludes the disability grant), the resulting cumulative distribution function is not linear — it is roughly lognormal.

For the assumed conditions, Figure 1 suggests that in 2004, there were about 12 million people whose monthly expenditure (from all sources of income) was about R150 or less per month, and about six million at the R100 mark. Beyond reassuring us that the distributions resulting from the assumptions used in the study are not outlandish (in that they differ inexplicably from those observed by other researchers), graphical representations such as that in Figure 1 do not contribute greatly to the present enterprise. We need, therefore, no longer concern ourselves with them.

Looking back on the description above of the purging (or imputation) process, we see that it involves the prior sorting of households into the four groups: receives no grants; receives child support grant; receives old age pension; receives both child support grant and pension. After this, it is necessary to (a) remove all households whose income includes some migrant remittances and (b) remove all households whose income does not include any remittances, from the expenditure category in which the unpurged data locates them, if that income exceeds the upper bound of the expenditure category. Such households are then relocated to the appropriate expenditure category. The final worksheet in the pods containing the purged data sets (one each for each expenditure category) thus contains four data sets, each of which is exported to the appropriate calculating engine, for conversion into a distribution.

Rough-and-ready as the purging (imputation) process described above undoubtedly is, the similarity of household distributions in recent independent surveys (two in 2003 and two in 2004) lends confidence to the assertion that its results are conservative. This should be confirmed when the imputations have been checked using one of the more sophisticated regression packages available, a task that will be tackled as soon as time permits.

5. Estimation of actual from nominal under-reporting errors

One of the features of Table 2 above is that as the assumed nominal under-reporting error increased, the poverty headcounts began to converge on the van der Berg *et al* figure of 15.4 million in 2004. When the error was large enough, they fell below 15.4 million. To make the poverty headcount of 20 million obtained using the September 2004 LFS figures fall to the van der Berg *et al* figure, the Panel 2 figures in Table 2 tell us that it is necessary to adjust consumption expenditure for all individuals in the distribution upwards by about 50 per cent. Not shown in the table, but easily ascertained, is the information that if the total value of disability grants is treated as part of the poverty alleviation package, the nominal error would fall to about 38 per cent.⁵³ Estimating the size of the aggregate or nominal

⁵³ These results are obtained using the assumption that the same nominal error level holds in each of the five expenditure categories. This is not essential — the model

correction factor that would have to be applied to the purged expenditure estimates to obtain the same number of poor as do van der Berg *et al* in simplicity itself. This is because the calculating engine constructed to churn out poverty headcounts was originally built to allow for the use of the Deaton (2003) under-reporting factors discussed below (or any other value).

There is, however, a world of difference between knowing that a blanket upward adjustment of consumption levels by 38 per cent, or 50 per cent, or whatever, is required to produce certain headcounts,⁵⁴ and knowing what this implies for the salaries and wages which form the basis on which consumption must rest. The purpose of this section of the paper is to explore the question of the relationship between the blanket adjustment, which I have called the nominal under-reporting error, and the adjustments made to earned incomes which I have called the actual, or effective errors. The technique described below is part of a quest to extract meaningful information on poverty from the household surveys. It looks as though it may help to make the search for a satisfactory method of performing upward adjustments of expenditure or income estimates to compensate for under-reporting, a little less onerous.

Stripped to its barest essentials, the (income) poverty controversy in South Africa is about how to deal with the problem of under-reporting. Similar problems have to be faced if the job of estimating headcounts and poverty gaps is approached from the expenditure side. There is much speculation about which groups in society are most likely to under-report income and expenditure. Obvious candidates include the well-off. Some argue that poor households systematically under-report income and expenditure. It has been suggested that women in poor households might feel a need to conceal their earnings from male partners. Evidence to back up such claims undoubtedly exists, but I have made no effort to track it down. If it were to be uncovered, it is not clear that the findings could simply be imposed on the South African data. To the best of my knowledge, no national survey data that attempt to grapple with the problem of under-reporting exist. Smaller surveys might have been done. It is not obvious, however, if such studies do exist, that their results can be generalised to national level.

allows for them to differ, one from the other. The impacts of allowing them to do so have not been explored.

⁵⁴ It would be possible to discover, as well, the adjustments required to equalise poverty gaps. It is not possible, however, to adjust the figures that I have used so that both the headcount P_0 and poverty gap P_1 are simultaneously equalised.

No matter how much one wriggles, the formidable problem of finding ways to compensate for under-reporting has to be tackled. In my early work in this field (Meth, 2004), I used Deaton's (2003) international exploration of discrepancies between national accounts and household survey estimates of income and expenditure as a guide. Following Deaton, I applied corrections of between 15 and 70 per cent to reported expenditure figures to compensate for under-reporting. Although I recognised at the time that such a rough and ready way of going about things was unlikely to yield very accurate results (all of my published estimates give wide ranges), the dual demands of expanding the calculating engine to take account of higher expenditure groups (especially when I was experimenting with poverty lines corrected for child costs and household economies of scale), and of devising ways to deal with information on earned income in the household, migrant income and social grants, acted as a barrier to more adequate analysis of the data. In short, in the past I concerned myself only with what was happening in the bottom two expenditure categories, and then only with what I describe as the 'unpurged' figures.

Crude as it may sound, I assumed that nearly all of the poor people in South Africa would be located within the boundaries of the bottom two expenditure categories. They appeared, after all, to contain almost half the population. That assumption, as is shown in Table 4, is not correct — there could be upwards of six million people below the R250 poverty line in the expenditure categories in the range R800–2499. As long as I limited the analysis to the bottom two expenditure categories, the question of the adequacy of a maximum correction of 70 per cent did not present itself. Then in a piece of work I did last year (Meth, 2005b), a need to examine the whole distribution arose. The results of the examination came as something of a shock — reported consumption levels in the survey with which I was working (the 2003 GHS) suggested that when raw, unpurged estimates of the numbers of households in the various expenditure categories were used, a correction of 100 per cent might be nearer the mark.⁵⁵ That set in motion a search for a more satisfactory way of compensating for under-reporting.

Instead of taking a pair of arbitrary error limits culled from the international literature and imposing them on the South African data, the approach now is to attempt to use the data to specify plausible upper and lower boundaries to

⁵⁵ Accordingly, in Meth (2005a) I used nominal error levels of 80 and 120 per cent (on unpurged data).

the poverty estimates. The upper boundary (the zero under-reporting error estimate) is determined from the purged data (the data with ‘corrected’ expenditure categories imputed from income data). To fix the lower boundary (the minimum headcount) an attempt has been made to specify limits beyond which under-reporting errors cannot reasonably stray.

The purging operation described above is the least that can be done to the data to yield poverty headcounts with a claim to respectability.⁵⁶ In the absence of evidence to the contrary (such as might result, for example, from a more ‘scientific’ process of imputation of expenditure category), the estimates emerging from the purged data in which earned income, remittances and social grants have been taken into account, are regarded as being the upper bounds of a range within which the ‘true’ headcount lies.

After the purging operation described above has been carried out, the possibility remains that both income and expenditure are still under-reported.⁵⁷ The question is, by how much? In other words, the problem to be confronted is that of finding a way to fix a lower bound to the poverty estimates. The approach adopted in the present paper exposes (in rather spectacular fashion) the extreme crudity of my previous attempts at finding pairs of estimates, a high and a low, within which the ‘true’ headcount was argued to lie — to obtain similar numbers of poor in the bottom two expenditure categories from the unpurged data as are obtained from the purged data would require huge correction factors. The present paper therefore backpedals away from it as fast possible, by standing the question on its head and asking: how large an adjustment has to be made to the

⁵⁶ Comparing income and expenditure, where the data make it possible, is such an obvious thing to do. Having looked, for example, at the big poverty study undertaken by Statistics South Africa (see Hirschowitz *et al*, 2000, p.55), I was well aware of the interesting results that an exercise of this type produces. It is not easy to understand why it did not occur to me to make use of this valuable resource before now. The embarrassment attendant upon the fact that some apparently quite good evidence has been perched right under my nose all along, is only partly diminished by the fact that nobody else appears to have made use of it to date.

⁵⁷ According to van der Berg *et al* (2005, p.27) Black wages in the LFS are so much lower than those in the October Household Surveys (OHSs) that the former are probably under-estimating Black wages. In the bottom expenditure categories, populated almost entirely by Africans (Blacks?) the LFSs and the GHSs yield similar estimates of income from employment. It is not obvious why both surveys (conducted independently) should under-report income.

expenditure (income) estimates in order to generate the van der Berg *et al* results? If it is implausibly large, we may then, by the principle of *reductio ad absurdum*, reject the possibility that their figures constitute a lower bound to the range within which the ‘true’ headcount lies. If it is not, we are left with two sets of contending results whose relative merits must be decided by the scientific community. Obviously, a way has to be found to determine what constitutes ‘implausibly large’.

Estimates of the size of the correction that would have to be made to the Table 2 results to bring them into line with the van der Berg *et al* results (the beginnings of an answer to the question of whether or not their figures are plausible), have been provided above. To make sense of those estimates (50 and 38 per cent) it is necessary to be more precise about the distinction between nominal and actual or effective errors. Suppose that everyone’s consumption (income) level is (arbitrarily) raised by a given proportion (everybody in a particular expenditure category, that is), as is done when survey estimates of total consumption are raised to match national accounts totals. (Or as would be the case if there was a 50 or 38 per cent upward adjustment to the purged figure from which the poverty estimates in Table 2 are drawn).

To repeat the figures given in the example in Section 3 of the paper, a nominal under-reporting error of 50 per cent means that if a household reported that they were in household expenditure category R0–399 per month, their actual maximum potential consumption level would be $R399 + R399 * 50/100$, which equals R598.5. In turn, this would mean that an individual in a two-adult household in this expenditure category, instead of consuming roughly R200 per month, would thus consume about R300 per month of goods and services. If the nominal error were 20 per cent, an individual in a two-adult household in this expenditure category, instead of consuming roughly R200 per month, would thus consume about R239 per month of goods and services.

Since the value of social grants accruing to any household is known (and fixed, i.e., cannot legitimately be raised),⁵⁸ the only possible source of the income required to finance the additional consumption is the earnings of the

⁵⁸ Objections may be voiced against the method used to allocate social grant income among households, but the likelihood is that it probably does not do too much violence to the truth.

employed (in-house workers and migrants).⁵⁹ Knowing the value of social grant income, it is possible to estimate the level of worker earnings required to make possible the increased consumption implied by the (assumed) nominal error. The ratio of these earnings (salary, wage or remittances) to the pre-adjustment earnings is the actual or effective error that corresponds to the nominal error (the proportion) by which consumption is increased.

Discovering precisely what actual or effective errors correspond to nominal errors of any magnitude (while we are at it, it would be silly not to look at errors at all levels) is quite difficult. It is, however, possible to make rough estimates of the approximate values of the ratios of actual to nominal error that are probably accurate enough for the use to which they will be put here. Accordingly, estimates of the relative sizes of these two at the nominal error levels used thus far, and the apparent nominal error level in the van der Berg *et al* piece, are presented in Table 10.

In each expenditure category, the calculations make use of that mythical entity, the average individual in the average household. First, using the value of the child support grant and the state old age pension, and the mean numbers of grants made per household, we estimate the value of social grant income by expenditure category. Mean individual consumption in the *i*th expenditure category equals the category mean (e.g., R300 in the category R0–399, or R600 in the category R800–1199) divided by the mean number of individuals per household. Subtracting social grant income from mean individual consumption yields the implied income of workers (of both in-house, and of migrants) distributed equally among all the individuals in any expenditure category. Multiplying this difference by the ratio of individuals to workers in that category gives the implied earnings levels of those actually employed. Applying the assumed nominal error to the mean individual consumption level, then subtracting from this the known social grant income, yields a revised estimate of implied worker incomes equally distributed among all households. The actual error corresponding to any assumed nominal expenditure under-reporting error equals the difference between the earnings implied by the assumed error and the earnings implied by the zero-error estimate.⁶⁰

⁵⁹ Dissaving (borrowing, drawing down previous savings) is ignored.

⁶⁰ It should be noted that this rough-and-ready way of going about things is necessitated mainly by the fact, as discussed above, that the LFS does not collect adequate information on social grants. If this were available, it would be a relatively simple matter to construct full income accounts for each household (in-house worker

As may be seen from the figures in Table 10, a 20 per cent nominal error in the bottom expenditure category translates into an actual or effective error in the earnings estimates of 45 per cent. What this means is that if an overall upward adjustment of 20 per cent is made to survey income (expenditure) estimates to bring them into line with national accounts figures, the earnings estimate imposed by the researcher on a worker who reported earnings income of R400 per month would be R580 per month ($R400 + R400 * 45/100$). Staying in the bottom expenditure category, we note that at the 38 per cent error level implied by the van der Berg *et al* results (treating disability grants as if they all went to poverty alleviation), the actual error is 85 per cent. So, in place of earned household income of R265 per month reported by survey respondents, the researcher inserts a value R491. Reducing the poverty alleviating impact of the disability grant pushes this ever higher.

A marked feature of the Table 10 results is the rapidity with which the ratio of actual to nominal error falls as the expenditure level rises. Shown in the last row in the table, the value of this ratio, which is the same for all assumed nominal error levels, is determined by two variables. These are the ratios of workers to individuals in each expenditure category, and the assumed category mean income (or consumption level). And thereby hangs a tale — objectively, the actual or effective errors entailed in a 38 per cent nominal error (my estimate of the correction applied by van der Berg *et al*) are not so far above the figures reported by Deaton (2003) as to suggest that they are unreasonable. With the exception of the first two expenditure categories, where the actual errors are 85 and 91 per cent, they apparently lie comfortably within the 15–70 per cent range referred to earlier.

The results in Table 10 are, however, obtained by sleight of hand — the trick, broadly hinted at above, was to import into the model assumed mean incomes (or consumption levels) that are probably not justified by the survey results. The remarks above about researchers imposing income values on top of those reported by survey respondents, thus applies with equal force to me. This imposition, like that by van der Berg and his

earnings, migrant remittances and social grant income), then to perform the operations described above not for imaginary ‘mean’ households, but rather for each household individually, taking account of its precise composition. For the present paper at least, we have to make do with this somewhat unsatisfactory way of proceeding.

colleagues, has the effect of reducing measured poverty. In defence of the imposition, however, it needs to be pointed out that explicit reference has already been made to the ‘generosity’ of the assumed mean incomes (consumption levels) in the discussion on the Table 8 figures.

Table 10 Converting nominal into effective error levels

	R0- 399	R400- 799	R800- 1199	R1200 -1799	R1800 -2499	R0- 2499
Total number of people in expenditure category (1000s)	5 746	10 731	5 816	4 980	3 437	30 711
Assumed nominal error level (%)	0					
Implied monthly worker earnings	265	274	526	844	1251	622
Assumed nominal error level (%)	10					
Implied monthly worker earnings	325	340	605	946	1387	710
Actual error level (%)	22	24	15	12	11	14
Assumed nominal error level (%)	20					
Implied monthly worker earnings	384	405	684	1048	1523	797
Actual error level (%)	45	48	30	24	22	28
Assumed nominal error level (%)	38					
Implied monthly worker earnings	491	523	826	1232	1768	956
Actual error level (%)	85	91	57	46	41	54
Assumed nominal error level (%)	50					
Implied monthly worker earnings	563	602	920	1354	1931	1061
Actual error level (%)	112	119	75	60	54	71
Ratio of actual to nominal error	2.24	2.39	1.50	1.21	1.09	1.41

Note: Implied monthly earnings are R/month in 2000 prices. Basic data are from the September 2004 LFS.

It is time now to see what happens when less generous assumptions about means are made, a step urged upon us with some force by the estimates of the numbers of workerless (zero earned income) households. The key here is to estimate, for each of the five expenditure categories, the total expenditure in each category that would give rise both to my estimates of the poverty headcount, and to those implied by the van der Berg *et al* poverty results. The poverty estimates presented in the present paper rely on the assumption that consumption among households in each category is linearly distributed about an assumed mean. Changing that mean will change the poverty headcount (and, obviously, the poverty gap). The figures in Table 11 show the adjustments necessary to yield the original poverty headcounts when assumed mean income (once again, I am conflating income with expenditure) is allowed to vary. Results in the table refer only to the bottom expenditure category. The figures in the first column of the table (those under the assumed mean income (expenditure) assumption of R300) yield the poverty headcount of 5 153 000 among 5 748 000 individuals in Table 4 above.

The first two rows in the table give household monthly mean income (expenditure) levels in current and constant 2000 prices. Dividing this by mean household size gives individual consumption levels. Multiplying mean individual consumption by the number of individuals gives total consumption for everyone in the category, assuming zero under-reporting error. Subtracting from this total, the value of social grants leaves the sum to be earned by the workers (migrant and in-house) in the category. Dividing this figure by the number of workers yields an implied monthly mean income. As may be seen, pushing assumed mean category income (expenditure) to R250 per capita lowers implied monthly earnings (identified by the letter 'a') from R265 per month to R166. Lowering mean income (expenditure) to R200 pushes implied worker earnings down to R67 per month.

Table 11 Impacts of varying mean income assumptions on effective error levels

	R0-399	R0-399	R0-399
Assumed mean hh consumption level (R/m) (2004 prices)	300	250	200
Assumed mean hh consumption level (R/m) (2000 prices)	242.3	201.9	161.5
Mean household size	3.18	3.18	3.18
Mean individual consumption	76.2	63.5	50.8
Total consumption at assumed mean income at 0% error	5 255	4 379	3 503
Grant income	2 911	2 911	2 911
Implied consumption from worker earnings (Rm pa)	2 344	1 468	592
(a) Implied monthly earnings of workers at assumed mean income	265	1 66	67
(b) Correction required to yield vd Berg total consumption	38	65.6	107
Implied v d Berg total consumption Rm pa in 2000 prices	7 252	7 252	7 252
Adjusted consumption from worker earnings (Rm pa)	4 341	4 341	4 341
(c) Implied monthly earnings of workers to obtain vd Berg results	491	491	491
Ratio of earnings (c/a)	1.85	2.96	7.33
Actual error % ((c-a)/a)	85	196	633
Ratio of actual to nominal error	2.24	2.98	5.91

Source: Own calculations using the September 2004 LFS data

At the start of this section of the paper, I pointed out that my estimate of the upward adjustment to consumption required to replicate the van der Berg *et al* poverty results, when the full value of disability grants went towards poverty alleviation, was 38 per cent. Increasing R5.255 billion by 38 per cent gives the target total of R7.252 billion, to which total consumption (expenditure or income) must be raised if the van der Berg *et al* poverty results are to be obtained. The row that contains the magnitude of the adjustments that have to be made to obtain zero error consumption totals is identified by the letter 'b'. To raise R4.379 billion to the target total requires an adjustment of 65.6 per cent. Raising R3.503 billion to the target of R7.252 billion requires an adjustment of 107 per cent. When these have been made, worker earnings, in each case, have to total R4.341 billion. Implied monthly earnings (in the row identified by the letter 'c') of R491 per worker are required to reach this total. Dividing *c* by *a* gives the ratio of adjusted to unadjusted earnings. This may also be expressed in the form of the actual error $((c - a)/a)$.

Spectacular would not be too strong a word to use for a ratio of earnings (*c/a*) of 7.33 or its associated actual error $((c - a)/a)$, of 633 per cent. To explain why these figures should be so high, it is necessary to delve into conditions in the expenditure category R0–399 per month, looking in particular at the ratio of individuals to workers. Among the 1.9 million households in the category (see Table 8), about 81 per cent contained no income earners. Roughly 24 per cent of the households received migrant remittances. That leaves more than 50 per cent without workers. It is not unreasonable, therefore, to suggest that for households in the category as a whole, the mean earned (or remittance) income level could be R200 per month (and now we are not conflating income and expenditure). If that were the case, van der Berg and his colleagues would implicitly be saying to the hypothetical mean worker in the bottom expenditure category: your earnings are not R67 per month as reported, rather, they are R491. Readers are welcome to form their own opinions as to the permissibility of upward adjustment on such a grand scale. As far as I am concerned, they are unacceptably high. Taking away part of the poverty alleviating effects of the disability grants would push them higher still. In my view, adjustments of this magnitude to income estimates are implausibly high.

Making one's way up the distribution, the error, as we saw in Table 10, falls. In the second expenditure category, if mean income (expenditure) were lowered from its assumed value of R600 per month to R550, the ratio

of earnings (c/a) would be 523/200. An assumed mean income of R500 would see this ratio increase to 523/165, to yield an actual error of 217 per cent. In the highest expenditure category (R1800–2499) a reduction in assumed mean income (expenditure) from R2 100 to R2 000 leads to an actual error of 49 per cent (reported income of R1 187 per month has to be raised to R1 786). The latter may be plausible — the former are not.

It remains but to refer once more to the conservative nature of the poverty estimates made in the present paper. If we use the Table 10 figures to estimate the ratios of nominal to actual errors, their full extent is camouflaged by the fact that sweeping upward adjustments have already been made in the purging process. Only by showing how dependent these are on assumed means is it possible to get a glimpse of the adjustment to the survey figures required to replicate the van der Berg *et al* results. It is expected that people will under-report both income and expenditure — it is entirely unreasonable to suggest that they will do so to the extent disclosed above. On these grounds, I am confident that my zero error under-reporting estimates represent a plausible upper limit to the poverty headcount. A nominal ten per cent upward adjustment would probably be large enough to get us to a plausible lower limit. If that is so, then the poverty headcount at the R250 poverty line (2000 prices) must be somewhere in the region of 19–20 million, depending on how one treats disabilities.⁶¹ As an extreme lower bound, I propose a 20 per cent nominal adjustment. It is incumbent upon anyone who claims that the nominal under-reporting error in South Africa is greater than ten per cent, to produce better evidence than that offered by van der Berg *et al* (2005), to support the assertion.⁶²

⁶¹ It is important to note that the technique for estimating actual, as opposed to nominal, under-reporting error corrections depends for its accuracy upon the reliability of the employment estimates. The numerous claims that the household surveys under-estimate employment (despite their origins in an officialdom desperately anxious to demonstrate progress in the field of job creation) must stand as a warning against too easy an acceptance of the test proposed above.

⁶² If actual (effective) errors reach improbable levels, if that is, they begin to climb above levels reported in the international literature (Deaton, 2003), it is necessary to raise questions about the quality of the data from which the estimates are prepared.

6. Distributional implications of the poverty estimates

Thus far, the discussion of the various poverty headcounts has ignored the distributional implications that each of them so clearly has. To check the plausibility of the distributions of the poor generated by the calculating engine used for the present paper, two sets of comparisons are necessary. One of them is an internal comparison, watching how shares change as the nominal error level is raised; the other is external: it requires that the results be confronted with the conventional wisdom on the matter, in the period before 2004. Unfortunately, as the review of the debate in South Africa over poverty and inequality in van der Berg *et al* (2005, pp.6–10) shows, there is no conventional wisdom — there are merely competing claims. Settling upon any particular one of these claims as comparator must inevitably lend an air of unreality to the attempt to assess distributions.

It is not unreasonable to ask at this point whether, if the comparison exercise is so difficult, there is merit at all in attempting to perform it? We could, for example, invoke the advice implicit in the conclusion that ‘a fragile inference is not worth drawing’, offered, if my memory serves me well, by the eminent econometrician, Ed Leamer. Ultimately, however, the question of differences between the distributional implications of the approach suggested in the present paper, and those discovered by others, is going to have to be confronted. This is as good a place as any to begin considering the problems that are going to have to be faced when a proper comparison is attempted. The first and most obvious requirement for conducting such an exercise is the inclusion of the missing expenditure data (that for the folk in expenditure categories R2500–4999 to R10 000 plus, in addition to that for the non-negligible categories of ‘Don’t knows’ and ‘Refuses’). Since we are working only with the data from the bottom five expenditure categories, the distribution is artificially thinned out at its upper end, by virtue of the absence from it of people who might reasonably be expected to form part of it. For the meanwhile, the only way around this problem, whose resolution lies somewhere in the future, is speculation of the sort attempted below.

Table 12 shows what happens when the relatively sparse distributional data generated by the calculating engine are pressed into service.⁶³ For both the

⁶³ Part of the problem is that of integrating the area under the curve in Figure 1

internal and external comparisons, the way the model works prevents us from fastening directly onto the results for, say, the second, third and fourth deciles.⁶⁴ Instead, it is necessary to select a population proportion closest to the group in which one is interested. So, for the zero error level, for example, the closest we can get to a quintile is 23.1 per cent. Unsatisfactory though this is, let us see if there is anything to be learned from the use of the figures in Table 12.

Table 12 Cumulative distributions of consumption & population – various errors

Error: 0%		Error: 10%		Error: 20%		Error: 50%	
Share of Pop- ulation	Con- sump- tion	Share of Pop- ulation	Con- sump- tion	Share of Pop- ulation	Con- sump- tion	Share of Pop- ulation	Con- sump- tion
11.0	0.7	9.8	0.6	8.7	0.5	6.2	0.4
23.1	1.9	21.4	1.8	19.8	1.7	15.3	1.3
26.9	2.4	25.0	2.2	23.3	2.1	18.7	1.8
34.7	3.6	32.8	3.4	31.1	3.3	26.2	2.9
40.9	4.8	39.2	4.6	37.6	4.5	32.5	4.1
44.7	5.6	43.0	5.5	41.5	5.4	36.6	5.0
45.8	5.9	44.2	5.8	42.7	5.7	37.9	5.3
49.3	6.8	48.1	6.8	46.7	6.7	42.3	6.4
52.9	7.8	52.1	8.0	51.2	8.0	47.6	8.0
54.9	8.5	54.3	8.7	53.5	8.8	50.5	9.0
56.2	9.0	55.8	9.3	55.2	9.5	52.9	9.9
58.0	9.8	57.8	10.2	57.5	10.5	55.9	11.3
59.7	10.8	59.9	11.4	59.9	11.9	59.2	13.2
60.5	11.4	60.7	12.0	60.9	12.6	60.8	14.4
61.3	12.3	61.7	13.0	62.0	13.7	62.3	15.9

Source: Own calculations using the September 2004 LFS data

without having access to a large enough number of observations for the job. Simpson's Rule is usually called upon to assist in such tasks. It has been used here.⁶⁴ The difficulties of comparison are heightened by the difficulties of extracting distributions from an unwilling calculating engine (the exertion pushes it up against the limits of its capabilities).

Culled from the poverty estimates that exclude the disability grant (i.e., the assumption that reduces the share of the poor to a minimum), the table presents distributional shares at four error levels. These form the bases for both the internal and external comparisons. As far as the former is concerned, one would expect the share of the poorest to increase as error levels rise. This appears to happen. In Table 12, this is most clearly visible in the area where population shares are about 42–43 per cent. The zero error level is mildly disabliging, but the other three all point to a steady increase in the share of this group, amounting probably to at least one percentage point (i.e., the group's total consumption, which, at the zero error level, already includes social grants and migrant incomes, grows by about R7–8 billion). Although not negligible, the change is not large, being roughly in line with what one would expect. Relatively small changes in distributional shares with increasing error levels would be consistent with a situation in which the proportion of workerless households is large, and in which the volume of grant income (relative to total final consumption) is not very large.⁶⁵

Together, the old age pension and child support grants (the two primary grants considered by the calculating engine) were worth about R27 billion in 2004 (in current prices), or about 2.4 per cent of final consumption expenditure. Distributed over about 20 million people, mainly those in deciles 2 to 6 from the bottom, it does not make a great deal of difference to household shares. Yet much is made by van der Berg *et al* of the extent of redistribution through social grants (2005, p.3). The real picture is less straightforward. Between 2000 and 2004, the number of old age pensioners grew by 2.6 per cent per annum (National Treasury, 2004, p.74), and the pension's real value had yet to reach the level it had attained in July 1995.⁶⁶ Not very much additional redistribution originated from this, the largest of

⁶⁵ It may be useful, at this point, to recall the finding reported in Table 4 above that (ignoring the disability grant) there were reportedly 10.2 million individuals in households receiving only the CSG, as opposed to five million in households receiving only the OAP. There were a further 3.8 million in households receiving both the CSG and the OAP.

⁶⁶ The real value of the CSG was lower at the end of 2001 than when it was introduced in 1998. Six years later, its real value was approximately 16 per cent higher than when it was first introduced (in current prices the values were R100 in July 1998 and R170 in April 2004). By April 2004, the OAP had yet to reach the real value it had attained in July 1995 (in current prices, the relevant values were R410 and R740).

all the social grants (in terms of total cost) between 2000 and 2004. The two big changes in the social grant system are the increases in the numbers of households receiving disability grants and child support grants. Uncritical treatment of disability grants as a form of poverty alleviation I have argued above to be unacceptable. Although the number of child grant recipients is large, the value of the grant is low. The grant's effect on poverty, although noticeable, will be small.

The external comparison presents different, possibly more extreme problems. For the purposes of the present paper, one of the desirable features of whatever distribution is consulted (apart from having some degree of reliability), is that it should offer a table giving the income distribution by percentiles. Calculations that would enable these results to be extracted from the Hoogeveen and Özler (2004) paper had obviously to be performed by the authors — the necessary table does not, however, appear in the paper. Similar conclusions apply to the Leibbrandt *et al* (2004) paper. In this case, however, the authors are close at hand, so the relevant figures have been requested. They have not yet been received, however, so we must make do, for the meanwhile, with the one readily available set, those for the year 2000 reproduced in Statistics South Africa's *Earning and spending in South Africa* (2002). This is a poor compromise — not only are the results disputed, they are also presented as household rather than individual distributions, and as distributions of income rather than of expenditure. Table 13 reproduces the distribution in question.

Except for households in the bottom quintile, mean household size is greater at the bottom end of the distribution than at the top.⁶⁷ This means that population shares of total consumption in deciles 2–3 to 6 will be smaller than household shares, while in the top four deciles they will be larger. It is unclear what effect the presumed under-reporting of income that is said to plague the Table 13 results, will have on the distribution. Also, it is assumed that consumption and income do not differ much, one from the other, at the bottom end of the distribution. For present purposes the problem of under-reporting is simply ignored.

⁶⁷ Mean household size for the whole distribution in the September 2004 LFS is 3.70. After purging, the bottom expenditure category contains 15 per cent of all households; the second category (R400–799) 22 per cent; the third category 11 per cent and the fourth category, 10 per cent. Corresponding mean household sizes are 3.22 for the bottom category; 4.02, 4.16 and 3.76 respectively for the others.

Table 13 Household shares of expenditure, 2000

Cumulative share of (%)		Cumulative share of (%)		
Households	Income	Households	Income	
10	0.42	60	15.01	
20	1.63	70	22.97	
30	3.48	80	35.13	
40	6.08	90	54.85	
50	9.72	100	100.00	

Source: Statistics South Africa (2002, p.47)

Comparing the Table 12 zero error figures with their closest counterparts in Table 13, and making some imaginative extrapolations, it would appear that at the 20 per cent level, relative shares do not differ all that much (23.1 per cent of individuals share 1.9 per cent of total consumption). At the 60 per cent level, the difference looks quite big, until it is recalled that some significant number of the people one would expect to find in the fifth and sixth expenditure deciles are missing because of the aforementioned fact that we are making use of only the information for those in the expenditure categories ranging from R0–399 to R1800–2499. Mean monthly per capita expenditure levels for those at the 60th percentile appear to be about R800. This means that any household near the bottom end of the R2500–4999 expenditure category containing three people is within striking distance of being counted somewhere near the 60th percentile. As a matter of interest, there are about 850 000 adult and child households in the R2500–4999 expenditure category. With somewhere in the region of 5.1 people per household, there appear to be about 4.3 million people in this category. It only needs 450 000 people to pull up the 10.8 per cent at the 60th percentile by one percentage point.⁶⁸ With possibly one-third of those in the category consuming below the R800 per month mark, the zero-error share creeps close to the Statistics South Africa figure in Table 12 (which, it should be recalled, is for households, rather than individuals).

Findings from an exercise such as that conducted above must, of necessity, be treated with the greatest possible circumspection. If there is any conclusion at all to be drawn from the excursion above, and it can be no more than a hint, it is that any of the distributions in Table 12 may represent

⁶⁸ I have not yet taken out purged figures for the R2500–4999 expenditure category. The results presented here are estimated from rough (unpurged) figures taken out for the distribution as a whole.

the actual distribution in 2004. More than that cannot be said. Obviously, a lot more work is required to overcome the obstacles described above.

That work has to address the weaknesses of the model used for extracting the results from the data. As it stands, not only is the existing model incapable of taking account of everybody's consumption (all expenditure categories), it is also unable to select deciles of the population for comparison purposes. At a push, it could possibly be modified to deal with the first of these shortcomings; there is nothing that can be done about the second of them. In short, it is going to be necessary to scrap it and to start again from scratch. Even after that has been done, the problem of finding suitably authoritative estimates of South Africa's expenditure distribution with which the resulting distributions may be compared, will remain.

7. Other methodological and conceptual problems

So much for the description of the technique for extracting poverty estimates from the household data — we turn now to some of the other potential hazards involved in doing so. High on the list must be the fact that the method of estimating individual consumption (or income) levels involves digging down very deep into the survey data, down to levels where there are sometimes few observations in the cells. Statistics South Africa does not publish, on grounds of unreliability, any weighted total smaller than 10 000. Applying this rule to individuals in the bottom expenditure category in the adult and child-type households in the 2004 GHS data would wipe out about 99 000 individuals, most of them from large households, from among a total of 6.9 million. The error involved here is a little over one per cent (1.4), hardly a matter over which sleep should be lost.

Since household compositions (types) determine individual consumption levels (employing the unsatisfactory assumption that consumption is equal among household members), errors in observed distributions of household types could affect poverty headcounts. If it is true that one would not expect distributions of household types (compositions) to change too dramatically over short periods of time, then one way in which to test the reliability of the method of measuring poverty proposed in the present paper is to check the stability of household composition. This is not, however, a task to be tackled lightly. Numerous problems would have to be solved along the way.

One of them involves changes in household composition caused by migration between expenditure (income) categories. We know that household composition varies with expenditure, with the very poorest households in expenditure category R0–399 being smaller than those in the category R400–799. Given the apparently significant movement between these categories between 2003 and 2004, some compositional change is to be expected. If the LFS can be used in the manner intended, i.e., as a longitudinal survey, then it should be able to answer questions of this nature. Use of the LFS to conduct longitudinal analysis is, however, in its infancy.

Clearly, this is an area requiring more research. Equally clearly, it is not going to receive the careful treatment it deserves here. Instead, I will merely report that a rough-and-ready check performed on the base data used in the present paper (the purged adult-and-child households in the bottom expenditure categories in all four surveys) suggests at least some consistency. In the 2003 surveys, about 85 per cent of all households were concentrated in households containing up to four children and up to four adults. For the 2004 surveys, the figure rises to 88 per cent. Individual categories, e.g., one adult and two children, do not differ from each other by more than a percentage point or two. One could test the sensitivity of the poverty headcounts to variations in the distributions of households by imposing the distributions generated by one survey on the total numbers uncovered by a different survey. As may be imagined, this is a big job, and not one that I intend to tackle here. Like van der Berg and his colleagues, I must go out on a limb and say that I trust the survey distributions, and, by implication, place reliance on very small numbers.

Another problem we share is that the primary data, respectively on expenditure in the case of the present paper, and on income in their paper, is collected in bands or categories rather than as point estimates (van der Berg *et al*, 2005, p.29). Thus although the technology they use to produce their poverty estimates is a lot more sophisticated than that applied here, in the final analysis they have to deal with the fact that this is the way that the AMPS surveys collect income data.⁶⁹ If we can treat income as a proxy for

⁶⁹ The LFSs and GHSs are slightly better in this regard, insofar as about two-thirds of respondents provide point, rather than category, estimates of income. As I noted at the beginning of this paper, I am deeply indebted to my colleague Daniela Casale. She gave me the STATA code required to extract incomes from the data sets, and a complete set of Do Files for dealing with the migrant module in the LFS. This

expenditure (and this only breaks down when savings or dissavings are significant) then what matters above all else is the fact that Table 14 below suggests that the AMPS surveys may produce a distribution of expenditure (income) that differs enough from that yielded by the GHSs and LFSs to begin explaining some of the differences between our respective poverty headcounts.

Table 14 Distributions (1000s) of individuals 16 years and over in three surveys

AMPS 2004 distribution			September 2004 LFS distribution			2004 GHS distribution	
Household Income Category	No. of Indivi- duals	% of total	Household Expenditure Category	No. of indi- duals	% of total	No. of indi- duals	% of total
R1-499	2 205	7.3	R0-399	3 716	12.2	3 134	10.4
R500-899	6 147	20.3	R400-799	6 876	22.6	6 968	23.1
R900-1399	3 829	12.6	R800-1199	3 774	12.4	3 910	13.0
R1400-2499	5 764	19.0	R1200-1799	3 478	11.4	3 792	12.6
R2500-3999	3 881	12.8	R1800-2499	2 480	8.1	2 653	8.8
R4000-6999	3 546	11.7	R2500-4999	4 102	13.5	3 867	12.8
R7000-11999	2 690	8.9	R5000-9999	3 176	10.4	2 956	9.8
R12000+	2 248	7.4	R10 000+	2 426	8.0	2 451	8.1
Total	30 310	100.0	Total	30 436	100.0	30 167	100.0
<R2500	17 945	59.2	<R2500	20 325	66.8	20 458	67.8

Source of AMPS data: AMPS Technical Report 2004, Table 4, p.29
Note: LFS and GHS figures do not total 100 per cent because 'Don't know' and 'Refuse' responses have been omitted.

Table 14 compares the distributions of persons aged 16 years or more by income category obtained from the three surveys, income in the case of the AMPS, and expenditure in those of the GHS and LFS. Results for the latter two are the result of transforming the unpurged distributions by multiplying each expenditure category by the ratio of the numbers of purged to unpurged households. Migrant remittances and implausible zeroes have not been taken into account. Crude though this procedure might be, it is good enough for the purpose intended, namely, to show that although there are

spared me a great deal of effort and saved precious time.

differences between the AMPS figures on the one hand, and the two sets of figures from Statistics South Africa on the other, these are not so large as to incline one to the belief that they are from different planets.

Thinner at the bottom end, and thicker in the R1400–3999 per month part of the distribution, the AMPS figures, if this is what went into the van der Berg *et al* results, will probably generate lower poverty levels than either the September 2004 LFS or the 2004 GHS, even after these two surveys have been purged.⁷⁰ To reconcile the Simulation 2 results in Table 2 with the van der Berg *et al* figures requires an error of 38 per cent. As we saw in the discussion to Table 11, if the assumed mean income (expenditure) is too generous, a nominal error of this magnitude can easily translate into implausibly high effective errors (more than 600 per cent in the bottom expenditure category, for example). So, even if we were to admit the disability grants into the argument, the difference between the headcount in the present paper, and that in van der Berg *et al* (2005) is likely to be at least three million. On that note, let us conclude.

8. Conclusions and recommendations

Very few people will read a paper such as the present one, or the van der Berg *et al* (2005) effort. At most, those who pick up one or the other might possibly glance at the abstracts and the conclusions. That being so, it behoves us to make the most of the opportunities this possibility presents. The most important thing to be said here relates to the significance of (a) the estimates of the absolute sizes of the poverty headcount ratios (P_0) and poverty gap ratios (P_1), and (b) the fact that the van der Berg *et al* (2005) paper attempts to track changes in P_0 and P_1 over time.

Surprising though it may sound, it seems to be possible to read the latter paper in so slapdash a manner as to conclude, as did one of the commentators on the present paper, that even if the absolute numbers in it were incorrect, the poverty ‘trend’ disclosed in it is important. A moment’s reflection shows how mistaken is such a conclusion. Implicit in the word

⁷⁰ These, as we observe in Table 2, generate remarkably similar unadjusted poverty headcounts (if that is what one may call the numbers one obtains before taking account of social grants and remittances), despite the differences in the R0–399 expenditure category.

‘trend’ is the possibility that past behaviour may be extrapolated into the future. Yet van der Berg *et al* (2005) do not subscribe to any such notion. For them the major cause of the fall in poverty was the expansion of the social grant system.⁷¹ Opportunities for the further reduction of poverty by this means, they suggest, are limited; the past, therefore, may not be extrapolated into the future.

Their finding that the number of poor declined between 2000 and 2004 is hardly surprising, given the billions that have been dispensed in the form of social grants. The Taylor Committee report estimated that in 2001, the existing social grant system was capable of lifting about 840 000 people above a poverty line of R401 per adult equivalent per month in 2001 prices, at full benefit take-up (2001, p.60). Accounting for the bulk of take-up in that year were about 1.9 million pensions, 630 000 disability grants and 970 000 child support grants. By 2004, pensions were being distributed to 2.1 million people, disability grants to 1.3 million, and child support grants to 4.3 million, the latter increase occurring chiefly because of a raising of the age limit for grant eligibility from 6 years to 14 years. It would be most alarming if such an expansion of the system did not result in reductions in poverty — if not in the headcounts, then at least in the poverty gaps. In short, the interesting features of the van der Berg *et al* poverty reduction figures are that (a) they move in the right direction, and (b) they provoke debate about the absolute values of P_0 and P_1 in the relevant years.

Given the state’s strong antipathy to the extension of the social grant system beyond its present boundaries, what matters for policymaking purposes now is the absolute level of poverty, both the headcount and the gap. The past can teach us but little about what the effects might be of achieving something close to full take-up of current entitlements. To predict these, reliable estimates of the current headcount and the poverty gap are required.⁷² The burden of the argument in the present paper is that the van der Berg *et al* (2005) paper does not provide good estimates of the current poverty numbers. Their 2004 poverty headcount of 15.4 million at the R250 per capita poverty line (in 2000 prices) is argued to be too low by somewhere between three and five million, i.e., there were probably about 18–20 million people below the chosen poverty line in 2004. The figure

⁷¹ They also make some highly speculative assertions about job creation, and its effect on the poor, based on a probit model using the current, somewhat suspect, employment data (2005, pp.21–22).

⁷² Making sure that all those entitled to social grants actually receive them requires good administrative data and good administrators.

finally settled upon depends on (a) how disability grants are treated, and (b) the view one takes on the extent of under-reporting error. Neither can be fixed without further research.

In essence, controversies over income poverty headcounts in South Africa turn, as they do elsewhere, on the question of the extent of under-reporting of income and expenditure in household surveys and censuses. If, as is argued in the present paper, the van der Berg *et al* estimates are indeed wrong, this will mean that the technique they use to address the problem of under-reporting of income — raising racial mean incomes derived from AMPS surveys to the means estimated from the national accounts — is faulty.

Welfare in the present paper is measured by maximum potential consumption of individuals in households in which monthly household expenditure was less than R2 500 per month (in the analysis, income and expenditure are conflated). The basic data are drawn from the Labour Force Survey for September 2004. To ensure that the September 2004 LFS results are not outliers, data from the September 2003 LFS and the GHSs for 2003 and 2004 are also examined. Under-reporting is dealt with by comparing reported total household expenditure with reported total household income. Estimates of the distributions of persons and households by expenditure category have been ‘purged’ in a way that ensures that there are no households in any particular expenditure category whose total income from all sources (earned, grant and remittance) exceeds the upper bound of that category. To ‘purge’ the data is to impute expenditure levels from income data. People’s reports of household incomes are taken as valid indicators of their consumption possibilities. Where income and consumption figures agree, there is no need for action. Where income exceeds consumption, households are relocated to the appropriate expenditure category.

When one wishes to approach poverty from the expenditure side (using either the LFS or the GHS), a process such as that described immediately above is necessary because no point estimates of consumption are available. Since the expenditure data are in category or band form, and because the bands are quite wide, a household income estimate fractionally above the upper bound of an expenditure category is sufficient to push the household into a higher expenditure category. This raises potential consumption levels, so it should have the effect of causing poverty to be understated, i.e., the imputation process is inherently conservative. If the assumptions underlying

the purging operation are rejected, then the results it generates must be invalid.

As part of the attempt to deal with the problem of under-reporting, a technique has been developed for estimating the implied corrections to income estimates, by household expenditure category, required to generate the van der Berg *et al* results. It makes use of the concepts of nominal and actual or effective under-reporting error. Nominal under-reporting error is the blanket correction made to everyone's expenditure (income) in any particular expenditure category to raise survey total expenditures (incomes) so that they approach the corresponding national accounts totals. It looks as though the overall correction in the van der Berg *et al* (2005) piece was in the region of 38 per cent. Since the values of social grants are known (and fixed), the income necessary to fund the increased consumption can only come from increased worker incomes. With some juggling, it is possible to estimate the magnitude of these corrections, called in the present paper the actual or effective under-reporting errors.

When relatively generous assumptions are made about mean expenditure in the different expenditure categories, e.g., the R300 assumed in the expenditure category R0–399 per month, actual errors of roughly the magnitude of those reported in the international literature are found. With even the slightest of relaxations of the assumptions made about mean expenditure levels, actual under-reporting errors start to climb, rapidly reaching implausible levels. It is concluded from this that van der Berg *et al* over-correct for under-reporting error at the bottom end of the distribution. This over-correction, it is argued, explains why their poverty headcounts are lower than those made by other researchers in the field.

Rudimentary though the technique of purging may be, it is no more crude than raising survey income totals until they match national accounts estimates of income. The van der Berg *et al* approach to the problem of under-reporting has the apparent advantage that the national accounts constitute a natural bound (a control total) to which income (or expenditure) totals have to be raised to account fully for under-reporting errors. Whether or not this bound is correct is, however, as Deaton (2003) points out, a major problem for those using this type of approach. In any case, the existence of this control total (correct or incorrect) cannot protect their results against a charge that the implicit assumptions they make about under-reporting of income by income (or expenditure) class, are arbitrary.

No attempt is made in the present paper to try the technique used in it, on, for example, the 2000 LFS data — that will have to wait until a number of questions about both the reliability of the technique itself and the quality of the 2000 data have been answered. It is thus not possible to subject the claim made by van der Berg *et al* about the extent of the poverty decline between 2000 and 2004 to thorough scrutiny.⁷³ All that is attempted in the paper is a comparison between my estimates of the poverty headcount for the year 2004, and the Hooegeveen and Özler (2004) figures for the year 2000. The comparison presented in Table 6 above suggests, very tentatively, that the reduction in the poverty headcount between 2000 and 2004 may not have exceeded 1.5–2.0 million. The line of inquiry this excursion opened needs much more exploration before any confidence can be attached to the finding.

In any case, even if data for the year 2000 were available in usable form, the margins of error in the poverty estimates made using the LFS data are so large (of the order of two million or so) that it would not be possible to use them to evaluate the van der Berg *et al* claims. Acknowledging that error margins (rather like confidence intervals) are large is not only an admission of the fact that the data, and the method used to extract information from it,⁷⁴ are too frail to support more precise conclusions; it is also a claim that the apparent precision in the van der Berg *et al* results is spurious. Like all other researchers in the poverty field, however, van der Berg *et al* have been forced to make many assumptions. Two obvious examples are how to treat disability grants, and how to deal with benefits captured by those outside the poorest two or three expenditure categories. There will be many others.

⁷³ If treating the household surveys as time series instruments were permissible (and the changes are statistically significant), we could say, on the basis of the estimates presented in Table 2 above, that the number of poor fell by almost two million between 2003 and 2004. That, however, seems excessive, given the generally small impact of the CSG. Further investigation is necessary to determine why the 2003 and 2004 LFS results differ so substantially.

⁷⁴ As noted earlier, the model used to extract the results given in the present paper is going to have to be scrapped in favour of something capable of dealing more efficiently with the mountain of information that has to be digested to churn out the desired results. Creaking under the strain of the demands placed on it, the calculating engine, originally designed to handle much simpler questions, has reached the limits of its capabilities.

Clearly, to produce single point estimates, as van der Berg *et al* do, a specific set of assumptions has to be applied. Until such time as we see all the assumptions used to obtain their results spelled out in detail, along with a detailed set of results of tests of the sensitivity of those results to variations in the assumptions, their report on the magnitude of the poverty reductions should be treated with circumspection.

Recommendations

To arrive at the findings presented in this paper, much time has been devoted to digging into household surveys. The surveys do not, at present, allow for as accurate a picture of poverty to be constructed as they could do if some fairly simple changes to them were made. The excavations have disclosed several areas in which there is scope for improvement; doubtless there are others. The purpose of this part of the paper is to suggest ways in which the surveys could be changed to produce the desired results. This is done within the context of a brief overview of the official statistics that have been, or may be, used in poverty studies.

In the past, the quinquennial Income and Expenditure Surveys (IESs), coupled with an October Household Survey (1995) or a Labour Force Survey (2000), served, however poorly, as the basis for a number of poverty and inequality studies. Previous IESs relied on the ‘recall’ method to collect data on income and expenditure. Data for the most recent IES has been collected using the ‘diary’ method. There is now no link between the IES and the LFS. With the changeover to the diary method of data collection, the facility of using the coupled surveys has been lost. Although it might be possible to make estimates of poverty and inequality using the new IES data, it will no longer be possible to link this to other economic and social phenomena. This reduces significantly the capacity of researchers to understand South Africa’s socio-economic path.

Because a population census is not to be carried out in 2006, a massive survey to gather some of the data that would have been gathered by the census, and some of the data presently gathered by the LFS and GHS, will now be carried out. The Community Survey (CS), as this proposed survey is known, will look at a couple of hundred thousand households. It has not been possible to form a view on the ability of the CS to serve as a vehicle for poverty studies, because not enough information about the instrument has been made public. There is no guarantee that the survey — a huge, costly exercise — will be repeated.

National Treasury is said to be interested in the question of how best to conduct surveys that will enable the numbers of poor, and the extent of their poverty, to be measured with greater precision. Statistics South Africa is currently doing some preliminary investigative work on the question. One possible outcome of such activity is a recommendation that a survey specially designed to measure poverty and inequality be instituted. Given Statistics South Africa's limited resources, such a step would not have much to recommend it. Far better, or so it would seem from the research conducted in the present paper, would be a series of revisions to the GHS and the LFS to render those surveys better able to address the poverty and inequality question. Researchers who are interested in poverty can make a contribution to this enterprise by digging deeply into the GHS and LFS to discover what the surveys can say, and equally importantly, what they cannot, about poverty and inequality.

The GHS

If there is merit in the argument that household surveys like the GHS (and the LFS) are potentially useful sources of information about poverty, consideration is required of the revisions necessary to each to enable them to fulfil their promise. Two major modifications to the GHS are required:

- A migrant module, similar to that used in the LFS, needs to be incorporated into the survey; and
- The expenditure component in the household section needs to be improved.

Without the migrant module, the GHS is unable to take account of the not inconsiderable contribution made by remittances to household wellbeing (existence?) at the bottom end of the income distribution. To accommodate the module without increasing the respondent burden disproportionately, a critical review of the questionnaire is required. Consideration could be given to the suggestion that certain sections of the survey only be used every second year. One candidate for such treatment is the medical section (questions 1.18 to 1.33 in the 2005 GHS questionnaire). Some of it appears, in any case, to be of dubious value. Analysis of some of the morbidity data sought by question 1.20, for example, yields the sort of statistical rubbish one would expect when one asks questions that are too intrusive, as are the questions, for example, on AIDS, substance abuse, and 'Other sexually transmitted disease' (I thought the PC term was 'infection'). It seems likely

that neither the incidence of epilepsy (and its treatment), nor the distribution of disabilities would change in the short-term. Nor, one suspects, would patterns of treatment in the various health care institutions.

The 2005 GHS went into the field asking questions on the Expanded Public Works Programme (EPWP) (2.19 to 2.21) that were similar to those asked in the September 2005 LFS (5.1 to 5.4). At a LFS user consultation organised by Statistics South Africa earlier this year,⁷⁵ it was agreed that these questions would only be considered (possibly as a bolt-on to one of the proposed quarterly LFSs) if a proper motivation for their inclusion was received. The questions, as they presently stand, are inadequate. In any case, there is no need to have them in both the LFS and the GHS. If the EPWP is to be taken seriously by either of the surveys, careful design is necessary to elicit the appropriate policy information. Consultation with the HSRC, which organisation, I understand, is responsible for designing the monitoring and evaluation systems for the EPWP, is desirable, before any attempt is made to incorporate EPWP questions into either the LFS or the GHS.

The ‘green’ questions (4.50 to 4.62), although both interesting and necessary, are another candidate for being asked every second year rather than each year.

Consultation with all those able to contribute to the design of a small but comprehensive expenditure component to the GHS (in addition to the questions asked in 4.80?) is necessary. If at all possible, the expenditure module should be redesigned so that a rudimentary check of the GHS figures against the IES figures can be performed (even if the former uses questions of the recall type, while the latter uses questions of the diary type).

The LFS

When this study commenced, it was thought that the LFS was potentially a contender in the quest for a household survey that could be used for measuring poverty. Earlier versions of the present paper drew attention to the possible damage that might be done to the LFS’s capacity to meet this need because of alterations made to it, following the conducting, last year, of an inquiry into the adequacy of the Labour Force Survey. Undertaken for Statistics South Africa by an expert furnished jointly by the IMF and the

⁷⁵ At this consultation, held on 18th March 2006 in Pretoria, details of the re-engineered surveyed were discussed at length.

World Bank, the (unpublished) report made a number of criticisms which affect the design of the surveys.

In the bad old days, Statistics South Africa's predecessor, the Central Statistical Service came in for a regular drubbing, both because of the poor quality of much of its output, and because of the long delays in releasing the information it had gathered. These days, researchers are beginning to drown in the flood of data. The present paper had reached an advanced stage when the September 2005 LFS was released (P0210, 24 January 2006). Having put an immense amount of effort into analysing the September 2004 LFS, I was not about to begin all over again with the September 2005 figures. Imagine my surprise, then, when I discovered, at the recent LFS user workshop, that the section on household information had been excised, in its entirety, from the LFS questionnaire. Removal of this section is part of a process of re-engineering of the survey, the intention of which is to make the LFS a more efficient instrument for understanding labour market dynamics.

This excision means that the capacity of the LFS to act as a source of information on poverty was, if not destroyed, then very badly damaged. One can understand the necessity, in the light of the re-engineering process of the LFS currently underway, to eliminate questions that have no obvious bearing on labour market dynamics. So sweeping, however, are the excisions, that the ability of the LFS to deliver the relevant information to researchers in its area of specific focus has unwittingly been damaged as well.

While there is no doubt that the questions on social grants in the LFSs (prior to September 2005) were inadequate, the information collected provided a rough and ready guide to conditions in households containing the unemployed, in particular, those in workerless households. That is now gone, swept away with the other questions in the household section of the LFS. As has been pointed out above, fixing the faulty questions on social grants would have been simplicity itself. Taking them out of the surveys altogether prevents researchers from constructing a complete picture of a household's income sources. This effectively prevents them from achieving the understanding of the labour market that the surveys seek to facilitate.

It is well known, for example, that the ability and the willingness of households to engage in search activities in the labour market are, in part, a function of household income. For one thing, job search is known to depend

on the ability of households to take the risks involved in such activity. If no information is collected on sources of income other than earnings of those in the household and on migrant remittances, then attempts to understand job search (or the lack thereof) must inevitably be frustrated. Many of the non-searching unemployed (the discouraged) are congregated in poor households. Since social grants (mainly intended for other purposes such as alleviation of poverty among the aged, support for children in poor households and meeting the special needs of the disabled) form an important component of the incomes of so many of these households, the failure of the LFS to gather information on social grants does fatal damage to the LFS's ability to say anything sensible about those whose (lack of) labour market attachment constitutes one of South Africa's most pressing problems. If the LFSs are to meet their express purpose, the section asking questions about receipt of social grants must be restored. These questions must be asked in the same way as they are presently asked in the GHSs, i.e., information must be gathered on individual grant recipients — the household is not an appropriate unit of analysis for this purpose.

Migrant remittances contribute significantly to the income of poor households, so constructing a fully-rounded picture of the ability of these households to sustain themselves requires that information on migrant remittances be collected. In earlier discussions with Statistics South Africa officials, I had suggested that consideration be given to moving the Migrant Section of the LFS to the GHS. On reflection (for the reason given immediately above) it is clear that this would not be a wise thing to do.

If the changes to the LFS suggested above (primarily the insertion of the GHS social grant questions) were made, the LFS would not fall far short of the GHS in its ability to serve as a poverty measuring instrument, at least as far as measuring income poverty is concerned. The only other addition required is the restoration of a question on household expenditure. This could be like the GHS questions, with one question designed to elicit broad expenditure categories, and the other to elicit a little more detailed information on half-a-dozen items of household expenditure (questions 4.79 and 4.80 in the 2005 GHS). The expenditure questions function as a useful device against which to check reported income (there is, of course, little to prevent respondents under- or over-reporting either, but that is something with which one has to live).

Pooling of data, with consequent increases in the reliability of the survey results, becomes possible if the questions on income and expenditure in the GHS and LFS are asked in an identical manner. This obviously adds to the duplication that is already a feature of the surveys. It also entails an increase in respondent and interviewer burden. If, however, the collection of data in two independent surveys, whose results can be pooled, improves the quality of the information about poverty, then the duplication and increased respondent/interviewer burden would seem to be a small price to pay.

Since it is intended that the LFS be conducted in future on a quarterly basis, the question of which of the surveys needs to gather all of the information discussed above will have to be addressed. There are two obvious candidates: the June and September surveys. There is probably some virtue in maintaining September as the month in which the full survey is conducted, even though it is slightly more distant in time from the GHS than would be the future June LFS.

Although neither the LFS nor the GHS were designed with the intention that they be used as instruments for measuring income poverty, until 2004 both offered useful, if somewhat incomplete or flawed, bases for doing so. If the recommendations made above are implemented, the surveys will acquire the ability, jointly and severally, to contribute significantly to our understanding of poverty. Pooling the survey data could make that contribution more significant still.

General

Official disbelief of the results of most poverty studies undertaken in South Africa (nearly all based on survey and census data supplied by Statistics South Africa) is quite marked. The work by van der Berg and his colleagues at the University of Stellenbosch, whose results are contested in the present paper, is an exception to this, both in the sense that it makes use of non-official (AMPS) as well as official (national accounts) data, and in the sense that its results have been welcomed by the state. There can be little doubt that this reception is a response to the van der Berg *et al* finding that income poverty fell dramatically between 2000 and 2004, and to the claim that the headcount stood at 15 odd million in 2004, as opposed to the 18–20 million in 2000 found by other researchers.

Independent verification of the van der Berg *et al* results, or rejection of them if they cannot be verified, is a task of the utmost importance. A

potentially good source of the data with which to tackle this task is Statistics South Africa, especially if the information yielded by the GHSs and LFSs can be pooled in the manner suggested above. The organisation aspires to the position of preferred supplier of statistics. If it is to lay claim to that mantle in the field of poverty studies, it is going to have to do considerably better than it has done up until now. The fact that the Babita *et al* (2003) paper on poverty has been sitting inside the organisation in draft form for about three years, smacks of paralysis in the face of an overwhelmingly important political goal.

Rather than allowing poverty estimates based on private sector surveys to rule the debate, surveys that are as unequal to the task, if not more so, than the data supplied by Statistics South Africa, the official statistics producer has to make its mark once more in this field. The recommendations offered above represent a beginning. It may well be that in the future, a special survey to measure poverty is required. In the meanwhile, however, the opportunity for researchers to place before the public respectable estimates of the extent of the problem of income poverty in South Africa lies within easy reach. Of course, the GHS allows researchers to delve into other aspects of poverty as well, and in some detail. A critical review of the ability of the additional data collected by the survey to answer the 'social wage' questions would also not be out of order.

As far as the construction of poverty measures themselves is concerned, there is still quite a bit of research required as to the best ways of extracting the poverty information from the survey data. The slight progress in tackling the problem of under-reporting made in the present paper stops far short of ensuring that the surveys deliver data on the basis of which reliable estimates of inequality could be made. We can probably pinpoint workerless households reasonably reliably; under-reporting of income and expenditure elsewhere in the distribution remains as difficult as ever.

References

Babita, Miriam; Özler, Berk; Shabalala, Nozipho and Thema, Harry. 2003. “Changes in Poverty and Inequality in South Africa: 1995-2000”, Unpublished manuscript, Statistics South Africa. Cited in Hoogeveen and Özler, 2004.

Barnes, Helen; Guitierrez-Romero, Roxana and Noble, Michael. undated, 2005. “Multiple Imputation of Missing Data in the 2001 South African Census”, Working Paper Number 4, Centre for the Analysis of South African Social Policy (CASASP), University of Oxford.

Bhorat, Haroon; Leibbrandt, Murray; Maziya, Musi; van der Berg, Servaas, and Woolard, Ingrid. 2001. *Fighting Poverty: Labour Markets and Inequality in South Africa*, Cape Town: UCT Press.

Deaton, Angus. 2003. “Measuring Poverty in a Growing World (or Measuring Growth in a Poor World)”, NBER Working Paper No. 9822, Cambridge, Massachusetts: National Bureau for Economic Research.

Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*, Baltimore: Johns Hopkins Press.

Community Agency for Social Enquiry (CASE). 1999. *We Also Count: The extent of moderate to severe reported disabilities and the nature of the disability experience in South Africa*. Disability Survey for the Department of Health. Cited in Dube, 2005.

Department of Social Development (DoSD). 2002. *Transforming the Present – Protecting the Future: Report of the Committee of Inquiry into a Comprehensive System of Social Security for South Africa*, (Consolidated Report), Pretoria. Referred to as the Taylor Report, or DoSD, 2002.

Dube, Andrew K. 2005. “The role and effectiveness of disability legislation in South Africa”, report prepared for Disability KaR (Knowledge and Research), March 2005. Downloaded from www.disabilitykar.net, 20th February 2006.

Havinga, Ivo; Kamanou, Gisele and Vu, Viet. n.d. “Bridging the Poverty Measures based on National Accounts and household Income and Expenditure Survey”, United Nations Statistics Division (no date but probably 2004 or 2005).

Hirschowitz, Ros (ed). 2000. *Measuring Poverty in South Africa*, Pretoria: Statistics South Africa.

Hirschowitz, Ros, Orkin, Mark and Alberts, Piet. 2000. “Key baseline statistics for poverty measurement”, in Hirschowitz, Ros (ed.). 2000, pp.53–81.

Hoogeveen, Johannes G and Özler, Berk. 2004. “Not Separate, Not Equal: Poverty and Inequality in Post-Apartheid South Africa”, Draft paper, World Bank, February 2004.

Ijumba, Petrida; Day, Candy and Ntuli, Antoinette (eds). 2004. *South African Health Review*, Health Systems Trust. Referred to as SA Health Review 2004.

Keswell, Malcom and Poswell, Laura. 2002. “How Important is Education for Getting Ahead in South Africa?”, Centre for Social Science Research, University of Cape Town, CSSR Working Paper No. 22, December 2002.

McClaren, Pam; Solarsh, Geoff and Saloojee, Gillian. 2004. “Disability and Disablement”, in SA Health Review 2004, pp.163-173.

Meth, Charles. 2004. “Half Measures: The ANC’s Unemployment and Poverty Reduction Goals”, DPRU Working Paper No 04/89, Development Policy Research Unit, University of Cape Town, December 2004. Referred to as Meth, 2004.

Meth, Charles. 2005a. “Half measures revisited: The ANC’s unemployment and poverty reduction goals”, forthcoming in *Poverty and Policy in Post-Apartheid South Africa*, edited by Haroon Borhat and Ravi Kanbur (extensively revised version of Meth, 2004). Referred to as Meth 2005a.

Meth, Charles. 2005b. “Reflections on the difficulties of identifying and protecting vulnerable children and orphans”, unpublished paper prepared for Save the Children UK, School of Development Studies, University of KwaZulu-Natal, June 2005. Referred to as Meth, 2005b.

Meth, Charles and Dias, Rosa. 2004. “Increases in poverty in South Africa, 1999-2002”, *Development Southern Africa*, 21(1), March 2004, pp.59–85.

Nattrass, Nicoli. 2006. “Disability and Welfare in South Africa’s Era of Unemployment and AIDS”, CSSR Working Paper No. 147, University of Cape Town, February 2006.

National Treasury. 2004. *Trends in Intergovernmental Finances: 2000/01–2006/07*, Pretoria, August 2004.

Seekings, Jeremy. 2002. “The Broader Importance of Welfare Reform in South Africa”, *Social Dynamics*, Winter 2002, 28(2), pp.1–38.

Simkins, Charles. 2004. “What happened to the distribution of income in South Africa between 1995 and 2002?” unpublished paper, University of the Witwatersrand, 22 November 2004.

South African Advertising Research Foundation. 2005. *AMPS 2004 Technical Report: Sampling method*, downloaded from www.saarf.co.za, 6th December 2005.

Statistics South Africa. 2002. *Earning and spending in South Africa: Selected findings and comparisons from the income and expenditure surveys of October 1995 and October 2000*, Pretoria: Statistics South Africa.

Statistics South Africa. 2004. *Census 2001: Primary tables South Africa, Census '96 and 2001 compared*, Report No. 03-02-04 (2001), Pretoria.

Statistics South Africa. Data sets for General Household Surveys 2003 and 2004, and Labour Force Surveys September 2003 and 2004.

The Presidency (PCAS). 2003. *Towards a Ten Year Review: Synthesis Report on Implementation of Government Programmes*, Policy Co-

ordination and Advisory Services (PCAS). Referred to as PCAS, 2003.

Van der Berg, Servaas; Burger, Ronelle; Burger, Rulof; Louw, Megan and Yu, Derek. 2005. "Trends in poverty and inequality since the political transition", Stellenbosch Economic Working Papers, No. 1 / 2005. Referred to as van der Berg *et al*, 2005.

Woolard, Ingrid and Leibbrandt, Murray. 2001. "Measuring Poverty in South Africa", in Borat *et al* (eds), 2001, pp.41–73.

The Southern Africa Labour and Development Research Unit

The Southern Africa Labour and Development Research Unit (SALDRU) conducts research directed at improving the well-being of South Africa's poor. It was established in 1975. Over the next two decades the unit's research played a central role in documenting the human costs of apartheid. Key projects from this period included the Farm Labour Conference (1976), the Economics of Health Care Conference (1978), and the Second Carnegie Enquiry into Poverty and Development in South Africa (1983-86). At the urging of the African National Congress, from 1992-1994 SALDRU and the World Bank coordinated the Project for Statistics on Living Standards and Development (PSLSD). This project provide baseline data for the implementation of post-apartheid socio-economic policies through South Africa's first non-racial national sample survey.

In the post-apartheid period, SALDRU has continued to gather data and conduct research directed at informing and assessing anti-poverty policy. In line with its historical contribution, SALDRU's researchers continue to conduct research detailing changing patterns of well-being in South Africa and assessing the impact of government policy on the poor. Current research work falls into the following research themes: post-apartheid poverty; employment and migration dynamics; family support structures in an era of rapid social change; public works and public infrastructure programmes, financial strategies of the poor; common property resources and the poor. Key survey projects include the Langeberg Integrated Family Survey (1999), the Khayelitsha/Mitchell's Plain Survey (2000), the ongoing Cape Area Panel Study and the Financial Diaries Project.

Southern Africa Labour and Development Research Unit
School of Economics
University of Cape Town
Private Bag, Rondebosch, 7701
Cape Town, South Africa

Tel: +27 (0) 21 650 5696
Fax: +27 (0) 21 650 5697

Email: brenda.adams@uct.ac.za
Web: www.saldru.uct.ac.za

