

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

# An examination of Labour Discrimination in South Africa 2005-2008.

---

Discussion and Sensitivity Analysis

**Luke Barnett**  
**BRNLUK001**

**Supervisor: Ingrid Woolard**

*This submission is a partial requirement for fulfilment of Master of Commerce degree.<sup>1</sup>*

Unemployment levels in South Africa are some of the highest in the world. Since the birth of democracy in 1994, the South African government has not only made employment a major focal point, but also the reduction in unemployment inequality. Using probit analysis, as well as non-linear decomposition techniques, this paper investigate the effects these policies have had on the labour market, in light of the recent economic recession experienced locally and abroad. Results indicate that racial labour discrimination still exists within the workplace, however whether this is due to differing endowments or pure discrimination is still a contentious issue.

<sup>1</sup> This copy has been edited post markers' comments May 2011

## Contents

Introduction .....	3
Literature Review and Contextual Framework.....	3
Datasets.....	5
Descriptive Statistics .....	8
Methodology .....	9
Probit Analysis .....	9
Decomposition of the Gap in Unemployment Probability and Sensitivity Analysis .....	16
Daymont & Andrisani (1984) .....	21
Reimers (1983) .....	21
Cotton (1988).....	22
Neumark (1988) .....	22
Conclusions and Recommendations.....	23
References .....	26
Figure 1: Incidence of Unemployment .....	9
Figure 2: The White-Black Unemployment Gap, (Source, GHS and own calculations) ....	12
Table 1: Unemployment Rates by Category .....	8
Table 2: Pooled Probit Regression, Source GHS 2005-2008 .....	13
Table 3: Probit Regression for Black Respondents only, Source GHS 2005-2008.....	14
Table 4: Probit Regression for White Respondents only, Source GHS 2005-2008.....	15
Table 5: Decomposition of the Unemployment Gap.....	20

## **INTRODUCTION**

Racial Labour discrimination is a well researched topic within South Africa. It has been on the agenda of national government since first it took power in 1994. According to the broad definition of unemployment, South Africa had an unemployment rate of over 43%<sup>2</sup>. Job creation and a reduction in unemployment of previously disadvantaged individuals is high on the government's agenda. Specifically, government aims to reduce past inequalities, and improve working conditions for all. Most notable is the Employment Equity Act (EEA), which specifically aims to eliminate unfair discrimination, as well as implement affirmative action measures to ensure equal representation of all members of South African society (Department of Labour). There have been numerous studies into discrimination within South Africa, namely Brookes & Hinks (2004), Kingdon & Knight (2004), and Burger & Jafta (2006), amongst others. The most recent work was again conducted by Burger and Jafta (2010), which includes examination of the topic up to 2006, however there has been no research post that period, and thus have not had the opportunity to examine discrimination during or after the recent economic recession. This paper therefore adds to the body of knowledge by following the previously established methodology to investigate labour market discrimination. The paper uses various econometric techniques in order to better understand South Africa's labour market position during this economic upheaval. The goal is to "update" previous work, in order that we can better understand if government legislation has had an effect within the workplace.

Although, a primary goal of the paper is to investigate the causes of unemployment probability, the study also investigates and decomposes discriminatory behaviour within the workplace. Specifically, sensitivity analysis of the well known Oaxaca-Blinder Technique is conducted, with the use of methods which extend this technique, and examine how labour discrimination has evolved over the period 2005-2008, and thus extend previous work by Kingdon & Knight (2004), as well as Brookes & Hinks (2004).

## **LITERATURE REVIEW AND CONTEXTUAL FRAMEWORK**

There has been much research conducted into employment inequality and the racial gap within the South African labour market, specifically post 1994 and the end of Apartheid. There has also been considerable interest surrounding the effects that skills differentials have within the labour market, with conclusions consistently indicating that low skilled individuals face less chance of employment than their highly skilled counterparts.

Fallon and Lucas (1998) find that that "the group most at risk from unemployment under the broad definition are African women", and even controlling for education, find

---

<sup>2</sup> See table 1

that “probability of unemployment, is still high among Blacks with 10 years of education.” With regard to skill levels, the same study indicates that “unskilled wages have been much less responsive to unemployment levels than those of the skilled.” Kingdon and Knight (2004) conduct a study using probit and decomposition analysis to determine the gap in employment probability. They conclude that “young uneducated Africans living in homelands and remote areas are most vulnerable to unemployment”, and “that expanding education and skills will reduce overall unemployment”.

Brookes & Hinks (2004) found that there has been little change in the employment gap, during the period 1995 to 2002, and conclude that “reverse discrimination employment hiring policies have been ineffective in curbing the labour market inequalities of the past”. They indicate that this has primarily been due to white “advantages” from the supply side, in the form of “greater education and experience, as well as better household composition” (Brookes & Hinks, 2004). They indicate within the study, that factors such as “access to labour market information, social network differences (with labour force participants tending to same-race network) and differences in labour market expectations” (Brookes & Hinks, 2004), must be taken into account, but also point out that this is particularly important. An important point is made, in that due to the sheer size and transparent nature of large firms, it is of far more interest to them that they adhere to new labour legislation, while small to medium sized firms can easily avoid penalisation (Brookes & Hinks, 2004).

Burger & Jafta (2006), as in this paper, undertook 3 differing decomposition techniques to track the changes in employment discrimination. Specifically “Oaxaca-Blinder decompositions were used to evaluate the evolution of discrimination at three stages of the employment process: employment, occupational attainment and wages” (Burger & Jafta, 2006). Conclusions indicate that “affirmative action may be influencing labour market outcomes by focusing attention on representation at the top end of the occupational ladder” (Burger & Jafta, 2006), and also that “quality” of education is of increasing importance. Case & Deaton (1999) undertook an interesting study into the effect that various schooling inputs can have on an individual’s outcome, and reiterate that simply controlling for education is perhaps not a robust measure rather than an indication of the quality of schooling should be considered. With regards to current government policy, Burger & Jafta (2006) indicate that although affirmative action policies have had an effect within the labour market, the impact has primarily been restricted to a very small group of workers, and restate that “any recommendation for policy-makers should include a focus on attempting to increase Black returns to education if they want to avoid a widening of the productivity gap” (Burger & Jafta, 2006).

The interest therefore lies in the effect which government legislation has had within the labour market, and whether affirmative action policies have had any effect on the racial gap in unemployment. Affirmative Action policies have been in practice within South

Africa since 1998. However, the more complex and perhaps comprehensive Black Economic Empowerment (BEE) policy was only introduced in 2004. Burger & Jafta (2006) highlight the three core components of BEE, *direct empowerment* (Ownership and control of businesses and assets), *human resource development* and *indirect empowerment* (Preferential procurement, enterprise development, profit- and contract-sharing by black enterprises). However, as Burger & Jafta (2006) point out, “A comparison of employment rates or average wages across time is by no means a controlled experiment, and identifying the exact effect of affirmative action without being able to observe the counterfactual will always be a matter of some degree of conjecture.” Furthermore, when the policy of BEE was implemented, South Africa experienced large economic transformations. These included the liberalisation of trade controls and an acceleration of privatization (Chabane et al, 2006), which ultimately led to “the start of the longest economic upswing in South Africa’s post-war history” (Burger & Jafta, 2006).

The South African economy is by no means stagnant, and is evolving all the time, as it is influenced not only by domestic events and politics, but also by events occurring outside of its borders. The most prominent of these events of late, is the recent global recession, triggered by the “subprime crisis” of 2007. “Between 2004 and 2007, [South Africa’s] economy grew by a perky 5% a year, after averaging a still respectable 3% over the previous decade” (The Economist Online, 31 January 2010). However, according to an article in *The Economist*, the economy came to a virtual standstill in the second half of 2008, indicating that the crisis had finally reached the shores of the South African economy. This dramatic change to the economic environment has direct effect within the labour market, as firms look to reduce costs, and move to the most efficient means of production. The question then arises, as to how the labour market will react. The South African labour market therefore finds itself in an interesting scenario, where government pushes for equality within the labour markets, while businesses look to survive the economic storm. The focus therefore of this study is to set the backdrop for discrimination analysis as the South African economy becomes part of the global recession, and give some exploratory indications as to how affirmative action policies have affected the labour environment.

## **DATASETS**

Data was used for the period 2005 to 2008. The General household Survey (GHS), for the aforementioned period was the dataset employed. The GHS is an annual household survey covering 5 broad areas: education, health, activities related to work and unemployment, housing and household access to services and facilities. Information is collected from over 30 000 households around South Africa, through face to face interviews. A multi-stage, stratified random sample was drawn using probability-

proportional-to-size principles (Stats SA, 2010). These samples are drawn from a master sample, using enumeration areas established during the demarcation of the Census for 2001.

The GHS dataset is split into 4 separate sections namely, “person”, “worker”, “tourism” and “house”. For the purposes of this study, the person, worker and household datasets were merged, such that controls on both an individual and household level can be used within the model. This allowed for the examination of the data using indicators which examine the relationship from an individual as well as household level.

The investigation was limited to respondents eligible to work (15-65 years old). Given the massive discrepancy in unemployment rates between black and white individuals, the sample was further reduced to examine only these two population groups. In line with previous studies surrounding the topic, the broad, as opposed to the narrow definition of unemployment is used. Kingdon & Knight (2000), investigate and conclude that the broad definition be employed, as “many unemployed persons do not search actively for work because they are discouraged workers.” Klaasen & Woolard (2000) come to similar conclusions, therefore justifying the use of the broad definition.

In order to calculate the measure of broad unemployment, constructed variables provided within the dataset are used for 2005, 2006 and 2007. However, such a variable is not provided for 2008, and thus this is duly constructed according to the definition of broad unemployment. In order to maintain consistency, and comparable results, the same batch of controls are utilised within each period. On an individual level, age, gender, race, marital status, whether an individual is head of the household, education and provincial location are duly controlled for. Furthermore, whether the house is owned, the number of children (<16) and the number of elderly (>65) are controlled for on a household level.

There are, however, some limitations with this dataset. As the GHS survey is a cross sectional study, there is an inability to track an individuals change over time. It gives us only a snapshot of a person’s employment at any particular point in time, and there is an inability to analyse over the period of interest. Furthermore, post 2008, Stats SA no longer asks employment related questions within the GHS questionnaire. Instead, all questions relating to household activity are asked in the GHS, while all labour related information is asked in the Quarterly Labour Force Survey (QLFS). As a result of this, all GHS datasets from 2009 onward contain no information regarding labour market activity, while the QLFS contains no information regarding household information. The inability in the QLFS to allow the researcher to examine household effects would give a skewed view of the individual, thus limiting the analytical power of the study. This therefore renders both datasets powerless for this study, and data analysis using the QLFS is not included within the results. Furthermore, the period of interest is limited to 2005-2008, and unfortunately cannot truly examine the effects which the recent global

recession has on the labour market within South Africa, post 2008, but rather gives an initial indication of movements. However, it is duly acknowledged that the topic of discrimination is one which must be examined over an extended period of time. It would be recommended that an extension to this piece would include more datasets, to examine structural changes in the topic at hand.

Having illustrated the limitations within the datasets, a further statistical limitation arises, that of omitted variable bias. Given that the QLFS is not used with statistical analysis, certain factors are likely underestimated, leading to the possibility of omitted variable bias. As a result, readers must be aware of the issue when examining the results.

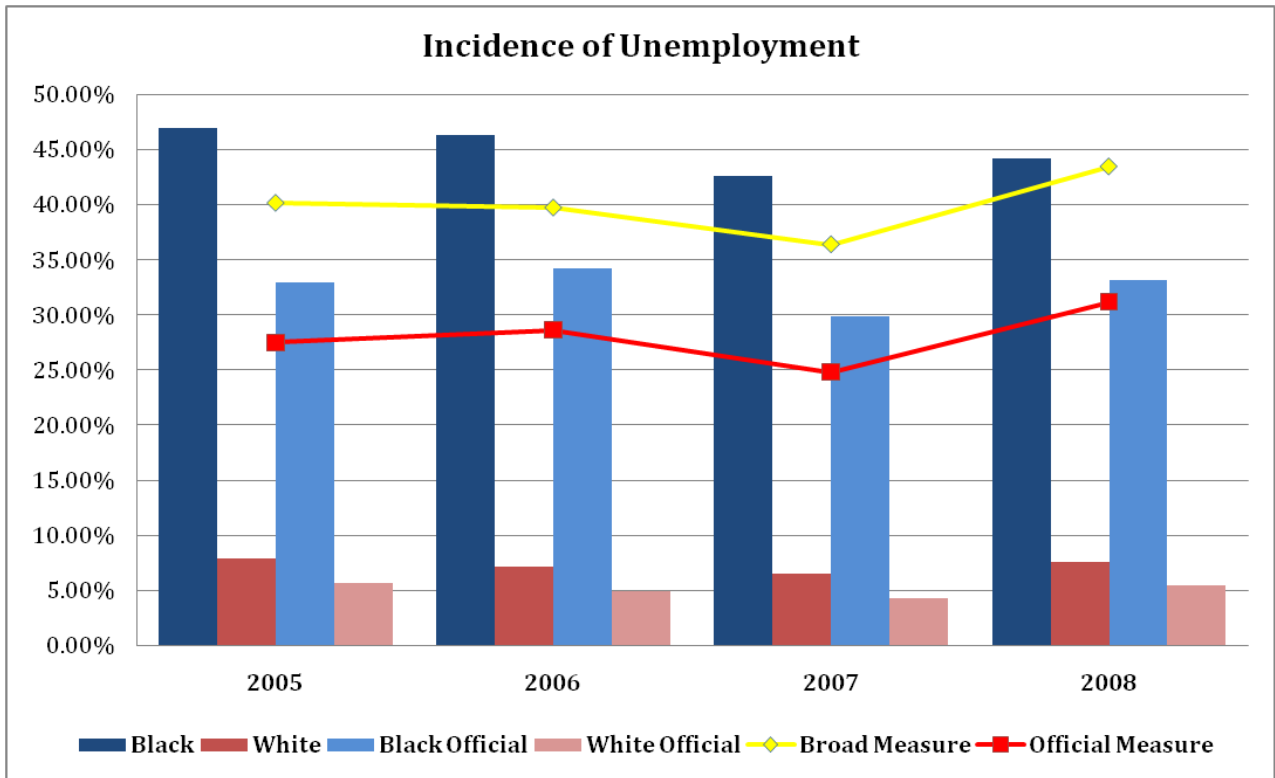
University of Cape Town

## DESCRIPTIVE STATISTICS

**Table 1: Unemployment Rates by Category<sup>3</sup>**

	<b>2005</b>			<b>2006</b>			<b>2007</b>			<b>2008</b>		
<b>Overall</b>	<b>Broad</b>	<b>Official</b>	<b>Gap</b>	<b>Broad</b>	<b>Official</b>	<b>Gap</b>	<b>Broad</b>	<b>Official</b>	<b>Gap</b>	<b>Broad</b>	<b>Official</b>	<b>Gap</b>
	<b>40.17%</b>	<b>27.50%</b>	<i>12.67%</i>	<b>39.75%</b>	<b>28.63%</b>	<i>11.12%</i>	<b>36.38%</b>	<b>24.82%</b>	<i>11.56%</i>	<b>43.51%</b>	<b>31.19%</b>	<i>12.32%</i>
<b>Age</b>												
16-24	68.58%	54.27%	<i>14.31%</i>	67.69%	55.18%	<i>12.51%</i>	62.68%	48.52%	<i>14.16%</i>	65.47%	53.58%	<i>11.89%</i>
25-35	41.89%	30.09%	<i>11.80%</i>	41.97%	31.58%	<i>10.39%</i>	38.75%	27.44%	<i>11.31%</i>	40.47%	31.14%	<i>9.33%</i>
36-45	28.49%	18.40%	<i>10.09%</i>	27.54%	18.40%	<i>9.14%</i>	25.56%	16.14%	<i>9.42%</i>	24.85%	17.56%	<i>7.29%</i>
46-55	21.19%	12.00%	<i>9.19%</i>	20.13%	12.93%	<i>7.20%</i>	18.39%	11.55%	<i>6.84%</i>	18.05%	11.70%	<i>6.35%</i>
56-64	13.18%	6.55%	<i>6.63%</i>	15.42%	8.57%	<i>6.85%</i>	12.50%	7.08%	<i>5.42%</i>	10.65%	6.07%	<i>4.58%</i>
<b>Education</b>												
None	37.50%	21.56%	<i>15.94%</i>	39.09%	24.64%	<i>14.45%</i>	32.03%	18.02%	<i>14.01%</i>	58.53%	41.79%	<i>16.74%</i>
Primary	45.10%	28.84%	<i>16.26%</i>	42.95%	28.65%	<i>14.30%</i>	39.40%	24.70%	<i>14.70%</i>	54.35%	41.84%	<i>12.51%</i>
Junior	44.67%	32.16%	<i>12.51%</i>	43.90%	32.98%	<i>10.92%</i>	41.12%	29.47%	<i>11.65%</i>	44.15%	34.06%	<i>10.09%</i>
Secondary	18.21%	13.33%	<i>4.88%</i>	18.13%	13.89%	<i>4.24%</i>	15.13%	10.83%	<i>4.30%</i>	23.48%	18.04%	<i>5.44%</i>
Higher	7.09%	4.91%	<i>2.18%</i>	4.88%	3.86%	<i>1.02%</i>	4.78%	3.00%	<i>1.78%</i>	11.66%	9.08%	<i>2.58%</i>
<b>Gender</b>												
Male	32.34%	23.30%	<i>9.04%</i>	32.77%	23.11%	<i>9.66%</i>	30.00%	20.18%	<i>9.82%</i>	30.96%	22.79%	<i>8.17%</i>
Female	48.76%	32.80%	<i>15.96%</i>	53.11%	40.38%	<i>12.73%</i>	48.48%	34.62%	<i>13.86%</i>	49.65%	37.57%	<i>12.08%</i>
<b>Race</b>												
Black	47.01%	32.90%	<i>14.11%</i>	46.37%	34.24%	<i>12.13%</i>	42.64%	29.83%	<i>12.81%</i>	44.21%	33.19%	<i>11.02%</i>
Coloured	29.23%	21.90%	<i>7.33%</i>	29.33%	21.25%	<i>8.08%</i>	26.76%	18.77%	<i>7.99%</i>	25.32%	19.46%	<i>5.86%</i>
Indian	19.83%	12.50%	<i>7.33%</i>	25.90%	18.71%	<i>7.19%</i>	14.09%	8.31%	<i>5.78%</i>	18.02%	14.16%	<i>3.86%</i>
White	7.91%	5.70%	<i>2.21%</i>	7.15%	4.92%	<i>2.23%</i>	6.52%	4.29%	<i>2.23%</i>	7.63%	5.43%	<i>2.20%</i>

<sup>3</sup> The values are calculated according to the Broad and Narrow definitions respectively, using the GHS datasets



**Figure 1: Incidence of Unemployment**

Table 1 and Figure 1 give an indication of the underlying data used with the study. It is evident that both the official and broad measures of unemployment steadily decrease from 2005, up until 2007, before both measures increase dramatically. This increase in unemployment can clearly be illustrated when examining the effects it has had in education, as individuals with lower levels of education appear to find it far more difficult to find employment in 2008. Females consistently find the workplace more difficult than their male counterparts.

## **METHODOLOGY**

### **Probit Analysis**

In order to get an indication as to the probability of unemployment for each race group, the technique of probit analysis is used. This multivariate technique gives a good initial understanding of the data, by giving an indication of an individual's probability of unemployment, holding other indicators constant. The benefit is a tentative indication as to the policy measures which may be taken to increase the chances of employment for individuals. The dependant variable takes on a value of "1" for unemployed, and "0" for employed, and as discussed, follows the definition of broad unemployment.

Table 2 shows the pooled probit models using GHS datasets for the years 2005-2008. This set of results illustrates how various controls have altered the probability of employment for all race groups over a 4 year period. Tables 3 and 4 follow the same analysis, however here each race group (black and white) is isolated, such that one is

able to specifically examine the controls which influence that particular group. Analysis is taken over the period 2005 and 2008, to get an indication as to how circumstances have changed over the period of interest. It provides a “snapshot” of the situation at one point, in comparison to another.

Kingdon & Knight (2004), discuss the limitations in the exclusive use of probit analysis in order to examine employment discrimination. They indicate that there lies an inherent difficulty in distinguishing between job-rationing reasons and worker preferences, furthermore these models “are largely unable to distinguish between the constraints on and preferences for employment since their effects are not readily separable” (Kingdon & Knight, 2004). For these reasons, the inclusion of probit analysis is merely used as a tentative indication in to employment probability, and is principally examined for their use in decomposition explanation.

Examining Table 2, it is clear that throughout the period 2005-2008, black individuals are more likely to be unemployed than any other race group, this is not surprising in light of earlier discussion. Table 2 indicates, even after controlling for both individual endowment effects, as well as household level variables, black individuals are 30% more likely to be unemployed than their white counterparts. This is a significant effect, and implies that the probability of unemployment for black individuals has increased from 20% in 1994 (See Kingdon & Knight, 2004), to over 35% in 2008. This is very disconcerting, as many would have hoped for better employment conditions with the arrival of a democratic society.

The probability of unemployment decreases with age but at a decreasing rate. As individuals grow older, so too will their knowledge, and experience of the workplace, thus increasing the chances of employment. There is a clear disadvantage for female labourers, as males face up to a 10% less chance of becoming unemployed in 2008, however the effect is far more apparent with an examination of regressions restricted to black and white individuals. The effect is not always significant, or the effect is very small amongst whites, however, amongst black respondents, there is a significant, and increasing difference between the probabilities of employment between male and female.

Unsurprisingly, “household head” status significantly reduces the probability of unemployment across all regressions (tables 2, 3 & 4), but once again the effect is more prominent within black households. This is true also for married individuals. These results are common to the economic theory that household heads, and married persons are faced with far more responsibility, and tend to be more mature and trustworthy, endowing them with a greater chance of employment opportunities<sup>4</sup>.

The influence of home ownership has been investigated by Oswald (1997). Specifically, it is proposed in this study that there are in fact two opposing effects which

---

<sup>4</sup> See Kingdon & Knight 2004, and also Oppenheimer (1997) and Hundley (2000)

homeownership could possibly have on unemployment. Firstly, homeownership could increase an individual's immobility, and could be considered a significant proxy for household wealth, and thus homeowners tend to demand higher wages, leading to an increase in the probability of unemployment (a positive coefficient). A more prominent and cited finding from the Oswald paper is "that homeowners are less willing than private renters to move to jobs when they become unemployed because owners have larger costs of moving" (Green & Hendershott, 2001). However, home ownership may have an endogenous influence, for example, being employed will determine one's chance of owning a house, and thus the coefficient will be negative (Kingdon & Knight, 2004). An examination of the pooled estimation (table 2) shows that homeownership will lead to an increase in the probability of unemployment, indicating that Oswald's hypothesis holds true in this case. However, when the model is split up according to race (tables 3 & 4), the picture becomes interesting. It would appear that homeownership for white individuals implies there may well be an endogenous effect at work, as owning a house will decrease the probability of unemployment. Conversely, the argument for black individuals is the opposite. In this instance, home ownership significantly increases the chance of unemployment, and this is a powerful indication that black workers are perhaps immobile, and would rather maintain their household than relocate and find employment. This however places these individuals at a disadvantage.

The number of children and elderly, within a household, although mostly statistically significant has a very limited effect when examining the coefficients. Across all probits, both coefficients are positive, signalling that individuals will reduce their job search when there are household members who need care or attention, and interestingly, the effect of having children has a greater effect amongst white respondents.

As would be expected, as the years of schooling increase the probability of unemployment decreases. Further education, and tertiary level learning, has by far the greatest effect on unemployment. This is particularly true for black respondents, as tertiary education has the effect of reducing unemployment by up to 50%. This effect however is reduced in 2008, but as descriptive statistics have shown, this can be accounted for by an increase in the overall level of unemployment. However, it may be a situation where the quality of one's schooling is a powerful determinant of whether one will become employed. Kingdon & Knight (2004) employ individual cognitive scores in their analysis; however such powerful data is not available within the GHS framework, but should be investigated with new data that is soon to be available within South Africa<sup>5</sup>.

Of particular interest to this study is the large gap between the unemployment rates between black and white individuals illustrated in Figure 2. This gap ranges between 36% and 39% when using the raw data, despite movements in the overall levels of unemployment. When controlling for both individual and household characteristics

---

<sup>5</sup> See the National Income Dynamic Study (NIDS)

using probit analysis however, the gap is reduced to between 32% and 35%. The difference in the unemployment gaps obtained from the raw data and that of probit analysis, indicate that there is some unexplained portion which cannot be controlled for with traditional regression (or probit) analysis. It is this gap which will be examined with this paper, and investigate the extent to which it exists as a result of differing endowments between groups, or if there is in fact a level of discrimination within the South African labour market.

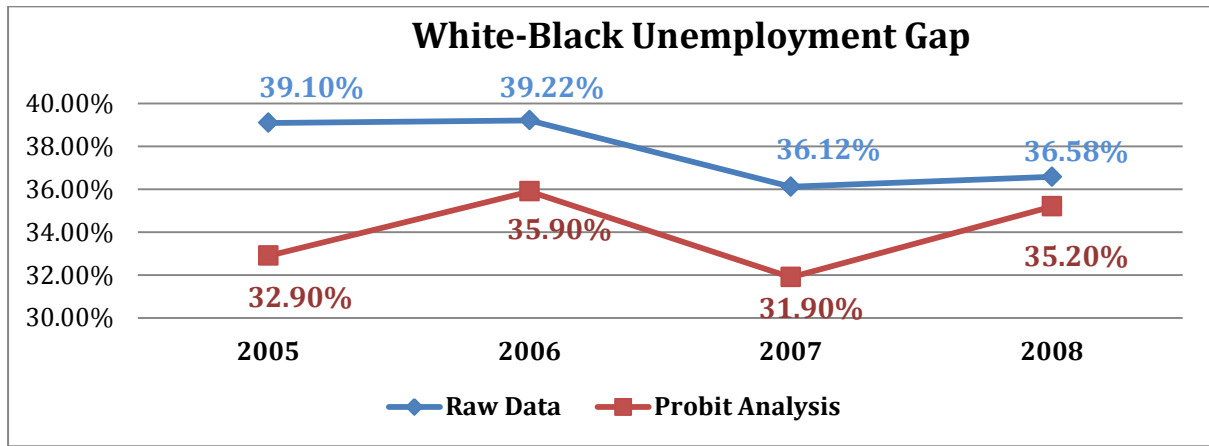


Figure 2: The White-Black Unemployment Gap, (Source, GHS and own calculations)

University of Cap

Table 2: Pooled Probit Regression, Source GHS 2005-2008<sup>6</sup>

VARIABLES <sup>a</sup>	2005			2006			2007			2008		
	Coefficient	Marginal Effect	T-Stat	Coefficient	Marginal Effect	T-Stat	Coefficient	Marginal Effect	T-Stat	Coefficient	Marginal Effect	T-Stat
<i>Age</i>												
age1	<b>-0.0700***</b>	-0.0262***	-9.96	<b>-0.0818***</b>	-0.0303***	-12.69	<b>-0.0764***</b>	-0.0271***	-11.11	<b>-0.120***</b>	-0.0440***	-20.00
age2	<b>0.000578***</b>	0.000216***	6.22	<b>0.000744***</b>	0.000275***	8.74	<b>0.000670***</b>	0.000238***	7.35	<b>0.00103***</b>	0.000376***	12.96
<i>Individual</i>												
Male <sup>^</sup>	<b>-0.200***</b>	-0.0749***	-8.09	<b>-0.221***</b>	-0.0817***	-9.40	<b>-0.217***</b>	-0.0768***	-8.89	<b>-0.281***</b>	-0.103***	-12.77
Head <sup>^</sup>	<b>-0.702***</b>	-0.263***	-30.47	<b>-0.717***</b>	-0.265***	-26.55	<b>-0.658***</b>	-0.233***	-25.80	<b>-0.0406*</b>	-0.0149*	-1.86
Married <sup>^</sup>	<b>-0.192***</b>	-0.0720***	-7.50	<b>-0.198***</b>	-0.0735***	-7.49	<b>-0.153***</b>	-0.0544***	-5.37	<b>-0.000188</b>	-6.87e-05	-0.01
<i>Household</i>												
Own <sup>^</sup>	<b>0.419***</b>	0.157***	12.96	<b>0.364***</b>	0.135***	10.24	<b>0.405***</b>	0.144***	10.52	<b>0.452***</b>	0.166***	14.43
totchild	<b>0.0716***</b>	0.0268***	9.60	<b>0.0774***</b>	0.0287***	9.51	<b>0.0831***</b>	0.0295***	10.72	<b>0.0577</b>	0.0211	1.62
totelder	<b>0.110***</b>	0.0411***	4.02	<b>0.0900***</b>	0.0333***	2.89	<b>0.0630**</b>	0.0223**	2.20	<b>0.240***</b>	0.0879***	7.70
<i>Race</i>												
Black <sup>^</sup>	<b>0.880***</b>	0.329***	14.07	<b>0.969***</b>	0.359***	14.73	<b>0.901***</b>	0.319***	11.69	<b>0.960***</b>	0.352***	14.30
Coloured <sup>^</sup>	<b>0.525***</b>	0.197***	6.36	<b>0.669***</b>	0.248***	7.80	<b>0.534***</b>	0.189***	5.38	<b>0.638***</b>	0.234***	8.07
Indian <sup>^</sup>	<b>0.343***</b>	0.128***	3.56	<b>0.603***</b>	0.223***	6.30	<b>0.157</b>	0.0556	1.15	<b>0.284***</b>	0.104***	2.61
<i>Education</i>												
Primary <sup>^</sup>	<b>-0.0132</b>	-0.00496	-0.30	<b>-0.0877*</b>	-0.0325*	-1.87	<b>0.0230</b>	0.00815	0.51	<b>-0.0171</b>	-0.00628	-0.43
Matric <sup>^</sup>	<b>-0.147***</b>	-0.0548***	-3.32	<b>-0.190***</b>	-0.0703***	-4.21	<b>-0.0782*</b>	-0.0277*	-1.65	<b>-0.178***</b>	-0.0653***	-4.50
Furthered <sup>^</sup>	<b>-0.729***</b>	-0.273***	-12.64	<b>-0.777***</b>	-0.288***	-12.41	<b>-0.683***</b>	-0.242***	-10.66	<b>-0.565***</b>	-0.207***	-10.62
Tertiary <sup>^</sup>	<b>-0.976***</b>	-0.365***	-9.38	<b>-1.216***</b>	-0.450***	-12.10	<b>-1.027***</b>	-0.364***	-7.77	<b>-0.719***</b>	-0.264***	-7.33
<i>Province</i>												
Westc <sup>^</sup>	<b>-0.214***</b>	-0.0799***	-3.66	<b>-0.399***</b>	-0.148***	-7.13	<b>-0.181***</b>	-0.0640***	-3.07	<b>-0.250***</b>	-0.0916***	-5.20
Eastcape <sup>^</sup>	<b>0.116***</b>	0.0434***	2.58	<b>0.00356</b>	0.00132	0.08	<b>-0.0433</b>	-0.0153	-0.65	<b>-0.219***</b>	-0.0803***	-4.85
Ncape <sup>^</sup>	<b>0.120**</b>	0.0449**	2.04	<b>-0.0377</b>	-0.0139	-0.72	<b>0.0101</b>	0.00358	0.17	<b>0.178***</b>	0.0652***	3.03
Fstate <sup>^</sup>	<b>0.0285</b>	0.0107	0.60	<b>-0.00784</b>	-0.00290	-0.17	<b>-0.0230</b>	-0.00814	-0.44	<b>0.00916</b>	0.00336	0.19
Kzn <sup>^</sup>	<b>-0.0114</b>	-0.00428	-0.28	<b>0.0296</b>	0.0109	0.64	<b>-0.111***</b>	-0.0392***	-2.61	<b>0.0382</b>	0.0140	0.98
Nw <sup>^</sup>	<b>0.0804*</b>	0.0301*	1.67	<b>0.00198</b>	0.000734	0.04	<b>0.00348</b>	0.00123	0.07	<b>0.0414</b>	0.0152	0.80
Mpu <sup>^</sup>	<b>-0.104**</b>	-0.0389**	-2.23	<b>-0.130***</b>	-0.0483***	-2.67	<b>-0.136***</b>	-0.0483***	-2.80	<b>-0.0448</b>	-0.0164	-1.04
Limp <sup>^</sup>	<b>0.333***</b>	0.125***	6.56	<b>0.305***</b>	0.113***	5.92	<b>0.305***</b>	0.108***	6.80	<b>0.238***</b>	0.0873***	4.89
Constant	<b>0.837***</b>		5.96	<b>1.067***</b>		7.47	<b>0.797***</b>		5.28	<b>1.611***</b>		11.24
<b>Population</b>	<b>19287056</b>			<b>19771112</b>			<b>19852255</b>			<b>20252243</b>		
<b>Observations</b>	<i>41,504</i>			<i>41,570</i>			<i>42,692</i>			<i>37,057</i>		

<sup>6</sup> ^ Indicates Indicator 1/0. Dependant Variable: Unemployed=1, Employed=0 according to the broad definition:  
Base case: *White, female, not household head, never married, no schooling, Gauteng* <sup>a</sup>(\*\*\* p<0.01, \*\* p<0.05, \* p<0.1)

Table 3: Probit Regression for Black Respondents only, Source GHS 2005-2008<sup>7</sup>

VARIABLES <sup>a</sup>	2005			2006			2007			2008		
	Coefficient (Black)	Marginal Effect	T-Stat	Coefficient (Black)	Marginal Effect	T-Stat	Coefficient (Black)	Marginal Effect	T-Stat	Coefficient (Black)	Marginal Effect	T-Stat
<i>Age</i>												
age1	-0.0650***	-0.0259***	-13.42	-0.0736***	-0.0293***	-15.26	-0.0649***	-0.0255***	-13.67	-0.108***	-0.0426***	-22.71
age2	0.000453***	0.000181***	7.09	0.000566***	0.000225***	8.95	0.000448***	0.000176***	7.13	0.000844***	0.000334***	13.43
<i>Individual</i>												
Male <sup>^</sup>	-0.185***	-0.0738***	-11.02	-0.223***	-0.0887***	-13.33	-0.205***	-0.0803***	-12.5	-0.268***	-0.106***	-16.27
Head <sup>^</sup>	-0.656***	-0.261***	-34.86	-0.657***	-0.262***	-35.14	-0.634***	-0.249***	-34.25	-0.0446**	-0.0177**	-2.42
Married <sup>^</sup>	-0.174***	-0.0693***	-8.37	-0.138***	-0.0551***	-6.67	-0.117***	-0.0459***	-5.66	0.0341*	0.0135*	1.76
<i>Household</i>												
Own <sup>^</sup>	0.641***	0.255***	29.52	0.637***	0.254***	28.99	0.616***	0.242***	28.88	0.563***	0.223***	25.68
totchild	0.0581***	0.0231***	11.86	0.0643***	0.0256***	12.52	0.0668***	0.0262***	13.53	0.0472*	0.0187*	1.88
totelder	0.121***	0.0482***	5.75	0.105***	0.0419***	5.07	0.0881***	0.0346***	4.48	0.246***	0.0973***	12.11
<i>Education</i>												
Primary <sup>^</sup>	-0.109***	-0.0434***	-3.4	-0.125***	-0.0499***	-3.91	-0.0532	-0.0209	-1.55	-0.0135	-0.00536	-0.4
Matric <sup>^</sup>	-0.191***	-0.0760***	-6.06	-0.223***	-0.0889***	-7.14	-0.167***	-0.0657***	-4.97	-0.163***	-0.0644***	-5.12
Furthered <sup>^</sup>	-0.822***	-0.327***	-18.51	-0.865***	-0.345***	-18.74	-0.857***	-0.337***	-18.49	-0.544***	-0.215***	-12.19
Tertiary <sup>^</sup>	-1.249***	-0.498***	-14.84	-1.408***	-0.561***	-15.42	-1.285***	-0.505***	-14.85	-0.822***	-0.325***	-11.02
<i>Province</i>												
Westc <sup>^</sup>	-0.345***	-0.137***	-7.47	-0.315***	-0.126***	-7.11	-0.283***	-0.111***	-6.66	-0.325***	-0.129***	-8.15
Eastcape <sup>^</sup>	0.0774**	0.0308**	2.44	-0.0348	-0.0139	-1.1	-0.0506	-0.0199	-1.63	-0.314***	-0.124***	-10.3
Ncape <sup>^</sup>	-0.180***	-0.0718***	-3.49	-0.123**	-0.0491**	-2.5	-0.106**	-0.0418**	-2.14	0.170***	0.0673***	3.88
Fstate <sup>^</sup>	0.0121	0.00481	0.35	0.00473	0.00188	0.14	0.00465	0.00183	0.14	-0.0556*	-0.0220*	-1.8
Kzn <sup>^</sup>	-0.0531**	-0.0212**	-1.98	0.0163	0.00650	0.61	-0.0427	-0.0168	-1.63	-0.0603**	-0.0239**	-2.28
Nw <sup>^</sup>	0.116***	0.0462***	3.54	0.114***	0.0454***	3.47	0.115***	0.0452***	3.53	-0.0460	-0.0182	-1.48
Mpu <sup>^</sup>	-0.191***	-0.0759***	-5.78	-0.152***	-0.0608***	-4.54	-0.160***	-0.0629***	-4.92	-0.146***	-0.0577***	-4.91
Limp <sup>^</sup>	0.238***	0.0947***	7.25	0.267***	0.106***	8.08	0.261***	0.103***	8.11	0.151***	0.0596***	5.04
Constant	1.620***		17.2	1.798***		19.21	1.512***		16.3	2.353***		25.49
<b>Observations</b>	<b>31,017</b>			<b>31,342</b>			<b>32,398</b>			<b>29,539</b>		
LR chi2(20) =	9839.79			9948.93			9889.1			6586.67		
Log likelihood =	-16568.32			-16742.249			-17347.336			-17072.972		
Pseudo R2 =	0.229			0.2291			0.2218			0.1617		

a(\*\*\* p<0.01, \*\* p<0.05, \* p<0.1)

<sup>7</sup> ^ Indicates Indicator 1/0. Dependant Variable: Unemployed=1, Employed=0 according to the broad definition: Base case: Female, not household head, never married, no schooling, Gauteng

**Table 4: Probit Regression for White Respondents only, Source GHS 2005-2008<sup>8</sup>**

VARIABLES <sup>a</sup>	2005			2006			2007			2008		
	Coefficient (White)	Marginal Effect	T-Stat	Coefficient (White)	Marginal Effect	T-Stat	Coefficient (White)	Marginal Effect	T-Stat	Coefficient (White)	Marginal Effect	T-Stat
<i>Age</i>												
age1	-0.0555***	-0.00640**	-2.78	-0.0662***	-0.00670***	-3.13	-0.0825***	-0.00749***	-3.76	-0.152***	-0.0152*	-6.85
age2	0.000587**	6.77e-05**	2.35	0.000804***	8.14e-05***	3.05	0.000983***	8.93e-05***	3.65	0.00154***	0.000154*	5.4
<i>Individual</i>												
Male <sup>^</sup>	-0.0952	-0.0110	-0.93	-0.376***	-0.0381***	-3.82	-0.0712	-0.00646	-0.65	-0.137	-0.0137	-1.17
Head <sup>^</sup>	-0.694***	-0.0800***	-8.61	-0.697***	-0.0706***	-7.7	-0.563***	-0.0511***	-5.97	0.160*	0.0160	1.72
Married <sup>^</sup>	-0.180**	-0.0208*	-1.97	-0.314***	-0.0318***	-3.18	-0.148	-0.0134	-1.37	0.0657	0.00657	0.66
<i>Household</i>												
Own <sup>^</sup>	-0.167**	-0.0193**	-2.3	-0.0256	-0.00259	-0.31	-0.312***	-0.0283***	-3.69	-0.0506	-0.00506	-0.5
totchild	0.0937**	0.0108**	2.54	0.0742*	0.00751*	1.78	0.0133	0.00121	0.29	0.168	0.0168	0.76
totelder	0.128	0.0148	1.37	0.0630	0.00637	0.63	0.124	0.0112	1.17	-0.0460	-0.00461	-0.38
<i>Education</i>												
Primary <sup>^</sup>	4.268	0.492	0.05	-1.369*	-0.139*	-1.69	-4.491	-0.408	-0.03	3.300	0.330	0.02
Matric <sup>^</sup>	3.719	0.429	0.04	-1.052	-0.107	-1.53	-5.224	-0.474	-0.04	3.214	0.322	0.02
Furthered <sup>^</sup>	3.359	0.388	0.04	-1.393**	-0.141**	-2	-5.588	-0.507	-0.04	2.905	0.291	0.02
Tertiary <sup>^</sup>	2.987	0.345	0.03	-1.617**	-0.164**	-2.29	-5.786	-0.525	-0.04	2.768	0.277	0.02
<i>Province</i>												
Westc <sup>^</sup>	-0.0596	-0.00688	-0.6	-0.271**	-0.0274**	-2.48	-0.145	-0.0132	-1.2	-0.207	-0.0207	-1.49
Eastcape <sup>^</sup>	-0.0457	-0.00527	-0.38	-0.245*	-0.0248*	-1.82	-0.162	-0.0147	-1.15	-0.349*	-0.0349	-1.9
Ncape <sup>^</sup>	-0.274*	-0.0317*	-1.91	-0.429***	-0.0434***	-2.75	-0.245	-0.0222	-1.4	-0.179	-0.0180	-0.96
Fstate <sup>^</sup>	-0.0524	-0.00604	-0.4	0.0274	0.00278	0.2	0.141	0.0128	1	0.170	0.0170	1.1
Kzn <sup>^</sup>	-0.106	-0.0123	-0.79	-0.0623	-0.00630	-0.46	-0.144	-0.0131	-0.92	0.00300	0.000301	0.02
Nw <sup>^</sup>	0.0292	0.00337	0.18	-0.317*	-0.0321*	-1.82	-0.290	-0.0263	-1.49	0.169	0.0169	0.97
Mpu <sup>^</sup>	0.0434	0.00501	0.29	-0.183	-0.0185	-0.92	-0.184	-0.0167	-0.91	0.0718	0.00718	0.43
Limp <sup>^</sup>	-0.0746	-0.00861	-0.4	-0.249	-0.0252	-1.22	0.460**	0.0418**	2.34	-0.418	-0.0418	-1.22
Constant	-3.238		-0.03	1.836**		2.34	5.995		0.04	-1.196		-0.01
<b>Observations</b>	<b>3,305</b>			<b>2,993</b>			<b>2,864</b>			<b>2,263</b>		
LR chi2(20) =	290.2			242.01			196.9			180.53		
Log likelihood =	-839.78675			-687.60424			-592.60536			-495.52163		
Pseudo R2 =	0.1473			0.1496			0.1425			0.1541		

<sup>a</sup>(\*\*\* p<0.01, \*\* p<0.05, \* p<0.1)

<sup>8</sup> <sup>^</sup> Indicates Indicator 1/0. Dependant Variable: Unemployed=1, Employed=0 according to the broad definition:  
Base case: Female, not household head, never married, no schooling, Gauteng

## **Decomposition of the Gap in Unemployment Probability and Sensitivity Analysis**

An examination of the descriptive statistics (Table 1) shows the raw unemployment rates for each race group. Furthermore, Figure 2 gives an indication of the unemployment gap between white and black individuals. Using the broad definition, and the raw data, the unemployment gap between white and black individuals' moves between 39% and 36%. Table 2, however shows that this gap is somewhat reduced when controlling for both individual and household characteristics using probit analysis. The difference in the raw unemployment gap, and the information obtained from probit analysis, gives an indication as to the level of discrimination. For example, using the raw data for 2005, the unemployment gap between blacks and whites was 39.1%, while using the pooled unemployment probit however; results indicate that the gap is reduced to 32.9%, a difference of 6.2%. This indicates that 16 percentage points<sup>9</sup> of the 39.1% of the employment gap can be attributed to differences in observed characteristics, while the remaining portion (84 percentage points) is attributed to differences in these characteristics, or discrimination. In summation, the unexplained residual (84% of the gap) can be explained by racial discrimination. A similar procedure can be applied to other years within table 2, as well as to tables 3 & 4.

However, there are limitations to this measure, and one cannot come to robust conclusions using probit analysis alone. Kingdon & Knight (2004), point out the "restrictive assumption that the probit of unemployment is identical for blacks and whites in all respects except the intercept", and thus little can be said in terms of labour market discrimination<sup>10</sup>. Therefore, in order to overcome this restriction within the data analysis, decomposition techniques proposed, initially deduced by Oaxaca (1973) and Blinder (1973), and extended by Daymont & Andriasani, Cotton (1988) and Reimers (1983) are employed.

Traditionally, the Oaxaca-Blinder technique makes use of linear regression analysis, as in this instance the decomposition only requires "coefficient estimates from linear regressions for the outcome of interest and sample means of the independent variables used in the regressions" (Fairlie, 2003). Specifically, the fitted line in a linear regression passes through the mean of the independent variables, and thus "the dot product of the vector of coefficients and the vector of mean variable values gives the mean of the dependent variable" (Kingdon & Knight, 2004). There is of course a difference when using discrete choice models such as probit analysis. The more prominent difference lies in the fact that the real mean of the dependant variable, and the mean obtained within the regression need not be the same. However, previous research by Blackaby et al (1998), Kingdon & Knight (2004), and Brookes & Hinks (2004), have all found this restriction to be negligible. Furthermore, Sinning, Hahn & Bauer (2008), have found that the use of "ordinary least squares (OLS) yields inconsistent parameter estimates and, in

---

<sup>9</sup>  $6.1/39.1=0.156=16\%$

<sup>10</sup> One could incorrectly conclude that discrimination has been reduced to 4% in 2008

turn, misleading decomposition results”, and therefore the extension of the Oaxaca technique to non-linear modelling is justified.

Analysis follows the methodology found within Kingdon & Knight (2004), as well as Daymont & Andrisani (1984), with extensions derived by Cotton (1988) and Reimers (1983). Bauer and Sinning (2008), derived an extension to the Oaxaca Blinder technique<sup>11</sup>, specifically for use with nonlinear models, and this is the model employed. It is clear from descriptive statistics, as well as traditional probit analysis, that white individuals have better employment prospects within the work place than any other race group. It is of interest therefore, to decompose the employment gap between white workers and their black counterparts, who face the highest levels of unemployment.

Before establishing the methodology used for the nonlinear case, it is important to understand the original model as Oaxaca & Blinder (1973) established.

Consider the linear regression, fitted separately for whites (**w**) and blacks (**b**): where:  $g = (w, b)$

$$Y_{ig} = X_{ig}\beta_g + \varepsilon_{ig}$$

Where  $i=1, \dots, N_g$  and  $\sum_g N_g = N$  Thus for these separate regressions, the following decomposition is applied

$$\bar{Y}_w - \bar{Y}_b = \Delta^{OLS} = (\bar{X}_w - \bar{X}_b)\hat{\beta}_A + \bar{X}_b(\hat{\beta}_w - \hat{\beta}_b) \quad [1]$$

Or  $A = B + C$

Where  $A = \bar{Y}_w - \bar{Y}_b$ ,  $B = (\bar{X}_w - \bar{X}_b)\hat{\beta}_A$ ,  $C = \bar{X}_b(\hat{\beta}_w - \hat{\beta}_b)$

And  $\bar{Y}_g = \frac{1}{N_g} \sum_{i=1}^{N_g} Y_{ig}$ ,  $\bar{X}_g = \frac{1}{N_g} \sum_{i=1}^{N_g} X_{ig}$

We can interpret the equation [1] as follows. **B** explains the difference between black and white employees due differences in observed characteristics, or the “endowment affect” Jann (2008), this is considered the “discrimination” component. **C** shows the difference due to a differential in coefficient estimates.

As already stated, it would not be appropriate to use this method when utilising the probit model, or any nonlinear model, as the conditional expectations  $E(Y_{ig}|X_{ig})$  may differ from actual mean  $(\bar{X}_g\hat{\beta}_g)$ . Thus, Bauer and Sinning (2008) propose the following version using conditional expectations. Here, the reference group is white individuals, and thus the decomposition is analysed according to their probit regression.

$$\Delta_w^{NL} = \{E_{\beta_w}(Y_{iw}|X_{iw}) - E_{\beta_w}(Y_{ib}|X_{ib})\} + \{E_{\beta_w}(Y_{ib}|X_{ib}) - E_{\beta_b}(Y_{ib}|X_{ib})\} \quad [2]$$

---

<sup>11</sup>The “nldecompose” function within STATA

Where  $E_{\beta g}(Y_{ig}|X_{ig})$  refers to the conditional expectation of  $Y_{ig}$  and  $E_{\beta g}(Y_{ih}|X_{ih})$  refers to the conditional expectation of  $Y_{ih}$  evaluated at the parameter vector  $\beta_g$ , with  $g = (w, b)$  and  $g \neq b$

The interpretation is the same as for the linear regression, in the form  $\mathbf{A} = \mathbf{B} + \mathbf{C}$ .

An extension by Daymont and Andrisani (1984) decompose the model into 3 parts as follows:

$$\bar{Y}_w - \bar{Y}_b = (\bar{X}_w - \bar{X}_b)\beta_b + \bar{X}_b(\beta_w - \beta_b) + (\bar{X}_w - \bar{X}_b)(\beta_w - \beta_b) = \mathbf{E} + \mathbf{C} + \mathbf{CE} \quad [3]$$

Where in general the components are given as follows;

$$\mathbf{E} = \{E_{\beta b}(Y_{iw}|X_{iw}) - E_{\beta b}(Y_{ib}|X_{ib})\}$$

$$\mathbf{C} = \{E_{\beta w}(Y_{ib}|X_{ib}) - E_{\beta b}(Y_{ib}|X_{ib})\}$$

$$\mathbf{CE} = \{E_{\beta w}(Y_{iw}|X_{iw}) - E_{\beta b}(Y_{iw}|X_{iw})\} + \{E_{\beta w}(Y_{ib}|X_{ib}) - E_{\beta b}(Y_{ib}|X_{ib})\}$$

Here  $\mathbf{E}$  represents the differential due to the difference in endowments which black and white respondents carry. Specifically, it is the employment gap accounted for by differences in the individual characteristics which each group is endowed. It is termed the “explained” portion of the gap, as it reveals how the employment gap would change of black individuals had the same characteristics as white individuals, but the way that the labour market treated each group did not change.  $\mathbf{C}$  is once more attributable to the difference in coefficients. It shows how the employment gap would change if the labour market rewarded both groups according the white regression. This portion is considered to be a measure of discrimination. It is often termed the “unexplained” component of the gap, as given the fact that both groups are rewarded according to the same regression; there is no reason why the gap should exist. The last term  $\mathbf{CE}$  explains the interaction between terms  $\mathbf{C}$  and  $\mathbf{E}$ , and accounts for the fact that “differences in endowments and coefficients exist simultaneously between the two groups” (Jann, 2008).

Oaxaca and Ransom (1994) illustrated the following generalised linear decomposition:

$$\bar{Y}_w - \bar{Y}_b = (\bar{X}_w - \bar{X}_b)\beta^* + \bar{X}_w(\beta_w - \beta^*) + \bar{X}_b(\beta^* - \beta_b) \quad [4]$$

In equation [4],  $\beta^*$  is defined as a weighted average of the coefficient vectors,  $\beta_w$  and  $\beta_b$ :

$$\beta^* = \Omega\beta_w + (I - \Omega)\beta_b$$

In the original model derived by Oaxaca (1973) and Blinder (1973), it the special case where  $\Omega$  is a null matrix or equal to the identity ( $I$ ) (Sinning et al, 2008). However, there are different assumptions regarding the weighting matrix. “Reimers (1983) and Cotton (1988) treat  $\Omega$  as a scalar matrix. Reimers (1983) proposes the weighting matrix  $\Omega = (0.5)I$ , while Cotton (1988) chooses the weighting matrix  $\Omega = sI$ , where  $s$

denotes the relative sample size of the majority group” (Sinning et al, 2008). Neumark (1988) and Oaxaca and Ransom (1994) propose a pooled model to gain an estimate of the coefficient vector,  $\beta^*$ . These methods are explained further on.

Having explained the model composition, it is now of interest to examine the results.

University of Cape Town

Table 5: Decomposition of the Unemployment Gap

<b>Results</b>	<b>GHS 2005</b>		<b>GHS 2006</b>		<b>GHS 2007</b>		<b>GHS 2008</b>	
	<b>Coefficient</b>	<b>Percentage</b>	<b>Coefficient</b>	<b>Percentage</b>	<b>Coefficient</b>	<b>Percentage</b>	<b>Coefficient</b>	<b>Percentage</b>
<b><u>Omega=1 (Daymont &amp; Andrisani)</u></b>								
Characteristics GHS	-0.2205016	<b>56.35%</b>	-0.2157243	<b>54.88%</b>	-0.2047578	<b>56.58%</b>	-0.1888039	<b>50.11%</b>
Coefficient GHS	-0.3365205	<b>86.00%</b>	-0.30594	<b>77.83%</b>	-0.2259679	<b>62.44%</b>	-0.3339557	<b>88.63%</b>
Interaction GHS	0.1657293	<b>-42.35%</b>	0.1285922	<b>-32.71%</b>	0.0688385	<b>-19.02%</b>	0.1459651	<b>-38.74%</b>
<b><u>Omega=0.5 (Reimers)</u></b>								
Prod	-0.1266611	<b>32.37%</b>	-0.1413786	<b>35.97%</b>	-0.1840761	<b>49.91%</b>	-0.0911762	<b>24.20%</b>
Adv	-0.0715201	<b>18.28%</b>	-0.0716022	<b>18.22%</b>	-0.0608695	<b>17.13%</b>	-0.0710118	<b>18.85%</b>
Disadv	-0.1931117	<b>49.35%</b>	-0.1800914	<b>45.82%</b>	-0.1169416	<b>32.97%</b>	-0.2146065	<b>56.96%</b>
<b><u>Omega = 0.90 (Cotton)</u></b>								
Prod	-0.0667363	<b>17.06%</b>	-0.0942346	<b>23.97%</b>	-0.142227	<b>38.47%</b>	-0.0482675	<b>12.81%</b>
Adv	-0.011204	<b>2.86%</b>	-0.0097438	<b>2.48%</b>	-0.0072912	<b>2.03%</b>	-0.0075067	<b>1.99%</b>
Disadv	-0.3133526	<b>80.08%</b>	-0.2890937	<b>73.55%</b>	-0.212369	<b>59.49%</b>	-0.3210203	<b>85.20%</b>
<b><u>Omega=Neumark</u></b>								
Prod	-0.2407861	<b>61.54%</b>	-0.2373215	<b>60.38%</b>	-0.2209016	<b>60.52%</b>	-0.2151551	<b>57.10%</b>
Adv	-0.1310988	<b>33.50%</b>	-0.1355272	<b>34.48%</b>	-0.1236746	<b>34.65%</b>	-0.1398251	<b>37.11%</b>
Disadv	-0.019408	<b>4.96%</b>	-0.0202234	<b>5.14%</b>	-0.0173111	<b>4.83%</b>	-0.0218142	<b>5.79%</b>

### ***Daymont & Andriasani (1984)***

Table 5 contains the decomposition results following the methodology initially proposed by Oaxaca (1973). The first set of results follows the methodology extended by Daymont & Andriasani (1984), where the weighting matrix is set to 1, and the decomposition is broken up into three components. Component 1 (“Characteristic”), as explained earlier is considered the “endowment affect”. It is the explained portion of the unemployment gap. Interpretation of the results from the GHS 2005 would indicate that, should black individuals carry the same skills as white individuals, the unemployment gap would decrease by around 22%. For 2005, this explained portion accounts for 56% of the gap. According to this method, the explained portion accounts for a consistently similar amount of the gap. However, in 2008, when the unemployment levels appear to increase, it marginally increases.

The second component, the “coefficient” portion, is the unexplained element, considered to be the measure of discrimination. If there were no discrimination, and each group were rewarded in the same way, the unemployment gap would decrease by 33% in 2005, and this accounts for 86% of the gap. This component exhibits a decreasing explanation of the unemployment gap between 2005 and 2007, but by 2008, it explains more of the gap than in 2005. It would appear therefore, that according to the methodology employed by Daymont & Andriasani (1984) that although observable differences between black and white within the labour market appear to be relatively stable, discrimination is still present and has in fact risen once more to levels seen in 2005.

### ***Reimers (1983)***

Reimers (1983) understood that “since employers' preference for the majority and their distaste for the minority no doubt distort both groups' wages, neither group's observed wage-offer function would be likely to exist in a non-discriminatory world.” As a result of this, the original formulation proposed by Oaxaca (1973), where the weighting matrix was simply  $\Omega = \mathbf{I}$ , is not relevant. Instead, Reimers proposed a weighting matrix of  $\Omega = (0.5) \mathbf{I}$ , which allowed the no-discrimination wage function to lie somewhere in between the black and white groups.

Under this methodology, we can interpret the “Prod” component in a similar fashion to the “endowment” component under the Daymont & Andriasani method, whereby it explains differences in productivity. In 2005, if black individuals were as productive, or had the same characteristics and endowments as white individuals, the unemployment gap would be reduced by 24%, and this accounts for 62% of the gap. The “Adv” component is the advantage experienced by white individuals, in terms of observable characteristics controlled for in probit analysis. In 2005, if black individuals were as advantaged as white individuals, the employment gap would be reduced by 13%, and accounts for 34% of the gap. Lastly, the “Disadv” portion measures how disadvantaged black individuals are in comparison to their white counterparts. In 2005, if black

individuals did not face the disadvantages which they do, the gap would be reduced by 2%, however, this component makes up 49% of the total gap, a considerable portion.

Results indicate that over time, the productivity differences between whites (the base group) and blacks fluctuates, moving from 32% of the gap in 2005 to 24% of the gap in 2008. It is interesting to note however, that the advantage experienced by white individuals remains relatively stable, and indication that no matter how unemployment levels change, white workers always have an advantage within the workplace. The third and last component, which accounts for a far larger portion of the gap, also fluctuates, in negative correlation to the first: endowment” component. This perhaps explains that there is some kind of relationship between the endowment differences between black and white respondents, and the disadvantages faced by black individuals. As the probability in unemployment increases due to the disadvantages which black individuals increase, so their productivity will be reduced, and thus reducing the probability of employment. The method proposed by Reimers, appears to indicate that the unemployment gap can consistently be accounted for by advantages enjoyed by white individuals, while other components, although carry a large weighting, are difficult to effectively measure or predict.

### ***Cotton (1988)***

Cotton extends the idea first proposed by Reimers, in that the no-discrimination wage function, should fall somewhere in between white and black individuals. However, instead of simply setting  $\Omega = 0.5$ , Cotton chooses the weighting matrix  $\Omega = sI$ , where  $s$  denotes the relative sample size of the majority group (Sinning et al, 2008). As black respondents far outnumber white respondents,  $s$  is set to 0.90 or close to it for all years of examination. This has a very strong influence on results.

Initially, in 2005 32% of the gap is attributed to productivity advantages of white individuals, and 49% to the disadvantages which black individuals face. We can now see a direct effect of the weighting matrix in the results, as the advantage of being white accounts for only 3% of the gap. In a similar fashion seen in the Reimers method, the productivity component steadily increases up until 2007, before a dramatic decrease in 2008, while the “disadvantage” experienced by black individuals steadily decreases until 2007, before a dramatic increase in 2008. Again, as seen with Reimers’ results, the component attributed to the “advantage” of white individuals remains relatively consistent throughout the period of interest, albeit a far smaller proportion. These results indicate once more that the advantages of white individuals are *consistent* in explaining the unemployment gap, but productivity or endowment differences, and disadvantages faced by black employees are more likely to be the reasons for large discrepancies.

### ***Neumark (1988)***

It has been discussed as to how the decomposition model has been extended by altering the weighting matrix. Neumark (1998), as well as Oaxaca & Ransom (1994) explain that

the decomposition can be explained in three parts. The first being “pure productivity differences” (Bauer & Sinning, 2008) between groups, the second considered the “white wage advantage” and the third the “black wage disadvantage” (Oaxaca & Ransom, 1994). The issue which these two researchers uncovered is how much weighting should be given to each component of the decomposition. The focal point of contention lies in the choice of the non-discriminatory wage structure, the methodologies of Reimers (1983) and Cotton (1988) have illustrated this. Oaxaca & Ransom (1994), illustrate that “employers with discriminatory tastes are willing to trade reduced profits for fewer minority (or black) employees”, an idea first postulated by Becker (1971). This they argue therefore has direct implications as to how one should interpret, and weight the discriminatory portion of the decomposition. Therefore, in order to derive the  $\beta^*$ , the counterfactual coefficient vector, they propose that a pooled model is estimated. This methodology is perhaps more rigorous than the previous two methods, as it attempts to bring employer preferences within the model, and is thus perhaps a more comprehensive approach.

In 2005, if black individuals were as productive as white individuals, the unemployment gap would be reduced by 24%, and this “explained” portion accounts for 62% of the gap. In the same year, if black individuals were as advantaged as white individuals, the gap would reduce by 13%, and this accounts for 34% of the total gap. Furthermore, if blacks did not face the disadvantages that they do, the unemployment gap would be reduced by 2%, and this only accounts for 5% of the total gap. The results obtained under this methodology are far more stable than any other method. Throughout the period, the endowment or productivity differences between black and white individuals lies at close to or around 60%, and has the potential to reduce the probability of unemployment by over 20%. Furthermore, the white advantage is responsible for around 35%, and interestingly, the disadvantage which black workers appear to have, can only describe around 5% of the gap.

## **CONCLUSIONS AND RECOMMENDATIONS**

It has become evident that the unemployment gap between black and white race groups is very much a part of the labour market within South Africa. Descriptive statistics, as well as multivariate analysis in the form of probit modelling has clearly illustrated how white individuals are far more likely to become employed than their white counterparts. Unfortunately, the individuals found to face the highest probability of unemployment are no different to the group established to be in this position 5 years ago, according to the study conducted by Kingdon & Knight (2004). Young black females are more likely to be unemployed than any other group. Furthermore, the earlier individuals terminate their schooling, the harder they will find it to become employed, and this is particularly true for black individuals.

After examining probit results, one might be quick to conclude that an improvement in the quality of education of the black population will improve the chances of employment. This however, will not necessarily achieve the desired effect; Kingdon & Knight (2004) indicate that “upgrading the education of Africans will at best change the composition of employment in their favour”. What is of fundamental importance is an increase in the overall level of jobs within the country, as if this is not the case, highly skilled white individuals will occupy more of the positions currently available. An extension beyond this paper, could involve an investigation into the probability of employment of skilled versus unskilled labour, and how this has evolved over time.

Analysis was undertaken to investigate the extent to which this unemployment gap was due to endowment effects, as opposed to issues of discrimination. Within this analysis, multiple techniques were employed. The traditional methodology first proposed by Oaxaca (1973), extended to the nonlinear case by Daymont & Andrisani (1984), indicates that discrimination within South Africa is very much in existence, and although it had appeared to be decreasing between 2005 and 2007, 2008 saw discriminatory levels rise to levels witnessed in 2005. Reimers (1983) and Cotton (1988) offered a new technique by adjusting the weighting of each group under examination. According to these techniques, there too is evidence of discrimination. The Reimers (1983) method indicates an initial increase, but eventual reduction in the “endowment” effect in 2008, and an initial decrease and eventual increase in the disadvantages faced by black workers. The Cotton (1988) methodology shows the very same pattern, however, due the difference in the weighting matrix, the disadvantages faced by black respondents appear to explain a far greater proportion in the unemployment gap. The last and perhaps most robust technique is that developed by Neumark (1988). This model, as with the previous two adjusts the weighting matrix, but also attempts to incorporate employer preferences. The results from this model are far more consistent through the period. Well over half of the unemployment gap is accounted for by differences in endowments, and over a third, to the advantages in characteristics faced by white individuals. Interestingly, only 5% of the unemployment gap, according to the Neumark model can be explained by disadvantages faced by black individuals.

After examining the battery of results produced through these four well established and researched techniques, it is still difficult to come to clear cut conclusions. While it is evident that there is in fact a gap in the probability of unemployment and it is clear which groups are at the lower bound, determining exactly where to focus policy and recommendations is not as simple as many interest groups would propose. The current policies of reverse discrimination have been found to be ineffective according to the results of this study. Despite continued efforts on the part of national government to address the issue through direct policy implementation, little effect has been witnessed. Although Results have indicated that although there is a definitive advantage for white workers within the workplace, one cannot out rightly conclude that it is due to

disadvantages faced by black employees, or due to an inherent difference in skill sets. Rather, further investigation must be taken to determine as to which sectors of the workforce are suffering greater job loss, and the onus falls on government to create jobs in sectors where there is greatest supply. Furthermore, government must re-examine the current affirmative action policies, which at present appear to only aid those individuals with skill sets demanded within the economy. The focus must be turned to long term solutions aimed at improving the skill sets of the disadvantaged group, whilst encouraging job creation for the currently unemployed.

University of Cape Town

## **REFERENCES**

- Bauer, T.K. & Sinning, M. 2008. An extension of the Blinder–Oaxaca decomposition to nonlinear models. *AStA advances in statistical analysis*. 92(2):197-206.
- Becker, G.S. 1971. *The economics of discrimination*. University of Chicago Press.
- Blackaby, D.H. and others. 1998. The ethnic wage gap and employment differentials in the 1990s: evidence for Britain. *Economics letters*. 58(1):97-103.
- Blinder, A.S. 1973. Wage discrimination: reduced form and structural estimates. *Journal of human resources*. 8(4):436-455.
- Brookes, M. & Hinks, T. 2004. The racial employment gap in South Africa. *South African journal of economics*. 72(3):573-580.
- Burger, R. & Jafta, R. 2010. Affirmative action in South Africa: an empirical assessment of the impact on labour market outcome. *Oxford CRISE working paper, no. 76*.
- Burger, R. & Jafta, R. 2006. Returns to race: Labour market discrimination in post-apartheid South Africa. *Working papers*.
- Burger, R. & Jafta, R. 2010. Affirmative action in South Africa: an empirical assessment of the impact on labour market outcomes.
- Case, A. & Deaton, A. 1999. School Inputs and Educational Outcomes in South Africa\*. *Quarterly journal of economics*. 114(3):1047-1084.
- Chabane, N., Roberts, S. & Goldstein, A. 2006. The changing face and strategies of big business in South Africa: more than a decade of political democracy. *Industrial and corporate change*. 15(3):549.
- Cotton, J. 1988. On the decomposition of wage differentials. *The review of economics and statistics*. 70(2):236-243.
- Daymont, T.N. & Andrisani, P.J. 1984. Job preferences, college major, and the gender gap in earnings. *Journal of human resources*. 19(3):408-428.
- Daymont, T.N. & Andrisani, P.J. 1984. Job preferences, college major, and the gender gap in earnings. *Journal of human resources*. 19(3):408-428.
- Fairlie, R.W. 2005. An extension of the Blinder-Oaxaca decomposition technique to logit and probit models. *Journal of economic and social measurement*. 30(4):305-316.
- Fallon, P. & Lucas, R. 1998. South African labor markets: adjustment and inequalities. *Informal discussion papers on aspects of the South African economy*. 12.

- Fallon, P.R. & Lucas, R.E.B. 2002. The impact of financial crises on labor markets, household incomes, and poverty: a review of evidence. *The World Bank research observer*. 17(1):21.
- Green, R.K. & Hendershott, P.H. 2001. Home-ownership and the duration of unemployment: a test of the Oswald hypothesis. *NBER working paper*.
- Hundley, G. 2000. Male/female earnings differences in self-employment: The effects of marriage, children, and the household division of labor. *Industrial and labor relations review*. 54(1):95-114.
- Jann, B. 2008. The Blinder–Oaxaca decomposition for linear regression models. *Stata journal*. 8(4):453-479.
- Kingdon, G. & Knight, J. 2000. *Are Searching and Non-searching Unemployment Distinct States when Unemployment is High?: The Case of South Africa*. CSAE publ.
- Kingdon, G.G. & Knight, J. 2004. Race and the incidence of unemployment in South Africa. *Review of development economics*. 8(2):198-222.
- Kingdon, G.G. & Knight, J. 2004. Race and the incidence of unemployment in South Africa. *Review of development economics*. 8(2):198-222.
- Kingdon, G.G. & Knight, J. 2004. Unemployment in South Africa: the nature of the beast. *World development*. 32(3):391-408.
- Klasen, S. & Woolard, I. 2009. Surviving unemployment without state support: unemployment and household formation in South Africa. *Journal of African economies*. 18(1):1.
- Neumark, D. 1988. Employers' discriminatory behavior and the estimation of wage discrimination. *Journal of human resources*. 23(3):279-295.
- Oaxaca, R. 1973. Male-female wage differentials in urban labor markets. *International economic review*. :693-709.
- Oaxaca, R.L. & Ransom, M.R. 1994. On discrimination and the decomposition of wage differentials. *Journal of econometrics*. 61(1):5-21.
- Oppenheimer, V.K. 1997. Women's employment and the gain to marriage: The specialization and trading model. *Annual review of sociology*. 23.
- Reimers, C.W. 1983. Labor market discrimination against Hispanic and black men. *The review of economics and statistics*. 65(4):570-579.
- Rospabéa, S. 2002. How did labour market racial discrimination evolve after the end of apartheid? *South African journal of economics*. 70(1):185-217.

Sinning, M., Hahn, M. & Bauer, T.K. 2008. The Blinder–Oaxaca decomposition for nonlinear regression models. *Stata journal*. 8(4):480-492.

Unknown 2009. *Tough times ahead But Africa’s biggest economy may still recover quickly from a looming recession*. [Online].

Available: <http://www.economist.com/node/13109968> [2010, 31 January].

University of Cape Town