



**DPRU**  
WORKING PAPERS

**Understanding Contemporary Household  
Inequality in South Africa**

**Murray Leibbrandt, Haroon Borat and  
Ingrid Woolard**

No 99/25  
ISBN: 0-7992-1965-7

May 1999

Development Policy Research Unit  
University of Cape Town

**Understanding  
Contemporary Household Inequality in South Africa**

**Murray Leibbrandt, Haroon Borat and Ingrid Woolard**

**Development Policy Research Unit  
Working Paper 99/25**

**May 1999  
ISBN 0-7992-1965-7**

## **Abstract**

This paper uses various decomposition techniques to understand the nature of household inequality in contemporary South Africa.

It examines, firstly, the importance of race in overall inequality; secondly, the contribution of major income sources to national inequality; and thirdly, the relationship between inequality, poverty and the labour market.

The paper finds that income inequality between the races in South Africa is amongst the highest in the world. Within-race inequality is also high with intra-African inequality being highest. The paper also shows the importance of differential access to wage income in driving household income inequality in South Africa.

## **Acknowledgements**

This paper is part of a larger one-year project on Labour Markets, Poverty and Inequality in South Africa. We would like to thank the African Economic Research Consortium (AERC) and the South African Department of Labour for their generous financial assistance.

## Table of Contents

Introduction	1
The Importance of Race in National Inequality	3
Measures and Estimates of Income Inequality	4
Sources of Income and National Inequality	7
A closer look at inequality, poverty and the labour market	16
Conclusion	18
References	19
Appendix 1	21
Appendix 2	25
A. "Decomposing" the Gini coefficient	25
B. The effect of a change in income on the Gini	26
C. Theil's Entropy Index (T)	28
D. Theil's second measure (L)	28
E. "Decomposing" Atkinson's index	29
F. Choice of decomposition technique	30
Appendix 3	31

## List of Tables

Table 1: Comparison of selected middle income countries	1
Table 2: Comparison of distribution measures	5
Table 3: Within-race contribution to overall inequality	5
Table 4: The Theil-T decomposition by race in Malaysia	6
Table 5: Decomposition of total national income by income shares	11
Table 6: Decomposition of total national income by income sources	
• below the poverty line	14
• above the poverty line	15
Table 7: Expanded definition of unemployment	17
Table 2.1: Comparison of distribution measures	21
Table 3.1: Within-race contribution to overall inequality	21
Table 5.1: Decomposition of total national income by income sources (PSLSD)	22
Table 6.1: Decomposition of total national income by income sources	
• below the poverty line	23

• above the poverty line	24
Table 7.1: Narrow definition of unemployment	31

## 1. Introduction

The dominant themes of South Africa's economic history are inequality and exclusion. Given this history, a key benchmark against which all contemporary economic planning must be assessed is the role of such plans in narrowing inequality and breaking down the barriers that exclude people from participating in the economy on the grounds of race, gender or location.

Such planning necessitates an information base detailing the dimensions of inequality and poverty in South Africa in the mid 1990s. This base needs to be nuanced enough to allow for assessment of programmes that are narrowly targeted at different segments of the South African population. For example, with policy focusing predominantly on the upliftment of the African group, an understanding is needed of the economic forces at work within this populace. An accurate information base is also a *sine qua non* for any more ambitious social development modelling that sets out to inform the policy selection process.

South Africa's Gini coefficient has always served as the starkest indicator of the country's unequal distribution of income.<sup>1</sup> For a long time, South Africa's Gini was the highest recorded in the world. Table 1 presents a comparison of South Africa's Gini coefficient and income shares to countries with similar income levels. It is clear that Brazil and South Africa are far less egalitarian societies than the other nations presented here<sup>2</sup>, but also that Brazil has a slightly higher level of income inequality compared to South Africa. Both these Gini values though are extremely high, indicating very skewed distributions of income. By comparison, Poland and Thailand have Gini coefficients of 0.27 and 0.46 respectively, showing that these economies have a significantly more equitable distribution of income.

**Table 1: Comparison of selected middle income countries**

	Poland	Thailand	Venezuela	Brazil	South Africa	Malaysia
GNP per capita US\$ (1994)	2 410	2 410	2 760	2 970	3 040	3 480
Gini	0.27	0.46	0.54	0.63	0.60	0.48
% share of income of poorest 20%	9.3	5.6	3.6	2.1	2.8	4.6
% share of income of richest 10%	22.1	37.1	42.7	51.3	41.9	37.9

Source: 1996 World Development Report and own calculations (South Africa).

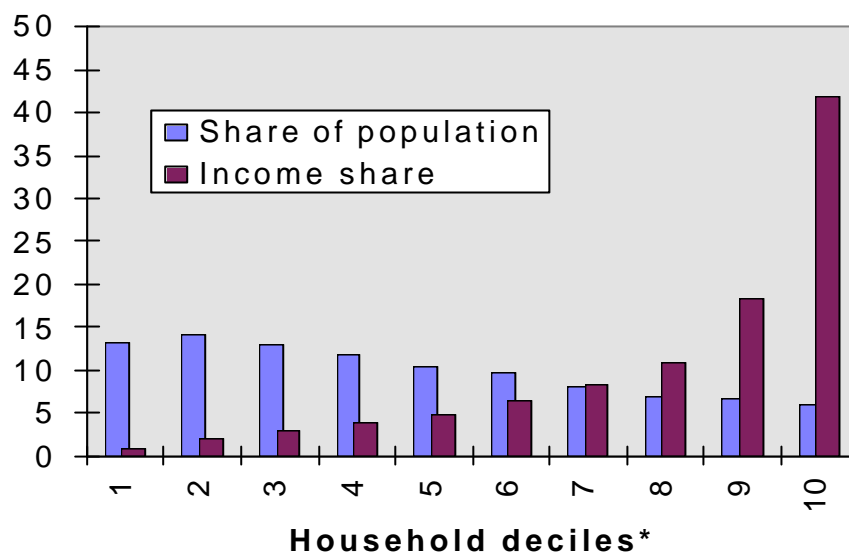
Another way to express the degree of inequality in a country is to examine the income shares of households by decile. From Figure 1, it is evident that a decile is a 10% segment of all households ranked according to income level. Using this measure, the degree of inequality is striking. The poorest four deciles (40%) of households, equivalent to 52% of the population, account for less than 10% of total income, while the richest decile (10%) of households, equivalent to just 6% of the population, capture over 40% of total income.<sup>3</sup>

<sup>1</sup> The Gini coefficient always has a value between zero and one. The bigger the number, the more inequality exists.

<sup>2</sup> Note that because of variability in the date of data collection and differing methodologies, these figures should be taken as indicative only.

<sup>3</sup> As the average household size is larger for the lower deciles, the inequality is worse than it would have been if no attention had been given to household size. However, in our calculations household incomes were ranked according to adult equivalent incomes rather than per capita incomes. These adult equivalents give explicit cognisance to the fact that children require less income than adults and that there are certain economies of scale associated with larger households. Following May (1995), we used an adult equivalence scale here and later in the paper, of the form:

$E = (A + 0.5K)^{0.9}$ , where E = number of adult equivalents, A = number of adults, and K = number of children.

**Figure 1: Inequality in South Africa**

Whiteford and McGrath (1998) have shown that, while the Gini coefficient remained static between 1975 and 1991, this masked the fact that the rich got richer while the poor got poorer. They found a similar pattern when taking each race group separately. In other words, they observed a widening of the gap between the richest Africans and the poorest Africans, the richest Whites and the poorest Whites. For example, the income share accruing to the poorest 40% of African earners fell by a disquieting 48%, while the share accruing to the richest 10% rose by 43%.<sup>4</sup>

Given the above background, this paper seeks to further unpack the nature of inequality in contemporary South Africa. The intention is to complement the recent historical review provided by Whiteford and McGrath (1998) - especially their between and within race analysis - and then to focus more explicitly on the link between the labour market and household inequality. There are three sections to the paper. The first section examines the racial fault line in South African inequality through the use of various categorical decomposition techniques. Once the aggregate importance of 'between racial group' versus 'within racial group' inequality has been examined, the second section uses a decomposition analysis of income inequality by income components to immediately focus attention on the key labour market, asset ownership and state welfare processes driving South Africa's inequality. This analysis suggests that the labour market is the key driver of household inequality. In the light of this finding, the final section focuses explicitly on the labour market.

In South African policy debates there is generally insufficient scrutiny of empirical results. There are two major maladies. First, there is scant recognition of the fact that different measurement techniques are going to generate different results. In contrast to this, the theoretical literature on inequality has paid a great deal of attention to the fact that different measures of inequality do not define inequality in exactly the same way and therefore will arrive at different estimates of inequality.<sup>5</sup> The importance of this literature lies in its questioning of the extent to which any inequality results are technique driven rather than neutral representations of the circumstances prevailing in that society. We control for this

<sup>4</sup> These figures are questioned by the Centre for Development and Enterprise (1995) who suggest that the distribution among black households is "more equal" than Whiteford and McGrath suggest. CDE Research No. 1, September 1995: "Post-Apartheid Population & Income Trends: A New Analysis".

<sup>5</sup> See Deaton (1997) and Cowell (1995) for recent reviews of this literature.

possibility by using a variety of techniques wherever possible. Conflicting results will then serve as an indication that the situation really is not as clear-cut as any of the techniques would have us believe. Indeed, a detailed discussion of how techniques differ and why these differences should have led to the measured differences, is itself a useful way to start an interrogation of the processes generating inequality in South Africa.<sup>6</sup>

Second, the South African literature gives too little recognition to the fact that different data are going to give different results. In the past this could be excused because of the shortage of usable data. However, there are now two national data sources available to cover contemporary South Africa; the 1995 *Income and Expenditure Survey* (IES) conducted by the South African Central Statistical Service and the survey undertaken as part of the *Project for Statistics on Living Standards and Development* (PSLSD) by the Southern Africa Labour and Development Research Unit at the University of Cape Town in late 1993. There are clear advantages to using IES data rather than the PSLSD: the data is more recent, the sample was much larger (almost 30 000 households compared with just under 9 000 in the PSLSD study), and the questionnaire was solely devoted to collecting income and expenditure data which should point to greater attention to detail and less respondent fatigue. The only disadvantage to using this data is that it does not provide information about small-scale agricultural production and consumption from own production.<sup>7</sup> We therefore focus our textual analysis on the IES data. However we have used the PSLSD data to reproduce all of our tables (see Appendix 1) and we highlight any discrepancies in results from the two sets of data as part of our discussion.

## 2. The Importance of Race in National Inequality

The literature on the decomposition of total inequality by sub-groups has a long lineage.<sup>8</sup> If we divide the population into mutually exclusive exhaustive sub-groups then there is a degree of inequality both *within* these sub-groups and *between* these sub-groups. It is desirable that one should be able to decompose a measure of overall inequality into the "within" and the "between" portion.<sup>9</sup> The value of decompositions is that "they gauge the relative importance of various sources and sectors in respect of overall inequality, and thereby direct our attention to potentially fruitful areas of research" (Fields, 1980: 438). Indeed South Africa's historical legacy makes a much stronger case than this, in that we are drawn to these tools because they allow for an explicit focus on race in driving inequality. This section concentrates solely on this racial question.

### Measures and Estimates of Income Inequality

The most commonly cited additively decomposable measure of inequality is the Theil T-statistic, derived directly from the notion of entropy in information theory (Fields, 1980: 103). The definition of the Theil-T is given in Appendix 2C.

The Theil-T can be decomposed as follows:

---

<sup>6</sup> To bring out such points requires that we spend a lot of time unpacking the details of the various techniques. We have done this in the paper partly in the hope that access to such information will be valuable to South African readers. However, we have relegated much of this discussion to the appendices wherever possible.

<sup>7</sup> While the questionnaire asked for considerable detail about these activities, the CSS failed to place a value on them.

<sup>8</sup> See Fields (1980) for a review.

<sup>9</sup> A generally decomposable or aggregative index is defined as one where the overall inequality level can be expressed as some general function of the sub-group means, population sizes and inequality measures. The most useful type of decomposability is additive decomposability. A measure is additively decomposable if it can be tidily expressed as the sum of a "between-group" term and a "within-group" term. The between group component is the value of the measure were every member assigned the group mean (i.e. there is assumed to be no inequality within the group). Similarly, the within-group component is the value of the inequality measure when all the between group inequalities are suppressed.



$$T = T_B + \sum q_i T_i$$

where  $T_i$  is the Theil-T inequality measure within the  $i$ th group,  $q_i$  is the proportion of income accruing to the  $i$ th group and  $T_B$  is the between group contribution.  $T_B$  is calculated the same way as  $T$ , but assuming that all incomes within a group are equal.

The Theil-L (defined in the Appendix 2D) decomposes in a similar way to the Theil-T, except that the group statistics are weighted by the proportion of households (not income) in each group, i.e.

$$L = L_B + \sum p_i L_i$$

where  $p_i$  is the population share of the  $i$ th group.

A second broad class of inequality measures is contained in the Atkinson measure. This starts from an additive social welfare function in order to derive the following inequality index::

$$I = 1 - \left( \frac{1}{N} \sum_{i=1}^N (y_i / m)^{1-\epsilon} \right)^{1/(1-\epsilon)}$$

which may be decomposed into between- and within-group inequality such that

$$I = I_B + I_W + Residual$$

As explained in Appendix 2E, the measure can be interpreted as the proportion of the present total income that would be required to achieve the same level of social welfare as at present if incomes were equally distributed (Atkinson, 1970:48). Atkinson explicitly introduces distributional objectives through the parameter  $\epsilon \geq 0$  which represents the weight attached to inequality in the distribution. By specifying different values of  $\epsilon$  one can vary the importance society attaches to mean living standards versus equality. If society is indifferent about the distribution, we will set  $\epsilon$  equal to zero. By increasing  $\epsilon$  we give more weight to inequality at the lower end of the distribution. At  $\epsilon$  equal to infinity, society is concerned only with the poorest household.

All three of the above decomposition techniques, the Theil-L, Theil-T and Atkinson's measure, would seem to have obvious relevance in South Africa. Yet, it is only recently that such decompositions have begun to be used in South Africa.<sup>10</sup> Table 2 presents the results of the decomposition of South Africa's total national income by race using the three decomposition techniques discussed above. We attach the household adult equivalent income to each individual in the household, thus we are comparing inequality amongst individuals, not households.

---

<sup>10</sup> See McGrath and Whiteford (1994) and Moll (1998).

**Table 2: Comparison of Distribution Measures**

MEASURE	Between Component	Within Component	Residual	Total
Theil-T	0.319 (39.7)	0.483 (60.3)		0.802
Theil-L	0.254 (36.0)	0.452 (64.0)		0.706
Atkinson $\epsilon = 0.5$	0.091 (32.8)	0.187 (67.0)	0.001 (0.2)	0.278
Atkinson $\epsilon = 1.5$	0.215 (36.8)	0.368 (63.0)	0.001 (0.2)	0.584
Atkinson $\epsilon = 2.5$	0.279 (38.5)	0.445 (61.5)	0.001 (0.01)	0.724

NOTE: The figures in brackets show the percentage contribution to total inequality.

All the indices point in a similar direction, i.e. that the "within" and "between" components are both important contributors to overall inequality, with within-group inequality accounting for more than three-fifths of overall inequality, by all the measures used. In addition, the more highly we value equality - i.e. the larger we set  $\epsilon$  in Atkinson's index - the between group inequality grows in significance as a contributor to overall inequality. This suggests relatively greater equality amongst African households as we move down the income distribution, and that a key inequality wedge is that between White and African households.

Table 3 further decomposes the within-group Theil measures, by race. Hence we determine the share of each racial group in explaining aggregate within-group inequality. It is immediately evident that the choice of the Theil-L versus the Theil-T index paints a very different picture of the contribution of different races to overall inequality.

**Table 3: Within-race contribution to overall inequality**

Measure	African	Coloured	Asian	White
Theil-T	0.594 [0.265] (33.2)	0.387 [0.027] (3.4)	0.400 [0.018] (2.2)	0.395 [0.172] (21.4)
Theil-L	0.486 [0.371] (52.6)	0.353 [0.030] (4.2)	0.350 [0.009] (1.3)	0.333 [0.042] (5.9)

NOTES:

1. The first row of figures show the measure when considering only the particular race group.
2. The figures in square brackets show the absolute contribution to total inequality.
3. The figures in round brackets show the percentage contribution to total inequality.
4. Atkinson's index is generally but not additively decomposable, hence we cannot apportion the within contribution amongst the race groups.

The Theil-T suggests that inequality among the White group is almost as large a contributor to overall inequality as inequality amongst the African group, yet the Theil-L suggests that African inequality contributes 52.6% to total inequality vis-à-vis a contribution of 5.9% from white inequality. The reason for the different Theil-T and Theil-L results can be found in the use of income as opposed to population weights. This stresses the importance of considering the nature of the decomposition measure before relying on any one statistic.

The results in Tables 2 and 3 utilising the PSLSD data set instead, are provided in Appendix 1. It is evident that a similar general result is found, namely that overall inequality is driven primarily by within-group inequality. However, there remain important differences in the figures obtained from the two surveys. Firstly, while the within-group component does drive overall inequality from the PSLSD, its share on average of 55%, is much smaller than the corresponding mean of 63% found in the IES. Hence, the IES data has a much stronger contribution from the within-group component to total inequality. Secondly, the Atkinson measure for the IES data showed that the higher the value of  $\epsilon$ , the greater the contribution of between-group inequality to the total.

For the PSLSD, however, the result is the opposite: the higher the value of  $\epsilon$ , the smaller the contribution of between-group inequality. This suggests that since most of the poor are African, inequality at the lower end of the income distribution will tend to be within-group. A third important difference in the data is in the within-race Theil measures. Here, the PSLSD also shows a rising contribution of African inequality, from 22.6% to 48.1%, when the Theil-L is used instead of the Theil-T. However, the Theil-T African and White contributions to overall inequality are almost the same, at 22.6% and 22.1% respectively, which is very different to the IES figures where the African and White contributions are much further apart.

One of the possible explanations for the above trends is that in the PSLSD survey, the variance in income amongst Africans is smaller given that more poor Africans relative to non-poor Africans were sampled. This would imply that the contribution of the within-group component in PSLSD to overall inequality is not as large as in IES. However, as there are relatively more poor than non-poor Africans in PSLSD, an increased concern about inequality through Atkinson's measure, meaning a move down the income distribution, leads to the within-group share re-asserting itself in the measure. Furthermore, this relatively higher number of poor Africans sampled may be reflected in the lower within-race Theil-T value, at 22.6%, as opposed to 33.2% for IES. Ultimately then, the comparison between the two surveys suggests that the sample income distributions differ, with the distribution in the PSLSD weighted more toward poor African individuals and households.

Despite these difficulties in trying to reconcile results from these two data sets the racial contribution to inequality in South Africa, within an international perspective, provides for a fairly clear, and indeed very powerful picture. The table below illustrates this. Here Anand's (1983) Theil-T decomposition analysis of Malaysian household income by race is provided. Malaysia offers a good comparative example, as it is also a society with a history of social and economic stratification by race.

**Table 4: The Theil-T Decomposition By Race in Malaysia**

	<b>Per Capita HH</b>
Malay	0.41
Chinese	0.42
Indian	0.54
Other	0.94
All (Total)	0.52
Within	0.45 (87%)
Between	0.07 (13%)

Source: Anand (1983:95-6)

Malaysia's between-group share in inequality at 13% though, is very low compared to South Africa's. When using personal income instead, Anand (1983:96) finds an even lower contribution of 9.2% due to between-group inequality. The baseline value for South Africa, either with the IES or PSLSD data, for the Theil-T measure is 36%. In the case of Malaysia then, between group inequality is not very helpful in explaining individual income inequality. In South Africa on the other hand, income inequality between the four racial groups, and particularly between African and White, is a crucial predictor of total income inequality in the society.

A Latin American example, provided by Fiszbein and Psacharopoulos (1995), helps to further illustrate just how important race is in South Africa. This study of seven countries at two periods of time finds that the *joint contributions* of Age, Employment Status and Education to worker per capita income ranges from 32.6% to 53% with a mean of 45.3 %. Given that these are the major variables that one conventionally looks at in explaining the individual incomes, this gives a sharp sense of perspective on the magnitude of the between race numbers in South Africa.

### 3. Sources of Income and National Inequality

The decomposition literature of the previous section has a much older vintage than the income source analysis of this section. However, over the last decade, a busy international literature has developed around the derivation and refinement of techniques for decomposing inequality measures (in particular the Gini coefficient) by income sources.<sup>11</sup> Such decompositions highlight those income sources that are dominating the distribution of income and, as such, offer a bridge between the description of inequality and the key economic processes generating inequality in a society.

Elsewhere (Leibbrandt *et. al.*, 1996) we have decomposed the Gini coefficient by income sources using the PSLSD survey. We report these results in Appendix 1. Here, we apply the same methodology to the IES. The application of such work to South Africa provides an immediate addition to our knowledge of South African inequality. The IES data set contains detailed information on all sources of income and, therefore, is an ideal data set to apply such analysis. Clearly, the level of aggregation that is chosen is determined by the context under consideration and the questions that the analysis is addressing. For the purposes of this paper, we want to distinguish between the relative importance of the major foci of policy attention. Total income for each household is therefore divided into five sources:

- (1) *Remittances* - remittances from absent family members and marital maintenance (alimony);
- (2) *Wage income* - regular and casual employment and value of benefits such as subsidised housing, transport and food;
- (3) *Capital income* - dividends, interest, rent income, imputed rent from residing in own dwelling and private and civil (contributory) pensions;
- (4) *State transfers* - social pensions, disability grants, poor relief, unemployment insurance and child maintenance grants;
- (5) *Self-employment* - informal and formal business activities.

Such a breakdown is still at a fairly aggregate level and any number of more disaggregated breakdowns are possible to answer more specific questions.

---

<sup>11</sup> The literature starts with Shorrocks (1983) and is most recently extended by Lerman and Yitzhaki (1994).

The key aspects of the decomposition technique are summarised here, while a more detailed description is provided in Appendix 2. If South African society is represented as  $n$  households deriving income from  $K$  different sources (i.e.  $K$  different income components), then the **Gini coefficient** ( $G$ ) for the distribution of total income within the group can be derived as:

$$G = \sum_{k=1}^K R_k G_k S_k$$

where:

$S$  is the share of source  $k$  of income in total group income (i.e.  $S_k = m_k/m$ ),

$G_k$  is the Gini coefficient measuring the inequality in the distribution of income component  $k$  within the group, and

$R_k$  is the Gini correlation of income from source  $k$  with total income.<sup>12</sup>

This equation tells us that the effect of source  $k$  income on total income inequality can be broken down into three components:

- the share of income component  $k$  in total income (captured by the term  $S_k$ );
- the inequality within the sample of income from source  $k$  (as measured by  $G_k$ ); and
- the correlation between source  $k$  income and total income (as measured by  $R_k$ ).

The larger the product of these three components, the greater the contribution of income from source  $k$  to total income inequality. However, it must be noted that whilst  $S_k$  and  $G_k$  are always positive and less than one,  $R_k$  can fall anywhere on the interval  $[-1,1]$ . When  $R_k$  is less than zero, income from source  $k$  is negatively correlated with total income and thus serves to lower the overall Gini measure for the sample.

Now, suppose that there is an exogenous increase in income from source  $j$ , by some factor  $s_j$  then it can be shown that the derivative of the Gini coefficient with respect to a change in income source  $j$  is:

$$\frac{\partial G}{\partial s_j} = S_j (R_j G_j - G)$$

If  $\partial G/\partial s_j$  is negative then a marginal increase in income component  $j$  will lessen income inequality. This will be the case either when:

- income from component  $j$  has either a negative or zero correlation with total income ( $-1 \leq R_j \leq 0$ ); or when
- income from source  $j$  is positively correlated with total income ( $R_j > 0$ ) and  $R_j G_j < G$ .

Alternatively, in order for a marginal increase in source  $j$  income to worsen income inequality it is necessary that  $G_j > G$  (i.e. income from source  $j$  must be more unevenly distributed than total income). However this condition alone is not sufficient for a change in income component  $j$  to worsen the overall income distribution as the sign of  $\partial G/\partial s_j$  will still be influenced by the strength of the Gini correlation between source  $j$  income and total income (Stark *et al.*, 1986: 260).

Table 5 presents the results of this decomposition for the total South African sample.<sup>13</sup> A few illustrative features of this table will be highlighted. It can be seen that wage income has a dominant share of

<sup>12</sup>  $R_k$  is a form of rank correlation coefficient as it measures the extent to which the relationship between  $Y_k$  and the cumulative rank distribution of total income coincides with the relationship between  $Y_k$  and its own cumulative rank distribution.

<sup>13</sup> It should be noted that the overall Gini coefficient in the table is 0.60 as opposed to the 0.65 of McGrath and Whiteford (1994) from the same data. McGrath and Whiteford re-weighted the sample to coincide with 1991 census population shares,

income (66%) and makes a similar contribution to inequality (67%). The reason for this is the high  $R$  of .88, implying that a household's rank in the distribution of wage income is strongly correlated with that household's rank in the distribution of total income. This strong correlation is more than enough to compensate for the fact that the Gini coefficient for wage income (.67) is the lowest of all income sources.

The Gini coefficient for a particular income source ( $G_k$ ) is driven by the inequality amongst those earning income from that source ( $G_A$ ) and the proportion of households who have positive income from that source ( $P_k$ ), or, changing the focus, *the proportion of households with no access* to a particular income source ( $1-P_k$ ). Then we see that, for example:

$$G_{\text{wage}} = 0.67 = P_{\text{wage}} G_A + (1-P_{\text{wage}}) = 0.37 + 0.30.$$

This brings us part of the way to apportioning the "blame" for Gini inequality between the inequality amongst earners and the inequality between those with some wage income and those with none. It would appear that almost half of what we have termed "wage inequality" is in fact driven by the 30% of households with zero wage income.

Remittance income has the smallest share of total income (2%) and makes a small, negative contribution to inequality (-.25%). This negative contribution arises because of the small negative correlation ( $R = -.07$ ) between the rank ordering of remittance income and the rank ordering of total income. This negative correlation would seem to imply that the fairly high Gini coefficient for remittances is due to the fact that remittance income is disproportionately distributed to those at the bottom of the total distribution relative to those at the top. In essence, this analysis is suggesting that factors that boost remittance income for current recipients would lower overall inequality.

The last column of Table 5 shows the effects of a 1% increase in a particular income component. We see that a change in state transfers, remittances or income from self-employment will have the greatest effect on the overall Gini. In the last case the Gini increases, but in the other two cases it decreases. The components which increase inequality correlate highly with total income rankings (i.e.  $R_k$  is high), which implies that an increase in these sources will primarily benefit the better off and thus aggravate the Gini. The sum of the absolute changes in the Gini coefficient is zero. This follows because increasing all components of income by 1% has no effect on the income distribution and therefore no effect on the Gini.

From the point of view of government policy, state transfers are of special interest. A well-targeted, redistributionist state expenditure programme would be evidenced by a strongly negative  $R$ . The value of  $R$  at  $-0.12$  suggests that state transfers serve to decrease the value of the overall Gini. Moreover, we see that an increase in state transfers of 1% will reduce the Gini by 0.04 (7%).

While  $G_k$  is the coefficient needed to calculate the contribution to inequality, a look at  $G_A$  is instructive.  $G_A$  is the Gini coefficient when considering only those households actually receiving income from that particular source. We see that there are large disparities in the incomes being earned from self-employment, capital income and wage income. This points to the dichotomous nature of the South African economy, where there is an immense gap between those engaged in high and low wage employment, formal versus informal self-employment and those earning income from interest and dividends versus those accruing a small capital benefit as a result of owning their dwelling.

A comparison with the PSLSD shows fairly similar patterns in the data. These include for example, a share of wage income in total income at 69% and a contribution to overall inequality of 69%, both figures slightly higher than those in the IES. What is interesting from the wage income data is that the PSLSD reports a  $G_{\text{wage}}$  decomposition that has a greater share (34%) of households in the sample that earn no wage income. Hence, the greater wage income contribution to overall household inequality ( $G_k$ ) is a function partly of the PSLSD reporting more households with no wage earners.

---

thereby giving more weight to white incomes and accentuating inequality. In our calculations we used the survey enumeration weights and used a slightly refined data set taking account of the errors in the social pensions data discovered by Pieter le Roux (University of the Western Cape).

In terms of remittances, the PSLSD captures a larger proportion of households receiving this type of income. This is an accepted difference between the two data sets, given that the PSLSD was more diligent in tracking down remittance-recipient households. Given the relative homogeneity of these households in terms of income levels, the percentage share in the overall Gini is stronger at -0.40 rather than -0.25 found using the IES. The IES yields a stronger result on the impact of state transfers, as the PSLSD reports a zero, rather than negative, contribution to inequality for these transfers. An interesting result in comparing the two data sets, is the effect on the Gini of a 1% change in the different income components. While the IES reports the strongest impact on the Gini from a 1% increase in an income source to emanate from self-employment income (0.047), the PSLSD is far less sensitive, with self-employment only increasing the Gini by 0.006. This may be partly due to the more structured and consistent efforts made in the IES to uncover individuals involved in self-employment across the entire income distribution.

A central point to note from Table 5 though, is that while wage incomes on their own are very important, remittances and self-employment are also incomes derivative of the labour market. If the cumulative impact of all three of these sources is considered, it can be seen that the labour market dominates South African income and income inequality. This is in line with comparisons of such studies in other countries. A quote from Fields (1980) will suffice:

“Individually and together, the results for Taiwan, Pakistan and Colombia give a common impression about the contribution of various income sources to overall inequality. The bulk of income inequality is attributable to labor income. The high factor inequality weights for labor incomes suggest that the principal inequality-producing factor is some people receiving a great deal more income for their work than do others. This has important implications both for the research (researchers should study the labor market) and for policy (policy makers should create more well-paying jobs). The intuitively prior notion that the most unequally distributed factors (property, gifts, etc) contribute the most to total inequality is found to be false in each case.” (p114).

**Table 5: Decomposition of Total National Income by Income Sources**

Income source	Proportion of households receiving income source ( $P_k$ )	Mean income from source	Share in total income ( $S_k$ )	Gini for income source for households receiving such income ( $G_A$ )	Gini for income source for all households ( $G_k$ )	Gini correlation with total income rankings ( $R_k$ )	Contribution to Gini coefficient of total income ( $S_k G_k R_k$ )	Percentage share in overall Gini	Effect on overall Gini of a 1% change in income component
Remittances	0.13	R64.81	0.02	0.48	0.93	-0.07	-0.001	-0.25	-0.015
Wage income	0.70	R1815.63	0.66	0.53	0.67	0.88	0.39	66.59	0.002
Capital income	0.18	R251.51	0.09	0.69	0.95	0.69	0.06	10.16	0.006
State transfers	0.33	R155.84	0.06	0.40	0.80	-0.12	-0.006	-0.94	-0.039
Self-employment	0.09	R451.02	0.16	0.71	0.97	0.89	0.14	24.44	0.047
Total		R2738.82	1.00				0.59	100.00	

Notes:

1.  $G_A$  is the Gini for the income source when we only consider households with positive income from that source.
2.  $G_k$  is the Gini of the income source when we consider all households. Lerman and Yitzhaki (1994) show that  $G_k = P_k * G_A + (1 - P_k)$



However, while the analysis of any of the income sources presented in the table is usefully indicative, it does not really reveal enough about what is going on at the lower end of the distribution relative to the top end. So, for example, it is quite possible for the same aggregate outcomes to result from an income source that is contributing exclusively to the very poor and very rich or exclusively to the middle of the distribution. This points to the need for some complementary sensitivity analysis. A particularly useful exercise would seem to be one that splits the population by a poverty line. This was done for the South African case and the results are presented in Table 6 below.

Inspection of this table shows that the data from Table 5 above is only a rough average of very different processes taking place above and below the poverty line.<sup>14</sup> As agriculture is a consistently low contributor to average income and to inequality in both the above and the below group, it will not be discussed further at this point.

We will consider wage income first. In the above group, this income source makes a large and stable contribution to average income (68%) and to the distribution of income (63%) in the above group. This distributional effect is the result of a low Gini coefficient (.58) being offset by a high  $R$  of .84. For the below group, the share of wage income in total income is far lower (49%) but, even within the poor, higher wage income is strongly correlated with higher total income ( $R = .60$ ) and this income source therefore still makes a high contribution to inequality (63%). It is clear from this breakdown of above and below groups that access to wage income is central to determining which households are able to avoid poverty and, even, the depth to which poor households sink below the poverty line. This re-asserts the importance of the labour market in understanding poverty: that the formal earnings capacity of households will either reinforce or shed their poverty status.

On the other hand, it is encouraging to see that state transfers are a much smaller part of the total income for the above group (4%) than the below group (33%). Moreover,  $R = -.15$  in the above group reveals that this income is not going to the higher income households in society. However, the fairly high Gini coefficient for state transfers in the below group (.67) and rank correlation ( $R = .31$ ) implies that it is the relatively better-off within the poor who are receiving state transfers.

There are two possible explanations for such an outcome. The first is that the targeting of state assistance is not that successful. The second is that the depth of poverty in South African society is so acute that access to some state assistance is sufficient to move a household away from the bottom of the poverty ranking. Apartheid-derived racial biases in welfare allocations are certainly a cause of inefficient allocation.<sup>15</sup> However, there is also some recent econometric evidence (Deaton, 1995) that state pensions are not badly targeted. In addition, studies of rural poverty (see May *et al*, 1995) have made it clear that 'claims against the state' are central to rural livelihoods. Thus, on balance, the second explanation is more likely to be true.<sup>16</sup> What can be said with more certainty is that the analysis of wage income and state transfers serves to confirm that, in South Africa, the poorest of the poor are those households that lack access to either wage income or state transfers.

The low share of remittances in total income (2%) and the negative correlation for remittances ( $R = -.09$ ) in the above group along with the very much higher share of remittances (12%) in the income of the below group indicate that remittances are much more important on average in the below group. In South Africa remittances generally flow from urban to rural areas and, to a large extent, this result is merely confirming that a large component of South Africa's poor are located in the rural areas (Whiteford *et al*, 1995). However, there is additional information to be gleaned as well. The low, positive rank correlation ( $R = .16$ ) in the below group results in a small contribution to inequality (5.6%). This implies that remittance income is well disbursed within the poor. So, while remittances are not

<sup>14</sup> For ease of expression, this paper will refer to those above the poverty line as the above group and those below the poverty line as the below group.

<sup>15</sup> See Borhat (1995) for the historical details of racial biases in pension allocations.

<sup>16</sup> The one clear contribution made by this analysis of state transfers is to illustrate how careful one has to be in adding interpretation to the empirics of the income decomposition analysis.

important enough to be a major discriminator of who lies above or below the poverty line, factors which might cause an increase in remittances would have a generalised positive impact on the poor. The converse is true for capital income. This income source is far more important to the above group than the below group both as an average share (10% and 4% respectively) and as a contributor to inequality (9.9% and 5.5% respectively).

A look at the "actual" Gini ( $G_A$ ) for the components when considering only those households actually receiving such income reveals much the same picture as the analysis for the total population. We would, perhaps, have anticipated lower Ginis in the below group, which might be expected to be fairly homogenous since everyone is, after all, technically "poor". This is, however, not the case. There are considerable deviations in the incomes earned from capital income and self-employment both in the above and below groups. Self-employment in the below group, however, produces less inequality amongst those engaged in these activities than in the above group. This is to be expected, since, all those in the below group are likely to be involved in marginal informal activities.

The comparison with the PSLSD yields some interesting results. Within the below group, for example, the percentage share of wage income in total inequality is smaller, at 51.6% as opposed to 63% for IES. This is due to a larger contribution from state transfers to total inequality, of 28.4% rather than 22% in the IES. More interestingly, the effect on the Gini of a 1% change in the income source, show different signed answers within the below group for capital income, state transfers and self-employment. While IES predicts an increase in the Gini by 0.005 with capital income, the PSLSD survey predicts a decrease of 0.005. This may be due in part to the fact that imputed rent as part of capital income, was more carefully collected in the PSLSD. In this case, more poor households were allocated this income, and hence the decline in inequality from an increase in this source. Likewise, state transfers were shown to increase rather than decrease the Gini, while self-employment was predicted to decrease the Gini. The state transfers figure, makes a claim that while state pensions may be well targeted at the poor, there is a smaller share of the poor receiving these transfers than not. Hence, a rise in the value of pension payouts to poor households, in the PSLSD survey, will exacerbate inequality. The self-employment numbers suggest that the returns to the poor, through working for themselves, may be more evenly distributed than the IES predicts. Using PSLSD, one may be more confident of the equity-generating impact of increased self-employment income to poor households.

For the above group, the percentage share of wage income in the Gini is higher (70.7%) than in the IES (63.3%). Further, the contribution of capital income to the overall Gini is much higher (16.6%), while self-employment income is far less important in explaining the Gini (7.82%) than in the IES (27.98%). The impact of a 1% change in the income sources on the Gini, reveal all the same-signed answers for the two data sets. Using the PSLSD then, one would be confident that the key drivers of total inequality for the above group were wage income, followed by capital income and self-employment income. Using the IES however, one would predict that wage income, income from self-employment and then capital income determined inequality in the above group. For households above the poverty line, the PSLSD places more importance on capital income than income from self-employment in generating inequality. It is true however, that in its concerted efforts to uncover those individuals working for themselves, the IES may be a more accurate predictor of the dynamics of self-employment in the above group, and indeed the below group as well.

Irrespective of the data set used, this sensitivity analysis serves to reinforce the finding that by both above and below groups, wage income is the key determinant of income inequality in the society. In earlier work Leibbrandt, Woolard and Woolard (1996) decomposed within African income using identical methods. This within-African analysis generated a picture that is quite different from the total income picture for the above and below picture. This illustrates some of the complexity of South Africa's income dynamics. However, despite these differences, wages have a more dominant influence on South Africa's inequality (79.44% in the IES and 82.25 in the PSLSD) in this group than in any of the other cuts. Thus, the importance of wage income and, by direct implication, the labour market is very clear across all cuts of South African households.

**Table 6: Decomposition of Total National Income by Income Sources, Below and Above the Poverty Line****a. Below the poverty line**

Income source	Proportion of households receiving income source ( $P_k$ )	Mean income from source	Share in total income ( $S_k$ )	Gini for source for households receiving such income ( $G_A$ )	Gini for income source for all households ( $G_k$ )	Gini correlation with total income rankings ( $R_k$ )	Contribution to Gini coefficient of total income ( $S_k G_k R_k$ )	Percentage share in overall Gini	Effect on overall Gini of a 1% change in income source
Remittances	0.23	R78.02	0.12	0.38	0.86	0.16	0.02	5.64	-0.020
Wage income	0.50	R308.00	0.49	0.31	0.65	0.60	0.19	63.0	0.044
Capital income	0.09	R23.65	0.04	0.59	0.96	0.47	0.02	5.51	0.005
State transfers	0.49	R206.81	0.33	0.33	0.67	0.31	0.07	22.03	-0.032
Self-employment	0.04	R17.24	0.03	0.42	0.98	0.44	0.01	3.86	0.003
Total		R633.72	1.00				0.30	100.00	

**b. Above the poverty line**

Income source	Proportion of households receiving income source ( $P_k$ )	Mean income from source	Share in total income ( $S_k$ )	Gini for source for households receiving such income ( $G_A$ )	Gini coefficient for income source ( $G_k$ )	Gini correlation with total income rankings ( $R_k$ )	Contribution to Gini coefficient of total income ( $S_k G_k R_k$ )	Percentage share in overall Gini	Effect on overall Gini of a 1% change in income source
Remittances	0.09	R58.81	0.02	0.50	0.95	-0.09	-0.001	-0.27	-0.01
Wage income	0.79	R2501.56	0.68	0.47	0.58	0.84	0.33	63.3	-0.02
Capital income	0.21	R355.19	0.10	0.67	0.93	0.58	0.05	9.9	0.002
State transfers	0.26	R132.65	0.04	0.45	0.85	-0.15	-0.005	-0.91	-0.02
Self-employment	0.11	R648.39	0.18	0.68	0.97	0.86	0.15	27.98	0.05
Total		R3696.60	1.00				0.52	100.00	

#### 4. A closer look at inequality, poverty and the labour market

The overriding message of the previous section was the dominance of wage income in driving household inequality in South Africa. The decomposition analysis also robustly suggested that the role of wage income is significantly influenced both by the fact that many households have no access to wage income and by the fact that wage income is very unequally distributed across those households that do have access to it. When we have presented similar results in the past (Leibbrandt, Woolard and Woolard (1996), Borat, Leibbrandt and Woolard (1995)), such a picture has been taken to imply empirical support for an insider/outsider model of the labour market in South Africa. While our empirics certainly do not preclude this possibility, such an implication is premature. As stated in the introduction, this paper is focussing on inequality at the household level. In contrast to this, individuals are usually the focus of attention in labour market studies. There is therefore an aggregation problem and an uneasy relationship between our inequality analysis and any labour market analysis. The non-wage earning households are particularly problematic as such households do not constitute a tight labour market category. For example, a household with two pensioners would be a non-wage earning household. A household containing a mother taking care of her children would also be a non-wage earning household. Neither of these households contain *any* labour market participants and they therefore do not imply anything about the operation of the labour market.

This mapping between individuals in the labour market and household-level poverty and inequality outcomes has proved to be problematic in all international studies such as ours.<sup>17</sup> In this concluding section we look for a tighter exploration of the labour market implications of our earlier decomposition work by focussing directly on the unemployed and their attachments to different households in the society. Table 7 seeks to highlight the differences between households when classified by the number of unemployed members resident in the household. The expanded definition of unemployment is utilised here, while figures based on the narrow definition are included in Appendix 3.

From the table it can be seen that over two-thirds of households (72%) have no unemployed members. This figure falls to 64% among African households. Amongst households with unemployed members, most contain only one unemployed person. Nevertheless, a significant number of households (approximately 800 000) contain two or more unemployed persons. Urban households are more likely to have no unemployed members, despite higher participation rates in urban areas.

The demographic section of the table is striking. Households where no one is unemployed are typically smaller and the members are significantly older. This has been explained (Klasen and Woolard, 1998) by the fact that the young unemployed generally remain with their parents or attach themselves to the households of other relatives. Once employment is found, they are able to form separate (and thus smaller) households. Not surprisingly, households with no unemployed persons are slightly better educated.

We see from the next section of the table that more than half of the unemployed are in households with two or more unemployed persons. The situation of these households is clearly particularly grim when one considers the average employment (or conversely unemployment) rates in households with two or more unemployed persons. While 47% of labour force participants in households with one unemployed member are formally employed or own-account workers, this figure falls to 27% in households with two unemployed members and to a dismal 17% in households with three or more unemployed.

The average household unemployment rate controls for household size, or more specifically, for the number of labour-market participants in households. It is therefore a tighter measure of the severity of unemployment at the household level. This row in the table shows that such average unemployment rates are higher than the more conventional unemployment rates that are discussed above. Using these average household rates, over half of the labour market participants are unemployed in all households with any unemployed members.

---

<sup>17</sup> The furore over the poverty impacts of a minimum wage in the United States is a good example. See Card and Krueger (1994) versus Neumark and Washer (1997).

Not surprisingly, income levels fall as the number of unemployed in the household increases. Incomes in households with no unemployed are almost twice those in households with one unemployed person, before taking account of the fact that households with unemployed members are significantly larger. If we compare the Theil-T contributions to inequality to the population shares (shown in the first line of the table) we see that households where no one is unemployed are the major contributors to inequality. Thus, most of the household-level inequality in South Africa is driven by income dynamics within households with no unemployed members. Thus labour market earnings rather than unemployment need to be highlighted when looking at labour market factors driving household income *inequality*. However, this does not imply that unemployment is unimportant. Indeed, one of the major reasons for this finding is that households with unemployed members are uniformly bunched in the low-income sections of the household income distribution. This is confirmed by the poverty decomposition analysis.

**Table 7: Expanded definition of unemployment**

Household Type	0	1	2	3+	Total	Column Shares
<b>A. General</b>						
ALL	71.8	19.1	5.9	3.2	8 801 992	100.0
African	64.0	23.8	7.7	4.4	5 950 904	67.6
Coloured	73.4	18.6	5.7	2.3	747 530	8.5
Asian	81.6	14.8	2.7	0.8	245 661	2.8
White	94.7	4.8	0.4	0.1	1 857 897	21.1
Rural	68.1	20.8	7.2	3.9	3 483 220	
Urban	74.2	18.0	5.0	2.8	5 318 772	
<b>B. Other demographics</b>						
Ave Age	31.8	26.3	27.1	27.8	30.4	
Ave Size	3.8	5.1	6.4	8.1	4.3	
Ave # of Children under 15	1.3	1.9	2.2	2.3	1.5	
Ave # of Adults	2.5	3.1	4.2	5.8	2.8	
Ave # of Labour Market Participants	1.3	1.9	2.7	4.3	1.6	
Ave adult years of education	7.0	6.0	5.6	5.6	6.7	
<b>C. Labour Market</b>						
% of total unemployment	0	43.4	28.0	28.5	100	
% of total self-employment	78.1	14.7	4.6	2.6	100	
% of total formal employment	78.2	15.6	4.0	2.3	100	
Ave HH Unemployment Rate	0	65.6	80.9	85.7	24.7	
Ave Unemployment rate	0	53.1	73.4	83.0	28.6	
Ave self-employment rate	14.5	6.5	4.4	2.7	10.3	
Ave formal employment rate	85.3	40.6	22.3	14.4	61.0	
<b>D. Poverty and Inequality</b>						
Ave hh income per annum	42 094	22 886	17 929	17 970	35 770	
(standard deviation)	(75493)	(48 283)	(21 675)	(18 559)	(67 662)	
Ave hh expenditure per annum	40 564	22 848	17 209	18 197	34 658	
(standard deviation)	(73 687)	(47 611)	(20 968)	(19812)	(66073)	
Theil-T ( <sup>18</sup> / <sub>18</sub> contributions to overall inequality.)	79.9	11.5	1.9	0.9	94.1	
Poverty shares:						
FGT(P <sub>1</sub> )	51.9	28.0	12.2	7.9	100.0	
FGT(P <sub>2</sub> )	48.1	28.7	13.9	9.3	100.0	
FGT(P <sub>3</sub> )	45.6	29.0	15.1	10.3	100.0	

<sup>18</sup> The figures sum to 94.1%. The remaining 5.9% is the "between group" inequality.

The incidence of poverty (measured by the Foster-Greer-Thorbecke  $P_0$  measure, more commonly known as the head-count index) clearly increases as the number of unemployed household members grows. While 72% of households have no unemployed members, they only make up 52% of the poor. Similarly, while only 3% of households have 3 or more unemployed members, they account for 8% of poor households. In addition to being more *likely* to be poor, poor households are also *more poor*. We see from the FGT  $P_1$  and  $P_2$  measures (which can be considered to measure the depth and severity of poverty respectively), that households with unemployed persons make up even higher proportions of poverty than when measured by the head-count index.

## 5. Conclusion

The oblique references to the differences generated by using two different data sets, or indeed even alternative inequality measures do not detract from a few simple, yet powerful observations made here about income inequality in South Africa. Firstly, income inequality between different races although smaller than the within-race contribution, is amongst the highest in the world – if not the highest. Secondly, the largest within-race contributor to inequality is amongst African households. Greater inequality exists amongst African households than any other race group. Thirdly, it is evident that the most important determinant of the Gini coefficient in South Africa is wage income, while self-employment income appears as a highly relevant inequality measure as well. Finally, the paper reinforces the fact that the labour market is central to our understanding of poverty in the society. Specifically, most household-level inequality is driven by income dynamics within households with no unemployed members because households with unemployed members tend to be crowded below the poverty line at the lower end of the household income distribution.

## REFERENCES

- Anand, S (1983) *Inequality and Poverty in Malaysia: Measurement and Decomposition*, (New York: Oxford University Press).
- Atkinson, A.B. (1977). *The Economics of Inequality*. Oxford: Clarendon Press.
- Bhorat, H., Leibbrandt, M. and Woolard, I. (1995). "Towards an Understanding of South Africa's Inequality". Paper presented at the African Economic Research Consortium Conference. Johannesburg, December
- Bhorat, H. 1995. "The South African Social Security Net: Past Present and Future", *Development Southern Africa*. 12(4).
- Bhorat, H and Hertz, T. 1995. "A Statistical Overview of the South African Labour Market". Input paper for the Presidential Labour Market Commission.
- Card, D and Krueger, A. 1994. "A Living Wage? The Effects of a Minimum Wage on the Distribution of Wages, the Distribution of Family Earnings, and Poverty", *Working Paper No. 333*, Industrial Relations Section Princeton University.
- Cowell, F. 1995. *Measuring Inequality*. (2nd Ed.). London: Prentice Hall.
- Central Statistical Services. 1995. The October Household Survey. Pretoria: Government Printer.
- Deaton, A. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore: Johns Hopkins Press.
- Deaton, A. 1995. "Large Transfers to the Elderly in South Africa". Unpublished Paper, Research Program in Development Studies, Princeton University.
- Fields, G.S. 1980. *Poverty, Inequality, and Development*. Cambridge: Cambridge University Press.
- Fiszbein, A and Psacharopoulos, G (1995) "Income Inequality Trends in Latin America in the 1980s". In Lustig, N (Ed.) *Coping with Austerity: Poverty and Inequality in Latin America*. (Washington, D.C.: The Brookings Institution).
- Hertz, T. 1995. "Jobs, Farms, Discrimination and Education in South Africa: Simulations and Regressions on Household Survey Data". Unpublished Paper. Department of Economics. University of Massachusetts at Amherst.
- Hindson, D., and Crankshaw, O. 1990. "New Jobs, New Skills, New Divisions: The Changing Structure of South Africa's Workforce". *South African Labour Bulletin* 15(1): 23-31.
- Klasen, S. and Woolard, I. 1998. "Unemployment, household formation, poverty and nutrition in South Africa." Unpublished paper. Port Elizabeth.
- Leibbrandt, M., Woolard, C. and Woolard, I.. 1996. "The contribution of income components to income inequality in South Africa: a decomposable Gini analysis." *Living Standards Measurement Survey Working Paper no. 125*. Washington: World Bank.
- Lerman, R.I. and S. Yitzhaki. 1994. "Effect of marginal changes in income sources on U.S. income inequality". *Public Finance Quarterly*, 22(4):403-16.
- May, J., Carter, M. and Posel, D. 1995. "The Composition and Persistence of Poverty in Rural South Africa: An Entitlements Approach". *Land and Agriculture Policy Centre Policy Paper No. 15*, Land and Agriculture Policy Centre.
- McGrath, M and Whiteford, A. 1994. "Inequality and the Size Distribution of Income in South Africa". *Occasional Paper No. 10*, Stellenbosch Economic Project.
- Moll, P. 1998 "Discrimination is Declining in South Africa, But Inequality is not". *Working Paper No. 5*, South African Network for Economic Analysis, Cape Town, 1998.
- Neumark, D and Wascher, W. 1997. "Do Minimum Wages Fight Poverty?" *Working Paper 6127*, NBER.



- Nygaard, F. and Sandstrom, A. 1981. *Measuring Income Inequality*. Stockholm: Almquist and Wicksell.
- SALDRU. 1994. *South Africans Rich and Poor: Baseline Household Statistics*. Project for Statistics on Living Standards and Development.
- Shorrocks, A. 1983. "The Impact of Income Components on the Distribution of Family Income". *Quarterly Journal of Economics*. 98: 311-26.
- Stark, O., Taylor, J. and S. Yitzhaki. 1986. "Remittances and Inequality". *Economic Journal*. 96: 722-40.
- Tarsitano, A. 1990. "The Bonferroni Index of Income Inequality" in Dagum, C. and Zenga, M. (Eds.) *Studies in Contemporary Economics: Income and Wealth Distribution, Inequality and Poverty*. Heidelberg: Springer-Verlag.
- Whiteford, A., Posel, D. and Kelatwang, T. 1995. *A Profile of Poverty, Inequality and Human Development*, Human Sciences Research Council.
- World Bank (1997) *World Development Report*, Oxford University Press:New York.

**APPENDIX 1**

**Table 2.1. Comparison of Distribution Measures<sup>19</sup>**

MEASURE	Between component	Within Component	Residual	Total
Theil-T	0.340 (48.2)	0.365 (51.8)		0.705
Theil-L	0.293 (40.8)	0.425 (59.2)		0.718
Atkinson $\epsilon = 0.5$	0.149 (50.0)	0.148 (49.7)	0.001 (0.3)	0.299
Atkinson $\epsilon = 1.5$	0.322 (46.0)	0.373 (53.2)	0.006 (0.8)	0.701
Atkinson $\epsilon = 2.5$	0.393 (41.0)	0.566 (59.0)	0.0001 (0.01)	0.959

NOTE: The figures in brackets show the percentage contribution to total inequality.

**Table 3.1. Within-Race Contribution to Overall Inequality**

Measure	African	Coloured	Asian	White
Theil-T	0.414 [0.159] (22.6)	0.276 [0.023] (3.3)	0.491 [0.028] (4.0)	0.326 [0.156] (22.1)
Theil-L	0.463 [0.345] (48.1)	0.325 [0.025] (3.5)	0.390 [0.011] (1.5)	0.295 [0.045] (6.3)

NOTES:

1. The first row of figures show the measure when considering only the particular race group.
2. The figures in square brackets show the absolute contribution to total inequality.
3. The figures in round brackets show the percentage of the total inequality.
4. Atkinson's index is generally but not additively decomposable, hence we cannot apportion the within contribution amongst the race groups.

<sup>19</sup> The PSLSD data-set includes 33 households (0.4%) with zero income. For technical reasons, the Theil-L measure cannot be calculated for a sample that includes households with zero income. Consequently, all three measures in Table 2 and 3 are calculated after dropping these 33 households. It is easily shown that this does not influence the results.

**Table 5.1: Decomposition of Total National Income by Income Sources (PSLSD)**

Income source	Proportion of households receiving income source ( $P_k$ )	Mean income from source	Share in total income ( $S_k$ )	Gini for income source for households receiving such income ( $G_A$ )	Gini for income source for all households ( $G_k$ )	Gini correlation with total income rankings ( $R_k$ )	Contribution to Gini coefficient of total income ( $S_k G_k R_k$ )	Percentage share in overall Gini	Effect on overall Gini of a 1% change in income component
Remittances	0.27	R68.07	0.03	0.52	0.88	-0.08	0.00	-0.40	-0.021
Wage income	0.66	R1427.94	0.69	0.53	0.69	0.92	0.44	73.50	0.021
Capital income	0.67	R285.55	0.13	0.82	0.87	0.81	0.09	15.20	0.015
State transfers	0.23	R97.38	0.05	0.26	0.83	0.00	0.00	0.00	-0.029
Agriculture	0.18	R79.68	0.04	0.92	0.99	0.79	0.03	4.70	0.008
Self-employment	0.11	R123.47	0.06	0.75	0.97	0.97	0.04	7.00	0.006
Total		R2082.04	1.00				0.60	100.00	

Notes:

1.  $G_A$  is the Gini for the income source when we only consider households with positive income from that source.
2.  $G_k$  is the Gini of the income source when we consider all households. Lerman and Yitzhaki (1994) show that  $G_k = P_k * G_A + (1 - P_k)$

**Table 6.1.: Decomposition of Total National Income by Income Sources, Below and Above the Poverty Line (PSLSD)**

**a. Below the poverty line**

Income source	Proportion of households receiving income source ( $P_k$ )	Mean income from source	Share in total income ( $S_k$ )	Gini for source for households receiving such income ( $G_A$ )	Gini for income source for all households ( $G_k$ )	Gini correlation with total income rankings ( $R_k$ )	Contribution to Gini coefficient of total income ( $S_k G_k R_k$ )	Percentage share in overall Gini	Effect on overall Gini of a 1% change in income source
Remittances	0.42	R80.62	0.18	0.44	0.75	0.19	0.03	6.71	-0.044
Wage income	0.35	R171.56	0.38	0.36	0.76	0.69	0.20	51.57	0.051
Capital income	0.72	R32.53	0.07	0.67	0.75	0.42	0.02	5.84	-0.005
State transfers	0.30	R119.78	0.27	0.21	0.75	0.55	0.11	28.42	0.006
Agriculture	0.31	R18.19	0.04	0.61	0.87	0.34	0.01	3.08	-0.004
Self-employment	0.11	R23.97	0.05	0.50	0.94	0.34	0.02	4.39	-0.004
Total		446.65	1.00				0.39	100.00	

**b. Above the poverty line**

Income source	Proportion of households receiving income source ( $P_k$ )	Mean income from source	Share in total income ( $S_k$ )	Gini for source for households receiving such income ( $G_A$ )	Gini coefficient for income source ( $G_k$ )	Gini correlation with total income rankings ( $R_k$ )	Contribution to Gini coefficient of total income ( $S_k G_k R_k$ )	Percentage share in overall Gini	Effect on overall Gini of a 1% change in income source
Remittances	0.16	R57.42	0.02	0.52	0.92	-0.11	0.00	-0.41	-0.012
Wage income	0.83	R2049.90	0.72	0.48	0.57	0.88	0.36	70.70	-0.006
Capital income	0.63	R379.86	0.13	0.75	0.84	0.75	0.08	16.61	0.017
State transfers	0.19	R88.85	0.03	0.29	0.87	-0.11	0.00	-0.60	-0.019
Agriculture	0.12	R101.45	0.04	0.94	0.99	0.86	0.03	5.92	0.012
Self-employment	0.11	R173.22	0.06	0.72	0.97	0.68	0.04	7.82	0.009
Total		R2850.71	1.00				0.51	100.00	

## APPENDIX 2

### A. "Decomposing" the Gini coefficient

Assume that within the chosen sample there are  $n$  households deriving income from  $K$  different sources (i.e.  $K$  different income components). Let  $y_i$  denote the **total income of household  $i$** , where  $i=1, \dots, n$  and  $y_{ik}$  the **income of household  $i$  from source  $k$**  where  $k=1, \dots, K$  (thus  $y_i = \sum_{k=1}^K y_{ik}$ ). Also, let the **distribution of total household income** be represented by  $\mathbf{Y}=(y_1, \dots, y_n)$  and the **distribution of income component  $k$**  be represented by  $\mathbf{Y}_k=(y_{1k}, \dots, y_{nk})$ .

Using this notation, the **Gini coefficient** ( $G$ ) for the distribution of total income within the group can be defined as:

$$G = \frac{(2cov[\mathbf{Y}, F(\mathbf{Y})])}{\mathbf{m}} \quad (1)$$

where  $\mu$  denotes the mean household income of the sample and  $F(\mathbf{Y})$  the "cumulative distribution" of total household income in the sample (i.e.  $F(\mathbf{Y})=(f(y_1), \dots, f(y_n))$  where  $f(y_i)$  is equal to the rank of  $y_i$  divided by the number of observations ( $n$ )).

Equation (1) can be rewritten and expanded into an expression for the Gini coefficient that captures the "contribution to inequality" of each of the  $K$  components of income:

$$\begin{aligned} G &= \frac{2}{\mathbf{m}} \sum_{i=1}^n (y_i - E(y_i))(f(y_i) - E(f(y_i))) \\ \Rightarrow G &= \frac{2}{\mathbf{m}} \sum_{i=1}^n \sum_{k=1}^K (y_{ik} - E(y_{ik}))(f(y_i) - E(f(y_i))) \\ \Rightarrow G &= \frac{2}{\mathbf{m}} \sum_{k=1}^K cov[\mathbf{Y}_k, F(\mathbf{Y})] \\ \Rightarrow G &= \sum_{k=1}^K \left( \frac{cov[\mathbf{Y}_k, F(\mathbf{Y})]}{cov[\mathbf{Y}_k, F(\mathbf{Y}_k)]} \right) \left( \frac{2}{\mathbf{m}_k} cov[\mathbf{Y}_k, F(\mathbf{Y}_k)] \right) \left( \frac{\mathbf{m}_k}{\mathbf{m}} \right) \end{aligned}$$

where  $\mu_k$  is the sample mean of income from source  $k$  and  $F(\mathbf{Y}_k)$  is the cumulative rank distribution of income from source  $k$  (i.e.  $F(\mathbf{Y}_k)=(f(y_{1k}), \dots, f(y_{nk}))$  where  $f(y_{ik})$  is equal to the rank of  $y_{ik}$  divided by the number of observations ( $n$ )).

Thus, the Gini coefficient can be written as:

$$G = \sum_{k=1}^K R_k G_k S_k \quad (2)$$

## B. The effect of a change in income on the Gini

Assume that there is an exogenous increase in income from source  $j$  by some factor  $\sigma_j$  (i.e.  $y_{ij}(\sigma_j) = (1 + \sigma_j)y_{ij}$  for  $i=1, \dots, n$ ). Thus the distribution of income from source  $j$  becomes  $\mathbf{Y} = ((1 + \sigma_j)y_{1j}, \dots, (1 + \sigma_j)y_{nj})$ . Let  $G$  be the Gini coefficient before the change in income and  $G(\sigma_j)$  the Gini coefficient after the change in income. Equation (2) gives the expression for  $G$ . However, in order to derive an expression for  $G(\sigma_j)$  after a change in income from source  $j$  by factor  $(1 + \sigma_j)$  it is necessary to look at how the change affects each of  $G_k$ ,  $R_k$  and  $S_k$  for  $k=1, \dots, K$ :

(a) Since we are dealing with a  $(1 + \sigma_j)$  times increase for in  $y_{ij}$ ,  $G_j$  does not change. Thus,  $G_k$  for all  $k=1, \dots, j, \dots, K$ .

(b) Assuming that the change in income from source  $j$  is small enough to leave the ranking of both total income and source  $j$  income unchanged,  $R_k$ , as a function of ranks of income, will remain unchanged.

(c) Since  $S_k$  measures income component  $j$ 's share in total income,  $S_k$  for  $k=1, \dots, K$  will obviously change if income from source  $j$  changes. Let us call each income component's new share in total income after the change in income component  $j$   $S_k(\sigma_j)$ .

Thus, we can write the Gini coefficient after the change in income component  $j$  as:

$$G(\mathbf{s}_j) = \sum_{k=1}^K S_k(\mathbf{s}_j) R_k G_k$$

By definition, for  $k \neq j$ :

$$S_k(\mathbf{s}_j) = \frac{m_k}{\sum_{k=1}^K m_k + s_j m_j}$$

while for income component  $j$ :

$$S_j(\mathbf{s}_j) = \frac{(1 + s_j) m_j}{\sum_{k=1}^K m_k + s_j m_j}$$

thus, the change in the Gini coefficient ( $\Delta G$ ) stemming from the exogenous change in income from source  $j$  can be written as:

$$\Delta G = G(\mathbf{s}_j) - G = \sum_{k=1}^K [S_k(\mathbf{s}_j) - S_k] R_k G_k \quad (4)$$

For  $k \neq j$ , the bracketed term in equation (4) can be rewritten as:

$$S_k(\mathbf{s}_j) - S_k = \frac{m_k}{\sum_{k=1}^K m_k + s_j m_j} - \frac{m_k}{\sum_{k=1}^K m_k} = \left( \frac{-s_j m_k m_j}{\left[ \sum_{k=1}^K m_k \right]^2} \right) / \left( \frac{\sum_{k=1}^K m_k + s_j m_j}{\sum_{k=1}^K m_k} \right)$$

which simplifies to:

$$S_k(\mathbf{s}_j) - S_k = \frac{-s_j S_k S_j}{1 + s_j S_j} \quad (5)$$

Similarly, for  $k=j$  it can be shown that:

$$S_j(\mathbf{s}_j) - S_j = \frac{\mathbf{s}_j S_j - \mathbf{s}_j S_j^2}{1 + \mathbf{s}_j S_j} \quad (6)$$

Substituting equations (5) and (6) into (4), a more detailed expression for  $\Delta G$  is obtained:

$$\begin{aligned} \Delta G &= \sum_{k \neq j} \frac{-\mathbf{s}_j S_k S_j}{1 + \mathbf{s}_j S_j} R_k G_k + \frac{\mathbf{s}_j S_j - \mathbf{s}_j S_j^2}{1 + \mathbf{s}_j S_j} R_j G_j \\ \Rightarrow \Delta G &= \sum_{k=1}^K \frac{-\mathbf{s}_j S_k S_j}{1 + \mathbf{s}_j S_j} R_k G_k + \frac{\mathbf{s}_j S_j}{1 + \mathbf{s}_j S_j} R_j G_j \end{aligned} \quad (7)$$

In order to find the derivative of the Gini coefficient with respect to  $\sigma_j$ , it is simply necessary to take the limit of equation (7) divided by  $\sigma_j$  as  $\sigma_j$  tends to zero:

$$\lim_{\mathbf{s}_j \rightarrow 0} \frac{\Delta G}{\mathbf{s}_j} = -S_j \lim_{\mathbf{s}_j \rightarrow 0} \sum_{k=1}^K \frac{S_k}{1 + \mathbf{s}_j S_j} R_k G_k + \lim_{\mathbf{s}_j \rightarrow 0} \frac{S_j}{1 + \mathbf{s}_j S_j} R_j G_j$$

Hence, it can be shown that the derivative of the Gini coefficient with respect to a change in income source  $j$  is:

$$\frac{\partial G}{\partial \mathbf{s}_j} = -S_j \sum_{k=1}^K S_k R_k G_k + S_j R_j G_j = S_j (R_j G_j - G) \quad (8)$$



### C. Theil's Entropy index (T)

This is given by

$$T = \frac{1}{N} \sum_{i=1}^N (y_i / \mathbf{m}) \log(y_i / \mathbf{m})$$

1

This may be decomposed into between- and within-group inequality

$$T = T_B + T_W$$

2

where

$$T_B = \sum_j q_j \log(m_j / \mathbf{m})$$

3

and

$$T_W = \sum_j q_j T_j$$

### D. Theil's second measure (L)

This is given by

$$L = - \frac{1}{N} \sum_{i=1}^N \log(y_i / \mathbf{m})$$

4

This may also be decomposed into between- and within-group inequality

$$L = L_B + L_W$$

5

where

$$L_B = - \sum_j p_j \log(m_j / \mathbf{m})$$

6

and

$$L_W = \sum_j p_j L_j$$

## E. "Decomposing" Atkinson's index

Atkinson's measure starts from the additive social welfare function:

$$W = \frac{1}{N} \sum_{i=1}^N \frac{y_i^{1-\epsilon}}{1-\epsilon}$$

7

or, when  $\epsilon = 1$ ,

$$\ln W = \frac{1}{N} \sum_{i=1}^N \ln y_i$$

8

Atkinson explicitly introduces distributional objectives through the parameter  $\epsilon \geq 0$  which represents the weight attached to inequality in the distribution. By specifying different values of  $\epsilon$  one can vary the importance society attaches to mean living standards versus equality. If society is indifferent about the distribution, we will set  $\epsilon$  equal to zero. By increasing  $\epsilon$  we give more weight to inequality at the lower end of the distribution. At  $\epsilon$  equal to infinity society is concerned only with the poorest household.

In other words, the ratio of marginal social utilities of two individuals is then given by:

$$\frac{\partial W / \partial y_i}{\partial W / \partial y_j} = (y_j / y_i)^\epsilon$$

Hence, if  $\epsilon = 0$ , marginal utility is the same for everyone. As  $\epsilon$  tends to infinity, the utility of the poorest dominates.

The inequality index (Atkinson's I) associated with this is given by:

$$I = 1 - \left( \sum_{i=1}^N \frac{1}{N} (y_i / \mathbf{m})^{1-\epsilon} \right)^{1/(1-\epsilon)}$$

9

10

Atkinson's index may be decomposed into between- and within-group inequality such that

$$I = I_B + I_W + \text{Residual}$$

11

where

$$I_B = 1 - \left( \sum_j p_j (\mathbf{m}_j / \mathbf{m})^{1-\epsilon} \right)^{1/(1-\epsilon)}$$

12

and

$$I_w = \left( \sum_j p_j (\mathbf{m}_j / \mathbf{m})^{1-\epsilon} A_j^{(1-\epsilon)} \right)^{1/(1-\epsilon)}$$

## F. Choice of decomposition technique

Different techniques are better suited to particular problems than others. The Theil, for example, gives a clear picture of the importance of each explanatory variable in determining inequality but lacks intuitive appeal. The Gini, while well-known and easily conceptualised, is not additively decomposable and therefore of limited use if one wishes to avoid being overly technical.

Three issues need to be considered in determining which technique is most useful. One must consider the properties (or basic nature) of the measure itself, the value of the information derived from the decomposition, and the suitability of the measure to the problem.

In terms of the "basic nature" of the measures, the Gini coefficient is popular since it is easily conceptualised in terms of the Lorenz curve. The Theil entropy indices, on the other hand, have no clear intuitive justification and are somewhat arbitrary formulae from the point of view that they say little in terms of welfare implications (Mishra *et al*, 1992: 228). The Atkinson index has easily interpretable welfare implications, but is, of course, derived from a particular welfare framework which is open to criticism.

With regard to the usefulness of the inequality measure, the Gini, Theil and Atkinson measures all satisfy Sen's axioms of *scale independence*, *symmetry* and *weak principle of transfers*:

*Axiom of income scale independence.* This means that measured inequality should not depend on the magnitude of total income, i.e. if everyone's income changes by the same proportion, the measure of inequality should not change.

*Axiom of symmetry (or principle of population).* This requires that inequality should not depend on the number of income receivers. If we measure inequality in a particular economy with  $n$  people in it and then merge the economy with another identical one, then the combined economy with  $2n$  people must have the same measure of inequality.

*Weak principle of transfers.* Consider two individuals, one with income  $y$  and the other with income  $y + \delta$  where  $\delta$  is positive. Then transfer a positive amount of income  $\Delta x$  from the richer to the poorer man, where  $\Delta x$  is less than  $\frac{1}{2}\delta$ . Inequality should then definitely decrease.

Another consideration is the sensitivity of the different measures to income changes at various points in the distribution. If our value judgement is that we should give greatest weight to the economic position of the poor, then we require a measure which is more sensitive to inequality at the lower end of the income distribution. The Theil indices are most sensitive to inequality amongst the very rich, while the Gini is most sensitive to inequality in the middle of the income distribution. By altering the value of  $\epsilon$  the Atkinson measure can be easily manipulated to concentrate on different parts of the distribution.

### APPENDIX 3

**Table 7.1 Narrow definition of unemployment**

Household Type	0	1	2	3+	Total	Column Shares
<b>A. General</b>						
ALL	85.5	10.4	2.7	1.4	8 801 992	100.0
African	82.8	12.0	3.4	1.9	5 950 904	67.6
Coloured	81.4	14.0	3.3	1.3	747 530	8.5
Asian	86.0	11.3	2.3	0.4	245 661	2.8
White	96.0	3.6	0.4	0.0	1 857 897	21.1
Rural	87.0	9.0	2.7	1.3	3 483 220	39.6
Urban	84.6	11.3	2.7	1.5	5 318 772	60.4
<b>B. Other demographics</b>						
Ave Age	30.9	26.9	27.2	27.8	30.4	
Ave Size	4.1	5.1	6.6	8.1	4.3	
Ave # of Children	1.4	1.8	2.1	2.2	1.5	
Ave # of Adults	2.6	3.3	4.4	5.8	2.8	
Ave # of Labour Market Participants	1.2	1.9	2.8	4.3	1.4	
Ave adult years of education	6.8	6.4	5.9	6.1	6.7	
<b>C. Labour Market</b>						
% of total unemployment	0	48.9	26.2	24.9	100	
% of total self-employment	88.5	8.1	2.2	1.2	100	
% of total formal employment	87.7	9.1	2.1	1.1	100	
Ave HH Unemployment Rate	0	63.9	78.9	84.4	13.1	
Ave Unemployment rate	0	51.7	71.0	81.4	16.4	
Ave self-employment rate	14.6	6.3	4.3	2.9	12.1	
Ave formal employment rate	85.2	41.9	24.7	15.6	71.3	
<b>D. Poverty and Inequality</b>						
Ave Income	37 979	24 752	20 206	19 592	35 770	
(standard deviation)	(71 717)	(38 952)	(24 867)	(16 230)	(67 662)	
Ave Expenditure	36664	24837	19766	20049	34 658	
(standard deviation)	(70172)	(36324)	(24432)	(16489)	(66 073)	
Theil-T (% contributions to overall inequality.) <sup>20</sup>	91.3	5.6	0.9	0.3	98.2	
Poverty shares:						
FGT(P <sub>0</sub> )	78.5	13.5	4.9	3.1	100.0	
FGT(P <sub>1</sub> )	77.8	13.5	5.4	3.3	100.0	
FGT(P <sub>2</sub> )	77.3	13.3	5.8	3.6	100.0	

<sup>20</sup> The figures sum to 98.2%. The remaining 1.8% is the "between group" inequality.