University of Cape Town

CENTRE FOR
SOCIAL SCIENCE RESEARCH

AHOY THE GOOD HOPE?
SOME BEARINGS AND SIGNALS
IN SELDOM-NAVIGATED WATERS -
ON INEQUALITY IN SOUTH
AFRICA'S COLOURED AND
AFRICAN POPULATION

Sten Dieden

CENTRE FOR SOCIAL SCIENCE RESEARCH

Social Surveys Unit

AHOY THE GOOD HOPE?

SOME BEARINGS AND SIGNALS IN SELDOM-NAVIGATED WATERS - ON INEQUALITY IN SOUTH AFRICA'S COLOURED AND AFRICAN POPULATION

Sten Dieden

CSSR Working Paper No. 107

February 2005

Sten Dieden received his Ph.D. in the Department of Economics, University of Gothenburg, Sweden, in early 2005. This paper was largely written whilst he was a Visiting Research Fellow in the Centre for Social Science Research (CSSR).

Direct any comments to sten.dieden@handels.gu.se.

Ahoy the Good Hope? Some Bearings and Signals in Seldom-Navigated Waters - on Inequality in South Africa's Coloured and African Population

Abstract

Previous studies have decomposed South African income inequality into inequality between and within the population groups defined by the apartheid regime's racial classification system. While a substantial fraction of total inequality can be attributed to differences in mean income levels between those population groups, the level of inequality within the racial groups has been found to contribute more to total inequality. Yet few investigations have attempted to elucidate inequality within these population groups. This study therefore explores the extent to which inequality in a joint sample of African and coloured individuals can be attributed specific labour-market related characteristics of their households or household heads. The analyses apply the Theil-L measure of inequality to the distribution of a consumption bundle in a household survey data set from 1995. The education level of household heads is the strongest single explanatory factor, followed by households' main income sources. The race, age categories, or gender of household heads do not account for large fractions of inequality in this sample.

1 Introduction

The apartheid regime in South Africa recognised and applied different extents of racial discrimination to four different population groups.[1] Due to *inter alia* the legacies from those discrimination policies, South Africa has the seventh highest

---

[1] The four main "racial classifications" recognised by the apartheid regime were: African (black), coloured, Asian/ Indian and white. The discrimination by race ran through all aspects of life and had tremendous effects on everyone's living standards. For these reasons, official statistics in South Africa still apply "racial" categories. The same practice is followed here and the categories will interchangeably be referred to as "population" or "racial", "groups" or "categories".

level of inequality in the world (World Bank (2004)). Previous studies have applied those racial groups in additive decomposition analyses and found that substantial fractions of the inequality are attributable to differences in average income levels between those groups (Whiteford and McGrath (1998, 2000), May (2000)), Leibbrandt, Woolard and Bhorat (2000), Leibbrandt and Woolard (2001)).[2] The fraction of total inequality originating from *within* the racial groups is however always found to be larger. Yet few investigations have analysed inequality within the racial groups.[3] This study of household survey data therefore explores the extent to which inequality within a joint sample of African and coloured individuals can be attributed to specific characteristics of their households or household heads.

Several previous studies on South Africa emphasise the importance of households' access to employment and *wage income* in explaining income inequality and in evading poverty (Carter and May (1999), Leibbrandt, Woolard, and Woolard (2000), Leibbrandt, Woolard, and Bhorat (2000), van der Berg (2000), Jenkins and Thomas (2000)). The analysis undertaken here utilises a household's main income source as an explanatory factor that reflects a household's labour market attachment. The other explanatory factors are geographical location in the rural-urban dimension and in provinces (Leibbrandt and Woolard (1999)), and the education level and racial group of household heads, which are all commonly applied determinants in the literature on individuals' wage earnings (Moll (1996), Kingdon and Knight (1999), Mwabu and Schultz (2000)).

The decomposition methodology utilised in this study is commonly applied and has been developed by Bourguignon (1979), Shorrocks (1980, 1984), Blackorby, Donaldson and Auersperg (1981), and Cowell and Jenkins (1995). The empirical analysis uses the Theil-L index of inequality, which allows total inequality to be unambiguously split into the contribution due to differences *between* subgroups and the contribution due to inequality *within* subgroups

---

[2] In the literature on additively decomposable income inequality, the inequality attributable to differences between mean incomes of population subgroups is considered "explained". By definition, the total level of inequality is reached by the adding to the former, the summed inequality in the income distribution around the means *within* each subgroup (Bourguignon (1979), Shorrocks (1980, 1984), Blackorby, Donaldson and Auersperg (1981), Cowell and Jenkins (1995)). Other measures of the centre of income distribution than the arithmetic mean can also be used.

[3] Two relevant exceptions are Leibbrandt, Woolard and Woolard (2000) and Leibbrandt and Woolard (2001). The first work decomposes the Gini index in the African population group and in the second a multivariate technique developed by Fields (2003) is applied to decompose the variance in households' log per capita income in the same population group.

(Shorrocks (1984)). The units of observations are individuals, to which their households' per adult-equivalent expenditure levels have been attached.

The paper proceeds from here to discuss the methodology and justify the choice of index in Section 2. In Section 3 the data and choice of welfare measure are introduced. Thereafter Section 4 justifies the scope of the study and describes the sample delimitation process. The relationship between each explanatory variable and welfare levels are illustrated with descriptive statistics in Section 5, after which follows the empirical results in Section 6. Conclusions are drawn in the final Section 7.

2 Methodology

Given some measure of welfare, a decomposition of South Africa's welfare inequality presupposes a picture that measures distribution as a readily expressible function of the inequality *between* certain groups and (some aggregation of) the inequality *within* the same groups. A wide range of measures of inequality exist (for example, Champernowne and Cowell (1998)), but standard methodology draws on results from Bourguignon (1979), Cowell (1980) and Shorrocks (1980, 1984)[4], from which it can be concluded that the most suitable measures are those ordinally equivalent to the measures in the general entropy class.

Denoting a measure of the latter class G(·), the ordinally equivalent group of measures can be defined as:

$$(1) \qquad I(x) = C[G_\alpha(x), n(x), \mu(x)]$$

where $n(\cdot)$ is the finite dimension of the vector x which represents the welfare distribution of which $\mu(\cdot)$ is the arithmetic mean. The function $C[\cdot]$ is a cardinalisation of $G(\cdot)$. The latter increases monotonically in the first argument and is defined by

$$(2) \qquad G_\alpha(x) := \frac{1}{\alpha^2 - \alpha} \left\{ \frac{1}{n(x)} \sum_{i=1}^{n(x)} \left[\frac{x}{\mu(x)} \right]^\alpha - 1 \right\}$$

where the parameter α can be assigned any real value. For high positive levels of α, the index is sensitive to welfare changes in the upper level of the distribution

---

[4] See for example, Tsaklogou (1993) for an early application and Gustafsson and Li (2002) for a more recent application.

and for an index which is sensitive to redistributions at the lower level, α should take on a negative value. Shorrocks (1984) shows that among the additively decomposable inequality measures, the index derived from α = 0, Theil-L index, is the most satisfactory and allows total inequality to be unambiguously split into the contribution due to differences between subgroups and the contribution due to inequality within each subgroup.

Given a set of S groups, the decomposition of $G(x)$ is undertaken using the group-means in the following general manner:

(3) $G(x) = G_W + G_B$

where

(4) $G_W := \sum_{s=1}^{S} u_s^\alpha w_s^{1-\alpha} G_s(x_s)$

and

(5) $G_B := G(\mu_s)$

where , μ_s is the mean welfare level in group s, with u_s and w_s as the income and population shares of group s respectively, and μ is the vector of S group means. [5]

The units of observation, subject to inequality, in this case are human beings. From that perspective, the $G_0(x)$ index has the appealing feature that inequality within each group is weighted by the population fraction in that group. Hence, the index is deemed the most suitable for the purposes of this analysis.

When α = 0, l'Hôpital's rule yields

(5') $G_B = \frac{1}{w_N} \sum_{s=1}^{S} w_s \ln \frac{\mu}{\mu_s}$

and

(4') $G_W = \frac{1}{w_N} \sum_{s=1}^{S} \sum_{i=1}^{n(s)} w_i \ln \frac{\mu_s}{y_i^s}$

---

[5] In the inequality literature, $G_{\alpha=0}()$ and $G_{\alpha=1}()$ are also commonly and respectively referred to as Theil-L and Theil-T measures of inequality. It follows from equation (4) that the two measures differ in how total within-group inequality is computed. When the $G_0(x)$ is applied, each group's inequality is weighted by the population fraction in that group, whereas when the $G_1(x)$ is applied, the weights constitute each group's share of total expenditures.

where w_i is the sampling weight attached to observation i and

$$w_s = \sum_{i=1}^{n(s)} w_i \quad \text{and } w_N = \sum_{s=1}^{S} w_s$$

Cowell and Jenkins (1995) derive a measure of explained inequality, R_B, as

(6) $\qquad R_B := \dfrac{G_B}{G} = 1 - \dfrac{G_W}{G}$

The same authors show that this concept of explained inequality extends to the analysis of more than one determinant of inequality at the time, since through the

specification of a refined subpartition, $\Pi_{a \text{ and } b}$, of an original partition Π_a or Π_b, it must be true that

$$R\left(\Pi_{a \text{ and } b}\right) \geq R\left(\Pi_a\right) \qquad \text{and} \qquad R\left(\Pi_{a \text{ and } b}\right) \geq R\left(\Pi_b\right)$$

Thus, a succession of subpartitions yields a consistent representation of the importance of the characteristics that define the consecutive subpartitions.

3 Data and Choice of Welfare Measure

Data

In 1995 Statistics South Africa undertook its annual October Household Survey with questionnaire-based interviews on a wide range of living standards issues using a stratified and clustered sample of 30 000 households, representing all households in the country and containing nearly 131 000 inhabitants (the "OHS" sample). Two months later, 28 585 of the same households were revisited in a more detailed Income and Expenditure Survey (the "IES" sample and henceforth the surveys or samples will jointly be referred to as the "OHS/IES 95" data.)[6]

---

[6] At the time of the writing, a similar, nationwide South African data set from the year 2000 had been released. However, since the reliability of the 2000 data was also still under evaluation by South Africa's Statistics Council and since the other analyses in this paper are undertaken using the 1995 data, the latter was deemed preferable to the current analysis. One reason for the disputed comparability of the two data sets, is that nominal incomes were lower in the four lowest per adult-equivalent income quintiles in the year 2000 than 1995, whereas

In the surveys a household is defined by "a person or a group of people dependent on a common pool of income who normally occupy a dwelling unit or a portion thereof and who provide themselves with food or the necessary supplies or arranged for such provision." A member resides four nights a week in the household. The sample for the surveys was stratified by province, urban and non-urban area and population group. Altogether, 3 000 enumerator areas (EAs) were drawn as primary sampling units, within each of which ten households were visited. The data concerning households were weighted by the estimated number of households in each stratum. (Statistics South Africa (1997)). The analyses in this study of a subsample of the full OHS/IES95 are conducted with the supplied household weights renormalised to sum to unity, as suggested by Deaton (1997) when faced with missing survey data. It should be noted that, given nine provinces in South Africa, two types of areas and four population groups, the full data sets may be considered representative of 72 groups, 36 of which are African or coloured. The inference to a population level of results based on partitions into larger numbers of groups is thus limited.

Welfare Measure

Welfare is a complex phenomenon that involves multiple dimensions of deprivation and lack of goods and services is only one of those dimensions (Sen (1985, 1987)). Even so, there is a good deal of consensus on the value of using a consumption aggregate as a welfare metric of living standards (Deaton and Zaidi (2002)). This study follows that tradition and uses a consumption-aggregate based on household expenditure data as a summary measure.[7] The aggregate is constructed according to the guidelines put forth by Deaton and Zaidi (2002) and contains the summed subtotals of household expenditures in 1995 Rand on the following categories as defined by Statistics South Africa (1997): food, beverages, tobacco, personal care, fuel and power, household operation, housing costs, remuneration for domestic workers, footwear, clothing, medical care, transport, telecommunication, education, and reading material.

Individuals are units of analysis rather than households since it is difficult to conceive of households experiencing welfare (Deaton (1997)). Attached to each

---

the total expenditure data do not display such characteristics (see Tables A1 and A2). Some of the indications of the robustness of the core results of this analysis are based on the 2000 data and are provided in Appendix 1.

[7] A common justification for the use of consumption is that current consumption is a function of permanent income (Slesnick (1993), but as pointed out by Deaton (1997), the empirical support for the permanent income hypothesis is at best mixed. See Slesnick(1993, 1998) and Chaudhuri and Ravallion (1994) for discussions of the choice between income or consumption as welfare metrics.

sampled individual is its household's total annual expenditure on the categories in the consumption aggregate, divided by the household size as calculated in terms of adult-equivalence, yielding the households "per-adult-equivalent expenditure".[8]

The assumption of equal division among household members (whether in the format of adult equivalents or not) is as shown by for example Haddad and Kanbur (1990) questionable. It is also recognised that that welfare measures in some respect are too limited and, as developed in Sen's (1985, 1987) work, other indicators such as life expectancy, infant mortality and literacy would be better (Ravallion (1996). However, both these latter issues are beyond the scope of this investigation.[9]

4 Target Group and Sample Delimitation

Target Group

This study aims to find explanations for inequality among individuals in the African and coloured population. The two population groups are defined as the individuals that live in households where the head belongs to either the African or coloured race group. The objective of the study is justified by the figures in Table 1, which shows the distributions of all individuals sampled by the OHS/IES95 by per-adult-equivalent expenditure quintiles and population group.

As can be seen, African and coloured individuals constitute more than 95 percent of the individuals in the three lowest brackets, while the corresponding shares of the total population are approximately 85 percent. The summed expenditures of the African and the coloured subsamples however, amount to just over 50 percent of the total. At the same time, the white population fraction is miniscule in the three first quintiles, in the neighbourhood of seven percent in the fourth and only becomes substantial in the highest quintile. While the Indian population fraction is small, the group is over-represented in the highest

---

[8] Leibbrandt and Woolard (2001) investigate several adult equivalence scales for South Africa using the OHS/IES95 data and impacts appear to be miniscule. The authors proceed using the scaled applied by May, Carter and Posel (1995) i.e.: $E=(A+0.5K)^{0.9}$, where E is number of adult equivalents, A is number of adults and K is the number of children 15 years old or younger. This study applies the same procedure. Information about the quintiles for the full samples based on this concept is displayed in Table A3 in Appendix 1.

[9] See Klasen (1997, 2000) for two multi-dimensional approaches to deprivation in South Africa.

brackets. Similarly, the Indian/Asian fraction of total expenditure is twice as large as its population share, whereas the expenditure share of the white sample is more than three times as large as its population share.

Table 1 Composition of per adult-equivalent household expenditure quintiles in the full OHS/IES95 sample, by population group.

Quintile	African	coloured	Asian/ Indian	white	All	Quintile cut-off Points (1995 Rand)	
						Lower	Upper
1	96.7	3.2	0.0	0.1	100.0	62	1 496
2	92.7	7.2	0.1	0.1	100.0	1 496	2 468
3	86.9	11.4	1.1	0.6	100.0	2 468	4 139
4	73.1	15.8	4.5	6.7	100.0	4 139	9 313
5	29.4	8.0	7.7	54.8	100.0	9 313	760 069
All	75.8	9.1	2.7	12.5	100.0	62	760 069
Number (millions)	30, 0	3, 6	1,1	4,9	39, 7		
Total expenditure share (%)	43.6	8.4	5.3	42.7	100.0		

Source: OHS/IES95, own computations, weighted figures. Absolute population numbers are weighted sample estimates in millions of individuals. n =125 112.

A closer investigation of inequality within the African and coloured subsample is warranted for at least three reasons; Firstly, the subsample represents the *overwhelming majority* of South Africans and virtually *all* individuals at the lowest end of the expenditure distribution. Secondly, the members of these groups face similar historical legacies. The identification of the factors that are associated with inequality within that sample may thus provide some insight into the nature of inequality at the lower segment of the expenditure distribution, where policy measures to reverse past injustices are most needed. Finally, using the same data, Leibbrandt, Woolard and Bhorat (2001) report the contributions to the total level of inequality in households' per-adult equivalent income from inequality within those groups to be 56.8 percent as measured by the same inequality measure. Hence, explained fractions of within-group inequality in the subsamples will add considerably to the total explained inequality in South Africa.

Sample Delimitation

For the purposes of this study and for the above reasons, only individuals that live in households where the head belongs to the African or coloured population groups were selected. Furthermore, for reasons which are motivated in the next

section, the origins of households' main sources of income by *inter alia* broad economic sectors are to be used as explanatory characteristics for inequality. Since the quality of the information on individuals' labour market characteristics were greater in the OHS module than in the IES, it was deemed desirable to extract that information from the former base.

Households in the two data sets are easily matched, since their unique codes were identical in both data sets. However, the within-household codes for individuals differed across the surveys. Persons that were captured with any amount of income in the IES module therefore had to be matched to the OHS data according to household, age, gender and race. By this procedure 97.5 percent of the utilised sample were matched. Another 773 earners was identified by allowing either age to mismatch by two years, with race and gender matching perfectly, or race to have been miscaptured, with age and gender matching perfectly. This procedure identified 30 906 earners in both data sets. The sample delimitation process is illustrated in Table 2. All results in the remainder of the analysis are weighted figures, based on the 86.5 percent (92 717) of individuals that resided in households that met the first criterion and where all income earners covered by the IES module were identified in both data sets.

Table 2) Sample delimitation process

Sample	Number of individuals	Share of total revisited sample	Share of revisited African and coloured sample
Total OHS/IES sample	125 112	100.0	
African and coloured OHS/IES sample	107 229	84.9	100.0
Final sample	92 717	74.1	86.5

Source: OHS/IES95, own computations, unweighted figures.

Quintiles based on per adult-equivalent expenditures in current Rands were designed for this sample and information about the expenditure in each quintile is presented in Table 3. A first impression of the welfare inequality in this sample is given by the ratio of the average expenditures in the *fourth* quintile to the first being 4.3, while the corresponding ratio is 11.4 for the *fifth* and first quintiles. Hence, the most distinct change in expenditure levels occurs between the two highest quintiles. In absolute terms the within-quintile expenditure span is by far the largest in the fifth. However, the range of relative expenditures is just slightly wider in the fifth quintile than it is in the first, with ratios of the highest to lowest expenditure at 23.3 and 22 respectively. The relative ranges are considerably narrower in the other three quintiles.

Table 3) Mean, minimum and maximum per adult-equivalent expenditure, by quintile (1995 Rand)

Quintile	Mean	Min	Max
1	984	62	1 369
2	1 738	1 369	2 141
3	2 646	2 141	3 241
4	4 248	3 241	5 663
5	11 255	5 663	133 037
All	4 174	62	133 037

5 Descriptive Statistics and Partition-Defining Characteristics

Given the importance of households' access to employment and *wage income* detected in previous research on South African inequality (Leibbrandt, Woolard, and Woolard (2000), Leibbrandt, Woolard, and Bhorat (2000), van der Berg (2000), Jenkins and Thomas (2000)), this study applies partitions into subcategories along, on the one hand, households' main income source category (henceforth "Main income source") - as a reflection of its labour market attachment – and, on the other, characteristics that are commonly used determinants for individuals' wage earnings (Willis (1986), Moll (1998), Kingdon and Knight (1999), Mwabu and Schultz (2000)).

While the concept of a household head is non-trivial, the definition used by Statistics South Africa enumerator's manual for the October Household Survey is applied: a head of household can either be male or female, and is the person who assumes responsibility for the household (Budlender (1997)). Assuming implicitly that the head is a significant earner of income, the implied determinants for the households wage or non-wage earnings characteristics are proxied by the population group, highest educational achievement, gender, and age category of the household head, as well as the household's location in rural or urban areas and province of residence (henceforth "Race", "Education", "Gender", "Location" and "Province" when referred to as explanatory variables).

As discussed by Leibbrandt and Woolard (2001), one can expect variables of this nature to "move together" in the South African setting. The reasons for the presumably high degree of correlation are found in the historical legacies of racially discriminatory practices which span across areas such as access to education, labour market regulations, migration, settlement and rights of landownership (e.g. Wilson and Ramphele (1989)).

Hence, a high degree of overlap in fractions of explained inequality by these characteristics would be expected in samples containing *all* the South African population groups. For an impression of the extent to which one can expect race to be of individual significance as an explanatory variable in this sample, the reader is referred to Table 4. The table shows that the fraction of coloured individuals constitutes less than ten percent throughout the third quintile and is just over 20 percent in the fifth. At R 6 253 per month, the mean expenditure in the same population group is some 50 percent higher than that of the African at R 3 920. Brief introductions to each of the five remaining explanatory factors for inequality are introduced below, with descriptive statistics that serve to justify their application.

Table 4) Percentage-wise composition of sample quintiles, by Race.

Location	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	All	Mean expenditures
African	97.2	93.5	90.1	84.9	79.9	89.1	3920
Coloured	2.8	6.6	9.9	15.1	20.1	10.9	6253
Total	100.0	100.0	100.0	100.0	100.0	100.0	4174

South African Households' Income Sources

The South African literature usually distinguishes, by one set of labels or another, between at least four broad groups of household income sources: private transfers, public transfers, self-employment, and wage income (e.g. Carter and May (1999), Leibbrandt, Woolard, and Bhorat (2000)). According to Leibbrandt, Woolard, and Bhorat (2000), income generation processes differ above and below the poverty line, in that the contributions of *wages* to total income are smaller among the poor and vice versa for remittances and state transfers. One conclusion made by the authors is that wage income is central in the determination of both poverty status and poverty depth. On the same note, Bhorat (2000) shows that households with earners that are exclusively either domestic workers or agricultural workers have relatively high poverty propensities. Also of high relevance to this study, van der Berg (2000) shows that the shares of remittance income decline in higher income-consumption quantiles and wage-income shares increase, both in general and as households' main sources of income.

With respect to the definition of relevant categories of income sources, in a study of poverty and labour market participation, Van der Berg (1992) decomposes the sectors of employment for the South African labour force into

three groups. The categorisation is based on the extent to which workers and dependants "participate in the modern consumer society" (*ibid*: 152). The three groups are:

- **the core economy** sectors – manufacturing, government, other industry and services
- **the marginal modern economy** – commercial agriculture, domestic services, mining
- **the peripheral economy** – subsistence agriculture, informal sector, unemployed

According to Van der Berg (1992: 152):

'… part of the labour force in the modern economy are to a larger degree no longer poor. Poverty in its most extreme form now mainly occurs in the peripheral sectors […], but is also widespread amongst workers and dependants relying on earnings from the primary and low-wage sectors'.

In this study, the classification of households' income sources are inspired by the above work, but categories within the marginal modern sectors have been created according to subsector origin and public and private transfer incomes implicitly represent household income generation in the "peripheral" segment. Here, the "core" thus includes all sectors *except the Primary sectors, Domestic services* and *Mining and quarrying*. The Core sector category furthermore encompasses households with capital income and all types of self-employment income as main income sources. In addition to these income sources, it is recognised that households also derive "indirect income" and "*diversifying*" households are defined as those without a unique main income source that meets a contribution requirement discussed below. The income source categories are described in greater detail and in as close approximation as possible to the wording in the IES95 questionnaire in Appendix 1.[10]

---

10 It has been noted by Leibbrandt *et al* (2000), that the IES95 data do not capture agricultural activities for own consumption well. In this study's sample from the IES95, 9.7 percent of all households were recorded with either slaughtered domestic animals or harvested crops in the year preceding the interview. Profit from agricultural activities should be registered in the IES questionnaire under "self-employment", but only 1.2 percent of the households that had slaughtered or harvested had records of any self-employment profits at all. The above figures presumably understate the importance of agriculture, which according to May (1996) assumes several important functions as *inter alia* a supplementary source of nutrition and as a safety net for vulnerable households in South Africa. But left with little choice other than taking the data at face value, agricultural production is not treated as a separate source of income. The individuals in the few households that would have agricultural

Main Income Source Definition

The definition of a main income source is not trivial.[11] One possible route is to construct the definition by the source's contribution to total household income. Some ambiguity necessarily enters the decision of where to draw the cut-off contribution line. This study uses a minimum contribution (regardless of the number of members that raise the income) of 66.7 percent to total household income, an appeal of which is that the main income source contributes at least twice as much as any other income source.

Table 5 shows the impacts on the distribution of individuals across the various main income categories from where the cut-off contribution is drawn. In the second row of the table can be seen that roughly 75 percent of the households had a main income source and that approximately half the main income sources originate in the core sectors. The second largest category is Public Transfers with 16 percent of the households, followed by Private Transfers with six percent. Both the Mining and Quarrying and the Domestic services categories are small with two percent each, while the Primary sectors and Indirect income groups contain five and four percent respectively.

The figures in Table 6 attest to the notion that income generation activities as defined by these categories vary across the expenditure distribution. The table shows the composition of the sample quintiles with respect to households' main income sources. The figures show, for example, that the fractions of households which rely on Core or Mining-and-Quarrying sector wage incomes increase dramatically from the lower income brackets to the higher. The opposite is true for the two categories of households that rely on Public transfers or Private transfers, as well as for households with main income from the Primary sectors and Domestic services.

Mean expenditure levels by Main income source reflect the above distribution, with annual averages in the neighbourhood of R 2 000-2 500 for individuals in households with either transfers or wage incomes from the Primary or Domestic service as main income sources. Members of households which rely on Indirect income, Core or Mining and Quarrying sector incomes are associated with mean expenditures in the range of R 5 000-6 000, while those in Diversifying households constitute a middle category with average annual expenditures just below R 3 500.

---

income as their main source are included in the core economy category along with other types of self-employment.

[11] For the analysis of livelihoods in a dynamic setting, Ardington and Lund (1996) raise a valid objection to the use of a "dominant source of income" in that such sources may be of a temporary nature.

Table 5) Percentage fraction of individuals in households by main income source category and various main-income cut-off contributions levels.

Main income contribution to total household income	*Main income source category*								
	No main income source	*Core sectors*	*Mining and quarrying*	*Primary Sectors*	*Domestic Services*	*Public transfers*	*Private Transfers*	*Indirect income*	*Sum*
50%	6	43	2	8	3	20	8	10	100
66.7%	**26**	**39**	**2**	**5**	**2**	**16**	**7**	**4**	**100**
75%	37	34	2	4	1	14	6	2	100
90%	55	25	1	2	1	11	5	0	100
100%	75	14	0	1	1	7	3	0	100

Table 6) Percentage-wise composition of sample quintiles by Main income sources

Main income Source	*Quintile 1*	*Quintile 2*	*Quintile 3*	*Quintile 4*	*Quintile 5*	*Total*	*Mean expenditures*
Diversifying	28.9	30.9	29.7	24.1	17.1	26.1	3435
Core sectors	13.7	23.4	36.0	**50.7**	**69.0**	38.6	5983
Mining & quarrying	1.0	1.2	1.9	**3.1**	**3.5**	2.2	5249
Primary sectors	**6.1**	**6.0**	6.4	4.3	0.8	4.7	2404
Domestic services	2.5	2.2	1.9	1.4	0.6	1.7	2506
Public transfers	**33.1**	**24.8**	13.2	6.9	1.6	15.9	1923
Private transfers	**11.4**	**8.6**	7.6	4.4	1.2	6.6	2232
Indirect income	3.3	3.0	3.3	5.1	6.3	4.2	5860
Total	100.0	100.0	100.0	100.0	100.0	100.0	4174

The Two Geographical Dimensions of Inequality

In the study, two sets of groups are defined by Province and by Location. Table 7 shows that 60 percent of the households in the sample are rural but also that the Location-wise composition of the quintiles differ considerably. In the lowest bracket, the fraction of rural individuals is nearly 85 percent while in the highest

bracket, the corresponding fraction is 28.3 percent. The differences in composition shift much more gradually across the three first quintiles while the rural fraction decreases by twenty percentage points between both the third and fourth and the fourth and fifth quintiles. It is also noteworthy that the urban mean expenditure at R 6 124 is more than twice that of the rural at R 2 878.

Table 7) Percentage-wise composition of sample quintiles by Location.

Location	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	All	Mean expenditures
Rural	84.3	75.6	65.5	46.6	28.3	60.1	2878
Urban	15.7	24.4	34.5	53.4	71.7	39.9	6124
Total	100.0	100.0	100.0	100.0	100.0	100.0	4174

From the figures in Table 8, it can be seen that two provinces, KwaZulu-Natal and the Eastern Cape, each contain fractions of almost 20 percent of the sample. In a second category of size are Limpopo and Gauteng with 14.3 percent and 12.3 percent respectively. The Northern Cape contains the smallest sample fraction with only 1.8 percent, while the remaining four provinces contain shares ranging from 8 to 9 percent. The differences in the provincial composition of the quintiles are perhaps best illustrated using, on the one hand, the poorest provinces of the Eastern Cape and the Free State, and on the other the richest, Gauteng and the Western Cape. The fractions of the poorest two provinces in Table 8 are considerably smaller in the highest bracket, at 9.5 percent and 4 percent respectively, than in the lowest with 29.7 percent and 12.2 percent. Vice versa applies to the two richest provinces, with 29.6 percent and 14.4 percent in the fifth quintile and 2.2 percent and 1.1 percent in the first.

Table 8) Percentage-wise composition of households' sample quintiles by Province.

Province	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	All	Mean expenditures
W Cape	1.1	4.6	8.2	14.2	14.4	8.5	5834
E Cape	29.7	24.9	17.6	12.3	9.5	18.8	2989
N Cape	1.7	2.3	2.2	1.9	1.1	1.8	3348
Free State	12.2	8.6	6.5	4.9	4.0	7.2	2814
KZN	15.5	21.1	24.8	21.8	15.4	19.7	3769
NW Province	11.9	11.4	9.3	7.1	7.1	9.4	3517
Gauteng	2.2	3.8	8.6	17.4	29.6	12.3	7558
Mpumalanga	7.8	9.4	9.6	8.2	5.2	8.0	3273
Limpopo	17.8	14.0	13.4	12.3	13.8	14.3	4122
Total	100.0	100.0	100.0	100.0	100.0	100.0	4174

In terms of average expenditures, the Eastern Cape and the Free State are both found at the bottom with less than R 3 000 per month. Slightly higher mean expenditure levels are found in Mpumalanga and the Northern Cape at approximately R 3 300. The average in the Northern Province is R 200 higher than the latter two and KwaZulu-Natal is higher than the Northern Province by the same amount. With over R 7 500, Gauteng is at a considerably higher level than that of the second highest province, the Western Cape, at just over R 5 800. The average expenditure in Limpopo is below the nationwide average by just over R 50.

Education of Household Heads

The association between the head's education level and the expenditure distribution is depicted in Table 9. As can be seen, almost 75.4 percent of the household heads in the sample have primary education or less and approximately two-fifths of those have no education at all. Almost 15 percent of all the heads have *some* secondary education, whereas only 11.5 percent have *completed* or above secondary education, out of which 5.2 percentage points have more than secondary education.

Table 9) Percentage-wise composition of sample quintiles by Education.

Education level of household head	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	All	Mean expenditures
None	46.8	37.6	30.8	18.8	7.9	28.4	2388
Primary	46.1	50.3	50.3	50.3	29.9	45.4	3304
Some secondary	5.5	9.5	13.6	20.4	24.4	14.7	5537
Complete secondary	1.5	1.9	3.3	6.6	18.4	6.3	8671
Above secondary	0.2	0.6	1.9	4.0	19.3	5.2	12177
Total	100.0	100.0	100.0	100.0	100.0	100.0	4174

The fraction of households that are headed by individuals with none or primary education each diminishes upwards in the income distribution from approximately 46 percent to around 9 percent and 30 percent, respectively, in the fifth quantile. The opposite is true for the fractions of households with better educated household heads that increase from 1.5 percent and 0.2 percent, in order of educational achievement, to 18.4 percent and 19.3 percent, respectively, in the highest bracket. Also here, the compositions within the three lower quintiles are reasonably similar and the change between the fourth and fifth quintiles is more dramatic than between the third and fourth.

There are also considerable differences in mean expenditure levels between the Education subgroups. The highest expenditures are found among households with heads that have more than secondary education, the average of which at R 12 177 is five times higher than for the category with none-educated heads. The mean expenditure of the second highest education category is found at R 8 671, which in turn is R 3 134 above the mean of the households with heads that have some secondary education. Households with heads that have only primary education have a mean expenditure level of R 3304.

Gender and Age of Household Heads

Table 10 illustrates that the fraction of female headed households host just over 35 percent of the sample, but the fraction decreases gradually by a total of almost ten percentage points, from 43 percent in the poorest quintile to the 33.8 percent in the fourth quintile. In the richest quintile however, the corresponding fraction is only 24.3 percent. The average expenditure level of male headed households at R 4 629 is almost 40 percent higher than that of the female headed at R 3 352.

Table 10) Percentage-wise composition of sample quintiles by Gender.

Gender of household head	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	All	Mean expenditures
Female	43.0	40.7	36.4	33.8	24.3	35.6	3352
Male	57.0	59.3	63.6	66.2	75.7	64.4	4629
Total	100.0	100.0	100.0	100.0	100.0	100.0	4174

With respect to age categories, the figures in Table 11 show that households with heads aged 41-59 years old contain 45 percent of the sample. The youngest and oldest categories, below 25 and above 60, each host 2.7 percent and the remaining two subgroups thus take in approximately one-quarter each. The fractions of the youngest and second oldest categories do not differ dramatically across the quintiles either and remain at approximately 2.5 percent and 45 percent. The fraction of individuals that live in households with elderly heads is reduced dramatically however, from almost 35 percent in the poorest quintile to just below 11 percent in the highest. The pattern is the diametrically opposite for the second youngest age category which increases from 17.4 percent in the first quintile to over 42 percent in the fifth. In both cases the shifts in composition are most dramatic from the fourth to the fifth quintile.

Table 11) Percentage-wise composition of sample quintiles by Age category.

Age category of household head	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	All	Mean Expenditures
≤ 25	2.6	2.7	2.8	2.9	2.2	2.7	4006
26 – 40	17.4	20.5	27.2	31.0	42.2	27.7	5513
41 – 59	45.0	43.6	43.3	46.1	44.8	44.6	4162
60 ≤	34.9	33.2	26.8	20.0	10.8	25.1	2741
Total	2.6	2.7	2.8	2.9	2.2	2.7	4174

With respect to expenditure levels, considerable differences exist between on the one hand the expenditures of households the heads of which are in the oldest age category and those in the second youngest category, in the age span of 26 to 40. The latter's average is found at R 5 513, while the former's average is at R 2 741. The expenditure level among individuals that live in households with heads in the second oldest age category is R 12 below the nationwide average and R 156 higher than that of the youngest age category.

Simultaneous Application of Several Explanatory Variables

Finally, reasons exist to believe that partitions both by Education and Main income sources to some extent may capture geographical variation; Firstly, during the apartheid era - when most heads in the sample were of schooling age – the access to and the quality of Education for Africans was subject to geographical variation (Wilson and Ramphele (1989)). Secondly, it is conceivable that households' Main income sources are determined *inter alia* by the education of household members that raise the income as well as by the household's geographical location. Also, Education may affect the level of earnings from individual income sources. In order to get an impression of the extent to which the partitions by Education and Main income source capture geographical variation and/or overlap each other, the results from four other partitions, based on combinations of the latter four explanatory characteristics will also be presented in the section with empirical results.

6 Empirical Analyses

Table 12 contains the results from the decomposition of the Theil-L index along the partitions defined by the various explanatory characteristics. The table is designed in the following manner: The total level of inequality is found in the

first column of the table. Each row contains the results from the partition along one specific characteristic. Within each row are found two sets of figures, the top ones are percentage fractions of total inequality and the lower ones contain absolute index values. The third column contains the fraction and index value of *explained inequality* pertaining to each partition, followed in the fourth column by the total *within-group* fraction and index value. The subsequent columns contain the contribution to total inequality from inequality within each group defined by the partition, as well as the index value for each group. (As an indication of the robustness of results, Table A4 in Appendix 2 contains the corresponding values for the decomposition of the $G_1(x)$ (Theil-T) index and Table A5 and A6 contain the corresponding decompositions for a data set from year 2000.)

The index value for total inequality in the sample is 0.393 and the rows in the table are found in rank order of each partition's fraction of explained inequality. Two points are of contextual interest in this respect. Firstly, Leibbrandt, Woolard and Bhorat (2000) report a Theil-L statistic of 0.706 for the *total* sample of individuals in the same survey – including also the Asian/Indian and white subsamples - with the identical adult equivalence scale applied to household *income*. Hence, inequality in this study's sample is considerably lower than in the full sample. Secondly, while the Theil-L and Gini-index of inequality do not in general yield identical results, South Africa has the seventh highest level of income inequality in the world as measured by the former applied to the full population (World Bank (2004)).[12]

As can be seen from the third column, the explained inequality from the gender of household head is the lowest at 2.9 percent, followed first by heads' population groups and then heads' age categories, with 3.0 percent and 7.9 percent respectively. The Province and Location partitions yield higher explained fractions at 13.1 percent and 17.9 percent, respectively, while differences in households' Main income source categories explain 23.5 percent of the inequality. The highest value is found for the education levels of household heads which account for 30.8 percent of the total inequality.

Equation (4') shows how the total within-group fraction of inequality for a given partition is a weighted-sum of the inequality within each subgroup and from equation (6), it follows that the percentage fraction of within-group inequality is always 100 minus the between-group fraction. Hence, within-group inequality is higher than 90 percent in the three first partitions, approximately 85 percent in the fourth, near three-quarters in the Main income source partition and just

---

[12] Leibbrandt, Woolard, and Woolard (2000) decompose the Gini-index for South Africa.

below 70 percent when subgroups are defined by Education. With the exception of the Province partition, the largest subgroups also contribute the largest fraction to within-group inequality. However, only for the partitions by Gender, Race, and Age are the largest subgroups also associated with the highest *levels* of inequality.

Among the smaller subgroups which display inequality levels that are distinctly higher than the nationwide figure are found the province of Limpopo and the Main income category Indirect income. The former contains 16.7 percent of the sample and has an inequality index value of 0.46, whereas 5.2 percent of the sample reside in households that belong to the latter income category, which includes highly varying types of income sources. With respect to Limpopo province, it is noteworthy that the average expenditure in the province was very close to the nationwide mean (see Table 8).

Several subgroups also display considerably *lower* inequality than the nationwide level. Among the groups with the very lowest inequality levels are the four provinces of the Western Cape, Northern Cape, Gauteng and Mpumalanga; the five Main income source categories, Mining and Quarrying, Primary sectors, Domestic services, and the two transfer categories, the two Education categories (Some primary and Post secondary) and finally the above-60 age category. The index value for all these groups is in the approximate range of 0.20-0.30. A plausible explanation for these low levels of inequality, in accordance with the previous section's descriptive statistics, is that the Western Cape, Gauteng, Mining and Quarrying, and the Post secondary education categories all predominantly contain observations at the *upper* end of the expenditure distribution, while the other subgroups contain observations clustered at the *lower* end.

The partition defined by Education category yielded the highest fraction of explained inequality above, followed by the partitions by Main income source, Location and Province. The results from four other partitions, that are defined along   more   than   one   dimension,   are   shown   in   Table   13.

Table 12) Inequality in the sample as measured and decomposed by the Theil-L index; percentage fractions of between- and within-group inequality partitions defined by one characteristic. Absolute index values in parentheses

Total index value	Partition	Between-group inequality	Total within-group	Within-group contributions and absolute levels of inequality							
0.393	**Gender**	2.9 (0.012)	97.1 (0.382)	**Female** 30.2 (0.333)	**Male** 66.9 (0.409)						
	Race	3.0 (0.012)	97.0 (0.381)	**African** 88.0 (0.389)	**Coloured** 9.0 (0.323)						
	Age	7.9 (0.031)	92.1 (0.362)	**≤25** 2.5 (0.375)	**26-40** 27.9 (0.396)	**41-59** 44.7 (0.395)	**60≤** 17.0 (0.265)				
	Province	13.1 (0.051)	86.9 (0.342)	**W Cape** 5.7 (0.265)	**E Cape** 17.8 (0.373)	**N Cape** 1.3 (0.281)	**Free State** 6.0 (0.327)	**KZN** 15.5 (0.308)	**NW Prov** 9.1 (0.383)	**Gauteng** 9.3 (0.299)	**Mpumalanga** 5.5 (0.269) · **Limpopo** 16.7 (0.461)
	Location	17.9 (0.070)	82.1 (0.323)	**Rural** 46.7 (0.306)	**Urban** 35.4 (0.349)						
	Main income source	23.5 (0.092)	76.5 (0.301)	**Diversifying** 22.5 (0.338)	**Core sectors** 32.4 (0.330)	**Mining & Quarrying** 1.5 (0.266)	**Primary sectors** 2.3 (0.192)	**Domestic services** 1.0 (0.234)	**Public transfers** 8.0 (0.197)	**Private transfers** 3.7 (0.219)	**Indirect income** 5.2 (0.486)
	Education	30.8 (0.121)	69.2 (0.272)	**None** 17.6 (0.244)	**Some primary** 31.4 (0.272)	**Some secondary** 11.3 (0.302)	**Complete secondary** 5.1 (0.314)	**Post secondary** 3.9 (0.292)			

Note: Within each row are found two sets of figures, the top ones are percentage fractions of total inequality and the lower ones contain absolute index values.

Table 13) Explained fractions of inequality; partitions defined by multiple characteristics.

Partition defining Characteristics	Explained inequality	Number of groups	
		Theoretical	*Observed*
Province and Location	25.2	18	18
Province, Location, and Main income source	39.8	144	141
Province, Location and Education	44.4	90	90
Province, Location, Education and Main income source	53.0	720	574

In the first row of results (in table 13), it can be seen that, when applied simultaneously to define 18 subgroups, Province and Location jointly explain 25.2 percent of the inequality in the sample. When each of these subgroups were refined by hypothetically eight Main income source subgroups each, 141 observed subgroups were returned, that jointly explain 39.8 percent of the inequality. Hence, 14.6 percentage points of explained inequality were added. If the 18 geographical subgroups were rather refined by Education, 90 groups were defined and observed, which added 19.2 percentage points to yield an explained fraction of inequality of 44.4 percent. Finally, refining further by applying both Education and Main income source to the combined geographical partition returned 574 observed groups and a fraction of explained inequality of 53 percent.

Hence, implications from the latter set of results are that, when applied solely, Education and Main income sources both capture some of the inequality explained by Location and Province. However, when partitions are defined by Location and Province jointly *and* either Education or Main income sources, both of the latter individually capture inequality that is not explained by geographical variation. In reality it is furthermore plausible that household heads' Education is a determinant of both (i) the households' type of main income source and (ii) the returns from that main income source. The results in the fourth row of Table 13 are perhaps most readily interpreted as evidence of point (i) but the results are likely to feature also inequality due to point (ii). The applied methodology can not resolve this issue.

7 Conclusions

Commenting on their results from the decomposition of income inequality in the US, Cowell and Jenkins (1995) consider their explained fractions in the ranges of 20 to 30 percent "not much". This study has utilised the Theil-L measure to decompose expenditure inequality in a sample of black and coloured South African individuals, sampled by a household survey from 1995. The results from partitions defined by *one* characteristic are higher than "not much" in only one case here, namely in that of the education level of household heads, which accounts for 30.8 percent of the inequality. However, partitions that take several factors into account return explained fractions up to 53 percent. Hence, at least some of the inequality in this sample can be explained.

Somewhat surprisingly, the explanatory power of race in these samples is relatively low, which is to some extent true for geographical location in both the rural-urban dimension and in provinces, when applied separately. However, the explanatory power of geography increases to 25.2 percent when province and rural-urban location are applied jointly, which suggests a more meaningful perspective on the spatial dimension of welfare in South Africa. When the same partition was further refined by households' types of main income source, explained fractions rose to 39.8 percent, while a refinement by household heads' education level, rather than main income sources, yielded explained fractions of 44 percent. Applied simultaneously to the joint geographical partition, education and main income sources yielded the abovementioned highest fraction of explained inequality.

Thus, among the explanatory variables applied in this study, the education levels of household heads stand out as the single most important associate of differing positions in the expenditure distribution. However, further research into the determinants of households' types of main income sources is also warranted, partly by its relatively high explanatory power in this setting. A further interpretation of the results is that main income sources add explanatory power to what is attributable to Education and geography. Hence, other factors in addition to the latter two may determine the allocation of main income sources to households.

Other researchers have shown that a substantial contribution to total inequality in South Africa arises from inequality within the African and coloured population. The results in this study show that a considerable fraction of that within-population group inequality *can* be explained by further refinement of partitions into a number of subgroups (a partition

by, for example, Race and Education applied to a national level would yield 20 subgroups). Currently most analyses of South African inequality are undertaken without investigation beyond the too narrow focus on population groups. Such an approach neglects several dimensions of inequality. The dimensions of households' core-economy integration and of education legacies from the apartheid era inform our understanding of the phenomenon. Relatively small differences in lengths of education affect the distribution of welfare among those worst off and this may be addressed by policy as may amounts and eligibilities for transfers.

References

Ardington, E., and Lund, F. (1996). "Questioning rural livelihoods", Lipton, M., De Klerk, M., and Lipton, M. (Eds), (1996) *Land, labour & livelihoods in rural South Africa*. Durban, South Africa: Indicator Press. Ch. 2.

Bhorat, H., Leibbrandt, M., Maziya, M., Van der Berg, S. & Woolard, I. (2001). *Fighting poverty – labour markets & inequality in South Africa*. Cape Town, South Africa: UCT Press.

Blackorby, C., Donaldson, D., and Auersperg, M. (1981) "A new procedure for the measurement of inequality within and among population subgroups", *Canadian Journal of Economics*, 14(4): 665-85.

Bourguignon, F. (1979). "Decomposable income inequality measures", *Econometrica*, 47: 901-20.

Budlender, D. (1997). "The debate about household headship." *Statistics South Africa Report*, Statistics South Africa, Pretoria.

Champernowne, D. G., and Cowel, F. A., (1998) *Economic inequality and income distribution*, Cambridge, Cambridge University Press, UK.

Chaudhuri, S., and Ravallion, M. (1994). "How well do static welfare indicators identify the chronically poor?", *Journal of Public Economics*, 53(3): 367-394.

Cowell, F. A. (1980). "On the structure of additive inequality measures", *Review of economic studies*, 47: 521-31.

Cowell, F. A., and Jenkins, S.P.(1995), "How much inequality can we explain? a methodology and application to the United States.", *Economic Journal*, 105(429): 421-430.

Deaton, Angus, and Salman Zaidi (2002). "Guidelines for constructing consumption aggregates for welfare analysis". *LSMS Working Paper* No. 135, World Bank, Washington.

Fields, G. (2003) "Accounting for income inequality and its change: A new method, with application to the distribution of earnings in the United States," *Research in Labor Economics*.

Gustafsson, B., and Li, S. (2002). "Income inequality within and across counties in rural China", *Journal of Development Economics*, 69(1): 179-204.

Haddad, L., and Kanbur, R., (1990). "How serious is the neglect of intra-household inequality", *Economic Journal*, 100: 866-881.

Hoogeveen and Özler (2004), *Not separate, not equal, poverty and inequality in post-apartheid South Africa*, unpublished World Bank mimeo.

Kingdon, G., and Knight, J. (1999). "Unemployment and wages in South Africa: A spatial approach." *Working Paper* WPS/99.12, Centre for the Study of African Economies, Oxford University.

Klasen, S., (1997). "Poverty, inequality and deprivation in South Africa: an analysis of the 1993 SALDRU survey", *Social Indicators Research*, 41.

Klasen, S. (2000). "Measuring poverty and deprivation in South Africa", *Review of Income and Wealth*, 46(1): 33-58.

Lanjouw, P., Prennushi,G. and Zaidi, S. (1996). "Building blocks for a consumption-based analysis of poverty in Nepal". The World Bank. Unpublished manuscript.

Leibbrandt, M. and Woolard, I. (1999) "A comparison of poverty in South Africa's nine provinces", *Development Southern Africa*, 16(1): 37-54.

Leibbrandt, M., Woolard, I., and Bhorat. H. (2000)."Understanding contemporary household inequality in South Africa", *Studies in Economics and Econometrics*, 24(3): 31-52.

Leibbrandt, M, Woolard, C, & Woolard, I, 2000. "The contribution of income components to income inequality: a decomposable Gini analysis." *Journal of African Economies*. 9(1): 79-99.

Leibbrandt, M & Woolard, I, 2001. Household incomes, poverty & inequality in a multivariate framework in Bhorat, H., Leibbrandt, M., Maziya, M., Van der Berg, S. & Woolard, I. (2001). *Fighting poverty – labour markets & inequality in South Africa*. Cape Town, South Africa: UCT Press.

Lipton, M., De Klerk, M., and Lipton, M. (Eds), (1996) *Land, labour & livelihoods in rural South Africa*. Durban, South Africa: Indicator Press.

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* 15. Land and Agriculture Policy Centre, Johannesburg, South Africa.

Moll, P., (1996). "The collapse of primary schooling returns in South Africa 1960-90", *Oxford Bulletin of Economics and Statistics*, 58(1): 29-55.

Mwabu, G., and Schultz, P. (2000). "Wage premia for education and location, by gender and race in South Africa", *Economic Development and Cultural Change*, 48(2): 307-334.

Ravallion, M.(1996). "Issues in measuring and modelling poverty", *Economic Journal* 106: 1328-43.

Sen, A. (1985). *Commodities and capabilities,* Amsterdam, North Holland.

Sen, A. (1987). *The standard of living*, Cambridge, Cambridge University Press.

Shorrocks, A. F. (1980). "The class of additively decomposable inequality measures", *Econometrica*, 48(3): 613-25.

Shorrocks, A. F. (1984). "Inequality decomposition by population subgroups", *Econometrica*, 52(6): 1369-85.

Slesnick, D. (1993). "Gaining ground: poverty in the postwar United States.", *Journal of Political Economy*, 101: 1-38.

Slesnick, D. (1998). "Empirical approaches to the measurement of welfare", *Journal of Economic Literature*, 36: 2108-2165.

Statistics South Africa, 1997. Documentation supplied with OHS & IES1995 raw data sets. Statistics South Africa, Pretoria.

Tsaklogiu, P. (1993), "Aspects of inequality in Greece", *Journal of Development Economics*, 40(1): *53-74.*

van der Berg, S, 1992. "Redressing economic inequality in South Africa" in *Redistribution and affirmative action – working on the South African political economy*, Hugo, P. (Ed.). Halfway House, South Africa; Southern Book Publishers.

van der Berg, S, 2000. "Social policies to address poverty" in Bhorat, H., Leibbrandt, M., Maziya, M., Van der Berg, S. & Woolard, I. (2001). *Fighting poverty – labour markets & inequality in South Africa*. Cape Town, South Africa: UCT Press.

Whiteford, A. and McGrath, M., (1998). "Income inequality over the apartheid years." *SANER Working Paper*, No. 6, South African Network for Economic Research, Cape Town, South Africa.

Whiteford, A. and van Seventer, D. E. (2000) "South Africa's changing income distribution in the 1990s", *Studies in Economics and Econometrics*, No 24:3: 7-30.

Willis, R. J. (1986). Wage determinants: a survey & reinterpretation of human capital earnings functions. In Ashenfelter, O. & Layard, R. (Eds), *Handbook of labor economics*, Volume I, Elsevier Science.

Wilson, F. and Ramphele, M. (1989), *Uprooting poverty: the South African challenge.* Cape Town, South Africa: David Philip Publishers

World Bank (2004) *World Development Indicators 2004*, World Bank, Washington DC.

Appendix 1

Income originating from the core economic sectors (henceforth "Core sector income"): *salaries and wages*[13] *from secondary sectors and tertiary sectors* including s*elf-employment income*, in the form of net profit from business or professional practice/activities conducted on a full time basis; and *capital income* from the letting of fixed property, royalties, interests, dividends and annuities.[14]

Primary sector income: *salaries and wages from agriculture, fishing, and forestry.*

Mining and quarrying sector income: *salaries and wages from mining and quarrying.*

Domestic services income: *salaries and wages from private households.*

Private transfers: alimony, maintenance and similar *allowances from divorced spouses or family members living elsewhere* and regular allowances from family members living elsewhere.

Pensions and public transfers: *pensions resulting from own employment, old age and war pensions, social pensions* or allowances in terms of *disability grants, family and other allowances, or from funds* such as e.g. the Workmen's Compensation, Unemployment Insurance, Pneumoconioses and Silicosis funds.

Indirect income: income derived from [i] hobbies, side-lines, part-time activities, or the sales of vehicles, property etc; [ii] payments received from boarders and other members of the household; [iii] the pecuniary value of goods and services received by virtue of

---

[13] In the "salaries and wages" concept is included bonuses and income from over time, commissions and directors fees, part-time work and cash allowances in respect of transport, housing and clothing.

[14] The secondary sectors encompass the Statistics South Africa (1997b) "Major sector divisions": Manufacturing, Electricity, gas and water and Construction. The tertiary sectors constitute the "Private services" and "Community, social and personal services" excluding "Private households with employed persons". "Private services" is made up of the major divisions: Wholesale and retail trade, repair of motor vehicles, motor cycles and personal and household goods, hotels and restaurants; Transport, storage and communication; and Financial intermediation, insurance, real estate and business services.

occupation; [iv] gratuities and lump sum payments from pension, provident and other insurance or from private persons; [v] 'other income' withdrawals, bursaries, benefits, donations and gifts, bridal payment or dowries and all 'other income'.

Appendix 2

This appendix provides indications of the robustness of the results in Table 12 and contains the results from the decomposition of inequality in the sample by the Thei-T measure. In addition to the latter, decomposition was undertaken with both measures to a data set from year 2000, generated by a nationwide questionnaire-based survey similar to the OHS/IES95, but with Statistics South Africa's biannual Labour Force Survey having taken the place of the then ceased annual October Household Survey. The comparability of the two data sets has been disputed in the South African research community. One reason for this uncertainty is that nominal incomes were lower in the four lowest per adult-equivalent income quintiles in year 2000 than 1995, whereas total expenditure data which do not display such characteristics. Tables A1 and A2 illustrate the latter issues.

The problem of matching individuals was much smaller with the LFS/IES2000 than with the OHS/IES95 data. In the former 96.2 percent of the observations that met the population group criterion are used. However, in addition to zero expenditures for 106 observations, 3438 observations were lost from the same data set due to either missing remuneration or industry data for wage earners or due to missing education or age data for household heads. The sample delimitation process is illustrated in Table A3.

Table A4 contains the decomposition results from the Theil-T index applied to the 1995 data. The decomposition results for the year 2000 data are shown in tables A5-6. (Results from the application of other inequality measures on both data sets yield similar results and are available from the author.) In the 2000 data the total levels of inequality are considerably higher. However, in all three decompositions the relative rank and approximate differences in explained inequality for the various partitions are similar to those in Table 12. In the Theil-T decomposition of the 1995 data the explanatory power of race and Main income source is however lower by approximately one-tenth and one-fifth respectively.

In both decompositions of the 2000 data the explanatory power of race is roughly 50 percent higher than in the 1995 Theil-L decomposition. Applying the same index to the 2000 data shows that the explanatory power of Gender is almost twice as high, while that of Location is around ten percent higher. The fraction of explained inequality yielded by Education is however lower by one-tenth. For the Theil-T decomposition the difference in inequality explained by Gender is approximately 60

percent higher than in the Theil-L decomposition for 1995. Both the explanatory power or Age and Main income source are also lower, in the former case by some 15 percent and in the latter by around one-fifth.

Table A1) Per adult-equivalent household income quintiles in the full OHS/IES95 and LFS/-IES2000 samples; weighted population sizes and annual incomes in nominal Rand.

Sample	Quintile	Weighted number of individuals	Mean Income	Minimum income	Maximum income
	1	8 014 923	1 434.1	148.8	2 098.3
	2	8 015 225	2 813.6	2 098.3	3 638.7
	3	8 015 152	4 874.6	3 639.1	6 511.6
1995	4	8 014 762	10 010.8	6 511.6	15 538.1
	5	8 015 588	41 255.5	15 538.1	2 657 998.0
	All	40 075 650	12 078.1	148.8	2 657 998.0
	1	8 358 799	1315.8	11.0	2 035.4
	2	8 359 303	2745.0	2 035.4	3 571.6
	3	8 359 271	4827.8	3 571.6	6 430.6
2000	4	8 358 789	9805.4	6 430.6	15 564.9
	5	8 359 527	46786.6	15 564.9	2203 030.0
	All	41 795 689	13096.5	11.0	2203 030.0

Note: n1995 =125 112   n2000=101 803.
Source: OHS/IES95 and LFS/IES2000, own computations, weighted figures.

Table A2) Per adult-equivalent household expenditure quintiles in the full OHS/IES95 and LFS/IES2000 samples; weighted population sizes and annual expenditures in nominal Rand.

Sample	Quintile	Weighted number of individuals	Mean Income	Minimum income	Maximum Income
	1	8 015 065	1 354.2	162.2	1 982.4
	2	8 015 152	2 685.9	1 982.4	3 504.0
	3	8 015 078	4 719.1	3 504.0	6 314.9
1995	4	8 014 905	9 725.5	6 314.9	15 104.2
	5	8 015 450	40 427.9	15 104.2	2 657 998.0
	All	40 075 650	11 782.8	162.2	2 657 998.0
	1	8 359 097	1 611.5	0.0	2 395.7
	2	8 359 166	3 149.0	2 395.7	3 977.1
	3	8 358 449	5 190.5	3 978.0	6 699.5
2000	4	8 359 738	9 894.5	6 699.5	15 215.0
	5	8 359 239	46 109.7	15 215.0	2 740 995.0
	All	41 795 689	13 191.2	0.0	2 740 995.0

Note: n1995 =125 112   n2000=101 803.
Source: OHS/IES95 and LFS/IES2000, own computations, weighted figures.

Table A3) The year 2000 sample delimitation process

Year	Sample	Number of individuals	Share of total revisited sample	Share of revisited African and coloured sample
2000	Total LFS/IES sample	101 803	100.0	
	African and coloured LFS/IES sample	93 842	92.2	100.0
	Final sample	90 298	88.7	96.2

Source: LFS/IES2000, own computations unweighted figures.

Table A4) Inequality in the 1995 sample as measured and decomposed by the Theil-T index; partitions defined by one characteristic

Total index value	Partition	Between-group inequality	Total within-group	Within-group contributions and absolute levels of inequality								
0.430	**Gender**			**Female**	**Male**							
		2.6 (0.01111)	97.4 (0,41840)	25.0 (0.37478)	72.4 (0.43589)							
	Race			**African**	**Coloured**							
		3.1 (0.01195)	96.9 (0,41612)	84.2 (0.43232)	12.7 (0.33308)							
	Age			**≤25**	**26-40**	**41-59**	**60≤**					
		6.9 (0.02981)	93.1 (0.39970)	2.6 (0.44073)	34.3 (0.40303)	45.0 (0.43462)	11.2 (0.29207)					
	Province			**W Cape**	**E Cape**	**N Cape**	**Free State**	**KZN**	**NW Prov**	**Gauteng**	**Mpumalanga**	**Limpopo**
		13.0 (0.05586)	87.0 (0.37365)	8.3 (0.30001)	14.2 (0.45358)	1.1 (0.32635)	4.0 (0.35644)	14.4 (0.34680)	8.1 (0.44060)	15.8 (0.30430)	4.2 (0.28473)	16.9 (0.51618)
	Location			**Rural**	**Urban**							
		16.4 (0.07056)	83.6 (0.35895)	34.5 (0.35768)	49.1 (0.35984)							
	Main income source			**Diversifying**	**Core sectors**	**Mining & Quarrying**	**Primary sectors**	**Domestic services**	**Public transfers**	**Private transfers**	**Indirect income**	
		19.6 (0.08419)	80.4 (0.34532)	20.0 (0.40029)	43.6 (0.33872)	1.6 (0.24794)	1.2 (0.19605)	0.5 (0.22480)	4.0 (0.23640)	2.0 (0.23741)	7.4 (0.54114)	
	Education			**None**	**Some primary**	**Some secondary**	**Complete secondary**	**Post secondary**				
		32.4 (0.13930)	67.6 (0.29021)	10.1 (0.10114)	24.3 (0.24332)	14.1 (0.31071)	9.1 (0.29793)	9.9 (0.27997)				

Note: n = 92 717.
Source: OHS/IES95, own computations, weighted figures.

Table A5) Inequality in the 2000 sample as measured and decomposed by the Theil-L index; partitions defined by one characteristic data.

Total index value	Partition	Between-group inequality	Total within-group	Within-group contributions and absolute levels of inequality								
	Gender			**Female**	**Male**							
		5.7 (0.028)	94.3 (0.454)	37.3. (0402)	57.0 (0.496)							
	Race			**African**	**Coloured**							
		4.8 (0.023)	95.2 (0.458)	87.7 (0.468)	7.4 (0.370)							
	Age			**≤25**	**26-40**	**41-59**	**60≤**					
		7.6 (0.036)	92.4 (0.445)	3.3 (0.401)	27.7 (0.478)	44.3 (0.503)	17.2 (0.322)					
0.485	Province	13.4 (0.065)	86.6 (0.417)	**W Cape** 6.0 (0.340)	**E Cape** 15.1 (0.460)	**N Cape** 1.3 (0.420)	**Free State** 5.7 (0.493)	**KZN** 17.5 (0.415)	**NW Prov** 7.4 (0.495)	**Gauteng** 18.2 (0.389)	**Mpumalanga** 4.9 (0.363)	**Limpopo** 10.5 (0.418)
	Location			**Rural**	**Urban**							
		20.1 (0.097)	79.9 (0.385)	31.9 (0.327)	48.0 (0.436)							
	Main income source			**Diversifying**	**Core sectors**	**Mining & Quarrying**	**Primary sectors**	**Domestic services**	**Public transfers**	**Private transfers**	**Indirect income**	
		23.7 (0.114)	76.3 (0.368)	19.9 (0.406)	33.1 (0.414)	1.3 (0.322)	1.6 (0.232)	2.2 (0.289)	6.8 (0.209)	5.5 (0.292)	6.0 (0.674)	
	Education			**None**	**Some primary**	**Some secondary**	**Complete secondary**	**Post secondary**				
		27.2 (0.131)	72.8 (0.351)	19.8 (0.385)	32.9 (0.340)	11.2 (0.337)	5.6 (0.348)	3.3 (0.328)				

Note: n = 90 298.
Source LFS/IES2000, own computations, weighted figures.

Table A6) Inequality in the 2000 sample as measured and decomposed by the Theil-T index; partitions defined by one characteristic data.

Total index value	Partition	Between-group inequality	Total within-group	Within-group contributions and absolute levels of inequality								
0.533	**Gender**			**Female**	**Male**							
		5.0 (0.027)	95.0 (0.506)	28.5. (0.454)	66.5 (0.496)							
	Race			**African**	**Coloured**							
		5.1 (0.027)	94.9 (0.505)	83.0 (0.534)	11.9 (0.367)							
	Age			**≤25**	**26-40**	**41-59**	**60≤**					
		6.5 (0.035)	93.5 (0.498)	3.2 (0.416)	35.4 (0.499)	42.6 (0.545)	12.3 (0.397)					
	Province			**W Cape**	**E Cape**	**N Cape**	**Free State**	**KZN**	**NW Prov**	**Gauteng**	**Mpumalanga**	**Limpopo**
		12.5 (0.066)	87.5 (0.466)	9.9 (0.360)	10.8 (0.543)	1.3 (0.488)	4.6 (0.537)	13.4 (0.484)	8.5 (0.670)	26.8 (0.432)	4.6 (0.404)	7.7 (0.519)
	Location			**Rural**	**Urban**							
		16.7 (0.089)	83.3 (0.444)	17.9 (0.179)	65.4 (0.654)							
	Main income source			**Diversifying**	**Core sectors**	**Mining & Quarrying**	**Primary sectors**	**Domestic services**	**Public transfers**	**Private transfers**	**Indirect income**	
		19.6 (0.104)	80.4 (0.429)	19.7 (0.526)	45.4 (0.417)	3.1 (0.502)	0.8 (0.226)	1.1 (0.278)	3.1 (0.241)	2.6 (0.321)	4.6 (0.692)	
	Education			**None**	**Some primary**	**Some secondary**	**Complete secondary**	**Post secondary**				
		29.6 (0.158)	70.4 (0.375)	13.0 (0.487)	25.3 (0.386)	12.1 (0.338)	10.8 (0.371)	9.1 (0.301)				

Note: N = 90 298.
Source: LFS/IES2000, own computations, weighted figures.

The Centre for Social Science Research
Working Paper Series

RECENT TITLES

du Toit, A. 2005. *Forgotten by the highway: Globalisation, adverse incorporation and chronic poverty in a commercial farming district.* Cape Town. CSSR Working Paper No. 101.

Perkins, P. Fedderke, J. Luiz, J. 2005. *An Analysis of Economic Infrastructure Investment in South Africa.* Cape Town. CSSR Working Paper No. 102.

Fedderke, J. Perkins, P. Luiz, J. 2005. *Infrastructural Investment in Long-run Economic Growth: South Africa 1875-2001.* Cape Town. CSSR Working Paper No. 103.

Seekings, J. 2005. *Prospects for Basic Income in Developing Countries: A Comparative Analysis of Welfare Regimes in the South.* Cape Town. CSSR Working Paper No. 104.

Fedderke, J. Szalontai, G. 2005. *Industry Concentration in South African Manufacturing Industry: Trends and Consequences, 1972-96.* Cape Town. CSSR Working Paper No. 105.

Ardington, C. Lam, D. Leibbrandt, M. Welch, M. 2005. *The Sensitivity of Estimates of Post-Apartheid Changes in South African Poverty and Inequality to key Data Imputations.* Cape Town. CSSR Working Paper No. 106.

The Centre for Social Science Research

The CSSR is an umbrella organisation comprising five units:

The Aids and Society Research Unit (ASRU) supports quantitative and qualitative research into the social and economic impact of the HIV pandemic in Southern Africa. Focus areas include: the economics of reducing mother to child transmission of HIV, the impact of HIV on firms and households; and psychological aspects of HIV infection and prevention. ASRU operates an outreach programme in Khayelitsha (the Memory Box Project) which provides training and counselling for HIV positive people

The Data First Resource Unit ('Data First') provides training and resources for research. Its main functions are: 1) to provide access to digital data resources and specialised published material; 2) to facilitate the collection, exchange and use of data sets on a collaborative basis; 3) to provide basic and advanced training in data analysis; 4) the ongoing development of a web site to disseminate data and research output.

The Democracy in Africa Research Unit (DARU) supports students and scholars who conduct systematic research in the following three areas: 1) public opinion and political culture in Africa and its role in democratisation and consolidation; 2) elections and voting in Africa; and 3) the impact of the HIV/AIDS pandemic on democratisation in Southern Africa. DARU has developed close working relationships with projects such as the Afrobarometer (a cross national survey of public opinion in fifteen African countries), the Comparative National Elections Project, and the Health Economics and AIDS Research Unit at the University of Natal.

The Social Surveys Unit (SSU) promotes critical analysis of the methodology, ethics and results of South African social science research. One core activity is the Cape Area Panel Study of young adults in Cape Town. This study follows 4800 young people as they move from school into the labour market and adulthood. The SSU is also planning a survey for 2004 on aspects of social capital, crime, and attitudes toward inequality.

The Southern Africa Labour and Development Research Unit (SALDRU) was established in 1975 as part of the School of Economics and joined the CSSR in 2002. SALDRU conducted the first national household survey in 1993 (the Project for Statistics on Living Standards and Development). More recently, SALDRU ran the Langeberg Integrated Family survey (1999) and the Khayelitsha/Mitchell's Plain Survey (2000). Current projects include research on public works programmes, poverty and inequality.