A Long Walk to Freedom and Toward Greater Inequality? A Household Level Review of Inequality in Post-Apartheid South Africa

June 2017

Dr. Robert Mullings

Economics Division
Newton Building
Nottingham Trent University
Nottingham
United Kingdom
NG1 4BU
robertmullings@ntu.ac.uk

Abstract

South Africa is still regarded as an unequal society more than two decades after the fall of apartheid. However, most studies exploring the dynamics and dimensions of the size distribution of income in post – transition South Africa use *either* individual *or* household level data. This paper shows that by committing exclusively to either level of analysis, key factors possibly driving the inequality could potentially be de-emphasized or obscured. This study represents a novel contribution to the literature since it provides a detailed analysis of South African Income and Expenditure Surveys (1995, 2000, 2005 and 2010) at *both* levels of aggregation so that a fuller picture of the dynamics of South African income inequality is achieved. Regression and decomposition analyses are combined at both levels; revealing rising inequality in South Africa, mostly benefitting non-black South Africans and the upper decile of the income distribution at the expense of the lower deciles and younger individuals. Starker trends are revealed from an analysis of the household level dataset, which reveals that jointly: gender, age, skill and educational attainment *structure* of the household also causally and significantly contribute to its location within the South African income distribution. Declining explanatory power is observed in both individual- and household level regressions, emphasising the need to incorporate un-modelled factors such as wealth, networks, historical and other institutional factors squarely within the discussion; as factors explaining inequality. This study argues for deep, consensus-based, economic, and political reforms in order to avert the plethora of problems a failure to address such deep inequality portends.

1. Introduction

Since the fall of apartheid, South Africa has not recovered from its reputation as a country of stark income and wealth inequality. This is remarkable given the watershed events of the 1990's which included the release of Nelson Mandela on February 11, 1990 from 27 years of incarceration and the subsequent relaxation of apartheid laws. For many after such a dramatic end to the apartheid system, a new dawn in the development of the economy and nation had begun providing an opportunity to ameliorate the problems of inequality and poverty; by establishing human rights, political and civil liberties, and other key institutions which have featured so heavily within the modern growth and economic development literature (North; 1981, 1990) as key ingredients for meaningful and sustainable economic development.

Since then, South Africa – the largest African economy- has undoubtedly experienced a measure of economic progress, with real GDP per capita increasing by 86% (Word Development Indictors, 2017)[1] over the review period (1995 to 2010) and average, annual levels of nominal economic growth of 3.24% per annum (World Development Indicators). Despite these achievements, a closer examination of alternative, multidimensional development indicators suggest far more modest gains in standard of living, with the country remaining within the medium human development range over the review period according to United Nations Human Development Index (HDI). This paper specifically focusses on the dynamics and sources of income inequality within South Africa over the post-transition period using 1995, 2000, 2005, 2010 editions of the Income and Expenditure Surveys since distributional issues are such a critical consideration to any accurate assessment of economic development[2].

Since most studies of income inequality, perform aggregated times series, individual level or household level analyses, the distinct contribution of this paper to the literature on South African poverty and inequality is that it: i) examines the effect of both individual and household characteristics on the income distribution ii) also explores how the contributions of both household and individual level factors vary over time and space iii) synthesizes the literature on South African inequality post-transition iii) provides insights which may be useful to researchers and policy-makers alike in light of the literature review and empirical findings on income inequality.

---

[1] According to the World Development Indicators (WDI) dataset, real world GDP per head increased by approximately 75% over the same period (1995-2010).

[2] Hicks (1997) while commending the multidimensional nature of the HDI in assessing development and standard of living improvements also notes that whereas the HDI helps "to put people back at the centre of development" the complementary question "which people" is also germane to the discussion.

Insights gained from this analysis are not only of interest to South African policymakers but have also much broader, even global significance at a time when the challenging issue of global inequality is once again, squarely at the forefront of the agenda. Picketty's (2014) tome which highlights the scale of international inequality,[3] from an historical, developed country perspective. Picketty (2014) frames the problem as an outcome of the capitalistic economic systems and proposes confiscatory taxes on income and wealth as a part of the solution to redress the widening inequalities. Bourguignon (2015), Atkinson (2015) and Stiglitz (2012, 2015) juxtapose the contemporaneous currents of decreasing inter-national inequality and increasing intra-national inequality and emphasize the role of globalization in contributing to these outcomes[4]. An in depth study of the causal factors contributing to inequality, its incidence and dynamics in the South African context is a useful addition to the body of literature[5] and is a germane issue in both developed and developing country settings

Section 2 of the paper provides a brief literature review after which *Section 3* examines developments in the evolution of the income distribution at both the individual and household levels within South Africa over the post-apartheid period and discusses descriptive statistics of the dataset. *Section 4*, then details the methodologies used to explore the data after which the dataset used for the analysis is fully described in *Section 5*. Detailed results and discussion of the regression and decomposition analyses are addressed in *Section 6*. The conclusion follows in *Section 7* which suggests a general policy direction on the basis of the empirical findings, in which the solutions to the inequality problem may lie.

2. Literature Review

The literature on income and earnings inequality in South Africa is voluminous and employs a combination of survey and national accounting data. Although voluminous in nature, most studies analyse trend indicators, and may perform decompositions of the data however there is a dearth of studies performing, causal analysis and –even more rare- household level analysis of inequality. Studies that carry out causal analysis are rare and tend not to address the issue of income inequality (see for example, Duflo; 2000). This is unfortunate since, as will be made clear in this study, data at

---

[3] Interestingly the first sentence of the first Chapter of Picketty's famous book mentions South Africa. In particular, the "Mariana Massacre": a labour market -related (strike) incident which took place in South Africa in 2012, which involved the use of force by the Government and subsequently led to deaths.
[4] Atkinson (2015), and Stiglitz (2012, 2015) also highlight these recent trends in the distribution of income.
[5] Indeed Barack Obama, the former President of the United States has referred to inequality as the "defining issue of our time".

the household level, and in particular- data relating to the structure of the household- can help to better understand the structure and dynamics of income inequality.

The analysis of inequality trends from survey data is exemplified by Wittenberg (2016a,2016b) who utilized the Post-Apartheid Labour Market Series PALMS dataset (version 2.1)[6] and decomposition techniques Shorrocks (1982) to investigate income inequality for the period 1994 – 2012 and finds increased wage inequality is South Africa since 1998. Finding increased wages at both higher and lower segments of the wage distribution, Wittenberg (2016a and 2016b) primarily attributes the increased inequality to differences in educational attainment and increased intra-racial inequality over time, especially among Black South Africans. The finding of increased inequality in South Africa was corroborated by also reflected in expenditure data for South Africa data (Bhorat and Van der Westhuizen; 2009). Hoogeveen and Ozler (2004) also observe increased inequality with declining expenditures at the lower end of the income distribution.

Van den Berg, Louw and Yu (2004), Leibbrandt and Levinsohn (2011) and Leibbrandt, Levinsohn and McCrary (2010) all report a fall in real incomes in South Africa in the sub-periods 1993-2008 and 1995- 2000 despite a rise in real GDP per capita, suggesting a role for inequality. Liebbrandt *et al* (2011) suggests that the fall in real wages could not be attributed to changes in endowments (socio-demographic characteristics) of the respondents. According the authors, the twin phenomena of rising per capita incomes and increased intra-racial inequality among "Black" South Africans appeared to have stymied a reduction in inequality over the period. Leibbrandt, Finn and Woolard (2012) and Leibbrandt, Levinsohn and McCrary (2010) posit that the brunt of the 40% decline which they report was borne by youth and Non-white South Africans.

In a comparative study, Gradín (2014) found that observed characteristics account for a mere one quarter of the (equivalized) racial income differential in South Africa compared to 50% in USA and Brazil. Gradín (2014) also concluded that i) of the three countries, South Africa has the strongest degree of association between the absolute gap and household characteristics and ii) that educational characteristics appear to be the main factor driving the gap in South Africa. In light of the apparent importance of household characteristics in this study, is somewhat surprising that more studies have not emerged in an attempt to identify specific household characteristics which matter the most for income distribution. The present study attempts to fill this void in the literature.

---

[6] which is an amalgam of data from the October Household Surveys (OHS) and Labour Force Surveys (LFS).

Although the discussion of inequalities in South Africa has usually been centred primarily on the racial divisions (Liebbrandt *et* al, 2012), a thorough survey of the economic literature on South African inequality identifies additional sources of inequality. In summary the sources which have been identified are: i) a health care system which disadvantages Blacks and Coloureds relative to Whites and Asians (Kon and Lakan; 2008) ii) a bifurcated education system with greater resources available for certain groups (Bhorat and Oosthuizen, 2008; Spaull, 2013; Levhinson, 2011; Branson *et al*, 2011; Rattsø and Stokke,2013;) iii)assets, wealth and network effects which perpetuate the structural nature of the inequalities (Adato, Carter and May 2006; Liebbrandt, 2017) iv) remittances which can have an exacerbating effect on inequality (Liebbrandt *et al*; 2012) v) gender effects (Casale, 2012; Tanaka, 2014) vi) geographical location which perpetuate inequalities in South Africa due to the mining economy and since certain regions act as growth poles and others lag behind (Bosker and Krugell, 2008) vi) labour market effects which reinforce inequality such as labour market segmentation, union power, lack of job creation and low levels of human capital (Liebbrandt *et al*, 2012; Kingdon and Knight, 2007; Oosthuizen, 2012) vii) international trade and skill-biased technological change (Jenkins Singh, 2006; Dhumale, 2000) viii) government failure to adequately address the issue[7] (Liebbrandt et al 2012; Tanaka, 2014). On the other hand, Albertyn (2011) notes some progress on the legislative front to address inequalities within South Africa.

3. The Distribution of Income in South Africa 1995 – 2010
3a. General Trends in Size Distribution of Income

Figure 1 below contains four *Panels (A) – (D)* which provide a broad graphical overview of developments in the South African income distribution since 1995. *Panels (A)* and *(C)* of the *Figure* present kernel density estimates of logged real income values *(A)* and a barplot *(C)* of the share of real income attributable to each quintile (ordered 20th percentile) of the South African income distribution based on the real income (2010 values)[8] of individual respondents of the 1995, 2000, 2005, 2010 Income and Expenditure Surveys. On the other hand, *Panels (B)* and *(D)* depict analogous figures using household level data from the same surveys[9].

 In constructing *Figure 1* and indeed for the remainder of the analysis, respondents falling within the top and bottom 1% of the income distribution for each survey year were excluded from the sample in

---

[7] Liebbrandt (2012) concludes that the doubling of state grants over the period 1993 – 2008 has had a small aggregate effect on inequality
[8] The real individual income values were derived by deflating nominal values using CPI index values for South Africa which were sourced from the World Development Indicators (WDI) dataset with a base year of 2010.
[9] Notice that the for Panel (a) it logged real values of average income per household that is used as the variable of interest.

order to eliminate outliers which could distort results, contributing to relatively large variances with unrepresentative point estimates. The income estimates derived from the surveys refer to regular income after deductions (taxes and transfers). Income was calculated in the 1995, 2000 and 2005/06 IES as income net from tax and unemployment insurance fund contributions (UIF) (total income, including income from work, grants and other allowances)[10]. Lustig and Higgins (2012) confirm that for the 2010/11 IES that individuals reported net incomes, instead of gross income; therefore income is taken as net of taxes. Secondly, respondents aged 60 or above were removed from the individual level dataset in order to provide an accurate profile of inequality within the general population, as despite no formal retirement age in South Africa, individuals aged 60 or above are eligible for a state pension (Gov. South Africa, 2015). Thirdly, individuals below 20 were removed from the individual level sample in an attempt to minimize excessive disparities in labour market experience and human capital likely to exist within the lowest age groups. This is consistent with human capital theory which suggests human capital accumulation is dependent upon the expected and actual time spent in the labour market (see Altonji and Blank, 1999). These adjustments to the dataset also hold for the individual level estimation below.
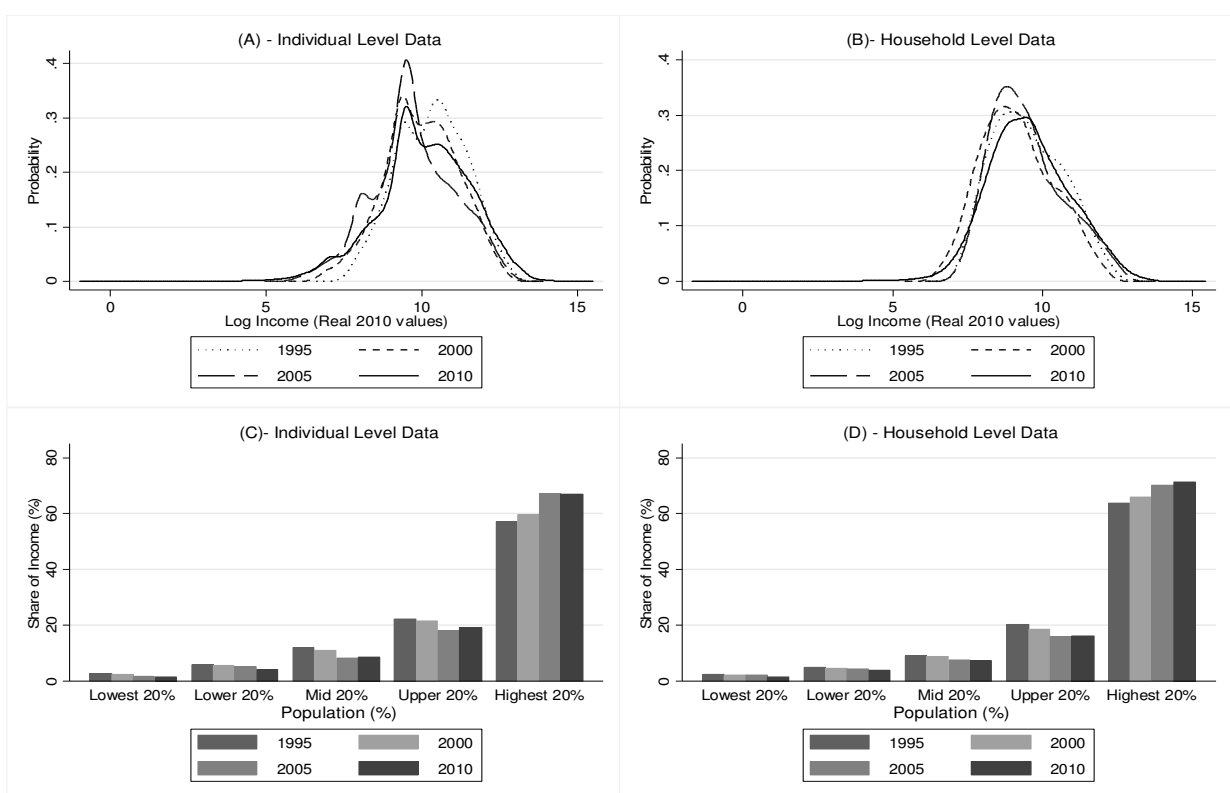

Figure 1: Income Distribution and Inequality in South Africa 1995, 2000, 2005, 2010.
Source: Author's computation for respective years of Income and Expenditure Survey.

---

[10] Yu (2008) presents a convenient, concise account of the components of regular income in the Income and Expenditure Surveys over time.

A comparison of the individual and household and level kernel density estimates of the income distributions in *Panels (A)* and *(B)* of the *Figure* reveals, quite unsurprisingly, that the household level density estimate is smoother in appearance and reveals that incomes the central mass of the distribution has shifted rightward, if only slightly, over the review period. *Panel A* reveals a noticeable variation in the (real) logged income distribution over time From the panel, a clear rightward shift in the distribution overtime is not apparent, implying that in general, real incomes in South Africa were not trending upwards over the review period. *Panels (C)* and *(D)* depict reveal the income distribution by quintile for the individual and household level datasets respectively Both figures paint a similar picture of the dynamics of the size distribution of size distribution of income in South Africa. According to the individual level data in *Panel (C)* the highest quintiles have consistently earned greater than 57 % of all income aver all survey years. The analogous threshold is, visibly, even higher (over 60%) when household level data is considered. By way of contrast, the same figures reveal that the lowest quintile whose share averages a mere 2% of income across all the survey years. Moreover, the *Panels* reveal, that the inequality worsened over the review period. By the end of the review period the lowest quintile accounted for a mere 1.48% of the income distribution (from household level data) while the highest quintile accounted for 71% of the income distribution. Furthermore the income progressively accruing to the top quintile appeared to have come at the expense of all other quintiles of the income distribution.
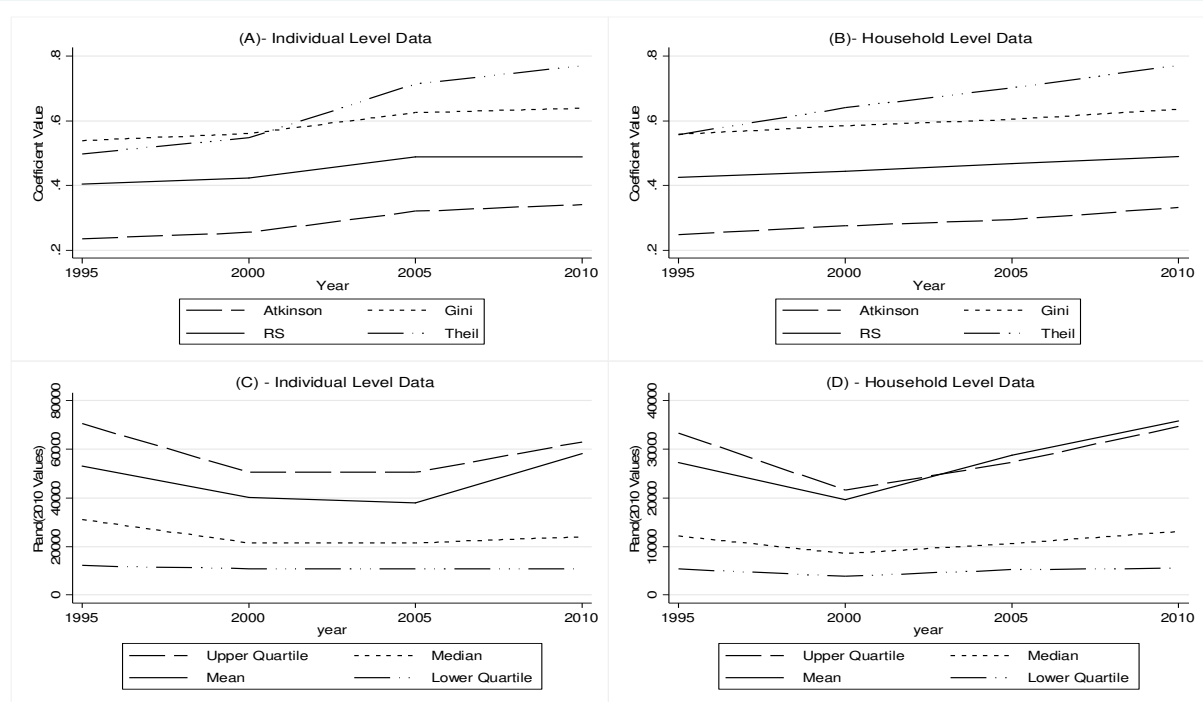

Figure 2: Inequality Summary Measures for Income Distribution and Inequality in South Africa 1995, 2000, 2005, 2010.
Source: Author's computation for respective years of Income and Expenditure Survey.

Panels (A) and *(B)* of *Figure 2* depict time series evolution of four common inequality measures: 1) the Gini coefficient 2) the Atkinson Index 3) the Theil Index and 4) the Ricci-Schutz coefficient which is also known as Pietra's or "Robin Hood" index for both individual- and household level distributions. The Gini coefficient, perhaps the most widely used measure of inequality, measures the deviation of the cumulative income function from the income equality line. The time series for the Gini coefficient in both *Panels (A)* and *(B)* suggests a general increase in inequality over time in South Africa; with inequality increasing at each 5 year interval up to 2005, before decreasing slightly to a value of 0.64 in 2010. The trend increase in the Gini is even more pronounced for household level data, as can be observed in *Panel (B)*. A similar pattern can be observed in the "Robin Hood" index which averages 0.46 for the entire period, implying that in real terms, and on average, almost half of the income of the higher-income South Africans would have had to be redistributed across the population during the review period to achieve perfect equality. The estimates are almost identical for the household level data in *Panel (B)*. The recurring theme of increasing inequality in South Africa over the review period is again self-evident. Both individual and household level data reveal increasing Atkinson indices over the period. Recall, that the Atkinson index measures the social utility to be gained by a complete redistribution of the income distribution also increased steadily over time. Examining the Theil index reveals a similar trend indicating that the entropic distance from an egalitarian distribution of incomes increased significantly between those two years.

Finally, *Panels (C)* and *(D)* of *Figure 2* depict time series of key measures of location and central tendency for both the individual-level and household income distributions. The y-axis of Panel (C) depicts individual income while the y-axis of Panel (D) depicts total household income. It is important to note that both *Panels* of the figure show that, in general median income remained flat over the period, after a slight but noticeable decline in 2000. The coincidence of increased inequality and stagnant real median incomes highlights the deteriorating circumstances of most South Africans relative to those individuals at the upper end of the distribution. The presence of high values in the rightmost tail of the South African income distribution is corroborated by the trend increase in the mean and the increasing positive skewness of the income distribution towards the latter survey years.

3b. Intra-Group Income Dynamics in South Africa

Tables 1 and *2* below provide information on intra-group, income dynamics and sample proportions for key socio-demographic groups at both the individual and household levels for IES years 1995, 2000, 2005 and 2010. The columns of the *Table* contain the variables and represent the respective survey years

Note that *Table 1* The *Table* allows for a more detailed understanding of the dynamics of income inequality within South Africa since it provides sample estimates of useful summary measures such as the median, mean and Gini coefficient of real (2010) income for sub-categories of the key socio-demographic groups, across all survey years. The topmost row of *Table 1* depicts column headings corresponding to the respective IES survey years. The rows of the Table correspond to variables of interest. Note the similarities and differences between the individual level variables in *Table 1* and household – level variables in *Table 2*. Variables capturing information regarding the household structure are introduced such as i) the size of the household and ii the proportion of female members within it. In every cell of the *Table 1* and *2* which contains statistics the (emboldened) share of the sample represented by the socio-demographic feature, median, mean and Gini coefficient for the sample sub-group is displayed.

.

Both Tables 1 and *2* reveal persistent and significant and persistent differences in the median real incomes of males and females across all survey years. Across all survey years, and at both levels of disaggregation males consistently enjoy higher incomes than females. This gender differential in median and mean incomes is greater towards the beginning and end of the review period and is especially pronounced for the household level sample represented in Table 2. Tables 1 and 3 also confirm the finding of Hoogeveen and Ozler (2004), Leibbrandt, Levinsohn and McCrary (2010) and Liebbrandt and Levhinsohn (2011) that there was indeed a fall in real post-transition income before a recovery towards the end of the sample. Closer examination of the Gini coefficient across both Tables reveals a rather nuanced picture. Overall intra-group inequality has worsened for South African males and male-headed households over the review period, however for male households the decline was not a consistent one but marked by a slight recovery in IES 2010.

The racial ordering of incomes within the South African income distribution of income since 1995 have been preserved across all survey years with Whites consistently earning more than respondents and households in the: "Other" (Indians, Asians), Coloureds and Black racial categories, in that order. It can be observed from *Table 2* that real incomes for White headed household have consistently increased over consecutive IES survey years. Moreover, while no other racial group in either *Table 1* or *Table 2* has consistently increased their median (or mean income, for that matter) across all periods incomes for it is clear that relative to 1995 IES all Non-black racial groups have improved their position. Tables 1 and 2 also reveal a divergent pattern in pairwise median income shares between racial groups. For example, whereas the ratio of median incomes of whites to Blacks in 1995

Table 1: Individual Level Shares and Central Tendency of Real Income (2010) and Gini Coefficients of By Sub – Group

Variable		IES 1995 (share)	IES 2000 (share)	IES 2005 (share)	IES 2010 (share)
Gender :	Male -median [mean] **(share)** (Gini)	36994.43 [62985.57] **(0.559)** (0.5332)	26464.34 [45408.64] **(0.551)** (0.5424)	20105.06 [45842.61] **(0.487)** (0.6012)	31200.00 [68255.45] **(0.509)** (0.6309)
	Female	24554.35 [40336.46] **(0.447)** (0.5192)	15804.41 [33656.81] **(0.449)** (0.5751)	13068.29 [30241.9] **(0.513)** (0.6363)	18000.00 [47867.75] **(0.49)** (0.6371)
Race:	"black"	23547.29 [35008.53] **(0.618)** (0.5051)	17163.16 [30374] **(0.74)** (0.5375)	13168.82 [27585.28] **(0.724)** (0.6052)	18000.00 [41989.17] **(0.731)** (0.622)
	White	102680 [120550.3] **(0.177)** (0.3723)	96542.77 [111290.6] **(0.089)** (0.3692)	90752.02 [115365.3] **(0.09)** (0.4346)	114000.00 [164704.0] **(0.097)** (0.4848)
	Coloured	26282.14 [38748.77] **(0.163)** (0.485)	23599.34 [38745.86] **(0.1406)** (0.5146)	16754.22 [35507.8] **(0.165)** (0.5689)	32400.00 [60758.59] **(0.147**) (0.5903)
	Other	68124.25 [87263.31] **(0.043)** (0.4113)	54350 [74947.14] **(0.0308)** (0.4539)	50262.66 [77343.25] **(0.021)** (0.5188)	72000.00 [105725.1] **(0.025)** (0.5173)
Age:	20 -29	28138.28 [42377.64] **(0.218)** (0.4947)	15203.7 [29007.53] **(0.2347)** (0.5494)	13096.21 [25285.92] **(0.243)** (0.6029)	21600.00 [39993.52] **(0.232)** (0.5847)
	30 -39	35543.09 [57012.29] **(0.337)** (0.5229)	25565.95 [44031.7] **(0.3366)** (0.5741)	16754.22 [41735.2] **(0.291)** (0.6189)	26400.00 [60028.55] **(0.294)** (0.5822)
	40 -49	35528.28 [59572.29] **(0.263)** (0.5702)	25744.74 [46036.03] **(0.2639)** (0.5925)	16754.22 [43987.49] **(0.268)** (0.6483)	27600.00 [68987.18] **(0.271)** (0.6247)
	50 - 59	23695.39 [48760.19] **(0.182)** (0.5704)	17733.48 [38547.11] **(0.1647)** (0.5866)	13291.68 [39183.57] **(0.198)** (0.6559)	18000.00 [62171.73] **(0.204)** (0.624)
Region:	Western Cape	33854.79 [56655.73] **(0.135)** (0.5296)	27890.13 [46210.55] **(0.1298)** (0.5256)	21780.48 [45851.67] **(0.144)** (0.5713)	36000.00 [75985.24] **(0.159)** (0.6029)
	Eastern Cape	20733.47 [40965.5] **(0.156)** (0.5586)	15017.76 [35624.28] **(0.1108)** (0.5904)	13068.29 [33822.48] **(0.123)** (0.6524)	15000.00 [47749.4] **(0.112)** (0.6640)
	Northern Cape	20733.47 [42663.84] **(0.047)** (0.5809)	17163.16 [41595.16] **(0.053)** (0.5860)	13794.31 [38572.27] **(0.073)** (0.6254)	22080.00 [58472.1] **(0.051)** (0.6295)
	Free State	15698.2 [39231.86] **(0.105)** (0.6063)	17163.16 [34741.04] **(0.0831)** (0.5749)	13179.99 [37097.92] **(0.082)** (0.7072)	15600.00 [45012.65] **(0.091)** (0.6583)
	Kwazulu/Natal	34728.56 [51675.56] **(0.183)** (0.5417)	18235.86 [37463.11] **(0.1679)** (0.6442)	13068.29 [28935.68] **(0.215)** (0.656)	21575.00 [50401.46] **(0.147)** (0.6259)
	North West	29619.24 [49672.73] **(0.078)** (0.6495)	25029.61 [39143.45] **(0.0994)** (0.6498)	13235.83 [32659.69]**(0.123)** (0.6439)	19200.00 [53569.35] **(0.079)** (0.6669)
	Gauteng	59238.48 [82370.48] **(0.135)** (0.4483)	32180.92 [50790.83] **(0.1648)** (0.506)	33438.63 [59378.54] **(0.123)** (0.5705)	36000.00 [74289.67] **(0.173)** (0.5956)
	Mpumalanga	29598.62 [51614.81] **(0.087)** (0.5416)	17163.16 [34061.16] **(0.0878)** (0.5762)	15078.80 [35441.4] **(0.078)** (0.6164)	24000.00 [55964.61] **(0.087)** (0.6346)
	Northern Province (Limpopo)	29619.24 [52381.93] **(0.075)** (0.5295)	13945.07 [34344.75] **(0.103)** (0.6054)	12733.21 [28200] **(0.079)** (0.6313)	13523.50 [43375.39] **(0.101)** (0.6765)

Source: Income and Expenditure Surveys 1995, 2000, 2005, 2010. Authors own calculations.

Note: the upper and lower 1% of the distribution of each survey year has been excluded from the estimates in order to remove outliers and for the sake of greater accuracy of the point estimates.

Table 2: Household Level: Shares and Central Tendency of Real Income (2010) and Gini Coefficients of By Sub – Group

Variable	IES 1995 (share)	IES 2000 (share)	IES 2005 (share)	IES 2010 (share)
Gender : Male hhead –Median [mean] **(share)** (Gini)	54440.16 [1118420.5] **(0.6904)** (0.5995)	34326.32 [77354.29] **(0.6054)** (0.6239)	44273.72 [118952.8] **(0.552)** (0.655)	50460.5 [124216.7] **(0.564)** (0.6309)
Female -hhead	29619.24 [53015.57] **(0.3096)** (0.5229)	18593.42 [36724.35] **(0.3946)** (0.5579)	26706.23 [53908.05] **(0.448)** (0.567)	29081 [62506.59] **(0.436)** (0.592)
Race of Head: Black **hhead**	31347.03 [56233.69] **(0.6522)** (0.531)	21834.76 [39265.2] **(0.794)** (0.545)	28109.39 [50314.17] **(0.757)** (0.5349)	31601 [65740.55] **(0.593)** (0.793)
White hhead	169080.2 [245831.5] **(0.1854)** (0.4943)	173419.4 [248541] **(0.08)** (0.483)	234476.7 [361749.3] **(0.105)** (0.503)	255722.00 [341578.9] **(0.081)** (0.449)
Coloured hhead	46398.54 [74559.64] **(0.1273)** (0.491)	40118.88 [68282.83] **(0.1036)** (0.5249)	45313.88 [85747.49] **(0.122)** (0.548)	65871.00 [123346] **(0.108**) (0.561)
Other -hhead	118462.1 [183186] **(0.0351)** (0.484)	102979 [142884.9] **(0.0217)** (0.476)	129594.6 [209620.5] **(0.016)** (0.539)	157318.00 [228492.7] **(0.018)** (0.483)
Age hh head: under 20	17425.99 [28770.41] **(0.0046)** (0.4841)	10369.41 [14695.62] **(0.014)** (0.474)	11566 [19488.37] **(0.013)** (0.5267)	7665.00 [15546.09] **(0.01)** (0.604)
20 to 29	38505.1 [79258.22] **(0.1005)** (0.562)	20381.25 [42699.82] **(0.125)** (0.595)	23023.79 [52671.06] **(0.106)** (0.621)	26474.50 [60900.4] **(0.119)** (0.634)
30 -39	53561.46 [113071.5] **(0.2376)** (0.5901)	31169.01 [68521.82] **(0.242)** (0.615)	35524.53 [93824.28] **(0.203)** (0.646)	39460 [100106.50] **(0.20)** (0.643)
40 - 49	53610.82 [119393.6] **(0.2327)** (0.6028)	32896.05 [73930.09] **(0.234)** (0.628)	40372.78 [108258] **(0.226)** (0.65)	48909.50 [118772.80] **(0.228)** (0.63)
50 - 59	47341.42 [107592.7] **(0.1747)** (0.6215)	28247.7 [67463.78] **(0.1633)** (0.641)	37861.74 [118790.2] **(0.179)** (0.686)	47866.00 [118829.70] **(0.198)** (0.641)
Over 60s (Incl.)	32798.37 [66538.59] **(0.2498)** (0.5686)	23170.26 [49038.23] **(0.223)** (0.585)	33392.55 [70454.35] **(0.272)** (0.57)	36823.50 [78699.54] **(0.245)** (0.578)
Structure: Mostly Female (Yes/No)	39092.46 [83596.07] **(0.3953)** (0.5885)	23170.26 [51955.82] **(0.4014)** (0.606)	30229.5 [76428.49] **(0.191)** (0.64)	37097.5 [90889.92] **(0.426)** (0.63)
Mostly Male (Yes/No)	44098.11 [98019.42] **(0.3411)** (0.6081)	28256.64 [59921.06] **(0.3771)** (0.6083)	35148.96 [89757.75] **(0.674)** (0.641)	38148.5 [91572.33] **(0.355)** (0.636)
Size: House size <= 2	38998.66 [94297.03] **(0.2633)** (0.6267)	22321.05 [52532.5] **(0.355)** (0.633)	28689.51 [83184.34] **(0.653)** (0.66)	27887.00 [74624.67] **(0.34)** (0.6583)
House size 3- 5	49502.39 [114340.3])**(0.454)** (0.6073)	28299.55 [74599.65] **(0.4065)** (0.638)	38975.9 [99142.36] **(0.3052)** (0.625)	44012.50 [117870.6] **(0.442)** (0.639)
House Size > 5	41466.93 [75821.16] **(0.283)** (0.542)	27328.75 [51184.04] **(0.239)** (0.550)	56683.01 [125891.2] **(0.042)** (0.591)	46739.5 [91006.79] **(0.218)** (0.56)
Region: Western Cape	69605.21 [130739.6] **(0.109)** (0.558)	48271.38 [93800.38] **(0.1298)** (0.5561)	65750.54 [167085.6] **(0.114)** (0.637)	79014.5 [157538.8] **(0.117)** (0.573)
Eastern Cape	26780.73 [66275.56] **(0.18)** (0.6144)	17145.28 [44653.39] **(0.1323)** (0.638)	27128.57 [68667.02] **(0.135)** (0.6341)	28298.00 [74262.24] **(0.132)** (0.648)
Northern Cape	35543.09 [81279.21] **(0.048)** (0.6061)	25204.81 [74561.53] **(0.05)** (0.6633)	34422.94 [80462.18] **(0.081)** (0.61)	45800.00 [100267.50] **(0.048)** (0.60)
Free State	26953.51 [63409.19] **(0.111)** (0.5943)	21789.17 [58867.41] **(0.0886)** (0.6786)	32515.05 [87148.56] **(0.088)** (0.0.651)	32953.00 [79370.98] **(0.086)** (0.625)
Kwazulu/Natal	52416.18 [101275.5] **(0.175)** (0.566)	23170.26 [54793.85] **(0.1685)** (0.6167)	28233.65 [61805.79] **(0.219)** (0.0.599)	36013.5 [85053.99] **(0.143)** (0.6263)
North West	36680.96 [98759.4] **(0.083)** (0.6634)	27800.74 [55363.6] **(0.107)** (0.596)	31852.56 [72107.82] **(0.073)** (0.603)	31944.5 [76695.81] **(0.10)** (0.637)
Gauteng	102420.9 [169436.1] **(0.118)** (0.5377)	40762.2 [84079.19] **(0.1508)** (0.596)	56794.01 [148970.9] **(0.118)** (0.6503)	61919.00 [139832.50] **(0.153)** (0.5956)
Mpumalanga	44552.27 [82458.18] **(0.086)** (0.5386)	24478.95 [49425.03] **(0.0869)** (0.568)	31447.67 [75461.47] **(0.08)** (0.64)	36841.5 [88830.28] **(0.091)** (0.637)
Northern Province (Limpopo)	36727.86 [89375.92] **(0.09)** (0.621)	17878.29 [43782.42] **(0.118)** (0.624)	27926.49 [53819.53] **(0.093)** (0.554)	27454.00 [62201.5] **(0.131)** (0.620)

Source: Income and Expenditure Surveys 1995, 2000, 2005, 2010. Authors own calculations.

Note: the upper and lower 1% of the distribution of each survey year has been excluded from the estimates in order to remove outliers and for the sake of greater accuracy of the point estimates.

was around 434% times, by 2010 the figure had grown to 633%. Coloureds and other races have also strengthened median value of incomes relative to Black South Africans. Intra-racial inequality, as measured by the Gini coefficient has also been increasing across all racial categories but especially among Blacks in both individual and household level dataset but is more nuanced in the household level dataset; decreasing within the White racial category, remaining constant in the Other category and increasing among Coloureds.

Examining the age distribution of income in *Table 1* and that of the household head in *Table 2* reveals that, consistent with human capital theory models of Ben-Porath (1967), Heckman (1976) and Weiss (1986), incomes appear to initially increase with tenure and experience across all survey years before declining in later years. The decline in later years is made even more stark in *Table 2* which includes summary statistics for household heads over 60 years of age. For most survey years the peak mean and median values are consistently realized within the 40-49 age group in South Africa. The *Table* also reveals that inequality, as measured by the Gini coefficient, increased across all survey years until 2005 with the subsequent survey year IES 2010 indicating a slight reversal in trend. The increased inequality among households headed by individuals younger than 20 years of age is also remarkable. *Table 1* also reflects homogeneity in the inequality dynamics between the two age categories – 40-49 and 50-59. *Tables 1* and *2* reveal that real median incomes for individuals and household-heads under 39 years have fallen significantly over the period.

Table 1 also reveals that intra-regional inequality within South Africa has risen relative to 1995 levels as measured by the Gini coefficient in 5 of the 9 administrative divisions (Western Cape, Northern Cape, North West, Gauteng and Northern Province). On the other hand intra-regional inequality derived from the household level data in *Table 2* reveals a more stable pattern with Kwazulu /Natal and Mpumalanga showing largest increases in inequality. Another interesting trend emerging from *Table 2* is that for provinces such as Kwazulu/Natal, Northwest, Gauteng, Mpumalanga and Northern Province median household income actually decreased between 1995 and 2010. It is also apparent that from examining statistics from *Table 1* that provinces such as Gauteng and Western Cape household inequality peaked in 2005. By 2010 the highest levels of inequality are observed in Western Cape, Gauteng, Kwazulu/Natal and Eastern Cape administrative regions. The theme of significant disparities in both median and mean incomes between provinces is shared across both the individual and household level data. In particular, administrative provinces such as Western Cape and Gauteng generally tended to realize relatively higher real mean and median incomes perhaps due to larger concentration of commercial and industrial activities in these regions. These disparities are

therefore likely to play a role in the inequality dynamics in South Africa (see Todaro and Harris, 1970; Bosker and Krugell, 2008).

4. Methodology

4.1 Regression Analysis

The empirical strategy adopted within the paper is multi-pronged. The first prong of the strategy involves the estimation of a series of linear regression models of real income on key socio-demographic statistics of South Africans IES survey years: 1995, 2000, 2005, and 2010. The second prong involves the application of decomposition analysis which explores the relative importance of individual and household characteristics in contributing towards income inequality – in particular racial income inequality-   at both household and individual levels over the review period. The decomposition analysis is treated in greater detail in *Section 4.2* which follows.

The naïve forms of the regression equation employed to conduct the empirical analysis is shown in Equations (1a) and (1b) below:

$$\log(Income)_i = \beta_o + \beta_1 gender_i + \beta_2 head_i + \sum_{j=3}^{5} \beta_i race_{j-2,i} + \sum_{j=6}^{8} \beta_j age_{j-5.,i} + \sum_{j=9}^{16} \beta_j geo_{j-8,i} + u_i \qquad (1a)$$

$$\log(Income)_i = \beta_o + \beta_1 gender\_head_i + \sum_{j=3}^{5} \beta_j race\_head_{j-2,i} + \sum_{j=6}^{10} \beta_j age\_head_{j-5,i} + \beta_{11} prop\_females_i$$

$$+ \sum_{j=12}^{18} \beta_j no\_within\_agegroup_{j-11,i} + \sum_{j=19}^{27} \beta_j geo_{j-18,i} \qquad (1b)$$

Equation (1a) is estimated on the individual level dataset. In Equation (1a), the dependent variable $\log(Income)_i$ is the log of the real value of income for each respondent within each survey year of interest. On the other hand, in Equation (1b) the dependent variable $\log(Income)_i$ represents the logged real value of average income per household[11]. Across both specifications, as in the previous

---

[11] Average income per household is simply: total household income divided by the number of persons within the household.

figures and tables, real income values were derived at by deflating nominal income values using CPI index values from the World Development Indicators dataset with 2010 as the base year.

The variables *gender* and *head* in Equation (1a) are binary indicators which take a value of one for males and household headship respectively while *age* is a similarly constructed set of binary indicators which indicate whether the respondent is within the *20-29,40-49* or *50-59* age groups. Note the omission of the *30-39* binary indicator which serves as the base category. Continuing with Equation (1a), the set of race variables ($race_{1i}$ - $race_{3i}$) represent binary indicators capturing information on the race of each survey respondent. The set of race indicators represents three racial categories: "Coloured", "Indian/Asian/Other", and "White" respectively. Note that the "Black" racial category is used as the reference category and is thus not explicitly included among the regressors. South Africa's administrative regions are also coded as set of binary indicator variables with Western Cape used as the base category. For this reason, variables geo_1 to geo_8 represent, in order: Eastern Cape, Northern Cape, Free State, Kwazulu/Natal, North West, Gauteng, Mpumalanga and Northern Province. Finally, note that Equation (1a) is estimated in turn for each Income and Expenditure Survey (1995, 2000, 2005 and 2010). The application of identical regression specifications across survey years makes it convenient to make intertemporal comparisons of the magnitude of the signs, marginal effects along and statistical significance of coefficients as captured by the key socio-economic variables included within the model.

Equation (1b) represents the naïve form of household level regressions which enable the effect of household level information such as household structure to be analysed. The variables: *gender_head, race_head* and *age_head* are binary indicator variables which capture key demographic characteristics of the head of household: gender, race and age respectively. Whereas the coding of the gender and race variables are identical to that in Equation (1a) the age of the household head is coded differently. The variable *age_head* is a set of 5 binary indicator variables which capture the age

groups: *under 20*, *20-29*, *40-49*, *50-59* and *over 60* with the *30-39* age-group serving as the base category. Equation (1b) also includes the variable *prop_females* which captures the proportion of females within each household while the variable set *no_within_age_group* captures simple counts of household members within the: *under 10*, *10-19*, *20-29*, *30-39*, *40-49*, *50-59* and *60 and over* age groups (inclusive) for each household. Finally, the variable *geo* is a set of binary indicator variables coded identically to those in Equation (1a) capturing information on the geographic location of each household.

Equations (2a) and (2b) below, though broadly similar to the formulation in *Equation (1)*, additionally allows for the inclusion of additional variables containing information regarding the skill levels of the respondents. Note the addition of the variables $skill_1$ to $skill_3$ in both formulations below, allowing for the marginal contribution of individual skill levels on expected log of real income to be ascertained, holding all other factors fixed.

$$\log(Income)_i = \beta_o + \beta_1 gender_i + \beta_2 head_i + \sum_{j=3}^{5}\beta_{j-2}race_{j-2,i} + \sum_{j=6}^{8}\beta_{j-5}age_{j-5,i} + \sum_{j=9}^{11}\beta_{j-8}skill_{j-8,i} + \sum_{j=12}^{19}\beta_{j-12}geo_{j-11,i} + u_i \quad (2a)$$

$$\log(Income)_i = \beta_o + \beta_1 gender\_head_i + \sum_{j=3}^{5}\beta_j race\_head_{j-2,i} + \sum_{j=6}^{10}\beta_j age\_head_{j-5,i} + \beta_{11}prop\_females_i$$
$$+ \sum_{j=12}^{15}\beta_j skill_{j-11,i} + \sum_{j=16}^{22}\beta_j no\_within\_agegroup_{j-15,i} + \sum_{j=23}^{30}\beta_j geo_{j-22,i} \quad (2b)$$

In the equation estimated on the individual level data, Equation (2a) the set of *skill* variables are coded as binary indicator variables indicating whether the individual respondent is *skilled* (technicians and associate professionals), *semi- skilled* respondents (service workers, skilled - agricultural and fishery, craft and trades, plant and machinery) or *low - skilled* workers (elementary labourers).The categorization of respondents into various categories was done on the basis of the 2000 South African Standard Classification of Occupations (SASCO) produced by Statistics South

Africa. In particular, variables related to skill levels included within the model are: *highly skilled* (professionals and managers). Politicians and senior civil servants were also included within the *highly skilled* category for the purposes of the analysis conducted here[12]. For the household level equations, the specification changes slightly - the proportion of semi-skilled, skilled and highly-skilled individuals within a given household is captured by the skill variables.

Since the skill-related variables are only available for the IES 2000, Equations (2a) and (2b) can only be feasibly estimated using IES 2000 data. Estimation of *equation (2)* is not feasible for the 1995 edition of the IES survey, for example, since it contains poorly recorded data regarding the professional occupations and skill levels of the respondents (occupational data for only 17 respondents were provided (or made available) in the entire dataset). On the other hand, the 2005 and 2010 editions of the IES do not elicit information on the professional occupations of respondents at all.

$$\log(Income)_i = \beta_o + \beta_1 gender_i + \beta_2 head_i + \sum_{j=3}^{5}\beta_{j-2}race_{j-2,i} + \sum_{j=6}^{8}\beta_{j-5}age_{j-5,i} + \sum_{j=9}^{16}\beta_{j-8}geo_{j-8,i} + \sum_{j=17}^{19}\beta_{j-16}edu_{j-16,i} + u_i \quad (3a)$$

$$\log(Income)_i = \beta_o + \beta_1 gender\_head_i + \sum_{j=3}^{5}\beta_j race\_head_{j-2,i} + \sum_{j=6}^{10}\beta_j age\_head_{j-5,i} + \beta_{11}prop\_females_i$$

$$+ \sum_{j=12}^{18}\beta_j no\_within\_agegroup_{j-11,i} + \sum_{j=19}^{j=21}\beta_j edu_{j-18,i} + \sum_{j=22}^{j=29}\beta_j geo_{j-21,i} + u_i \quad (3b)$$

Equations (3a) *and* Equations (3b) above add yet another dimension to the analysis. The inclusion of information regarding the highest level of educational attainment of survey respondents (for the individual level data) and the corresponding household proportions (for the household level dataset) allows for the effects of these variables on expected, real incomes over survey years 2005 and 2010 to be examined. In particular, Equation (3a) and Equation (3b) introduces the binary indicator

---

[12] Interestingly SASCO does not attribute a skill-level for politicians and senior civil servants. However for the sake of completeness, and hopefully not controversially, politicians and senior civil servants are included within the highly-skilled category.

variables- edu_1 to edu_3 which capture the highest level of educational attainment of each. Unfortunately, information on the highest level of educational attainment was not elicited from survey respondents before IES 2005. This implies that *equation (3)* can only be estimated for the 2005 and 2010 editions of the IES survey. In preparation for estimation, each respondent to the surveys is classified into categories reflecting their highest levels of educational attainment. The categories are: 1) no schooling, 2) primary 3) secondary diploma and national technical certificate (I, II and III) and finally 4) degree and higher - which includes individuals which have attained bachelor's degree, bachelor's degree and diploma, honours degree, higher degree (masters and doctorate and other higher degrees). The base class for the set of educational attainment variables is the "none" category which corresponds to "no schooling". All regression equations are estimated using ordinary least squares with robust standard errors.

4.2 Blinder - Oaxaca Decomposition analysis.

The decomposition analysis is carried out using the Blinder – Oaxaca method. The Blinder – Oaxaca technique was developed by Blinder (1973) and Oaxaca (1973). As noted by Jann (2008), more traditional applications of the methodology have tended to focus on decomposing mean differences in log wages on linear regression models by gender (Stanley and Jarrell 1998; Weichselbaumer and Winter-Ebmer, 2005) and race (e.g., Darity, Guilkey, and Winfrey,1996; Kim 2010). However, the Blinder- Oaxaca method has more recently been used to explore regional disparities in consumption in China (Bin et al; 2016). Non-linear variants of the model have also been applied to empirical analyses of migrant behaviour (Aristei; 2013). In general, there are two versions of the Blinder –Oaxaca decomposition: i) three-fold decomposition and ii) two-fold decomposition. The two-fold approach divides the (log) wage differential into an explained part which is based on group differences and an unexplained part which is attributable to discrimination but which, as Jann (2008) cautions, can also be due to the potential effect of unobserved variables. On the other hand the threefold decomposition,

which is applied to the South African IES data, divides the differential in the outcome variable (R)[13] into differences attributable to endowments (E), coefficients (C) and interactions (I) elaborated below.

The design matrices of explanatory variables implied by Equation (1a) and Equation (1b) take the following form respectively:

$$X_f = \begin{bmatrix} 1 & gender & head & age & geo \\ n_a \times 1 & n_a \times 1 & n_a \times 1 & n_a \times 3 & n_a \times 8 \end{bmatrix} \tag{4a}$$

$$X_f = \begin{bmatrix} 1 & gender\_head & race\_head & age\_head & prop\_females & no\_within\_agegroup & geo \\ n_b \times 1 & n_b \times 1 & n_b \times 3 & n_b \times 5 & n_b \times 1 & n_b \times 7 & n_b \times 8 \end{bmatrix} \tag{4b}$$

In expressions (4a) and (4b) above n_a and n_b, represent the size of the estimation sample for individual level data and household level data respectively. Following on from this, **1** in (4a) represents $n_a \times 1$ column vector of ones which is needed to estimate the constant term. Similarly, *gender*, *head*, *age* and *geo* represent matrices of earlier defined variables of specified dimensions. On the other hand expression (4b) corresponds to the Equation (1b). In general, therefore the dimensions of the matrices as well as the variables represented by their columns are consistent with the specifications presented in Equations (1a) and (1b).

Denoting the design matrix associated with either the individual or household level specification as X_f allows Equation (1a) and Equation (1b) to be concisely written in the general form:

$$Y_f = X_f^{'} \beta_f + u_f \tag{5}$$

with the subscript f taking on the values of "non-black" (*nb*) and "black" (*b*). The racial categories are chosen in this paper along the lines of what is arguably *the* major racial division within South

---

[13] In this case the outcome variable is logged real individual income for the individual level data and logged real average income (or income per head) for the household level dataset.

Africa[14]. Given the notational formalities, the threefold Oaxaca-Blinder decomposition can be written as:

$$R = E + C + I \qquad (6)$$

Where $R = E(Y_{nb}) - E(Y_b) = E(X_{nb})'\beta_{nb} - E(X_b)'\beta_b$ denotes the difference in the expected log of real income for the two categories of interest: "non-black" and "black" South Africans. As can be observed from *Equation (6),* the decomposition of difference (R) is comprised of three key quantities of interest: $E = \{E(X_{nb}) - E(X_b)\}\beta_b$, $C = E(X_b)'[\beta_{nb} - \beta_b]$ and $I = \{E(X_{nb}) - E(X_b)\}(\beta_{nb} - \beta_b)$. Note that E can be characterised as the part of the difference due to differences in endowments between "non-black" and "black" South Africans. On the other hand, C can be characterised as part of the income differential due to differences in group coefficients and part I is considered the part of the difference due to interaction between differences and endowments and coefficients. One convenient feature of the Blinder-Oaxaca framework is that each of the major components of the decomposition (*E, C, I*) can be further decomposed by independent variable which allows for more in-depth analysis of how covariates affect the income differential.

An important technicality of the decomposition technique is that the contribution of any set of related binary categorical predictors to the differential in logged real income between the "Black" and "Non-black" South Africans is dependent on the choice of the (omitted) base category[15]. In order to overcome this ambiguity, a transformation technique due to Gardeazabal and Ugidos (2004), Yun (2005) and Suits (1984) referred to as the "deviation contrast" technique is employed. The "deviation contrast" approach relies on the transformation of the standard regression coefficient estimates (Suits 1984; Yun 2005) so that the resulting decompositions are equal to the simple average of all decompositions which would have resulted from alternative specifications of the base category (Yun

---

[14] The results are qualitatively similar when white / non-white classification categories are used. For the sake of brevity only results for the "black" / non-"black" decomposition is presented

[15] Recall from *equation (1)* that the base category is the female, non-head of household who is "black", aged between 30 – 39 (inclusive) and residing in Western Cape. The standard application of the Oaxaca-Blinder decomposition would produce quantitatively different results if alternative base categories had been chosen.

2005). Robust standard errors are displayed for both individual level and household-level Blinder - Oaxaca decomposition estimates.

5. Data

The study uses all four (4) editions of the South African Income and Expenditure Surveys (IES); namely the 1995, 2000, 2005/2006, 2010/2011 IES surveys. The IES's are nationally representative surveys collected by Statistics South Africa – the national statistical agency of South Africa. The mode of collection of the data for the 1995 and 2000 editions of the survey is the *recall* method whereas the later 2005 and 2010 editions of the survey use both *recall* and *diary* methods. Despite the differences in collection methods across survey years, no difference in accuracy is expected regarding the standard demographic variables such as gender, age and geographic location. Combined, the IES provides a rich source of nationally representative[16], socio-demographic data on income and spending in the South African economy.

This study exploits IES data by making use of information on socio-demographic characteristics of the respondents which is consistently collected across all survey years thus allowing for inter-survey comparisons. This consistency is exploited for the estimation of Equations (1a) and (1b) since the model specifications are identical across all surveys allow for comparability across time periods across both the individual level and household levels of disaggregation. In addition this study exploits additional information made available in some surveys but not in others such as data pertaining to skill levels (IES 2000) and educational attainment (IES2005 and IES 2010). As mentioned earlier both the uppermost and lowest percentile income records from all IES survey years were removed to reduce the impact of outliers. This procedure was performed for both the individual and household level dataset. Secondly, the problem of missing observations and incomplete records, which is typical

---

[16] The IES 1995, 2000, 2005 and 2010 elicited information from 29582, 26263, 21144 and 25 328 households across all 9 administrative regions of South Africa, respectively.

in studies using microeconometric level survey data, is also partially addressed by adjusting the estimation sample to include only respondents aged between 19 and 60 years (exclusive). Note here that the problem of zero income values, quite conveniently, does not arise at when analysis is carried out at the household level. Beyond this, for individual level estimation, only complete records (with respect to the dependent and independent variables) are used.

An additional issue to be considered when working with South African IES survey data is the apparent inconsistency in recording zero income observations across survey years. In the 1995 edition of the IES, zero values are used to code missing values in the dataset which makes it impossible to distinguish missing values from *bona-fid*e zero income values. For individual level estimation purposes therefore they cannot be used as doing so would produce biased estimates. Zero values are alternatively and appropriately coded in IES 2000, 2005 and 2010[17].

Of possibly greater import to the analysis undertaken in this paper is the potential issue of under-representing demographic groups within the estimation sample. Ozler (2007) first raised this problem in relation to the IES, in a study utilizing both 1995 and 2000 editions of the IES suggested that the IES 2000 survey over-represented "black" Africans whilst under-representing Whites[18]. *Table 1* provides sample proportions of each socio-demographic group of interest across all survey years of the IES. In IES 1995, the sample proportions for "Black", "White", "Coloureds" and "Other" races are 0.618, 0.177, 0.163 and 0.043 respectively. However in IES 2000 blacks account for 74% and the proportion of whites had halved. While these aforementioned proportions have informed Ozler's (2007) conclusion, taking a broader view across all survey years (1995-2010) and taking into account

---

[17] Yu (2008a, 2008b) provides a useful discussion of the differences between IES 1995, IES 2000 with IES 2005.

[18] An alternative source of data is the population census, arising from the admission of Statistics South Africa that the first two IES surveys were not directly comparable (Statistics South Africa, 2002). The census despite reducing sampling issues, suffers from a high number of respondents reporting zero income or offering no response altogether (Van der Berg *et al.*, 2008). A new source of data used by Van der Berg *et al.* (2008), the All Media and Product Survey (AMPS), despite providing a more frequent review of poverty and inequality, also falls short as a suitable source, generally designed to allocate advertising expenditure rather than for determining changes in inequality over time.

census statistics, the opposite conclusion seems more likley. In fact, it is the 1995 sample proportions which seem appear conspicuous. For this reason, results presented below for both regression and Oaxaca-Blinder results employ the use of probability weights which equate racial sample proportions from IES 1995 to those from 1996 South African Census data. It should be noted, however, that both the sign and significance of coefficient estimates are qualitatively similar irrespective of whether compensatory probability weights were applied.

6. Empirical Results

6a. Regression

Table 3 below depicts coefficient estimates resulting from applying *Equations 1a, 2a, and 3a* to the estimation sample for individual level analysis. Columns *(1), (2)*, *(4)* and (6) of *Table 3* contain coefficient estimates of Equation (1a) estimated using IES data for 1995, 2000, 2005 and 2010 respectively. The reference category in the regression formulation is the Black, female from Western Cape, aged 30 – 39, who is *not* a head of household.

Starting with column *(1)* of *Table 3,* the results show that, on average, the income of males are higher than the income for the base category. The highly statistically significant coefficient estimates in columns *(2)*, *(4)* and *(6)* confirm that across every survey year, males enjoyed on average a higher income relative to females. The magnitude of the marginal effect of gender on real log income fluctuates over the survey years, reaching its zenith in 2005. Overall, in real monetary terms the premium remains relatively stable in the R\$6000 – R\$7000 throughout the period.

The coefficient on the household head binary indicator variable is also statistically significant at the 1% level of significance across all surveys implying that across all survey years, household headship yields a premium in real income in South Africa. Noticeably, however the coefficient size of this variable falls markedly between IES 1995 and IES 2010 in *Table 3*. This result is a possible

Table 3: Dependent Variable: Log of Real Income - Individual Level

Variables	(1) IES 1995	(2) IES 2000	(3) IES 2000: Occu.	(4) IES 2005	(5) IES 2005 :Edu.	(6) IES 2010	(7) IES 2010:Edu.
Gender : Male	0.278***	0.270***	0.296***	0.384***	0.392***	0.305***	0.332***
	(0.0117)	(0.0140)	(0.0128)	(0.0151)	(0.0144)	(0.0172)	(0.0163)
Head of Household	0.158***	0.245***	0.252***	0.218***	0.214***	0.0597***	0.0453***
	(0.0123)	(0.0151)	(0.0137)	(0.0158)	(0.0151)	(0.0181)	(0.0171)
Race: Coloured	0.257***	0.363***	0.368***	0.345***	0.339***	0.411***	0.380***
	(0.0162)	(0.0255)	(0.0235)	(0.0262)	(0.0247)	(0.0289)	(0.0273)
Indian/Asian/Other	0.920***	1.075***	0.800***	1.163***	0.894***	1.051***	0.755***
	(0.0221)	(0.0370)	(0.0340)	(0.0511)	(0.0489)	(0.0547)	(0.0535)
White	1.414***	1.581***	1.200***	1.623***	1.210***	1.558***	1.119***
	(0.0123)	(0.0199)	(0.0211)	(0.0256)	(0.0272)	(0.0269)	(0.0280)
Age: 20 -29	-0.256***	-0.398***	-0.323***	-0.399***	-0.469***	-0.290***	-0.305***
	(0.0150)	(0.0172)	(0.0159)	(0.0199)	(0.0194)	(0.0225)	(0.0217)
40 - 49	-0.0675***	0.0102	0.0249	0.0228	0.145***	0.0584***	0.155***
	(0.0139)	(0.0170)	(0.0154)	(0.0198)	(0.0189)	(0.0223)	(0.0213)
50 -59	-0.248***	-0.233***	-0.168***	-0.184***	0.0505**	-0.128***	0.138***
	(0.0153)	(0.0195)	(0.0180)	(0.0217)	(0.0214)	(0.0238)	(0.0236)
Skill: Low	-	-	-0.202***	-	-	-	-
			(0.0147)				
Semi-skilled	-	-	1.150***	-	-	-	-
			(0.0221)				
High	-	-	1.229***	-	-	-	-
			(0.0250)				
Region: Eastern Cape	-0.108***	-0.207***	-0.250***	-0.279***	-0.277***	-0.416***	-0.346***
	(0.0203)	(0.0296)	(0.0272)	(0.0309)	(0.0297)	(0.0337)	(0.0316)
Northern Cape	-0.396***	-0.299***	-0.250***	-0.215***	-0.167***	-0.201***	-0.0953**
	(0.0267)	(0.0336)	(0.0313)	(0.0320)	(0.0303)	(0.0396)	(0.0376)
Free State	-0.454***	-0.257***	-0.255***	-0.214***	-0.213***	-0.328***	-0.249***
	(0.0232)	(0.0327)	(0.0304)	(0.0355)	(0.0341)	(0.0366)	(0.0346)
Kwazulu/Natal	0.144***	-0.102***	-0.112***	-0.212***	-0.155***	-0.165***	-0.0705**
	(0.0194)	(0.0278)	(0.0259)	(0.0283)	(0.0275)	(0.0330)	(0.0315)
North West	0.0297	0.0547*	0.0151	-0.205***	-0.191***	-0.183***	-0.0994***
	(0.0256)	(0.0306)	(0.0286)	(0.0357)	(0.0342)	(0.0385)	(0.0367)
Gauteng	0.484***	0.253***	0.186***	0.318***	0.241***	0.196***	0.163***
	(0.0196)	(0.0270)	(0.0252)	(0.0316)	(0.0308)	(0.0311)	(0.0299)
Mpumalanga	0.0324	-0.111***	-0.119***	-0.0117	0.0143	-0.0316	0.0611*
	(0.0224)	(0.0322)	(0.0300)	(0.0349)	(0.0338)	(0.0372)	(0.0357)
Northern	0.228***	-0.140***	-0.211***	-0.234***	-0.237***	-0.380***	-0.358***
	(0.0260)	(0.0323)	(0.0292)	(0.0358)	(0.0345)	(0.0381)	(0.0360)
Edu: Primary	-	-	-	-	0.133***	-	0.136***
					(0.0225)		(0.0332)
Secondary	-	-	-	-	0.761***	-	0.905***
					(0.0233)		(0.0323)
Higher	-	-	-	-	1.769***	-	2.066***
					(0.0476)		(0.0464)
Constant	9.820***	9.631***	9.528***	9.427***	8.846***	9.839***	8.969***
	(0.0198)	(0.0277)	(0.0264)	(0.0290)	(0.0344)	(0.0309)	(0.0428)
Observations	33,526	25,095	25,095	25,384	25,384	28,124	28,124
R-squared	0.291	0.239	0.363	0.227	0.299	0.159	0.246

Individual Level Dataset
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

indication that over the 15-year review period there has been a shift away from the model of household head being the major income earner possibly resulting in a greater degree of labour market participation by non-head members of the household. Another possible and related explanation for this result in the South African case is significant urbanization which may result in more diverse household types where income earning roles are more diffuse as the home economy dynamically adjusts to both the opportunities and costs of urbanization.

An examination of the set of race indicator variables reveals a consistent pattern across all IES survey years: racial categories have a highly statistically significant impact on expected real incomes in South African income distribution. The magnitude of the coefficient estimates corroborate earlier results from *Tables 1* and *2* which reflect a stable ordering of incomes by racial group over time. The income disparity between Blacks and non-black racial groups appears to widen somewhat throughout the entire period for Coloureds and for most of the period in the case of respondents within the White or Other racial categories. Whites consistently earn significantly more than all other racial groups across all survey years.

The negative and highly statistically significant coefficient estimates on the age-related binary indicator variables in *column* (1) of *Table 3* implies that on average, and holding other factors fixed, the highest incomes were earned by individuals within the 30-39 (the base category) and 40-49 age groups. Note that across *columns (2)* and *(4)* the coefficient on the 40-49 age group indicator is not statistically significantly different from that of the base category while all other age-related coefficients are negative. The results are consistent with an inverted U-shaped rising wage-tenure (age) profile; a feature well documented within the literature.

Given the differences in the wage distributions across administrative regions depicted in *Table 3,* it is perhaps not surprising that coefficients estimates on the geographical location indicator variables across the columns of *Table 3* are highly statistically significant. The results imply for example, that across all survey years, expected income of respondents located in Gauteng were generally higher than those of workers located in Western Cape which is the base geographical region. This is an intuitive result since most of the territory of Gauteng is defined by two (2) of the nation's largest cities – Pretoria and Johannesburg. On the other hand, expected real incomes in Free State and Northern Cape have been consistently lower than that in Western Province over the review period with opposite results observed for Eastern Cape and Northern Province.

Columns *(3), (5)* and *(7)* of *Table 3* depict coefficient estimates of Equations 2(a), and 3(a). In column (3) coefficient estimates of the set of skill-related binary indicator variables are provided. Recall from 2(a) that the base category for the specification in *column (3)* is: "Black", Skilled, Female from Western Cape, aged 30 – 39, who is **not** a head of household. From *column 3* it becomes clear that semi-skilled and highly-skilled individuals earn more than skilled respondents and all three categories command higher expected real incomes than low skilled individuals. These results imply a skill premium in South Africa favouring all levels of skill, but in particular, high skill levels and practical skills. A similar pattern emerges when educational attainment variables are added in *columns (5)* and *(7)* in there is a robust reward for educational attainment in South Africa across both survey years. In fact, the coefficient on higher education has the largest magnitude across both regressions indicating that high levels of education can be rewarded even more highly than race within South Africa across both survey years and may have the potential to have a meaningful impact on standard of living. The marginal effect of secondary education on expected income is approximately half of the comparable marginal effect of higher education, but still positive. Primary educational attainment also has a statistically significant, positive effect on real income. The magnitude of educational attainment level-related marginal effects increases between IES 2005 and IES 2010 suggesting increasing returns.

Two additional, key observations can be made from *Table 3*. The first observation is that the *R-squared* value declines over consecutive comparable IES surveys. The *R-squared* value of 29.1% which results from using IES 1995 data in *Table, column (1)* declines to 15.9 by IES 2010. This result implies that standard socio-demographic information, traditionally collected in nationally representative surveys such as the IES collectively play a progressively smaller role in determining real expected incomes in South Africa.

Table 4: Dependent Variable: Log of Real Income – Household Level

Variables	(1) IES 1995	(2) IES 2000	(3) IES 2000: Occu.	(4) IES 2005	(5) IES 2005: Edu.	(6) IES 2010	(7) IES 2010: Edu.
Male hh-head	0.212***	0.226***	0.209***	0.594***	0.384***	0.362***	0.327***
	(0.0136)	(0.0155)	(0.0146)	(0.0160)	(0.0155)	(0.0174)	(0.0163)
Coloured hh-head	0.300***	0.372***	0.362***	0.207***	0.263***	0.376***	0.369***
	(0.0177)	(0.0250)	(0.0234)	(0.0256)	(0.0232)	(0.0281)	(0.0262)
Indian/Other hh-head	0.979***	1.083***	0.867***	1.075***	0.990***	1.026***	0.837***
	(0.0274)	(0.0422)	(0.0378)	(0.0540)	(0.0498)	(0.0556)	(0.0534)
White hh-head	1.640***	1.794***	1.427***	1.873***	1.427***	1.691***	1.211***
	(0.0141)	(0.0226)	(0.0237)	(0.0232)	(0.0244)	(0.0249)	(0.0259)
Age of hh-head: < 20	-0.107	-0.502***	-0.423***	-0.337***	-0.592***	-0.757***	-0.862***
	(0.0663)	(0.0444)	(0.0449)	(0.0548)	(0.0545)	(0.0886)	(0.0890)
20 -29	-0.0868***	-0.120***	-0.100***	-0.0479	-0.188***	-0.0277	-0.131***
	(0.0235)	(0.0264)	(0.0249)	(0.0359)	(0.0326)	(0.0338)	(0.0321)
40-49	0.0317	-0.0409	-0.00539	-0.0838**	0.0391	0.00262	0.0475
	(0.0219)	(0.0286)	(0.0267)	(0.0347)	(0.0324)	(0.0332)	(0.0307)
50-59	-0.0108	-0.163***	-0.0924***	-0.163***	0.0486	-0.0602*	0.0647*
	(0.0259)	(0.0349)	(0.0334)	(0.0386)	(0.0351)	(0.0364)	(0.0342)
Over 60	-0.238***	-0.306***	-0.248***	-0.286***	0.0771**	-0.139***	0.0452
	(0.0257)	(0.0319)	(0.0302)	(0.0316)	(0.0307)	(0.0324)	(0.0310)
Prop. of Female	-0.244***	-0.218***	-0.208***	0.634***	0.136***	0.249***	0.108***
	(0.0255)	(0.0249)	(0.0234)	(0.0205)	(0.0240)	(0.0257)	(0.0248)
Prop. Semi -Skilled	-	-	0.0484**	-	-	-	-
			(0.0217)				
Prop. Skilled	-	-	1.995***	-	-	-	-
			(0.0526)				
Prop. Highly Skilled	-	-	2.101***	-	-	-	-
			(0.0555)				
No. of under 10s	0.179***	-0.269***	-0.258***	-0.273***	-0.194***	-0.271***	-0.146***
	(0.00775)	(0.00615)	(0.00598)	(0.00727)	(0.00646)	(0.00671)	(0.00677)
No. aged 10-19	-0.0640***	-0.233***	-0.222***	-0.235***	-0.225***	-0.235***	-0.203***
	(0.00970)	(0.00551)	(0.00532)	(0.0104)	(0.00967)	(0.00673)	(0.00630)
No. age 20 -29	0.0750***	-0.0990***	-0.104***	-0.0132	-0.0847***	-0.0641***	-0.126***
	(0.00997)	(0.00706)	(0.00674)	(0.0120)	(0.0111)	(0.00817)	(0.00790)
No. age 30-39	0.119***	-0.00951	-0.0246**	0.0860***	0.0294**	0.110***	0.0443***
	(0.0117)	(0.0112)	(0.0105)	(0.0142)	(0.0130)	(0.0116)	(0.0108)
No. age 40-49	0.133***	-0.0252	-0.0415***	0.0944***	0.0357**	0.141***	0.0804***
	(0.0134)	(0.0157)	(0.0147)	(0.0155)	(0.0156)	(0.0159)	(0.0144)
No. age 50 -59	0.0810***	-0.0720***	-0.101***	0.0634***	-0.0270	0.103***	0.0383**
	(0.0164)	(0.0215)	(0.0208)	(0.0185)	(0.0167)	(0.0186)	(0.0177)
No. age over 60s	0.0573***	-0.0225	0.00322	0.0542***	-0.119***	0.0683***	0.0280*
	(0.0163)	(0.0196)	(0.0184)	(0.0131)	(0.0141)	(0.0154)	(0.0150)
Region: Eastern Cape	-0.296***	-0.413***	-0.437***	-0.200***	-0.186***	-0.328***	-0.264***
	(0.0206)	(0.0278)	(0.0263)	(0.0278)	(0.0262)	(0.0306)	(0.0288)
Northern Cape	-0.390***	-0.377***	-0.311***	-0.192***	-0.161***	-0.229***	-0.126***
	(0.0261)	(0.0324)	(0.0308)	(0.0295)	(0.0270)	(0.0375)	(0.0350)
Free State	-0.362***	-0.332***	-0.308***	-0.0690**	-0.0528*	-0.205***	-0.133***
	(0.0222)	(0.0321)	(0.0308)	(0.0320)	(0.0295)	(0.0332)	(0.0315)
Kwazulu/Natal	0.0313	-0.196***	-0.191***	-0.295***	-0.227***	-0.171***	-0.118***
	(0.0209)	(0.0269)	(0.0255)	(0.0271)	(0.0252)	(0.0310)	(0.0295)
North West	-0.0787***	-0.0992***	-0.109***	-0.131***	-0.0906***	-0.258***	-0.175***
	(0.0254)	(0.0297)	(0.0284)	(0.0333)	(0.0306)	(0.0345)	(0.0327)
Gauteng	0.383***	0.131***	0.0993***	0.217***	0.205***	0.209***	0.201***
	(0.0215)	(0.0269)	(0.0255)	(0.0305)	(0.0285)	(0.0300)	(0.0286)
Mpumalanga	-0.159***	-0.0816***	-0.0688**	-0.119***	-0.0418	-0.0700**	0.00966
	(0.0230)	(0.0296)	(0.0280)	(0.0318)	(0.0302)	(0.0344)	(0.0328)
Northern	-0.0201	-0.240***	-0.273***	-0.140***	-0.0744***	-0.271***	-0.227***
	(0.0261)	(0.0295)	(0.0276)	(0.0306)	(0.0285)	(0.0324)	(0.0306)
Proportion: Primary	-	-	-	-	0.373***	-	0.0756***
					(0.0213)		(0.0287)
Secondary	-	-	-	-	0.912***	-	0.844***
					(0.0207)		(0.0277)
Higher	-	-	-	-	1.612***	-	2.260***
					(0.0611)		(0.0600)
Constant	7.975***	9.729***	9.595***	8.974***	8.791***	9.475***	8.938***
	(0.0811)	(0.0352)	(0.0365)	(0.0362)	(0.0341)	(0.0394)	(0.0434)
Observations	29,595	26,252	26,054	21,003	21,071	25,164	25,164
R-squared	0.518	0.450	0.521	0.464	0.542	0.362	0.438

Household Level Dataset
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4 shows the results of all variations of the regression model equations *(1b)*, *(2b)*, *(3c)*. The more careful modelling of the characteristics of the household head afforded by these specifications corroborate earlier findings i) male-headed household earn significantly higher expected incomes ii) there is an inverted U-shape of the income-tenure profile with incomes peaking in the 30-39 (base category) and/or 40-49 age categories iii) racial categories of the household head reflect the results from individual level data and confirms the structural nature of the racial inequalities. The first variable which captures the structure of the household is the proportion of females within the household. Across all comparable specifications in columns *(1)*, *(2)*, *(4)* and *(7)* the variable is always statistically significant. However the sign is negative for survey years 1995 and 2000 and turns positive for the latter years of the survey. This implies that there has been a shift over time in the gender composition of the household on expected incomes. One possible explanation for this phenomenon is the increasing participation of women within the economy.

Household level analysis of the survey data also reveals an interesting finding related to the number of individuals within the household within respective age groups. Examining the signs and significance of these variables across comparable specifications of the model in columns *(1)*, *(2)*, *(4)* and *(7)* reveals that household comprised of greater numbers of individuals within the less than 19 age group earn a significantly lower expected real incomes. This result reflects that fact that these individuals would be dependents, unemployed or lacking in skills and experience. Although the results are mixed for higher age-groups columns *(1)*, *(2)*, *(4)* and *(7)* reveal mostly positive and statistically significant coefficient for household with larger number of individuals in within the 30-39 and 40-49 age-groups corroborating results from the regression analysis on individual level data in *Table 3*. Including variables capturing the proportion of skilled and educated individuals within each household as depicted in columns *(3)*, *(5)* and *(7)* confirm early findings that there is a premium to occupational skills and an increasing premium to educational attainment within the South African context between 2005 and 2010 versions of the IES.

6b. Oaxaca Blinder Decomposition

Figures 3 – 6 provide a graphical depiction of the Oaxaca-Blinder decompositions for the South African IES surveys for the years 1995, 2000, 2005, 2010. Part *(A)* of each Figure shows the decomposition for the household level analysis while part *(B)* of each figure presents the analogous decomposition using individual level data. Recall from *Section 4.2* that the Oaxaca-Blinder decomposition is performed along racial lines with the racial composition of the sample divided into two groups: Black and Non- Black South Africans. The error bars within the bar plots show robust standard errors. Part *(A)* of *Figure 3)*, shows the components of the differences in real expected log incomes for the household level data. The bar labelled "Difference" shows the positive difference in expected log real incomes between Non- Black and Black households in South Africa across all survey years. The difference has not declined perceptibly over time. *Part (A)* of Figure *(4)* also makes it clear that most of the difference is due to coefficients and not due to endowments such as socio-demographic statistics. Thus pattern is corroborated and even exaggerated by the results from the individual level analysis where the contribution of socio-demographic characteristics to the income differential is negligible. This implies that systematic and institutional factors are driving the differential.

Figure 4 helps to identify the specific factors that are contributing to the racial income inequality. Both panels *(A)* and *(B)* of the figure show that although the age structure of the household and geographical location both contribute to the part of the differential due to coefficients, there is a large proportion that is simply unexplained by the model, which is accounted for by the constant term. This finding adds credence to the literature exploring, *inter alia*, institutional factors, wealth, networks in affecting inequality in South Africa. *Figure 5* shifts the focus to the contribution toward the income differential between Non-Blacks and Blacks due to endowments. The figure highlights the nuances of understanding inequality in South Africa as different factors appear to be driving the differential at the household and individual levels over time. Panel *(A)* of Figure 5 reveals that, for

the household level analysis in the earliest survey year the age of household head, overall age and gender structure of the household and geographic location of the household were the major contributing factor to the difference in income differential between Non-Black and Black households due to endowments. However, as time progressed, the contribution to differential due to geography was concentrated in various provinces such as Western Province, Gauteng and Kwazulu./Natal. The age structure of the households continued to play a role in the income disparity and there is a growing effect of proportion of females within the household. While the corresponding analysis at the individual level in Panel (B) corroborates the role of age in affecting the racial income differential, there appears to be a noticeable, statistically significant difference in the incomes enjoyed by Non-Black and Black individuals in Northern Province across IES 2000, 2005 and 2010.

For the sake of completeness, Panels *(A)* and *(B) of Figure 6* displays the contribution of the interactions between the coefficients and endowments. Although the interaction plots are more difficult to interpret, they highlight the fact that coefficient and endowment effects can have a moderating and offsetting effect on the racial income differential as can be observed from the negative sign of the estimates. They also underline the importance of geographical location in driving the dynamics of inequality in real expected incomes due to race in South Africa over the post-transition period 1995 – 2010.

7. Conclusion

This paper argues for the complementary use of both household and individual level data to gain a richer understanding of the dynamics of inequality from micro-level surveys over time. In particular, it adds to the empirical literature on inequality in post-Apartheid South Africa by examining the size distribution of income through the lens of the full suite of, nationally representative, Income and Expenditure Surveys (IES) surveys for years 1995, 2000, 2005 and 2010 using a combination of descriptive, regression-based modelling and decomposition analyses on both individual and

household-level data. By employing this multipronged approach which analyses the data, in parallel, at both individual and household levels, key socio-demographic characteristics and stylized household characteristics which contribute to inequality and its dynamics over time are identified. The manner in which the intrinsic relationships change over time is also explored. A key feature of the paper is the manner in which it exploits changes in the structure of the IES instrument over time; incorporating data on occupational skills and educational attainment, as they are made available, to understand the interplay between these variables on incomes and inequality in South Africa.

The study reveals similarities in the household and income distribution of income in South Africa, although the household size distribution is relatively more positively skewed. By every measure employed income inequality has increased since 1995, although the individual level data shows a slight yet perceptible reduced rate of increase in inequality 2005-2010. The uppermost decile of both the household and individual income distributions has been benefitting from an ever increasing share of income which by 2010 stood at 71% and 67% respectively at the expense of all other quintiles, meanwhile median incomes have remained more or less stagnant.

A further disaggregation of inequalities *within* socio-economic groupings using the individual level dataset reveals, rather pessimistically, that income inequality within every socio-demographic group of interest inequality has increased relative to 1995 levels, especially among both gender groups, Blacks, young South Africans and respondents located in Western Cape, Gauteng, Mpumalanga and Northern Province. Analysis of the household level analysis confirms the trend of increasing inequality across most categories except among over-40 aged, White- and Other (Indian and Asian)-headed households and individuals located in Western Cape, Northern Cape and North West Province, Northern Province. It can therefore be observed that within the developing country context, separate inequality dyanamics can co-exist at household and individual level of analysis. At both

levels of analysis, real median incomes remained more or less flat over the period after a relative decline in the 2000 and 2005 survey years excepting for median incomes for Non-black racial groups, and relatively more urbanized provinces which increased over the period. On the other hand, Blacks, females and youth suffered declines in median incomes from the individual level data. A contrasting dynamic emerges at the household level, were incomes more stable among gender and racial groupings, declining median incomes of households headed by younger South Africans was corroborated by the data.

The regression analysis to the IES data allows us to unmask two key features of South Africans who occupy the extremities of the South African income distribution for both individual and household level data. For the individual data, at the right-most tail of the income distribution is the white, male, household head, aged 30-49 years of age, living in Gauteng province. By way of contrast, at the left – most tail of the distribution is a "black" female, who is not a household head, in the 20-29 age group living in either Free State (1995) , Northern Cape (2000) , or Eastern Cape (2005,2010) depending on the IES survey year under review. The household level data reveals additional nuances. While on the basis of individual and household level data a gender premium in expected income exists for males and male-headed households, household level regression analysis reveals that household consisting of mostly females on average earn a higher income towards the end of the sample. The coefficients on geographical variables over time reflect the changing fortunes of the administrative provinces and the urbanization and dynamic adjustments of individuals to the significant disparities in incomes across provincial boundaries.

Twin key observations emerging from the regression analysis are: i) occupational skill levels and especially educational attainment are key dimensions to any solution to the income inequality problem in South Africa and ii) over time, traditional socio-economic variables diminish in their explanatory power to explain the variations in real incomes in South Africa. The latter stylized fact

was also confirmed through the application of the Oaxaca-Blinder decomposition method technique which revealed that the household level of analysis was more explanatory of the income dynamics since it entails the racial, gender and age composition of the household. The decomposition analysis attributes the major source of differences in the expected real income differential between Blacks and Non- Blacks to the coefficients component, revealing the structural nature of the income inequality.

In conclusion, the analysis highlights limited progress to bring about greater equality in South Africa in the post-transition period. By augmenting individual level analysis of multiple Income and Expenditure Survey (IES) years with household level analysis a much richer picture of income inequality dynamics is observed. Deep structural inequalities continue to exist more than 15 years after the fall of apartheid. Uneven geographical development, rapid urbanization, skill-biased technical change are but a few of the contributory factors identified within the literature. Skills training, and educational reform and human capital investment robustly appear as key aspects of the solution to the problem. The racial and structural nature of the entrenched inequalities, however imply, that much deeper that broad-based political, social, educational and economic reforms are needed to address the problems of uneven development which continue to stymie South Africa's attempts to realize its promise of hope since the end of apartheid. The realization of the hope that inspired the "long walk" to freedom remains elusive but can be achieved with appropriate, data-led and sincere efforts in the long-term interests of the South African nation.

References

Adato, Michelle, Carter, R. Michael, May, Julian (2006). "Exploring Poverty Traps and Social Exclusion in South Africa using Qualitative and Quantitative Data", *The Journal of Development Studies*.42 (2) pp. 226-247.

Albertyn, Catherine. (2011) "Law Gender and Inequality in South Africa" *Oxford Development Studies*. 39 (2) pp.139 -162.

Altonji, J. and Blank, R. (1999). "Race and Gender in the Labor Market", *Handbook of Labor Economics*. 3, p.3164-3180.

Arestei, David (2013) "A Blinder-Oaxaca Decomposition for Double-Hurdle Model with an Application to Migrants' Remittance Behaviour". *Applied Economics Letters*. 20 (18) pp.1665-1672.

Atkinson. 2015. "Inequality: What Can be Done?" Harvard University Press, Cambridge Massachusetts, London, England.

Ben-Porath Y. (1967) "The Production of Human Capital and the Life Cycle of Earnings". *Journal of Political Economy* 75 (4) pp.352-365.

Bhorat Haroon and Oosthuizen, Morne (2008) "Determinants of Grade 12 Pass Rates in the Post-Apartheid South African Schooling System". *Journal of African Economies*, 18 (4) pp. 634-666.

Bhorat, H. and Van der Westhuizen, C. (2009). Income and Non-Income Inequality in Post-Apartheid South Africa: What are the Drivers and Possible Policy Interventions? DPRU Working Paper 09/138, Development Policy Research Unit, Cape Town.

Blinder, A. (1973). "Wage Discrimination: Reduced Form and Structural Estimates". *The Journal of Human Resources*, 8 (4), pp. 436-455.

Bosker, Maarten and Krugell Waldo (2008). Regional Income Evolution in South Africa after Apartheid. *Journal of Regional Science*. 48 (3) pp. 493-523.

Bourguignon, François (2015) "The Globalization of Inequality", Princeton University Press. Princeton and Oxford.

Branson, N., Garlick, J., Lam, D., Leibbrandt, M. (2012). "Education and Inequality: The South African Case". Southern Africa Labour and Development Research Unit Working Paper Number 75. Cape Town: SALDRU, University of Cape Town.

Casale, Daniela M. (2012). "Indirect Taxation and Gender Equity: Evidence from South Africa". *Feminist Economics* 18(3) pp.25-54.

Darity W, Guilkey DK, Winfrey, W. (1996). "Explaining Differences in Economic Performance Among Racial and Ethnic Groups in the USA", *American Journal of Economics and Sociology*, 55(4), 411-425.

Duflo, Esther, 2003, "Grandmothers and Granddaughters: Old Age Pensions and Intra-household Allocation in South Africa," *World Bank Economic Review*, 2003, 17 (1),1-25.

Gardeazabal, J., and A. Ugidos. (2004). "More on Identification in Detailed Wage Decompositions". *Review of Economics and Statistics* 86(4) pp.1034–1036.

Gov South Africa. (2015). 'Old Age Pension'. http://www.gov.za/services/social-benefits-retirement-and-old-age/old-age-pension (last accessed 11 February 2015)

Gradín, Carlos (2014). "Race and Income Distribution: Evidence from the USA Brazil and South Africa". *Review of Development Economics* 18 (1) pp.73-92.

Harris R. John and Michael P. Todaro. (1970). "Unemployment and Development: A Two Sector Analysis" *The American Economic Review*. 60(1) pp. 126 – 142.

Heckman. J. (1976) "A life Cycle Model of Earnings Learning and Consumption". *Journal of Political Economy,* 84 (4) pp.11-44.

Hicks, Douglas (1997) "The inequality adjusted Human Development Index: A Constructive Proposal" *World Development,* Vol. *25, No. 8,* pp. 1283-1298

Hoogeveen, Johannes and Berk Ozler (2004), "Not Separate, Not Equal: Poverty and Inequality in Post-Apartheid South Africa," February. Unpublished draft.

Jann, B. (2008). "The Blinder-Oaxaca Decomposition for Linear Regression Models" *The Stata Journal*, Number (8) 4, pp. 453–479.

Jenkins, Rhys. (2006). "Trade Technology and Employment in South Africa". *Journal of Development Studies.* 44 (1) pp. 60-79.

Kim, C, (2010). "Decomposing the Change in the Wage Gap Between White and Black Men Over Time, 1980-2005: An Extension of the Blinder-Oaxaca Decomposition Method." *Sociological Methods Research*, 38(4), pp. 619-651.

Kingdon, Geeta and John, Knight (2007) "Unemployment in South Africa, 1995-2003: Causes, Problems and Policies" *Journal of African Economies*. 16 (5) pp. 813-848.

Klasen, Stephen and Woolard, Ingrid. (2008) "Surviving Unemployment without State Support: Unemployment and Household Formation in South Africa". *Journal of African Economies*. 18(1) pp.1-51.

Kon, Z. R., and Lackan, N. (2008). "Ethnic Disparities in Access to Care in Post-Apartheid South Africa". *American Journal of Public Health*, 98 (12), pp. 2272-2277.

Leibbrandt, M. and Levinsohn, J. (2011) "Fifteen years on: Household Incomes in South Africa". *National Bureau of Economic Research*, Working Paper No. 16661.

Leibbrandt, M., Finn, A. and Woolard, I. (2012). "Describing and Decomposing Post- Apartheid Income Inequality in South Africa". *Development Southern Africa*, 29 (1), pp.19-34.

Leibbrandt M, Levinsohn James, and Justin McCrary (2010). "Incomes in South Africa After the Fall of Apartheid", *Journal of Globalization and Development*, 1(1) DOI**:** https://doi.org/10.2202/1948-1837.1023

Lustig, N. & Higgins, S., 2012. "Commitment to Equity Assessment (CEQ): Estimating the Incidence of Social Spending, Subsidies and Taxes Handbook", *Tulane University, Department of Economics. Tulane Economics Working Paper Ser*ies. Working Paper 12.

Magruder, Jeremy R. (2010). "Intergenerational Networks, Unemployment, and Persistent Inequality in South Africa". *American Economic Journal: Applied Economics* 2(1) pp. 62-85.

North, D. (1990), Institutions, Institutional Change and Economic Performance, Cambridge: Cambridge University Press.

North, D.C. (1981). *Structure and Change in Economic History*, New York, USA: Norton.

North, D. (1990), *'Institutions, Institutional Change and Economic Performance'*, Cambridge, UK: Cambridge University Press.

Oaxaca, R. (1973). "Male-Female Wage Differentials in Urban Labour Markets". *International Economic Review*, 14 (3), pg.693-709.

O 'Gorman, Melanie (2010) "Racial Earnings Inequality in South Africa: An Assessment of Policy Options". *Journal of Policy Modelling,* 32 (6), pp. 703-718.

Oosthuizen, Mourné. (2012) "Low Pay in South Africa". *International Labour Review,* 151(3) pp. 173-192.

Otherofer, Anna (2016) "Wealth inequality in South Africa: Evidenve from Survey and Tax Data." REDI3x3 Working Paper 15. June 2016

Ozler, B. and Hoogeveen, J. (2007). "Not Separate, Not Equal: Poverty and Inequality in Post-Apartheid South Africa". *Economic Development and Cultural Change*, 55(3), pp.487–529

Philip, Kate. (2012) "The Rationale for an Employment Guarantee in South Africa". *Development Southern Africa* 29(1) pp. 177 – 190.

Piketty, T. 2014. Capital in the Twenty-First Century, Cambridge, MA: Harvard University Press.

Rattsø, Jørn and Hildegunn E. Stokke (2013). "Trade, Skill Biased Technical Change and Wage Inequality in South Africa". *Review of International Economics* 21(3), pp. 419 -431.

Shorrocks, A. (1982). "Inequality Decomposition by Factor Components". *Econometrica*, 50(1), pp.193-211.

Shupp, Franklin R. (2002) "Growth and Income Inequality in South Africa". *Journal of Economic Dynamics and Control.* 26 (9-10), pp.1699-1720.

Singh, A. and Dhumale, R. (2000) Globalization, Technology and Income Inequality: a critical analysis. Working Paper No. 210, World Institute of Development Economics Research.

Spaull, Nicholas (2013). "Poverty and Privilege: Primary School Inequality in South Africa" *International Journal of Educational Development.*33 (5) pp. 436 – 447

Stanley, T. D. and S. B. Jarrell. (1998) "Gender Wage Discrimination Bias? A Meta - Regression Analysis". *Journal of Human Resources* 33 (4), pp. 947–973.

Statistics South Africa (2002). "Earning and Spending in South Africa: Selected Findings and Comparisons from the Income and Expenditure Surveys of October 1995 and October 2000". Pretoria: Government Printer.

Stiglitz. J. (2012). "The Price of Inequality" *W.W.Norton and Company*. New York and London.

Stiglitz. J. (2015). "The Great Divide: Unequal Societies and What We Can Do About Them". *W. W. Norton and Company*. New York and London.

Suits, D. B. (1984). "Dummy Variables: Mechanics V. Interpretation. *Review of Economics and Statistics.* 66 (1) pp.177–180.

Tanaka, Shinsuke. (2014) "Does Abolishing User Fees Lead to Improved Health Status? Evidence from Post- Apartheid South Africa" *American Economic Journal: Economic Policy* 6(3) pp. 282-312.

Ulriksen, Marianne S. (2012) "How Social Security Policies and Economic Transformation Affect Poverty and Inequality: Lessons for South Africa" *Development Southern Africa.* 29(1),pp. 3-18

Van der Berg, S., Louw, M., Yu, D. (2008). "Post-transition Poverty Trends Based on an Alternative Data Source". *South African Journal of Economics* 76 (1), pp.58–76.

Van Der Berg Servaas and Megan Louw (2004). "Changing Patterns of South African Income Distribution: Towards Time Series Estimates of Distribution and Poverty". *South African Journal of Economics.* 73(3) pp. 546-572.

Weiss Y. (1986). "The Determination of Life Cycle Earnings: A Survey," in Ashenfelter, O., and Layard R. eds. Handbook of Labour Economics, Vol. 1 chap.11 pp. 603-640, Amsterdam: Elsevier Science/ North Holland

Weichselbaumer, D., and Winter-Ebmer, R. (2005) "A Meta-Analysis of the International Gender Wage Gap". *Journal of Economic Surveys.* 19 (13). pp. 479–511.

Wittenberg, Martin (2016a) "Wages and Wage Inequality in South Africa 1994-2011: Part 1- Wage Measurement and Trends", *Southern African Journal of Economics*, *DOI: 10.1111/saje.12148*

Wittenberg, Martin (2016b) "Wages and Wage Inequality in South Africa 1994-2011: Part 2- Wage Measurement and Trends", *Southern African Journal of Economics*, *DOI: 10.1111/saje.12147*

Wittenberg, M. and Leibbrandt, M. (2017), "Measuring Inequality by Asset Indices: A General Approach with Application to South Africa". *Review of Income and Wealth*. DOI:10.1111/roiw.12286

Yu, D., (2008a). "The Comparability of Income and Expenditure Surveys 1995, 2000 and 2005/2006". *Stellenbosch Economic Working Papers,* 11/2008.

Yu, Derek., (2008b). "Factors Influencing the Comparability of Poverty Estimates across Household Surveys". *Development South Africa. Stellenbosch Economic Working Papers*: 11/08.

Yun, Myeong-Su (2005). "A Simple Solution to the Identification Problem in Detailed Wage Decompositions" *Economic Inquiry.* 43 (4) pp. 766-772.
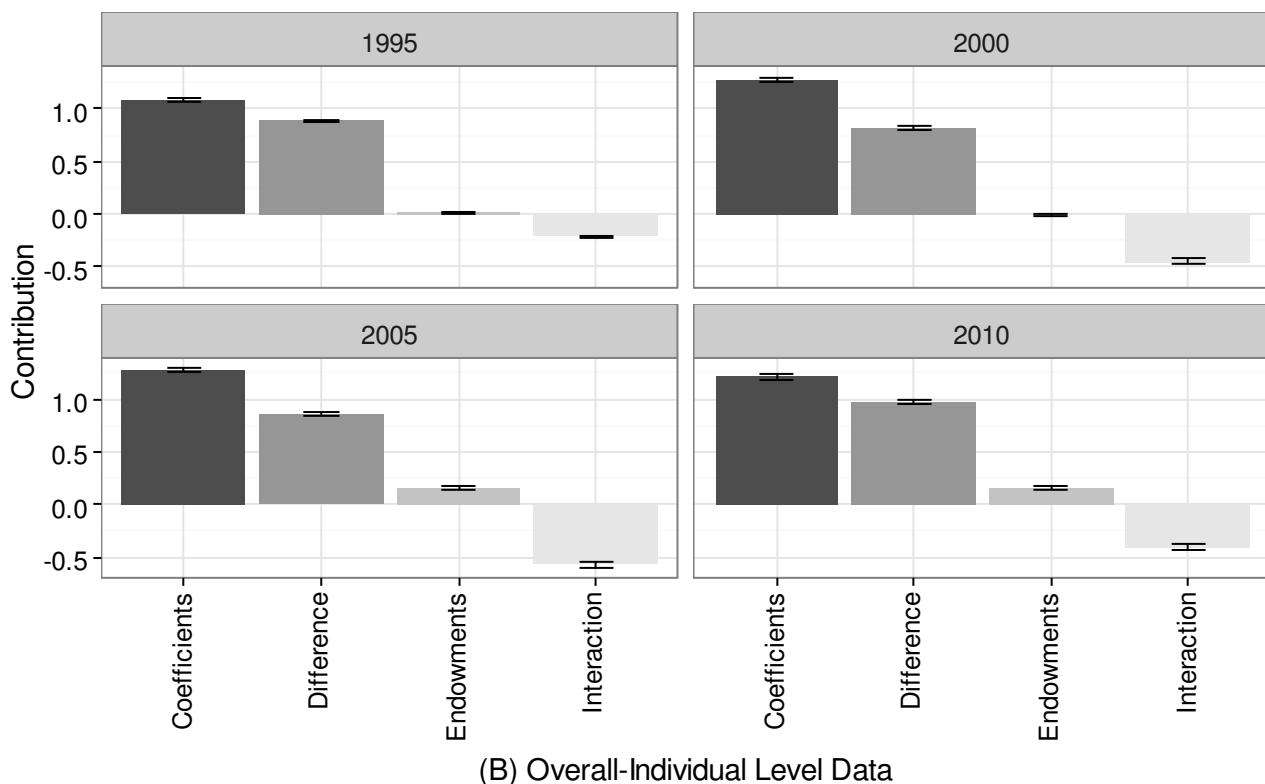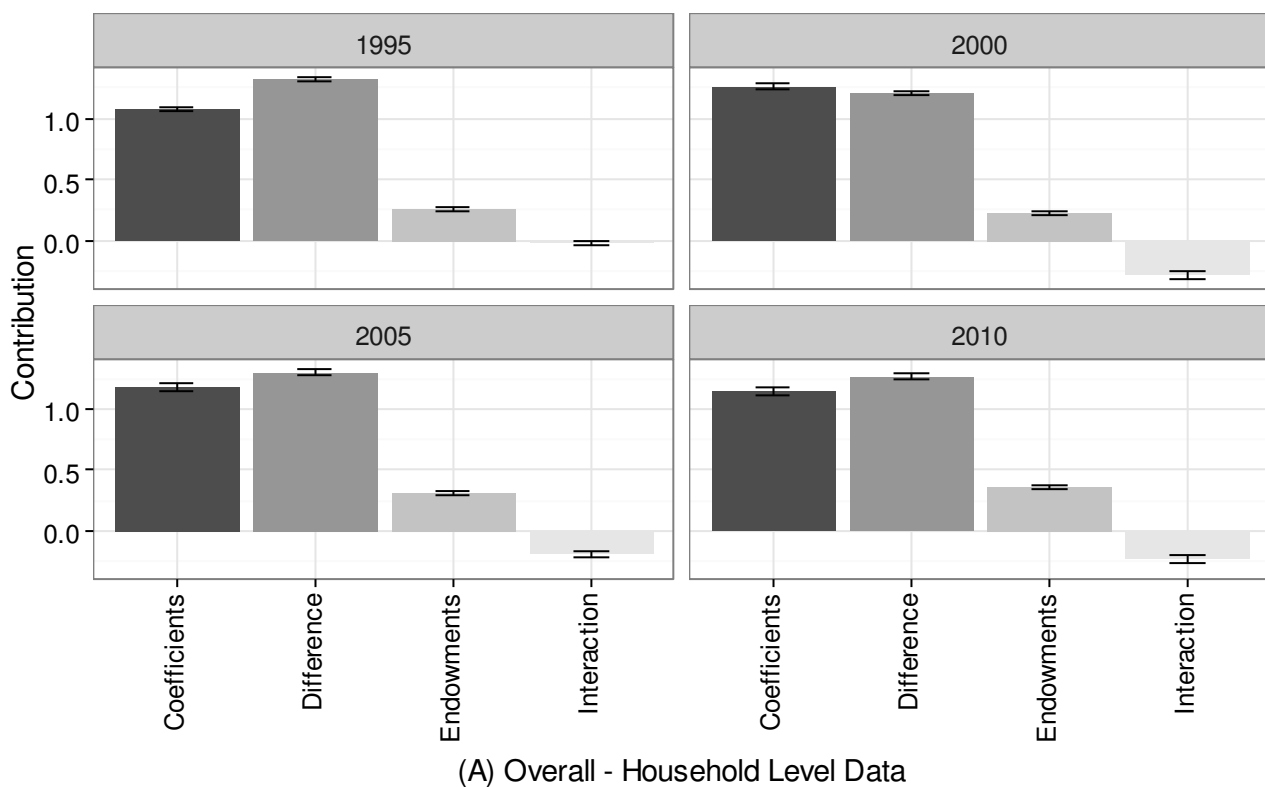
Figure 3: Overall: Oaxaca-Blinder Decomposition for South Africa - 1995, 2000, 2005, 2010.
Source: Author's computation for respective years of Income and Expenditure Survey.
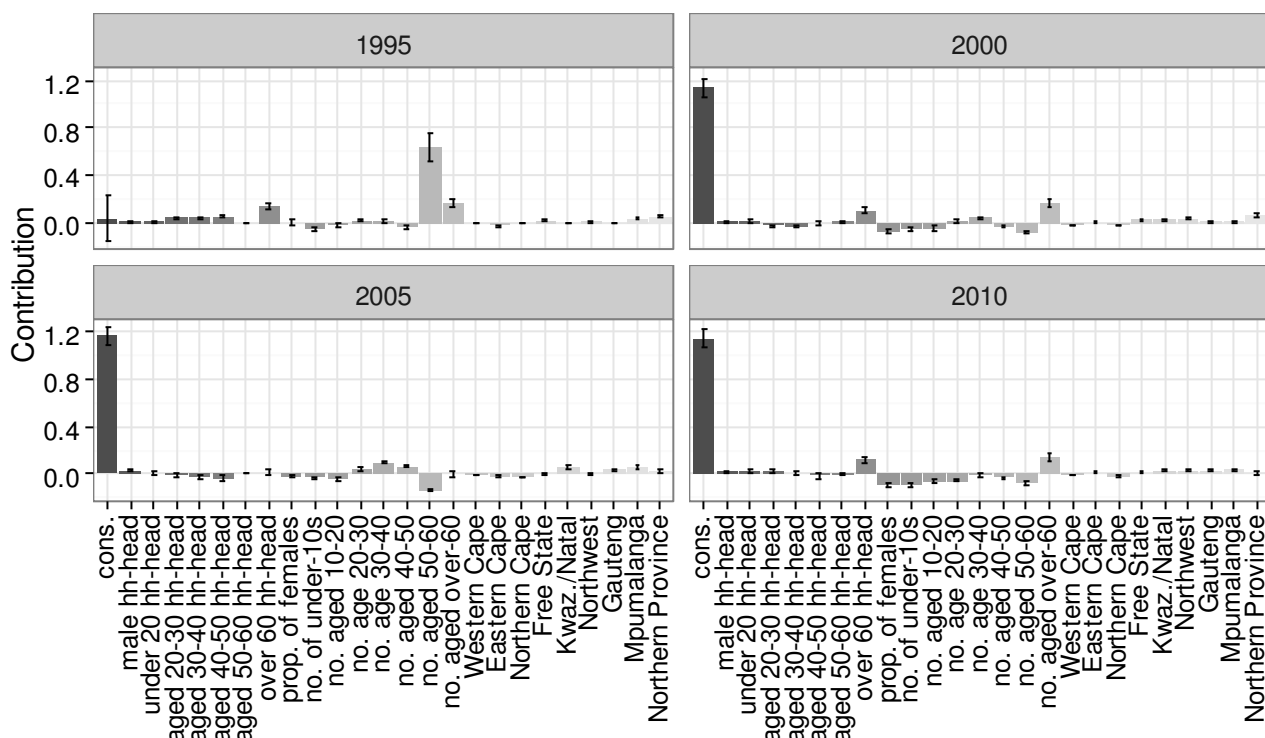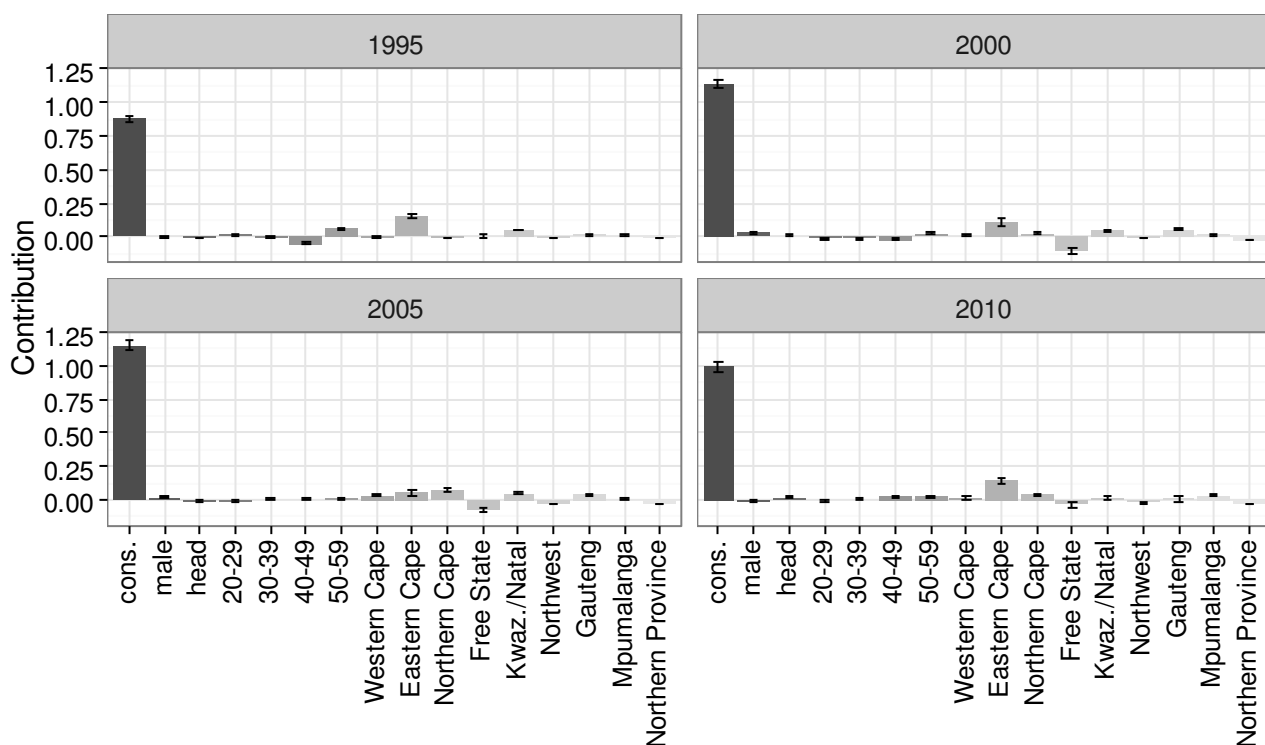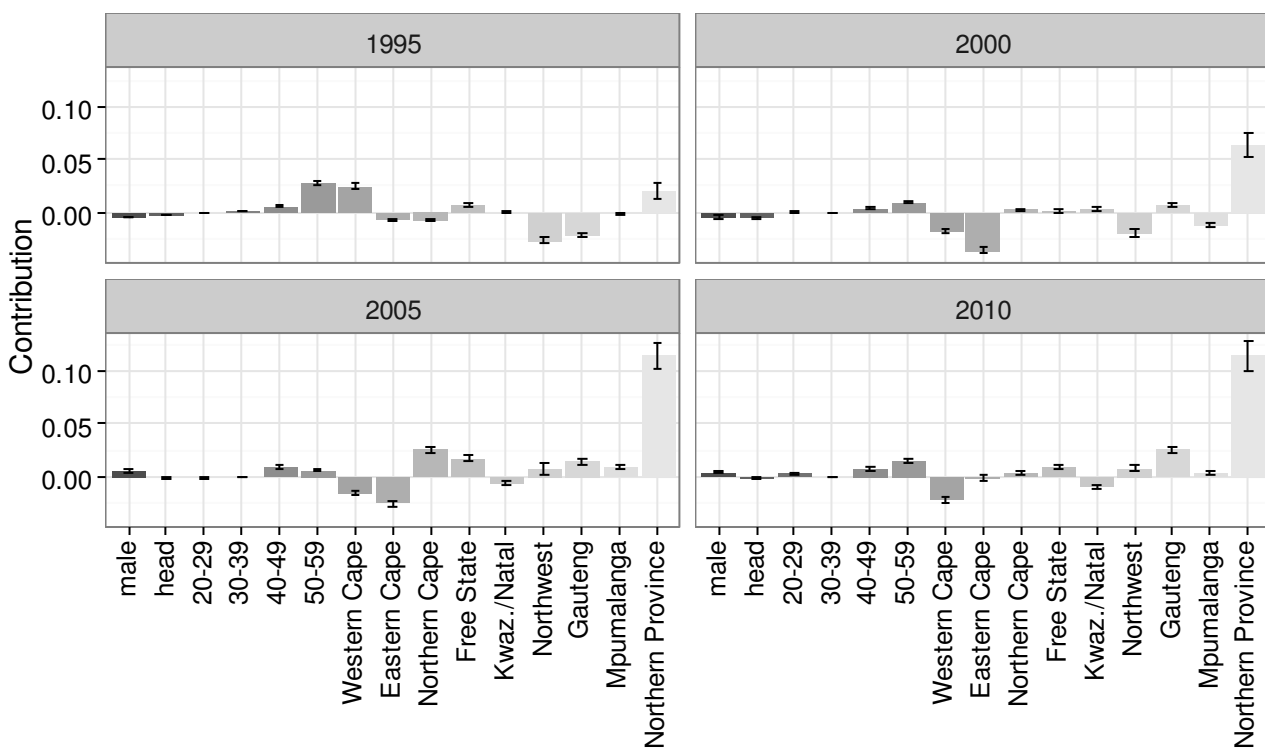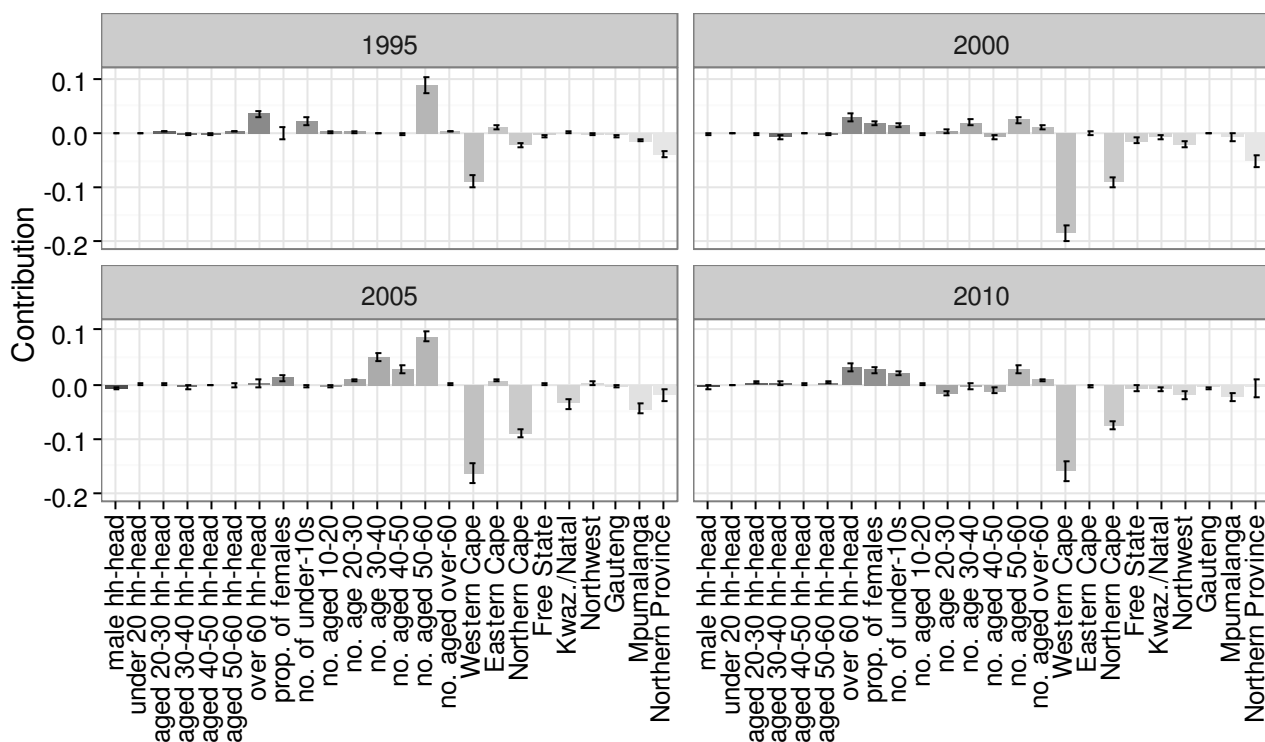
Figure 4: Coefficients: Oaxaca-Blinder Decomposition for South Africa - 1995, 2000, 2005, 2010.
Source: Author's computation for respective years of Income and Expenditure Survey.
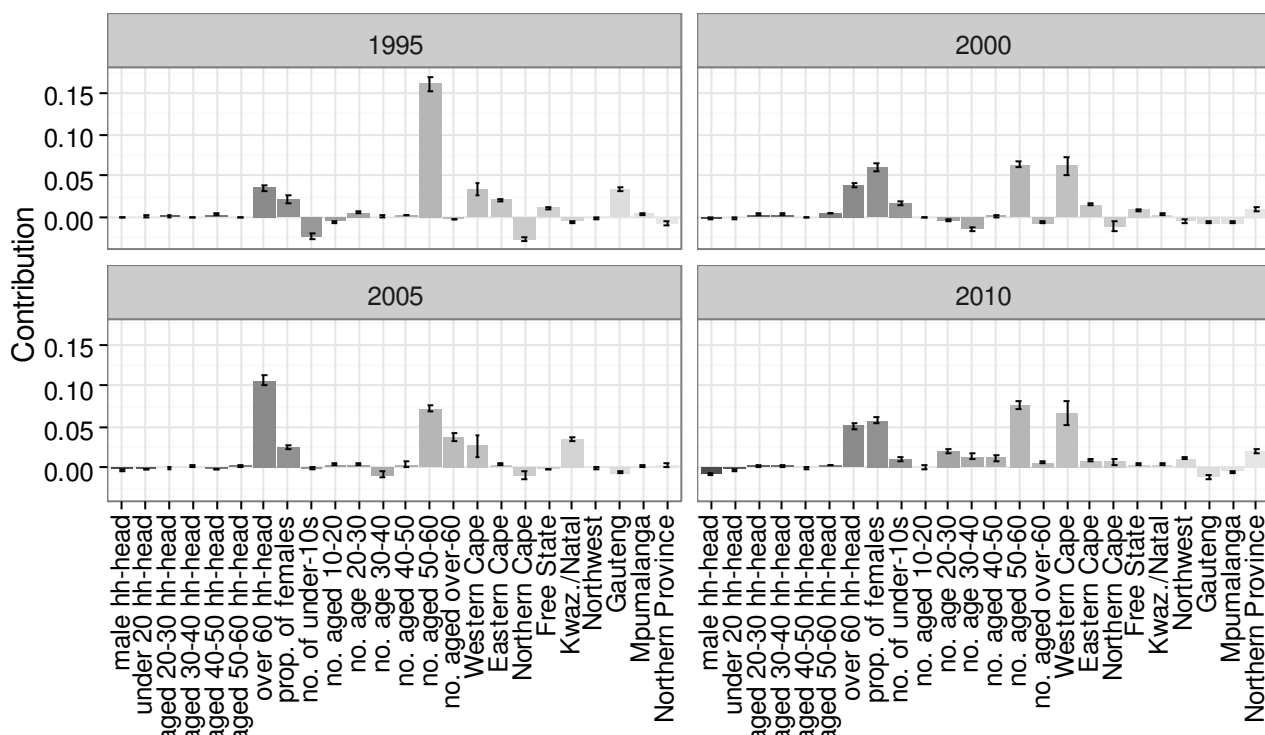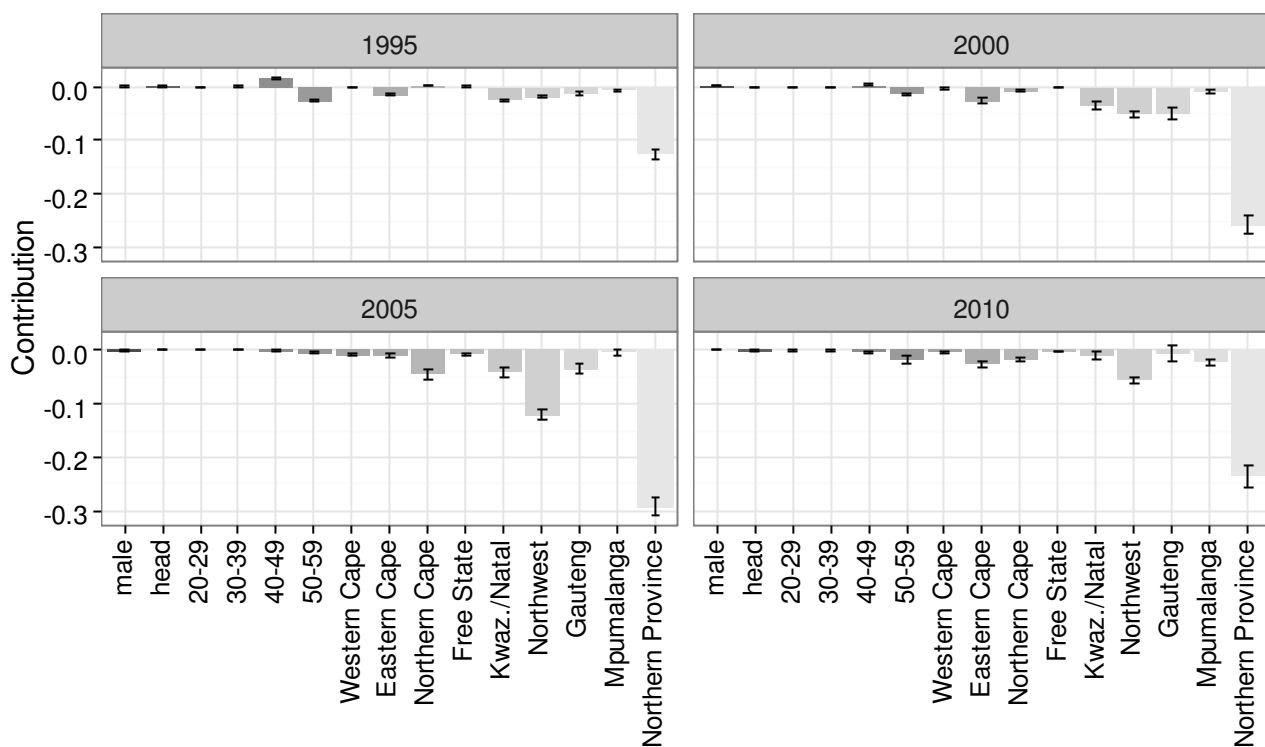
Figure 4: Endowments: Oaxaca-Blinder Decomposition for South Africa - 1995, 2000, 2005, 2010.
Source: Author's computation for respective years of Income and Expenditure Survey.

Figure 6: Interactions: Oaxaca-Blinder Decomposition for South Africa - 1995, 2000, 2005, 2010.
Source: Author's computation for respective years of Income and Expenditure Survey.