An empirical investigation of the gender wage gap in South Africa

Phumzile Ncube and Fiona Tregenna

INTRODUCTION

The South African labour market has for a long time been characterised by high wage or earnings inequality. The gender wage gap is one important dimension of earnings inequality. The gender wage gap is the difference in earnings between men and women, which is usually calculated by using the male earnings as a benchmark. Racial differences have received the most attention in research on the inequalities in South Africa because apartheid legislation mandated labour market, locational and economic segregation based on race.[1] However, women were also disadvantaged greatly as they had the lowest levels of education and consequently ended up with the lowest-paying jobs (Makgetla, 2004). With the advent of democracy in 1994, there was an expectation that all forms of inequality would begin to be eradicated. Since apartheid, black people's earnings, and especially those of black women, have increased. However, even though women's earnings have increased, gender wage inequalities still persist (Burger and Yu, 2006: 7). It is important then to assess the reasons for the gender wage gap in South Africa.

The bulk of the research on the gender wage gap, especially in South Africa, has been focused on assessing it at the mean. That is, the research has looked at differences in average earnings between men and women. Calculating and decomposing the gender wage gap at the mean provided a useful starting point for assessing gender wage inequalities. The main advantage of this is that it offers a simple method of assessing gender earnings inequalities. However, that method has significant limitations. Firstly, looking at the gender wage gap only at the mean does not take into account that returns to human capital may not be equal at all points of the wage distribution. Secondly, people may experience different levels of discrimination based on characteristics such as race, gender and provincial location, which may also be related to their position on the wage distribution. Consequently, the gender wage gap at the mean cannot be assumed to be the same across the entire wage distribution. The South African context of particularly high levels of inequality necessitates the analysis of the gender wage gap at points other than the mean, as is undertaken here.

[1] The racial categories as defined during the Apartheid era were African, Coloured, Indian and White. This study uses these categories for analytical purposes, taking cognizance of their political connotations. In this respect, instead of referring to 'Africans', this study refers to this group as 'black people'. Moreover, to take into account that people of Asian origin are a broader group than Indians, this study will refer to 'Asians'.

Assessing the gender wage gap across the wage distribution provides a method of comparing the earnings of people who are likely to have similar labour force characteristics. This method helps to bring out elements of the gender wage gap that are not obvious when looking at it at the mean. For example, men and women at lower ends of the wage distribution are likely to have lower levels of education than those at higher points. Examining the gender wage gap at different points of the wage distribution as opposed to the mean suggests that these two groups will not be conflated. This is especially important in a country such as South Africa where high inequality is experienced not only in earnings but also in education.

As such, the gender wage gap was assessed using quantile regression and decomposition techniques at the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} quantiles.

LITERATURE REVIEW

Since this study will be conducted using quantile regression methods, this literature review will concentrate only on the international studies that also used the quantile regression method. However, it must be noted that it is not the only way that analysis of the gender wage gap at different points of the wage distribution can be carried out. That said, this is the only method known that will estimate each quantile of the wage distribution as opposed to the mean. The South African studies reviewed will cover a broader spectrum of methodologies in order to fully contextualise this study.

Recent international research on the gender wage gap has been focused on assessing the gender wage gap across the wage distribution (see for instance Garcia *et al*, 2007; de la Rica, Dolado & Llorens, 2008). Much of this research has been conducted in order to determine whether there is a 'glass ceiling effect' (where the gender wage gap widens at the top of the wage distribution) and/or a 'sticky floor effect' (where the gender wage gap widens at lower levels of the wage distribution).

The results from some of these studies suggest that in some developed European countries, the gender wage gap widens at higher points of the wage distribution (Gardeazabal & Ugidos, 2005; Garcia *et al*, 2007; de la Rica, Dolado & Llorens, 2008), while in developing countries such as Vietnam, Chile and Mauritius, the gender wage gap is wider at lower points of the wage distribution, becoming narrower at higher points of the distribution (Montenegro, 2001; Sakellariou, 2004; Pham & Reilly, 2007). While general

conclusions about the gender wage gaps in these regions cannot be reached based on these studies alone, they do suggest that there is a glass ceiling effect in the European countries in the aforementioned studies and a sticky floor effect in countries such as Vietnam, Chile and Mauritius. The glass ceiling effect observed in the European countries and the sticky floor effect in these studies of developing countries indicate that differential returns to male and female characteristics are more likely to be found at lower quantiles in developing regions, while the opposite happens in developed regions. Thus, the South African results might be similar to those of other developing countries.

Early studies of the gender wage gap in South Africa were single-year studies (see Isemonger and Roberts (1999) and Winter (1999). The earliest post-apartheid studies of the gender wage gap were those by Isemonger and Roberts (1999) and Winter (1999). Winter's (1999) study found that in 1994, although women had similar levels of education as men within population groups, they received lower earnings. Women earned roughly 13% less than men. In contrast, Isemonger and Roberts's (1999) estimate showed that women earned 25-36% less than what men earned. Isemonger and Roberts's (1999) estimates of discrimination showed that it accounted for over 35% of gender differentials.

Another single-year study by Rospabe (2001) found that earnings increased as more education was obtained. However, there were no diminishing returns to education for both men and women. This means that at higher levels of education, returns to education also increased. Overall, the gender wage gap was 29%. 44% of the overall gender wage gap was explained by different human capital characteristics, and the rest was attributed to different returns to these characteristics.

With an increase in the availability of data, multi-year studies of the gender wage gap in South Africa were conducted. Using October Household Survey (OHS) studies from 1995 to 1999, Grün (2004) decomposed the gender wage gap among black people. The results indicate that, in general, earnings increase up to a point with age. The drop in earnings with age is more pronounced for black women than for black men.

Muller (2009) compared the gender wage gap among part-time workers to that found among full-time workers between 1995 and 2006. Overall, Muller (2009) found that the mean gender wage gap increases from 1995 to 1999 and it decreases between 2001 and 2006. Muller's results suggest that the increase in the gender wage gap between 1995 and

1999 was due to a depression in wages for part-time workers. For full-time workers the increase in the gender wage gap observed during this period was partly due to a depression in wages and a deterioration of observed characteristics for women relative to men. The decrease in the gender wage gap between 2001 and 2006 was due both to an increase in women's productive characteristics and to a reduction in differential returns to men and women's characteristics – in other words, a decrease in discrimination.

More recently, there have also been studies of the gender wage gap in South Africa at different points of the wage distribution. Using quantile regression, Ntuli (2007) estimated the gender wage gap for black people in the formal sector at the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles for 1995 and 1999. She employed the decomposition technique formulated by Machado and Mata (2005), which conducts the Oaxaca decomposition technique at quantiles. She found that the gender wage gap among black workers was highest at the 10^{th} quantile, decreasing from 60% at the 10^{th} quantile in 1995 to 29% at the 50^{th} quantile and 17% at the 90^{th} quantile in the same year. In 2004, the gender wage gap decreased from 52% at the 10^{th} quantile to 44% at the 50^{th} quantile and 13% at the 90^{th} quantile. The main driver of the gender wage gap was the coefficients component, constituting 107% of the gender wage gap at the 10^{th} quantile in 1995, 81% at the 50^{th} quantile and 154% at the 90^{th} quantile in the same year. In 2004, this number decreased slightly to 79% at the 10^{th} quantile and 73% at the 50^{th} quantile but increased to 260% at the 90^{th} quantile. The wider gender wage gap at the 10^{th} quantile indicates that there is a sticky floor effect among male and female black workers in South Africa.

Another study that used the quantile regression approach was that by Casale and Posel (2010). They studied the gender wage gap at the 20^{th}, 50^{th} and 80^{th} percentiles between 2000 and 2006 comparing unionised and non-unionised workers. They found that the gender wage gap in the unionised sector is only marginally lower than that in the non-unionised sector. The authors attribute the wage gap in the unionised sector to higher returns to men's human capital characteristics. When they estimated the earnings equations by educational attainment, they found that the union premium decreases as educational attainment increases but this decrease is faster for women than for men. Finally, they found that the union premium decreased as wages increased but that this happened at a higher rate for women than for men.

The last set of studies – Goga (2008) and Shepherd (2008) – looked at the gender wage gap both at the mean and at quantiles, albeit using different methods. For the assessment of the gender wage gap at the mean, both used the Ordinary Least Squares (OLS) method but in the method of analysing the gender wage gap along the wage distribution Goga used quantile regression while Shepherd used the Juhn-Murphy-Pierce (JMP) technique.

Shepherd studied the gender wage gap using the OHS from 1996 to 1999 and the LFS from 2000 to 2006. While the study of the average gender wage gap was for all race groups in the formal sector, the JMP decomposition was only conducted for black people in the formal sector. Shepherd found that the gender wage gap decreased between the 10^{th} and 90^{th} quantiles between 1996 and 2004. The characteristics components were all in favour of women, ranging from 6% to 24% between the 10^{th} and the 90^{th} quantiles in 1996 and 8% and 29% for the same quantiles in 2004. The gender wage gap was thus largely attributed to higher returns to male characteristics, as the coefficients components ranged from 10% at the 10^{th} quantile in 1996 to 12% at the 90^{th} quantile. In 2004 the coefficients components ranged from 17% at the 10^{th} quantile to 7% at the 90^{th} quantile. Shepherd (2008) suggests that since the coefficients components decrease at higher points of the wage distribution, there is a sticky floor pattern among male and female black workers.

These results are similar to those found by Ntuli (2007), but the stark differences in the actual results are due to sampling differences. While Ntuli (2007) sampled all black workers in the formal sector, Shepherd (2008) excluded domestic workers, agriculturists and the self-employed from her sample. Since there is a high concentration of black women employed as domestic workers, the higher equality of earnings observed among black workers due to their exclusion is understandable, given the relatively low wages received by domestic workers.

As previously stated, Goga (2008) used OLS and quantile regression to assess the gender wage gap. In the OLS regressions, Goga (2008: 27) found that on average women earned less than men. In 2001 women earned 18% less than men and in 2005 this increased to 20%. When she decomposed the gender wage gap, she found that in both years the unexplained portion was greater than the explained portion. Moreover, the unexplained portion increased between 2001 and 2005 from 55% to 69% of the gender wage gap, while the explained portion decreased from 45% to 31%. In order to assess the gender

wage gap at different parts of the wage distribution, Goga used quantile regression at the 10th, 50th, 75th and 90th percentiles. She found that the gender wage gap was 13% at the 10th quantile in 2001 and 18% at the 90th quantile. In 2005, she found that the gender wage gap had increased to 17% at the 10th quantile and 21% at the 90th. Unfortunately, no decompositions were provided for the quantile regression results, although the result of increasing gender wage gaps at higher points of the wage distribution suggests that there is a glass ceiling instead for this sample, which was the entire working age population in South Africa, excluding those whose earnings were missing, who had zero earnings and who had not reported hours worked.

These gender wage gap studies indicate that, depending on the sample chosen, there is evidence of a sticky floor effect in some cases and a glass ceiling effect in others. Either way, the gender wage gap in South Africa has been shown to be largely driven by unequal returns to male and female characteristics.

A number of reasons have been put forward for the persistence of the gender wage gap in South Africa. Firstly, growth in the formal economy focused largely on capital-intensive sectors in which women are not well represented. Although women's labour force participation has increased substantially since the end of apartheid, it is mostly in informal, low-paying employment (Casale, 2004). Secondly, women are less likely than men to be in unions. Even with the increased unionisation of women, the gender wage gap has persisted. Casale and Posel (2010) found that the gender wage gap among union members is bigger than that among non-union members. Finally, women are still the primary caregivers to children in the home, even if they also engage in paid work (Makgetla, 2004). This means that they have to split their time between labour market work, for which they will be paid, and household work, for which they do not get paid.

Other reasons still could be related to differences in skills and levels of training. In earlier international studies, such as in that of Suter and Miller (1973), the gender wage gap was almost fully explained by the differences in levels of education, work experience and tenure, which are referred to as human capital characteristics (Suter & Miller, 1973: 963). However, women's labour force participation has increased and so has the quality and quantity of their education. Moreover, the representation of women in more 'male-dominated' industries and occupations has increased, and there is legislation in most developed countries that stipulates that women and men are to be paid the same for work

of equal value (Plantenga & Remery, 2006). Despite all these changes, the gender wage gap is still persistent in developed countries. The world's largest economy, the United States of America (USA), has women in almost half of the labour force (The Economist, 2009). However, it has one of the most unequal wage structures, suggesting that there is more to the gender wage gap than differences in human capital characteristics (Blau & Kahn, 2003).

The gender wage gap has been explained using different theories, a number of which relate it to women's intermittent careers. While this may have been a valid explanation many years ago, discrimination also played a role, and, as seen in the results of the studies that were reviewed, continues to play an important role in the gender wage gap. The studies above suggest that although the gender wage gap has decreased slightly, this has not been a sustained decrease. Moreover, the South African studies indicate that depending on the sample the gender wage gap is widest at the lower parts of the wage distribution, pointing to a sticky floor effect and at times wider at higher points of the wage distribution, pointing to a glass ceiling effect.

METHODOLOGY

Data

The research analyses the gender wage gap across the wage distribution from 2001 to 2010 using a combination of the Labour Force Survey (LFS) from 2001 to 2007 and the 2010 Quarterly Labour Force Survey (QLFS). These surveys were conducted by Statistics South Africa (Stats SA), which is South Africa's statutory statistics agency. South Africa's official labour market statistics are based on these datasets.

A few abnormalities in some of the datasets led to their omission from the analysis. Firstly, this study starts from 2001 because the dataset from the year 2000 was fraught with measurement errors, including extreme outliers (Burger & Yu, 2006). Secondly, the QLFS from the year 2008 was not included in the analysis because there was no survey question and thus no data on earnings in that year. The first time a question on earnings was included in the QLFS was in the third quarter of 2009 – which was the September round of the survey. Therefore, to maintain comparability in the data, only the September datasets were used. Finally, the QLFS from September 2009 did not contain a question about whether or not the respondents were part of a trade union. As a respondent's trade union membership status is an important variable in the econometric model, the year 2009 was

left out. If every possible year were used for the econometric analysis, there would have been a gap of three years in a part of the analysis, that is, between 2007 and 2010. Thus, to maintain uniformity in the analysis, the years used for the econometric analysis were 2001, 2004, 2007 and 2010.

In preparing the datasets for analysis, particular preparation was required for the earnings variable. Firstly, missing earnings were imputed using the hotdeck multiple imputation method as used by Tregenna (2011). The decision to impute missing earnings rather than to remove the affected observations from the sample was taken after an assessment of the missing earnings data revealed that 35% of the missing earnings were for white people, who constitute roughly 9% of the population. Moreover, most people who did not report their earnings responded to other questions in the survey. For that reason, it is likely that the data on earnings were “missing not at random” and thus the observations could not be dropped, as this could lead to a bias in the empirical analysis. Therefore, it was important to impute the missing earnings in such a way that the imputed earnings were as realistic as possible.

Secondly, the earnings of those who reportedly had “zero earnings” were also imputed using the hotdeck imputation method. It was not likely that a large number of these people would be zero earners. Approximately 6 500 people were recorded as being zero earners, with around 4 000 of these people working as skilled agricultural workers in the agricultural industry across all the years. This was a peculiarity found in the datasets that were received from Stats SA after linking to allow the LFS and QLFS studies to be comparable. This was likely to be an error, because in the unlinked datasets there were only three zero earners, with only one of them being a skilled agricultural worker in the agricultural industry. It was therefore important to impute the earnings of specifically the zero earners who were skilled agricultural workers.

Finally, a few observations were dropped from the sample because, based on individual inspection, they exhibited unrealistically high earnings. Observations with very high earnings data that is in fact erroneous would have a much more significant distortionary effect on the results than would erroneous observations elsewhere in the earnings distribution, making the screening for such observations particularly important.

Sample

The population chosen consists of those who were classified as being employed by someone else in the formal sector – that is, those who are not self-employed. The total number of employed persons over the four separate years is 63 384 (unweighted) and 34 011 947 (weighted).[2]

Due to the extremely high rate of unemployment, studies of South African earnings have had to take into account sample selection of people first into labour force participation and then into employment (Bhorat & Leibbrandt, 1999; Goga, 2008; Ntuli, 2007). Accounting for sample selection is important for some studies, because the sample of employed individuals is not necessarily random, given that those who have some level of education, especially high levels, are more likely to find employment.

However, in this study, sample selection was not accounted for, because the focus is on those who are employed, not necessarily the working age population as in some studies, and specifically determining whether there is gender wage discrimination among them. Since wage discrimination can occur only among those people who are actually working for someone else (as opposed to the self-employed), accounting for how they managed to find that employment was not necessary.

Empirical analysis methodology

This research explores and seeks to explain the gender wage gap along the wage distribution using the Oaxaca method of decomposition at the quantiles. In 1973, Ronald Oaxaca published a paper in which he proposed a method of decomposition that is now known as the Oaxaca decomposition. This decomposition method allows one to estimate and decompose the gender wage gap into the part that is due to differences in human capital or other observable characteristics – the explained part or the characteristics component – and into the part that is due to different returns to one's characteristics – the unexplained part or the coefficients component. The unexplained part has generally been used as a proxy for discrimination, albeit with caution. The unexplained portion may also contain other characteristics such as innate ability or talent which cannot necessarily be captured in survey data (Shepherd, 2008: 13). However, because of the history of gender

[2] Weights used are the official weights as provided by Stats SA in order to make the sample nationally representative.

discrimination and because of the persistence of the gender wage gap even after the observable characteristics have been controlled for, it is generally assumed that the residual or the unexplained portion also contains discrimination (Oaxaca, 1973; Hinks, 2002; Grün, 2004). When the decompositions are calculated at different quantiles, they not only shed light on the gender wage gap at different parts of the wage distribution, but also show whether the gender wage gap gets wider at the top or at the bottom of the wage distribution.

Quantile regression is also used to explore male and female earnings functions separately and to assess whether the different control variables in the model affect them differently. In this way, quantile regression will be used to analyse gender inequality in the returns to characteristics. The method of quantile regression was pioneered by Koenker and Bassett (1978) and has since been used by labour economists to estimate earnings differentials. As opposed to OLS, which estimates the conditional mean of the dependent variable given the independent variables, quantile regression estimates the conditional percentiles of the dependent variable given the independent variables. One of the assumptions of the OLS model is that the relationship between the dependent and independent variables is linear in parameters (Gujarati, 2002). In contrast, quantile regression does not impose this condition on the quantiles in the model, thus allowing for a more accurate estimation. This is particularly important for the South African context of high inequality, where the relationship between earnings and the independent variables is not likely to be linear at each quantile necessitating its use for this study.

The quantile regressions and the decomposition analysis are two independent methods of analysis that were used concurrently in this study. The two methods are certainly sufficient if either of them is used on its own as seen in previous studies. However, when used together, they add more value to the analysis than they would separately. While the gender wage gap decompositions show the contribution of the two components to the gender wage gap, the quantile regressions help to explain the size of the coefficients component. Consequently, in this study, the following methodology was used.

Econometric models

For the quantile regressions separate male and female wage equations were run to account for the fact that the male and female earnings functions were not the same. Moreover, separate male and female wage equations served to account for the possible

interaction of gender with variables such as education level and occupation, especially in cases where occupational segregation occurs. The decompositions were run as a counterfactual decomposition that imposed the female characteristics on the male earnings distribution.

Gender wage gap decompositions

The first step was the calculation and the decomposition of the gender wage gap at each quantile being studied. The quantile gender wage gap is expressed as the sum of two parts:[3]

$$\mathbf{W}_m - \mathbf{W}_f = \quad (1)$$
$$(\beta'_{m\theta} - \beta'_{f\theta})E(\mathbf{X}_f \mid \mathbf{W}_f = \mathbf{W}_{f\theta}) \; + \quad (2)$$
$$\beta'_{m\theta}(E(\mathbf{X}_m \mid \mathbf{W}_m = \mathbf{W}_{m\theta}) - E(\mathbf{X}_f \mid \mathbf{W}_f = \mathbf{W}_{f\theta})) \quad (3)$$

where $\mathbf{W}_m$ and $\mathbf{W}_f$ are the (log) earnings for males and females respectively; $\mathbf{X}_m$ and $\mathbf{X}_f$ are the male and female vectors of human capital characteristics, and $\beta'_{m\theta}$ and $\beta'_{f\theta}$ measure the marginal return of each male and female characteristic to earnings at the θ^{th} quantile. Line (1) shows the raw difference between male and female earnings. Since $\beta'_{m\theta}$ and $\beta'_{f\theta}$ measure the marginal returns to male and female characteristics, line (2) measures earnings differences due to returns to characteristics. Line (3) measures differences in human capital characteristics. Thus this is the part of the gender wage gap that is explained by differences in factors such as education, occupation and industry. In the Oaxaca decompositions of the gender wage gap, line (3) is the 'explained' component. Line (2) constitutes the 'unexplained' component. Discrimination is considered to be part of this unexplained component due to the general inability to observe and measure it. If the model is fully specified and there is still an unexplained component, then that component would indeed accurately measure discrimination.

Quantile regression

[3] Both the gender wage gap decomposition and the quantile regression equations are adaptations of those presented in the study by Gardeazabal and Ugidos (2005: 168-169).

In order to elucidate line (2), the differing returns to characteristics, quantile regressions were run. For the quantile regressions, a general Mincerian wage equation was estimated and took the following form:

$$\mathbf{W}_{g\theta} = \beta'_{g\theta} E(\mathbf{X}_g \mid \mathbf{W}_g = \mathbf{W}_{g\theta}) + E(\mathbf{U}_g \mid \mathbf{W}_g = \mathbf{W}_{g\theta})$$

where $\mathbf{W}_g$ are the (log) earnings, $\mathbf{X}_g$ is the vector of human capital characteristics, β'_g measures the marginal return of each characteristic to earnings, the subscript *g* denotes gender and $\mathbf{U}_g$ denotes the error term. This specifies the wage equation at the θ^{th} quantile as a conditional function of the different characteristics observed at that quantile. The vectors of human capital characteristics used in both the gender wage gap decompositions and in the quantile regressions are populated by the variables that will be elaborated on in section 0.

The standard errors in both the quantile regressions were constructed using the bootstrap method with 50 repetitions. The bootstrap method is preferred in quantile regression analysis and is used in cases where the distribution is not necessarily normally distributed and thus the errors may not be homoscedastic (Rogers, 1992).

Variables used in the models

In order to build a model to estimate the gender wage gap and its causes, one has to decide on the variables to be chosen. The main variables identified by Mincer and Polachek (1974: 80) are years of schooling, age, total years of work, years of work before and after first child, tenure at current job, total home time, and years of residence in the country. To incorporate the specific characteristics of the South African labour market, the example of Ntuli (2007), Goga (2008) and Shepherd (2008) was followed, and thus the following variables were used as regressors: race, education, potential experience, potential experience squared, hours worked, marital status, whether the person is employed in the public or private sector, union membership, industry, occupation and province[4]. These variables were used in both the decomposition of the gender wage gap and the quantile regression estimates and are expressed as the vectors of human capital characteristics in both models.

[4] Appendix A shows the dummy variables and their base categories.

It must be noted that for the variables occupation and industry, two of the categories in each group had to be combined in order to avoid multicollinearity, as this could distort the results of the gender wage gap analysis and the returns to the different variables. Because all the people who were listed as domestic workers in the occupation list worked in the "private households" industrial category, domestic workers were combined with elementary workers, as these are the two most similar categories and domestic workers could in any event essentially be regarded as elementary workers. Similarly, the Community, Social and Personal Services and the Private Households industries were also combined, given that domestic work in private households could be seen as fitting well within the Community, Social and Personal Services industry.

ECONOMETRIC ANALYSIS

This section presents the results of the quantile regression and decomposition analysis for the years 2001, 2004, 2007 and 2010. The first section presents the results from the decomposition of the gender wage gap and the analysis of the characteristics while the second section presents the results from the quantile regression analysis by gender, also for those who are employed in the formal sector. The final section concludes.

Analysis of decomposition results

Table 1 below shows the decomposition results from 2001 to 2010. The results are shown for the 10^{th}, 25^{th}, 50^{th} (median), 75^{th} and 90^{th} quantiles. The gender wage gap, which is the raw difference, has been split into two components, namely, the characteristics component and the coefficients component. The characteristics and the coefficients components add up to the raw difference as opposed to reflecting proportions of the gender wage gap.

The raw difference is the overall gender wage gap as a percentage. For example, the raw difference at the 10^{th} quantile in 2010 is 23.6%. That is, given that male earnings are used as the benchmark, females earn 23.6% less than males. Alternatively, women earn 76.4% of what men earn. The gender wage gap is statistically significant at the 1% level of significance at all except the 50^{th} and 75^{th} quantiles. This does not mean that earnings are equal at these quantiles; it merely shows that if women were paid as men are they receive equal earnings. The gender wage gap decreases until it is reversed in favour of women at the 75^{th} quantile, although this difference is not statistically significant. After that, it

increases again at the 90th quantile, but does not reach as high a magnitude as that found at the 10th quantile.

Looking at the gender wage gap and its components across the years by quantile yields an interesting story. Before doing so, it must be noted that the coefficients and the characteristics components are both statistically significant regardless of the corresponding result for the gender wage gap. The gender wage gap at the 10th quantile was initially not the highest in 2001, with the 90th quantile exhibiting a higher gender wage gap. However, over the years, the gender wage gap at the 10th quantile increased faster than that found at the other quantiles, and was eventually the highest. This increase was driven by a rapid increase in differing returns to male characteristics, an increasing coefficients component, a decrease in the difference between male and female characteristics, and a decrease in the characteristics coefficient. As the gender wage gap increased at both ends of the wage distribution, this indicates that there was both a glass ceiling and a sticky floor effect in all the years analysed. However, the sticky floor effect worsened much more rapidly than the glass ceiling effect. This implies that the earnings of women at lower points of the wage distribution deteriorated much faster than those at higher points of the wage distribution.

The lack of statistical significance of the gender wage gap at the median for most years indicates that, overall, the difference between the median earnings of men and women is not statistically significant, the different characteristics and unequal returns to these characteristics notwithstanding. It must also be noted that the coefficients components are generally at their lowest at the median. In other words, unequal returns to characteristics are at their lowest at the 50th quantile.

Another important point to note is that the characteristics components at all quantiles are positive and statistically significant. This result is similar to that found by Shepherd (2008), although the samples are different. At the 50th quantile, the coefficients components, which are almost mirrored by the characteristics components, yield a gender wage gap that is not statistically significant. Thus in order for the gender wage gap not to be statistically significant, there has to be a high female advantage in characteristics in order to counter the high inequality in the returns to those characteristics.

A similar result at the 75th quantile re-affirms this position, as this is where the characteristics components are at their highest in favour of females. In other words, for female earnings to break even with male earnings, they have to be in much higher positions, be in better paying industries and/or have much higher qualifications than males do. Similarly, the decrease in significance of the gender wage gap (with a female advantage) at the 75th quantile indicates that the difference between the respective 75th quantiles of male and female earnings became less statistically significant. Moreover, the increasingly higher coefficients components at the 75th quantile, coupled with fluctuation of the characteristics components, indicate increasingly higher returns to male characteristics which further undermined the female advantage.

At the lower and upper quantiles, while still positive, the characteristics components are much smaller than those found at the median and the 75th quantile. At the 10th quantile, the characteristics component initially decreases between 2001 and 2007 – the difference between male and female characteristics becomes smaller – but increases again in 2010. At the 25th quantile, the characteristics components actually do not change much with time, except in 2007.

The finding that all characteristics components are positive, yet most of the raw differences are negative, can only be because the coefficients components are negative and larger (in absolute terms) than the characteristics components. The coefficients components are all negative and statistically significant at the 1% level of significance. They are highest at the 10th and the 90th quantiles, which also happen to exhibit the highest gender wage gaps. This implies that males receive higher returns for the same characteristics than females at all quantiles, and these disparate returns are highest where the gender wage gap is also highest. Unlike the characteristics components, which generally decreased with time (except at the 50th and 75th quantiles), the coefficients components increased with time. This means that, with time, males received increasingly higher returns to their characteristics than females did. At each quantile except the median, the coefficients component peaks in 2007.

Table 1: Gender wage gap decompositions by quantile

		2001	2004	2007	2010
10th quantile					
	Raw difference	-0.102***	-0.118***	-0.187***	-0.236***
	Characteristics	0.129***	0.106***	0.056***	0.097***
	Coefficients	-0.231***	-0.224***	-0.243***	-0.333***
25th quantile					
	Raw difference	-0.084***	-0.079***	-0.139***	-0.15***
	Characteristics	0.149***	0.144***	0.085***	0.14***
	Coefficients	-0.233***	-0.223***	-0.224***	-0.29***
50th quantile					
	Raw difference	0.028	-0.003	-0.078***	-0.02
	Characteristics	0.212***	0.21***	0.135***	0.2***
	Coefficients	-0.184***	-0.213***	-0.213***	-0.22***
75th quantile					
	Raw difference	0.037*	0.036*	-0.062**	0.013
	Characteristics	0.228***	0.243***	0.197***	0.228***
	Coefficients	-0.191***	-0.207***	-0.259***	-0.215***
90th quantile					
	Raw difference	-0.139***	-0.122***	-0.183***	-0.165***
	Characteristics	0.129***	0.17***	0.154***	0.168***
	Coefficients	-0.268***	-0.292***	-0.337***	-0.333***

****1% level of significance; **5% level of significance; *10% level of significance*

The results in Table 1 are presented again in Figure 1 and Figure 2 below. While Figure 1 shows the raw differences, Figure 2 shows the characteristics and coefficients components. As shown in the key, the columns in Figure 1 represent the 10th, 25th, 50th, 75th and 90th quantiles, in that order. As it can be seen, the raw difference is negative at the 10th, 25th and 90th quantiles in all the years except 2007, where all the raw differences were negative. The raw difference decreases as earnings increase and either becomes positive or remains as a small negative at the 50th and 75th quantiles before increasing again at the 90th quantile.

Figure 1: Raw gender wage gap by quantile, 2001–2010

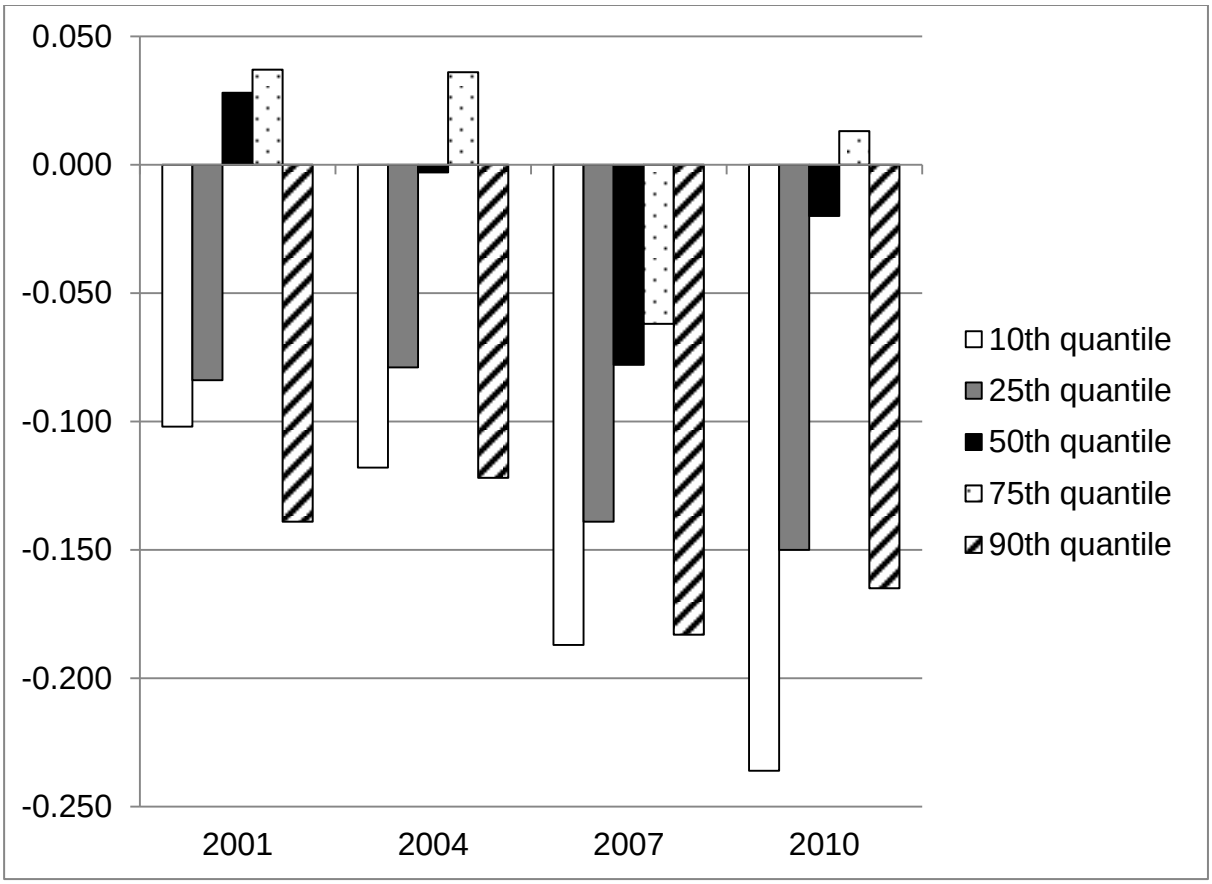
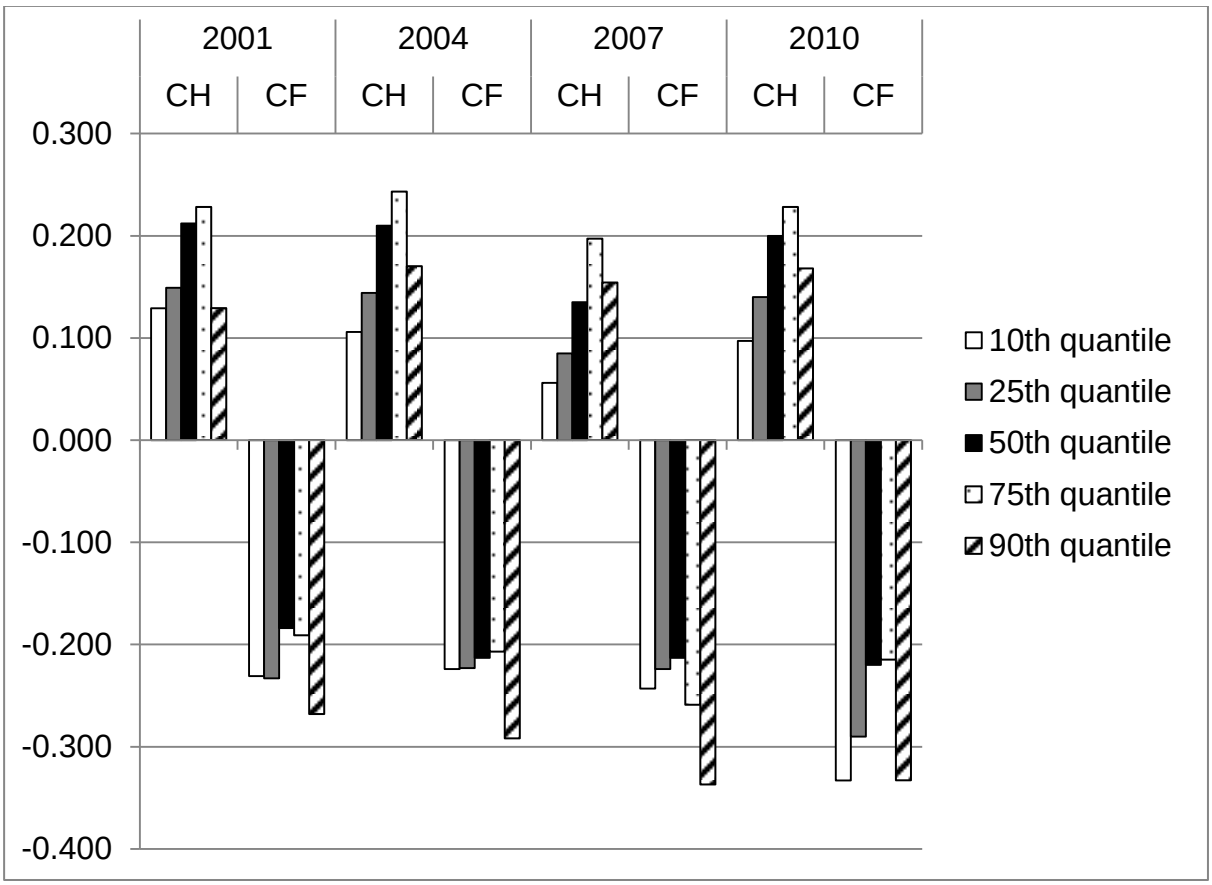


Figure 2 shows the characteristics and coefficients components, indicated as CH and CF respectively. The two components for each year in Figure 2 below add up to the raw difference represented in Figure 1 above. As it can be seen, the characteristics are always positive in each year. Moreover, they are highest at the 75th quantiles followed by the 50th quantile in most years. Conversely, the coefficients components are always negative and are generally at their lowest at the 50th and 75th quantiles. More importantly, the characteristics and coefficients components are roughly equal at the 50th and 75th quantiles, which is reflected in the low or positive raw differences in Figure 1. The coefficients components are largest at the 10th and the 90th quantiles. At all but the 50th and 75th quantiles, they are larger than the characteristics components. The coefficients components in 2004 are roughly equal from the 10th to the 75th quantiles, although the gender wage gaps are much lower at the 50th and 75th quantiles. This emphasises the importance of the size of the coefficients component for the gender wage gap, especially as it is reflected in the persistently negative gender wage gaps at the 10th, 25th and 90th quantiles, as reflected in Figure 1.

Figure 2: Gender wage gap decompositions, 2001–2010



A striking feature of the decomposition results was that the characteristics coefficients were positive and statistically significant at all the quantiles analysed. This implied that there was a female advantage in characteristics for this sample. Closer analysis of the characteristics and their distribution by gender showed that this could be attributed to the higher concentrations of women than men at higher educational levels and working in the public sector[5]. This is especially so at the 50th and 75th quantiles, where the characteristics components are at their highest. While the counterfactual decomposition of the gender wage gap does not indicate that women earn more than men due to their characteristics, these results suggest that women would get paid more than men for their characteristics were the returns to characteristics equal for women and men, thus supporting the human capital explanation of the gender wage gap.

The lower characteristics components at the 10th quantile could be driven by the high degree of homogeneity of characteristics between males and females, especially in their distribution by education level, occupation group, population group and industry. At the 90th quantile, the characteristics components are still positive, but lower than those at the 50th and 75th quantiles. This is likely due to higher heterogeneity in occupations between males and females – especially the high concentration of males in managerial positions compared to the female concentration in technical occupations. Even with high homogeneity at the 10th quantile, the gender wage gap was persistently negative and high. A similar observation was made at the 90th quantile.

As previously mentioned, the positive characteristics component values are outweighed by the negative and statistically significant coefficients components. Thus, where the gender wage gap is in favour of men and statistically significant, the factor driving this is the unequal returns to characteristics. The next section analyses the quantile regression results. The quantile regressions coefficients represent the returns to male and female characteristics, and thus this analysis tries to elucidate the persistently high coefficients components.

Quantile regression

After examining the gender wage decompositions and accounting for the characteristics components, the next step is to account for the coefficients components. Running separate quantile regressions for males and females enables us to observe which

[5] Tables containing these figures can be requested from the authors.

characteristics carry different returns for men and women, thus accounting for the coefficients components of the gender wage gap. This section presents the analysis of the quantile regression results which are found in Appendix B. The quantile regressions show the different returns to the control variables – the characteristics – for males and females. The objective was to see which characteristics yielded the biggest differences between male and female returns. If all the variables had consistently stark differences indicating a male advantage, then it could be concluded that it was the differing returns to observable characteristics that led to the gender wage gap.

As with the decomposition analysis, the quantile regressions were calculated at the 10^{th}, 25^{th}, 50^{th}, 75^{th} and the 90^{th} quantiles. For the dummy variables, the interpretation of the coefficients follows convention, that is, each coefficient represents the percentage difference in log hourly earnings between that particular category of the control variable and the base category.

When the male and female coefficients are compared, however, it is important to note the following: while male returns may be higher than female returns for some categories and not at all for others, the overall picture is more important. Because male earnings are generally higher than female earnings at all quantiles, higher male returns are likely to explain the size of the coefficient components. However, if female returns are higher than male returns to a given category, then they are likely to have a dampening effect on the coefficients component, pointing to a lower contribution to the coefficients component for that particular category. Moreover, in some cases such as the managerial occupation category, higher female returns are likely to be a result of an increase of females in a category that previously had few females. As Mwabu and Schultz (1996: 335) note in their study of racial returns to education, “a group will have higher returns to education if there are initially fewer of them with that particular level of education and this declines as they increase”.

For the dummy variables, each coefficient shows the percentage difference in log hourly earnings between each category and the relevant base category. For the continuous variables the coefficients just show the relationship between the variable and log hourly earnings.

The variables with the most obvious male advantage in returns were race, higher levels of education, work experience, public sector employment at higher quantiles, marital status and living in the rural areas. These variables are thus all likely to play a significant role in the coefficients component. These results are interesting, especially where they are related to the returns to education and to public sector employment. A quick scan of the composition of the characteristics by gender indicates that these were some of the primary drivers of the positive characteristics components. Thus while there was a higher female concentration at higher levels of education and in the public sector, particularly at the 50th and 75th quantiles, the same characteristics yielded higher returns for males than for females.

While the aforementioned variables are likely to have contributed significantly to the coefficients component, this did not occur for all the characteristics included in the model. It also did not occur at all quantiles. This observation indicates that the unequal returns to characteristics could also have male advantages in unobservable characteristics such as higher ability and intelligence. It could also indicate that there is a level of discrimination between male and female workers in the formal sector. This is especially so at the 10th quantile, where there is a higher homogeneity of male and female characteristics. At higher quantiles where the characteristics components are also positive but even higher, this indicates that discrimination plays an even larger role. As the positive characteristics indicate a female advantage if females were paid as males are, higher characteristics components that are accompanied by even higher coefficients indicate that discrimination is likely to play an even larger role than when the characteristics are lower.

The results in this section show that there seems to be both a sticky floor effect as well as a glass ceiling effect largely due to differences in characteristics such as education level and occupational distribution and discrimination. While these results cannot be directly compared to previous quantile regression analyses of the South African gender wage gap, namely the results of Goga (2008), Ntuli (2007) and Shepherd (2008), the decomposition analysis yielded similar results. Both Ntuli (2007) and Shepherd (2008) found that the coefficients components were higher than the characteristics components and that there was a sticky floor effect, which is also one of the findings of this study. In all the studies mentioned above, the gender wage gap was wider for the black population group at the bottom of the wage distribution than at the top. Since the 10th quantile was largely populated by black males and females, this result is in line with the results from previous

studies. Compared to the international results, at lower parts of the wage distribution South African results are similar to those found in developing countries such as those presented in Montenegro (2001), Sakellariou (2004) and Pham and Reilly (2007). Conversely, at higher parts of the wage distribution South African results are similar to those in developed countries such as those presented in Gardeazabal and Ugidos (2005), Garcia *et al.* (2007) and de la Rica *et al.* (2008).

DISCUSSION AND CONCLUSION

The gender wage gap in South Africa has received increasing attention over time, with more recent studies focusing on the gender wage gap along the wage distribution. Given the level of inequality in South Africa, this move has led to a broader understanding of the different dynamics that drive gender wage inequality at different points of the wage distribution.

Unlike previous literature, this study examined the entire sample of South African employees, as opposed to one racial group, as in the case of Ntuli (2007) and Shepherd (2008), or the entire working age population, as in the case of Goga (2008). The sample chosen in this study was important because in the analysis of gender wage discrimination the earnings of self-employed workers cannot be included, as they do not entirely depend on someone else for their earnings. Therefore, the factors that determine their earnings as well as the presence and nature of any discrimination which they face are different to those faced by people who are not self-employed.

The decomposition of the overall gender wage gap yielded three striking results: consistently positive and statistically significant characteristics components, persistently negative and statistically significant coefficients components and unequal gender wage gaps along the wage distribution.

The first striking finding was that the characteristics component was always positive and statistically significant. That is, the component of the gender wage gap that is attributed to differences in male and female characteristics was positive. This implies that the differences in characteristics are such that if women were paid as men are, and if their earnings were related only to the characteristics in the model, women would earn more.

In order to understand why the characteristics components were always positive, an analysis of the distribution of the characteristics was carried out by gender at some quantiles. For example, the number of men and women living in each province was analysed at each quantile to get a sense of where men and women were concentrated. If it had so happened that there were much more women than men living in Gauteng or the Western Cape, provinces that generally exhibit high earnings, this could have been a factor that explained the positive characteristics component.

The analysis of the characteristics component revealed that at lower quantiles there were more women concentrated at higher levels of education than men. This means that women at the 10^{th} and 25^{th} quantiles were located at higher levels of education than men at the 10^{th} and 25^{th} quantiles of the male earnings distribution. In addition to this, at the 50^{th} and 75^{th} quantiles, women were largely concentrated in the public sector in general and relative to men. At these quantiles, women were also found at clerical and technical occupations, while men were mostly operators and craft workers. While the distributional analysis of the characteristics is crude and does not provide conclusive evidence of the reasons for the positive characteristics components, it still points to the possible reasons for these. Moreover, the seemingly narrow focus on education levels and the public sector reflects where the largest difference in characteristics of men and women lay.

The second striking finding was that the coefficients components were always negative, statistically significant and larger than the characteristics components, bar for some instances of the 75^{th} quantile. Consequently, the gender wage gaps at most quantiles were also negative, with the quantiles that had the highest difference between characteristics and coefficients components having the highest gender wage gaps. As found by Ntuli (2007), the size and the sign of the coefficients component played a significant in the size and the sign of the gender wage gap. The bigger coefficients components imply that men received higher returns for similar characteristics than women. These higher returns could hypothetically be due to unobserved traits such as ability or superior on-the-job training.

However, while ability and intelligence are unobservable and are thus assumed to be contained in the “unexplained” component – part of the coefficients component – of the decomposed gender wage gap, not all ability and intelligence can be said to be “unexplained” or strictly unobservable. Having high levels of education and holding high-level jobs should imply high levels of ability and intelligence. Thus, that ability and

intelligence, though unobservable, are embedded in the characteristics component as well as the 'coefficients' component. If we observe that a person has tertiary education, it is impossible to separate the ability that comes with the tertiary qualification from the fact that the person holds the qualification. It may then be fallacious to assume that the ability is contained only in the coefficients or unexplained component. This implies that the higher returns to male characteristics cannot be viewed only as an indicator of higher ability or intelligence for men, but could point to more systematic discrimination.

The third striking point is that the gender wage gaps were not equal along the wage distribution. In particular, at the 50^{th} and 75^{th} quantiles, the gender wage gaps were either not statistically significant or, as in some cases at the 75^{th} quantile, were in favour of women. This result was driven by the high characteristics components, which were almost equal to the coefficients components. At these quantiles, the suggestion that factors such as ability and training are the "unobservable" in the coefficients components does not seem to apply, because the characteristics of women such as high education levels and being found in professional and technical occupations would imply high levels of ability. Thus upon first looking at the gender wage gap one would assume that there is no discrimination. However, an examination of the decompositions shows that both the characteristics and the coefficients components are large.

A corollary of the decomposition of the gender wage gap is the assessment of gender wage discrimination. As gender wage discrimination cannot be directly observed, especially given the abolition of laws that allowed for such kinds of discrimination, it is deemed to be a part of the coefficients component. This is because the coefficients component captures the differing returns to similar characteristics. Barring factors such as unequal ability, intelligence and different on-the-job training, one of the reasons for that could be put forward for unequal returns to similar characteristics is discrimination. In this sense, since the coefficients component was always negative it may be concluded that there was discrimination. However, this discrimination can manifest in different ways. At the 10^{th} quantile, the coefficients components were particularly high, but they were accompanied by relatively low characteristics components. Thus, as there were low differences between male and female characteristics at lower quantiles, it is likely that there was some level of direct gender wage discrimination at lower quantiles. At the 90^{th} quantile, the coefficients components were also particularly high. However, the characteristics components were higher than those found at the 10^{th} quantile. Since they

were positive, this means that women at higher quantiles should have an even greater advantage, but because of high coefficients, the gender wage gap is negative. Thus, the coefficients component, and consequently discrimination, also plays an important role in the gender wage gap at the top of the wage distribution.

By analysing the gender wage gap among employees alone, this research has provided a more specific focus and thus helped to build conclusions that are more directly related to wage discrimination among people who actually receive wages, employees. In particular, this research has shown that the gender wage gap is largely driven by differing returns to similar characteristics for men and women, regardless of the size and the sign of the characteristics components. The assessment of the gender wage gap at different points of the wage distribution rather than the mean was vindicated in this study, as it was shown that the gender wage gap had different magnitudes at different points of the wage distribution.

REFERENCES

Bhorat, H. & Leibbrandt, M. (1999). Modelling vulnerability and low earnings in the South African labour market. *DPRU working paper 99/32*

Blau, F. D. & Kahn, L. M. (2003). Understanding international differences in the gender pay gap. *Journal of Labor Economics,* 21(1):106-144

Burger, R.P. & Yu, D. (2006). Wage Trends in Post-Apartheid South Africa: Constructing an Earnings Series from Household Survey Data. *Stellenbosch Economic Working Papers: 04/06.* Stellenbosch: Stellenbosch University.

Casale, D. (2004). What has the feminization of the labour market 'bought' women in South Africa? Trends in labour force participation, employment and earnings, 1995-2001. *Development Policy Research Unit Working Paper number 04/84*

Casale, D. & Posel, D. (2010). Unions and the gender wage gap in South Africa. *Journal of African Economies,* 20(1): 27-59

De la Rica, S., Dolado, J.J. & Llorens, V. (2008). Ceilings or floors? Gender wage gaps by education in Spain. *Journal of Population Economics,* 21(3): 751-776

Garcia, J., Hernandez, P.J. and Lopez-Nicolas, A. (2001). How wide is the gap? An investigation of gender wage differences using quantile regression. *Applied Economics,* (26): 149-167

Gardeazabal, J. & Ugidos, A. (2005). Gender wage discrimination at quantiles. *Journal of Population Economics,* 18(1): 165-179

Goga, S. (2008). Understanding the gender earnings gap in the post-apartheid South African labour market. *Unpublished Master's thesis, University of Kwazulu-Natal*

Grün, C. (2004). Direct and indirect gender discrimination in the South African labour market. *International Journal of Manpower,* 25(3):321-342

Gujarati, D. (2002). *Basic Econometrics.* New York. McGraw-Hill.

Hinks, T., (2002). Gender wage differentials and discrimination in the New South Africa. *Applied Economics.* 34(6): 2043-2052

Isemonger, A. G. & Roberts, N. J. (1999). Post-entry gender discrimination in the South African labour market. *Journal for Studies in Economics and Econometrics*, 23(2): 1-25.

Juhn, C., Murphy, K.M. & Pierce, B. (1993). Wage inequality and the rise in returns to skill. *The Journal of Political Economy,* 101(3): 410-442

Koenker, R. & Bassett G. (1978). Regression quantiles. *Econometrica,* 46(1): 33-50

Machado, J.A.F. & Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics,* 20(4): 445-465

Makgetla, N. S. (2004). Women and the economy. Available: http://www.genderstats.org.za/ [Accessed: 31 March 2010]

Melly, B. (2007). Rqdeco: A stata module to decompose differences in distribution. Mimeo

Mincer, J. (1974). *Schooling, experience and earnings*. Massachusetts, NBER

Mincer, J. & Polachek, S. (1974). Family investments in human capital: Earnings of women. *The Journal of Political Economy,* 82(2): 76-108

Montenegro, C. (2001). Wage distribution in Chile: Does gender matter? A quantile regression approach. *World Bank Gender and Development Working Paper No. 20*

Muller, C. (2009). Trends in the gender wage gap and gender discrimination among part-time and full-time workers in post-apartheid South Africa. *University of KwaZulu-Natal working paper number 124*

Mwabu, G. & Schultz, T.P. (1996). Education returns across quantiles of the wage function: Alternative explanations for returns to education by race in South Africa. *The American Economic Review,* 86(2): 335-339

Ntuli, N. (2007). Exploring gender wage discrimination in South Africa 1995-2004: A quantile regression approach. *IPC Working Paper Series Number 56*

Oaxaca, R. (1973). Male-female wage differentials in urban labour markets. *International Economic Review,* 14(3):693-709

Pham, T-H. & Reilly, B. (2007). The gender pay gap in Vietnam, 1993 – 2002. A quantile regression approach. *Journal of Asian Economics,* 18(5): 775-808

Plantenga, J. & Remery, C. (2006). The gender pay gap. Origins and policy responses. A comparative review of thirty European countries. *The co-ordinators' synthesis report prepared for the Equality Unit, European Commission*

Rogers, W. (1992). Quantile regression standard errors. *Stata Technical Bulletin,* 9(1992): 12-14

Rospabe, S. (2001). An empirical evaluation of gender discrimination in employment, occupation attainment and wage in South Africa in the late 1990s. *DPRU working paper*

Sakellariou, C. (2004). Gender-earnings differentials using quantile regression. *Journal of Labor Research,* 25(3): 458-468

Shepherd, D. (2008). Post-Apartheid Trends in Gender Discrimination in South Africa: Analysis through Decomposition Techniques. *Bureau of Economic Research working paper number 06/08.*

The Economist. (30 December 2009). Women in the workforce: Female power. *The Economist.* Available from: http://www.economist.com/node/15174418 (Accessed: 8 June 2012)

Tregenna, F. (2011). Earnings inequality and unemployment in South Africa. *International Review of Applied Economics,* 25(5): 585-598

Winter, C. (1999). Women workers in South Africa: Participation, pay and prejudice in the formal labor market. World Bank Informal Discussion Paper Series number 19752

APPENDICES

Appendix A

Variable	Base category
Province	Gauteng
Race	Black
Level of education	Secondary education completed
Occupation	Plant and machine operators
Industry	Manufacturing
Sector (public vs. private)	Private sector
Unionisation (member vs. non-member)	Trade union member
Marital status	Married
Area (urban versus rural)	Urban

Appendix B: Quantile regressions: Table B1: Quantile regression results, 2001

	MALE					FEMALE				
Variable	10th	25th	50th	75th	90th	10th	25th	50th	75th	90th
PROVINCE										
WC	-0.026	-0.064	-0.047	-0.131***	-0.108*	-0.149**	-0.096	-0.108**	-0.010	-0.011
EC	-0.309***	-0.283***	-0.209***	-0.154***	-0.231***	-0.379***	-0.368***	-0.247***	-0.192***	-0.159***
NC	-0.411***	-0.329***	-0.222***	-0.163***	-0.149***	-0.563***	-0.422***	-0.376***	-0.298***	-0.169**
FS	-0.366***	-0.377***	-0.243***	-0.198***	-0.139***	-0.583***	-0.499***	-0.341***	-0.264***	-0.182***
KZN	-0.025	-0.101***	-0.052*	-0.084***	-0.074*	-0.244***	-0.226***	-0.220***	-0.202***	-0.134***
NW	-0.163***	-0.155***	-0.132***	-0.093**	-0.084**	-0.407***	-0.294***	-0.292***	-0.163***	-0.103*
MP	-0.083	-0.094**	-0.052	-0.013	0.124*	-0.265***	-0.268***	-0.207***	-0.157***	-0.179***
LMP	-0.327***	-0.299***	-0.179***	-0.08	-0.067	-0.467***	-0.295***	-0.228***	-0.227***	-0.128***
POPULATION GROUP										
Coloured	0.367***	0.323***	0.295***	0.321***	0.277***	0.329***	0.304***	0.296***	0.148**	0.152***
Asian	0.311**	0.392***	0.332***	0.363***	0.433***	0.462***	0.357***	0.286***	0.292***	0.197***
White	0.623***	0.696***	0.690***	0.709***	0.767***	0.488***	0.510***	0.507***	0.420***	0.469***
LEVEL OF EDUCATION										
No school	-0.186**	-0.189***	-0.161***	-0.163***	-0.214***	-0.224	-0.153*	-0.060	-0.084	-0.129
Primary not complete	-0.069	-0.098***	-0.080***	-0.084***	-0.104***	-0.088	-0.116*	-0.023	0.002	-0.033
Secondary not complete	0.1**	0.111***	0.152***	0.226***	0.249***	0.171**	0.084*	0.126**	0.189***	0.275
Secondary complete	0.369***	0.347***	0.390***	0.459***	0.52***	0.465***	0.310***	0.353***	0.437***	0.460***
Tertiary	0.556***	0.614***	0.701***	0.745***	0.786***	0.811***	0.544***	0.606***	0.674***	0.752***
OCCUPATION										
Manager	0.677***	0.664***	0.756***	0.745***	0.930***	0.834***	0.924***	0.959***	1.069***	0.928***
Professional	0.656***	0.537***	0.595***	0.679***	0.986***	0.671***	0.805***	0.852***	0.793***	0.699***
Technical	0.335***	0.316***	0.323***	0.306***	0.370***	0.431***	0.465***	0.544***	0.549***	0.416***
Clerk	0.157***	0.125**	0.214***	0.141***	0.127***	0.308***	0.310***	0.393***	0.412***	0.278***
Service	0.029	0.000	-0.001	0.013	0.054	-0.025	-0.008	0.075	0.147	0.101
Skilled agriculture	-0.192	-0.006	-0.031	0.088	0.405**	-0.168	-0.519	0.022	0.309	0.392
Craft	0.084*	0.087**	0.049**	0.029	0.049	-0.038	-0.016	-0.003	0.057	-0.065
Elementary/ domestic	-0.132***	-0.147***	-0.162***	-0.212***	-0.180***	-0.049	-0.090	-0.075	-0.067	-0.133*
INDUSTRY										
Agriculture	-0.637***	-0.669***	-0.790***	-0.733***	-0.715***	-0.277***	-0.520***	-0.578***	-0.650***	-0.613***
Mining	0.231***	0.161***	0.102***	0.046	0.057	0.158	-0.051	0.058	0.171	-0.096
EGW	0.009	-0.007	0.061	0.097	0.091	0.116	0.064	0.149	0.114	0.154
Construction	-0.075	-0.117**	-0.192***	-0.166***	-0.142**	-0.172	0.142	0.110	0.089	0.057
Wholesale	-0.197***	-0.183***	-0.188***	-0.232***	-0.247***	-0.256***	-0.266***	-0.233***	-0.258***	-0.239***
Transport	0.008	-0.035	-0.032	0.011	0.031	0.262	0.156*	0.089	0.071	0.063
Financial	-0.021	-0.009	0.009	-0.016	-0.047	-0.010	0.105	0.094**	0.101	0.183***
CSPHH	-0.114	-0.038	-0.053	-0.079*	-0.046	-0.082	-0.025	0.007	-0.057	-0.013
OTHER VARIABLES										
Hours worked	0.000	0.000	0.000	0.000***	0.000***	0.000	0.000	0.000	0.000	0.000
WE	0.032***	0.026***	0.026***	0.030***	0.024***	0.029***	0.021***	0.015***	0.012***	0.014***
WE squared	0.000***	0.000***	0.000***	0.000***	0.000***	0.000**	0.000***	0.000*	0.000	0.000
Public	0.249***	0.197***	0.179***	0.168***	0.101*	0.170***	0.269***	0.179***	0.192***	0.088*
Union	0.327***	0.289***	0.224***	0.156***	0.144***	0.335***	0.298***	0.256***	0.190***	0.131***
Married	0.160***	0.131***	0.150***	0.107***	0.151***	0.109**	0.053*	0.061**	0.092***	0.115***
Rural	-0.056	-0.077***	-0.110***	-0.157***	-0.143***	-0.175***	-0.107**	-0.128***	-0.147***	-0.130***
Intercept	-0.343***	0.157***	0.474***	0.798***	1.074***	-0.494***	0.049	0.356***	0.654***	0.933*

**** 1% level of significance; ** 5% level of significance; * 10% level of significance; WE = work experience; The province abbreviations are explained as follows: WC = Western Cape; EC = Eastern Cape; NC = Northern Cape; FS = Free State; KZN = KwaZulu-Natal; NW = North West; MP = Mpumalanga; LMP = Limpopo*

Table B2: Quantile regression results, 2004

	MALE					FEMALE				
Variable	10th	25th	50th	75th	90th	10th	25th	50th	75th	90th
PROVINCE										
WC	-0.139*	-0.097**	-0.077**	-0.094**	-0.185***	-0.045	-0.070	-0.141***	-0.168***	-0.222***
EC	-0.311***	-0.267***	-0.138***	-0.150***	-0.120**	-0.372***	-0.282***	-0.263***	-0.186***	-0.204***
NC	-0.427***	-0.324***	-0.213***	-0.164***	-0.082	-0.410***	-0.355***	-0.330***	-0.243***	-0.251***
FS	-0.399***	-0.366***	-0.302***	-0.253***	-0.269***	-0.510***	-0.438***	-0.368***	-0.343***	-0.363***
KZN	-0.113**	-0.115**	-0.050	-0.025	-0.019***	-0.137**	-0.105**	-0.136***	-0.110**	-0.175***
NW	-0.181***	-0.155***	-0.142***	-0.127***	-0.192***	-0.140**	-0.200***	-0.129*	-0.044	-0.074
MP	-0.189***	-0.195***	-0.140***	-0.135***	-0.013	-0.438***	-0.329***	-0.320***	-0.253***	-0.248**
LMP	-0.333***	-0.214***	-0.164***	-0.158***	-0.204***	-0.226***	-0.219***	-0.128**	-0.166***	-0.209***
POPULATION GROUP										
Coloured	0.317***	0.306***	0.295***	0.323***	0.319***	0.085	0.243***	0.292***	0.279***	0.259***
Asian	0.460***	0.426***	0.454***	0.488***	0.357***	0.242**	0.207***	0.181***	0.180**	0.299**
White	0.674***	0.699***	0.705***	0.657***	0.669***	0.475***	0.523***	0.577***	0.563***	0.593***
LEVEL OF EDUCATION										
No school	-0.058	-0.101	-0.153***	-0.145**	-0.222***	0.047	-0.165	-0.193**	-0.118	-0.075
Primary not complete	-0.042	-0.071	-0.085**	-0.072	-0.089	0.057	-0.035	0.024	0.014	0.087
Secondary not complete	0.138	0.110**	0.160***	0.169***	0.208***	0.323***	0.224***	0.179***	0.210***	0.323***
Secondary complete	0.361***	0.322***	0.380***	0.448***	0.477***	0.511**	0.431***	0.370***	0.385***	0.529***
Tertiary	0.665***	0.594***	0.643***	0.767***	0.996***	0.790***	0.680***	0.600***	0.571***	0.754***
OCCUPATION										
Manager	0.732***	0.763***	0.744***	0.774***	0.801***	0.750***	0.854***	1.075***	0.996***	1.110***
Professional	0.520***	0.509***	0.575***	0.589***	0.744***	0.660***	0.873***	1.014***	0.974***	1.171***
Technical	0.262***	0.312***	0.319***	0.355***	0.463***	0.368***	0.511***	0.701***	0.689***	0.692***
Clerk	0.091	0.133**	0.171***	0.195***	0.225***	0.190*	0.312***	0.497***	0.483***	0.545***
Service	-0.044	-0.101**	-0.099***	-0.066	-0.024	-0.187	0.053	0.219***	0.141**	0.198**
Skilled agriculture	-0.202	-0.269***	-0.292***	-0.231**	-0.144	-0.485	0.092	0.210	-0.046	-0.173
Craft	-0.055	-0.034	-0.017	0.016	0.072	-0.161	-0.069	0.067	-0.071	0.067
Elementary/ domestic	-0.161***	-0.218***	-0.218***	-0.253***	-0.207***	-0.198**	-0.142***	-0.049	-0.104*	-0.032
INDUSTRY										
Agriculture	-0.253***	-0.406***	-0.552***	-0.601***	-0.562***	-0.130	-0.146**	-0.334***	-0.465***	-0.547***
Mining	0.466***	0.348***	0.304***	0.250***	0.253***	-0.305	-0.156	-0.056	0.413	0.562*
EGW	-0.052	-0.118	-0.121	-0.123	-0.206	0.380**	0.091	0.102	0.384	0.908**
Construction	-0.127*	-0.207***	-0.199***	-0.200***	-0.251***	-0.264	-0.333	-0.122	-0.025	-0.128
Wholesale	-0.230***	-0.269***	-0.268***	-0.252***	-0.217***	-0.191**	-0.293***	-0.322***	-0.284***	-0.339***
Transport	0.114	-0.010	-0.037	-0.033	-0.077	0.234**	0.145*	0.168**	0.230**	0.294**
Financial	0.031	-0.026	-0.075*	-0.098**	-0.044	0.117	0.020	0.048	0.060	-0.042
CSPHH	-0.058	-0.036	-0.038	-0.097**	-0.072	-0.077	-0.120*	-0.138**	-0.115	-0.232***
OTHER VARIABLES										
Hours worked	0.000***	0.000	0.000	0.000	0.000	0.001	0.001	0.001	0.001	0.003***
WE	0.023***	0.021***	0.020***	0.019***	0.023***	0.021***	0.018***	0.015***	0.012**	0.009*
WE squared	0.000***	0.000***	0.000***	0.000**	0.000**	0.000*	0.000	0.000	0.000	0.000
Public	0.186***	0.223***	0.305***	0.316***	0.272***	0.192***	0.244***	0.269***	0.250***	0.218***
Union	0.330***	0.273***	0.208***	0.180***	0.137***	0.458***	0.379***	0.297***	0.229***	0.244***
Married	0.130***	0.102***	0.070***	0.096***	0.151***	0.006	0.050*	0.038	0.045	0.065**
Rural	-0.213***	-0.164***	-0.099***	-0.099***	-0.068*	-0.314***	-0.278***	-0.181***	-0.132***	-0.123***
Intercept	-0.048	0.446***	0.733***	1.023***	1.183***	-0.281	0.057	0.349***	0.735***	0.781***

**** 1% level of significance; ** 5% level of significance; * 10% level of significance; WE = work experience ; The province abbreviations are explained as follows: WC = Western Cape; EC = Eastern Cape; NC = Northern Cape; FS = Free State; KZN = KwaZulu-Natal; NW = North West; MP = Mpumalanga; LMP = Limpopo*

Table B3: Quantile regression results, 2007

	MALE					FEMALE				
Variable	10th	25th	50th	75th	90th	10th	25th	50th	75th	90th
PROVINCE										
WC	0.029	-0.017	-0.014	-0.015	-0.128*	0.029	0.057	-0.064	-0.036	-0.239***
EC	-0.313***	-0.223***	-0.151***	-0.121***	-0.197***	-0.348***	-0.267***	-0.268***	-0.249***	-0.322***
NC	-0.263***	-0.230***	-0.151***	-0.064	-0.018	-0.252***	-0.282***	-0.278***	-0.237***	-0.305***
FS	-0.258***	-0.275***	-0.179***	-0.117***	-0.188***	-0.353***	-0.314***	-0.298***	-0.235***	-0.355***
KZN	-0.152**	-0.137***	-0.112***	-0.092**	-0.141***	-0.234***	-0.220***	-0.177***	-0.216***	-0.296***
NW	-0.222***	-0.210***	-0.197***	-0.162***	-0.178***	-0.256***	-0.215**	-0.148***	-0.178**	-0.269***
MP	-0.032	-0.055	-0.034	0.004	0.045	-0.208***	-0.225***	-0.175***	-0.228***	-0.337***
LMP	-0.211***	-0.206***	-0.171***	-0.213***	-0.212***	-0.382***	-0.249***	-0.178***	-0.191***	-0.279***
POPULATION GROUP										
Coloured	0.137**	0.195***	0.235***	0.259***	0.383***	0.230***	0.190***	0.240***	0.170***	0.288***
Asian	0.297***	0.280***	0.359***	0.457***	0.683***	0.519***	0.570***	0.485***	0.537***	0.702***
White	0.636***	0.631***	0.731***	0.766***	0.804***	0.620***	0.501***	0.553***	0.501***	0.512***
LEVEL OF EDUCATION										
No school	-0.071	-0.114***	-0.155**	-0.113*	-0.141*	-0.465	-0.410*	-0.281**	-0.071	0.105
Primary not complete	-0.156	-0.037	-0.060	-0.051	-0.069	-0.212	-0.066	-0.121*	-0.060	-0.124
Secondary not complete	0.173**	0.194***	0.166***	0.224***	0.205***	0.100	0.179**	0.155***	0.199***	0.201**
Secondary complete	0.384***	0.416***	0.418***	0.463***	0.479***	0.249***	0.328***	0.319***	0.343***	0.329***
Tertiary	0.788***	0.868***	0.863***	1.046***	1.209***	0.826***	0.795***	0.764***	0.779***	0.773***
OCCUPATION										
Manager	0.561***	0.504***	0.675***	0.586***	0.743***	0.517***	0.709***	0.864***	1.042***	0.885***
Professional	0.346***	0.403***	0.397***	0.517***	0.637***	0.179	0.378***	0.507***	0.610***	0.534***
Technical	0.196***	0.197***	0.210***	0.163***	0.127*	0.169	0.275***	0.433***	0.593***	0.443***
Clerk	-0.005	0.129**	0.112**	0.134**	0.071	0.266***	0.365***	0.398***	0.491***	0.342**
Service	-0.144**	-0.165***	-0.151***	-0.145***	-0.118*	-0.117	0.028	0.150**	0.215***	0.100
Skilled agriculture	0.069	-0.020	-0.026	0.012	0.004	0.230	0.077	0.093	-0.025	-0.218
Craft	-0.046	0.009	0.078**	0.017	-0.007	-0.010	0.091	0.158**	0.141	0.013
Elementary/ domestic	-0.205***	-0.198***	-0.175***	-0.207***	-0.199***	-0.168*	-0.081	0.003	0.018	-0.079
INDUSTRY										
Agriculture	-0.185***	-0.332***	-0.482***	-0.574***	-0.527***	0.097	0.068	-0.156***	-0.395***	-0.442***
Mining	0.378***	0.300***	0.290***	0.260***	0.249***	0.263*	0.398***	0.331***	0.402**	0.658**
EGW	-0.058	0.086	0.159	0.104	0.196	-1.208*	-0.014	0.291	0.329	0.421
Construction	-0.080	-0.075	-0.133***	-0.079*	-0.059	-0.207*	-0.267***	-0.221**	-0.225*	-0.270**
Wholesale	-0.245***	-0.220***	-0.218***	-0.166***	-0.183***	-0.070	-0.014	-0.105*	-0.274***	-0.232***
Transport	-0.088	-0.184***	-0.146***	-0.133**	-0.130	-0.783**	-0.215	0.078	0.119	0.090
Financial	-0.061	-0.067	-0.054	-0.058	0.061	0.075	0.194***	0.061	-0.023	0.155
CSPHH	-0.096	-0.044	-0.005	-0.007	-0.011	-0.135*	0.043	-0.042	-0.156	-0.084
OTHER VARIABLES										
Hours worked	0.000	0.000	0.000	0.000	0.000	0.000**	0.000	0.000**	0.000	0.000
WE	0.023***	0.018***	0.015***	0.014***	0.010*	0.006	0.002	0.010***	0.013**	0.012*
WE squared	0.000***	0.000**	0.000*	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Public	0.197***	0.231***	0.181***	0.189***	0.201***	0.238***	0.310***	0.359***	0.398***	0.276***
Union	0.182***	0.214***	0.188***	0.160***	0.207***	0.485***	0.380***	0.313***	0.201***	0.198***
Married	0.117***	0.108***	0.117***	0.167***	0.217***	-0.027	0.023	0.036	0.050	0.099**
Rural	-0.211***	-0.173***	-0.165***	-0.117***	-0.114***	-0.287***	-0.233***	-0.178***	-0.127***	-0.106**
Intercept	0.038	0.330***	0.667***	0.921	1.233	-0.165	0.022	0.270***	0.657***	1.144***

**** 1% level of significance; ** 5% level of significance; * 10% level of significance; WE = work experience; The province abbreviations are explained as follows: WC = Western Cape; EC = Eastern Cape; NC = Northern Cape; FS = Free State; KZN = KwaZulu-Natal; NW = North West; MP = Mpumalanga; LMP = Limpopo*

Table B4: Quantile regression results, 2010

	MALE					FEMALE				
Variable	10th	25th	50th	75th	90th	10th	25th	50th	75th	90th
PROVINCE										
WC	0.114*	0.047	-0.026	-0.059	-0.136**	0.091	-0.055	-0.096**	-0.157*	-0.141**
EC	-0.224***	-0.169***	-0.153***	-0.104**	-0.077	-0.180*	-0.154***	-0.154***	-0.177***	-0.091
NC	-0.136*	-0.123*	-0.126***	-0.060	-0.162**	-0.251*	-0.212***	-0.263***	-0.235***	-0.154**
FS	-0.094	-0.115***	-0.151***	-0.130***	-0.165***	-0.318***	-0.306***	-0.284***	-0.258***	-0.221***
KZN	-0.149***	-0.084*	-0.066**	-0.084*	-0.092*	-0.103	-0.116**	-0.176***	-0.198***	-0.158**
NW	0.080	0.043	-0.008	-0.056	-0.074	-0.261**	-0.168***	-0.142***	-0.173***	-0.133
MP	0.039	0.087	0.096**	0.123**	0.163***	0.088	0.027	-0.051	-0.062	-0.059
LMP	-0.037	-0.003	-0.035	-0.026	-0.049	-0.299**	-0.172**	-0.079	-0.119**	-0.057
POPULATION GROUP										
Coloured	0.085	0.134***	0.135***	0.155***	0.214***	0.223***	0.210***	0.172***	0.166***	0.179***
Asian	0.368***	0.441***	0.438***	0.570***	0.622***	0.236**	0.192**	0.170**	0.187***	0.145
White	0.668***	0.639***	0.697***	0.708***	0.772***	0.523***	0.426***	0.348***	0.370***	0.311***
LEVEL OF EDUCATION										
No school	-0.293**	-0.105	-0.100	-0.029	-0.127	-0.084	-0.192	-0.241	-0.281**	-0.112
Primary not complete	-0.112	-0.101*	-0.115**	-0.028	-0.173**	0.054	-0.147*	-0.131	-0.112	-0.058
Secondary not complete	0.011	0.090	0.123***	0.174***	0.114	0.158	0.117	0.070	-0.031	0.026
Secondary complete	0.220**	0.308***	0.362***	0.472***	0.435***	0.448***	0.429***	0.376***	0.259***	0.338***
Tertiary	0.507***	0.670***	0.816***	0.940***	0.913***	1.066***	0.930***	0.780***	0.585***	0.629***
OCCUPATION										
Manager	0.717***	0.585***	0.533***	0.597***	0.833***	0.667***	0.954***	0.919***	1.081***	1.188***
Professional	0.687***	0.507***	0.510***	0.694***	0.969***	0.740***	0.890***	0.790***	0.850***	1.002***
Technical	0.372***	0.265***	0.190***	0.213***	0.159***	0.394***	0.684***	0.641***	0.609***	0.522***
Clerk	0.337***	0.195***	0.171***	0.133***	0.054	0.424**	0.531***	0.459***	0.519***	0.434***
Service	-0.084	-0.153**	-0.142***	-0.130***	-0.149***	0.035	0.172**	0.082	0.200*	0.201*
Skilled agriculture	0.004	-0.069	0.317	0.264**	0.276	0.158	0.353	0.097	0.363	0.447
Craft	0.050	0.036	-0.001	-0.005	0.119*	0.177	0.276***	0.207*	0.327*	0.344***
Elementary/ domestic	-0.125*	-0.195***	-0.232***	-0.198***	-0.222***	0.071	0.125	-0.044	-0.066	-0.187
INDUSTRY										
Agriculture	-0.127	-0.190***	-0.378***	-0.642***	-0.669***	-0.009	-0.098*	-0.190***	-0.410***	-0.499***
Mining	0.272***	0.386***	0.369***	0.315***	0.391***	0.159	0.307	0.434***	0.476***	0.716*
EGW	-0.677*	-0.205	0.064	0.001	-0.018	-0.089	0.239	0.214	0.056	0.090
Construction	-0.088	0.019	-0.106**	-0.195***	-0.213***	-0.229	-0.500***	-0.412***	-0.298*	-0.195*
Wholesale	-0.088	-0.078*	-0.091**	-0.222***	-0.168***	-0.045	-0.071	-0.126**	-0.219***	-0.173**
Transport	-0.119	0.065	0.047	-0.028	0.091	0.019	0.091	0.195**	0.192*	0.280**
Financial	0.042	0.022	-0.038	-0.170***	-0.087	0.099	0.129**	0.106**	0.095	0.125
CSPHH	0.050	0.102*	-0.029	-0.140**	-0.121*	-0.108	-0.066	-0.013	-0.123**	-0.049
OTHER VARIABLES										
Hours worked	0.003	0.002	0.000	0.001	0.000	0.012***	0.010***	0.006***	0.000	-0.002
WE	0.013**	0.015***	0.016***	0.022***	0.022***	0.020***	0.019***	0.020***	0.011**	0.007
WE squared	0.000*	0.000***	0.000***	0.000***	0.000***	0.000**	0.000**	0.000**	0.000	0.000
Public	0.078	0.222***	0.323***	0.336***	0.337***	0.163**	0.195***	0.213***	0.255***	0.191***
Union	0.394***	0.362***	0.321***	0.235***	0.207***	0.299***	0.360***	0.361***	0.333***	0.209***
Married	0.062	0.063**	0.079***	0.109***	0.102**	0.082	0.051*	0.018	0.020	0.077**
Rural	-0.162***	-0.153***	-0.123***	-0.148***	-0.185***	-0.181***	-0.189***	-0.201***	-0.248***	-0.212***
Intercept	0.033	0.388	0.823	1.107	1.534***	-1.060***	-0.574***	0.164	1.084***	1.533***

**** 1% level of significance; ** 5% level of significance; * 10% level of significance; WE = work experience; The province abbreviations are explained as follows: WC = Western Cape; EC = Eastern Cape; NC = Northern Cape; FS = Free State; KZN = KwaZulu-Natal; NW = North West; MP = Mpumalanga; LMP = Limpopo*