Understanding the Gender Earnings Gap in the Post-Apartheid South African Labour Market

Sumayya Goga
201500851

Supervisor: Dorrit Posel
Faculty of Management Studies
University of KwaZulu Natal
2008

Submitted in partial fulfilment of the requirements for the degree Masters in Economics

Declaration

I, Sumayya Goga, declare that this thesis represents original work that has not been previously submitted in any form to any university. Where use has been made of the work of others, this has been acknowledged and referenced in the text.

Sumayya Goga

Abstract

In this thesis, I analyse the gender earnings gap in South Africa using Labour Force Survey (LFS) data, for the period between 2001 and 2005. In addition to providing estimates of the gender earnings differential at the mean of the wage distribution (using a pooled regression), I also provide quantile regression estimates to account for the gap at different points of the distribution. To further explore reasons for the gender earnings gap, I separate the male and female earnings equations and employ a decomposition technique. This allows me to determine the proportion of the gap that is not explained by differences in observable characteristics between men and women. The 'unexplained' part of the earnings gap is suggestive of gender discrimination in the labour market.

Using Ordinary Least Squares (OLS) the pooled regression (controlling for sample selection), indicates an increase in the gender pay gap between 2001 and 2005. In turn, the quantile regression estimates for the period also illustrate a widening gender earnings differential throughout the distribution, except at the mean. By contrast, the descriptive statistics and the separated male and female earnings estimations show a decrease in the earnings gap over the period. Given that the pooled regression assumes the same returns to observable characteristics for males and females, which I reject through the use of a Chow test, the results from the separated estimations hold more weight.

The Oaxaca (1973) decomposition on the separated male and female earnings estimations illustrates that the 'unexplained' component of the gap accounts for a greater proportion of the gap than the 'explained' component in both years. Furthermore, the 'unexplained' proportion of the gap increased in the period, while the 'explained' proportion decreased. Thus, if the 'unexplained' part of the gap is considered to be a measure of discrimination, then the data indicates an increase in discrimination in earnings between the sexes over the period 2001 to 2005, even though there was a narrowing of the gender earnings differential.

Acknowledgements

I would like to particularly thank the following people:

Dori Posel, my supervisor, for her help, patience, and meticulous reading.

Haroon Bhorat, my director, for his help, and the occasional push.

My parents and family, for the constant reminders and support.

Contents

Chapter One: Introduction

There is a body of studies in the international literature that focuses on the gender earnings gap, its evolution and determinants. By contrast, due to South Africa's historically repressive racial policies, much of the work on earnings inequality in South Africa has focused on the race dimension. One of the markers of the post-apartheid South African economy has however been the feminisation of the labour force (Casale & Posel, 2002). Simultaneously, employment equity legislation in the post-apartheid period has placed greater focus on the representation and treatment of women in the labour market. Thus, the role and status of women in the labour market has taken on increased importance. Given the interest, both internationally and nationally, in women's experiences in the labour market as well as the dearth of studies on the gender earnings gap in South Africa, this thesis analyses the gender earnings differential in the post-apartheid period.

In Chapter Two of the thesis I undertake a review of some of the national and international literature on the gender earnings gap, focusing on methodology and data issues. Furthermore, I motivate for why I have chosen to use the Labour Force Surveys (LFSs) in my analysis of the gender earnings gap in South Africa. Studies on the gender earnings differential in South Africa have typically been single year studies (see for instance Hinks, 2002; Rospabé, 2001). 'Over time' comparisons have been complicated by comparability issues, since the October Household Surveys (OHSs) (administered from 1995 to 1999) and the LFSs (introduced in 2000) are not comparable based on how they measure employment (Casale, Muller and Posel, 2005). While some studies have considered the gender earnings gap over time, none of the recent studies for South Africa has used just the LFSs in considering the gender earnings differential.

In Chapter Three I first descriptively analyse changes in the gender pay gap over a range of covariates using mean earnings. Thereafter, in order to obtain a more nuanced picture of the gender earnings differential, I analyse changes in the gap along the distribution by calculating the gender earnings differential at the 10^{th}, 50^{th} and 90^{th} percentiles. I then turn to a multivariate analysis of earnings inequality in Chapter Four.

Gender earnings differential studies in South Africa have commonly calculated the gap at the mean of the wage distribution. Thus, in Chapter Four I first use Ordinary Least Squares (OLS) to do the same. While OLS regressions measure the impact of different covariates at the mean, quantile regressions take into account the fact that there may be different behaviour patterns in the coefficients at the conditional earnings distribution. Thus, in light of the fact that the international literature commonly uses quantile regressions to estimate the gap across the distribution, I furthermore also make use of this technique. In effect, with the quantile regression approach I can test whether the gender earnings gap is higher at the top or bottom end of the earnings distribution, and further, I can assess whether the gender earnings differential has been narrowing or widening across the earnings distribution between 2001 and 2005. For both the OLS estimation as well as the quantile regressions, I use a pooled dataset, and account for sample selection.[1]

The use of a pooled dataset assumes that the returns to observable characteristics are the same for both males and females. In the second part of Chapter Four, I relax the assumption of the same returns to observable characteristics for males and females by estimating separate earnings equations for both sexes. International studies on the gender earnings gap focus strongly on the drivers of the gap; specifically whether differences in men and women's earnings are driven by discrimination or other factors. This is most commonly done by using a decomposition technique such as the Oaxaca (1973) decomposition. The separation of the earnings equations by gender allows me to use the Oaxaca (1973) decomposition, and therefore to test whether the 'unexplained' proportion of the gender earnings gap, which is thought to be a measure of discrimination, has been increasing or decreasing over time.

In Chapter Five, I summarise the key findings of this study on the gender earnings gap in post-apartheid South Africa.

[1] Estimates that are uncorrected for sample selection for both the OLS estimation as well as the quantile regressions are presented in Appendix 3 and 4.

Chapter Two: The Gender Earnings Gap in a National and International Context: A Review of the Literature

Women throughout the world earn less than their male counterparts on average, and this is a consistently observed phenomenon (Polachek & Xiang, 2006). In looking at gender earnings gaps, researchers are interested on the one hand in the level of difference in earnings between men and women and how this has changed over time, and on the other hand in understanding the reasons for the differences between male and female earnings. Regarding the latter, the literature commonly uses the Oaxaca (1973) decomposition methodology to isolate the effect of the discrimination component on the overall gender earnings gap (see for example Beblo et al, 2003; Gonzales et al, 2005; Pena-Boquete et al, 2007). Thus, examining the gap in earnings between men and women is in itself interesting, but it is also useful as far as understanding gender-based discrimination is concerned.

Earnings gap and earnings discrimination studies in South Africa have predominantly focused on the race dimension (see for instance Mwabu & Schultz, 1996; Allanson et al, 2001; Erichsen & Wakeford, 2001; Rospabé, 2002; Chamberlain & Van der Berg, 2002; Burger & Jafta, 2006), with few studies having taken the gender dimension into account (see for example Isemonger & Roberts, 1999; Hinks, 2002; Rospabé, 2001; Gruen, 2004 and Ntuli, 2006). While this is unsurprising given South Africa's history of apartheid, studies on the gender earnings differential are beginning to take on increased importance, particularly since there is a large and developed international literature on the topic.

The international research on gender earnings gaps generally concentrates on country analyses, and consists of various different estimation techniques and methodologies. Weichselbaumer & Winter-Ebmer (2005) in a *Meta-Analysis of the International Gender Wage Gap* found that there are a variety of methods of estimating the gap, of which the dummy variable from the Mincer equation and the Oaxaca (1973) decomposition are popular. The dummy variable method entails estimating a pooled earnings regression, and including a female dummy in the earnings equation to isolate the difference between male and female earnings. This approach assumes that gender has only a shift or an intercept effect on earnings. The Oaxaca (1973) methodology in contrast, involves estimating separate male and female earnings equations, thereby assuming that the earnings equations for men and women may also have different slope coefficients. The Oaxaca decomposition then divides the earnings gap into two components: i) that component which can be accounted for by differences in individual productivity-enhancing characteristics such as schooling, occupation, industry and experience; leaving ii) the 'unexplained' component which cannot be accounted for by differences in observed characteristics and which typically is thought of as capturing discrimination in the labour market. Although the Oaxaca (1973) decomposition methodology is widely used, the limitations of this approach are acknowledged. One of the main criticisms of the Oaxaca (1973) decomposition has been the fact that the 'unexplained' component may also capture differences in unobserved characteristics (see for instance Kunze, 2007).

Both the dummy variable method and Oaxaca (1973) decomposition have been extended in various ways in the international literature. For instance, many country analyses using either of these methods include quantile regression estimates of the gap (see for instance Albrecht et al, 2004; Montenegro, 2001; Pham & Reilly, 2006; de la Rica et al, 2005). Quantile regressions are popular since they allow researchers to look at whether the pay gap and discrimination has been increasing or decreasing over the entire earnings distribution rather than simply at the mean of the distribution. In this way, it is possible to identify 'glass ceiling' and 'sticky floor' effects.[2] Quantile regressions were first undertaken by Buchinsky (1994) who used it to look at earnings disparities among women. However, his work prompted a large body of research using quantile regressions in the gender earnings gap literature.

The international literature that employs quantile regression techniques to understand the gender earnings gap more commonly uses separate male and female equations instead of a pooled estimation of earnings. Examples include Pham & Reilly (2006), Montenegro (2001) and de la Rica et al (2005), who employed the technique for Vietnam, Chile and Spain respectively. Many developing countries however, have tended to favour the dummy variable method. Example of these include studies by Hyder and Reily (2005) who looked at gender earnings disparities in Pakistan, Ajward and Kurukulasuriya (2002) who did the same for Sri Lanka, and Nielsen and Rosholm (2001) for Zambia.

More recently the international literature has begun to recognise that, aside from gender-specific factors such as differences in educational attainment and discrimination which aid in explaining gender earnings gaps, the overall earnings structure can also have a major impact on the relative earnings of different sub-groups in the labour market. Thus, differences in the level of the gender earnings gap between countries are increasingly explained by collective bargaining and compressed male earnings structures. For instance Blau and Kahn (2003) in a paper analysing wage structure and earnings differentials internationally found that the higher level of earnings inequality in the United States relative to other countries is the primary reason for its relatively high gender pay gap.

There are few studies in South Africa that deal directly with the gender earnings gap, and most studies have typically calculated gender discrimination only at the mean of the distribution (see for instance Hinks, 2002; Rospabé, 2001; Gruen, 2004). Only Ntuli (2006) who looked at gender earnings gaps for Africans in formal employment[3] in South Africa considered the gap and discrimination across the distribution through the use of quantile regressions. She estimated separate

---

[2] Glass ceilings generally refer to cases where earnings gaps are higher at the top of the distribution, while sticky floors refer to larger earnings gaps at the bottom of the earnings distribution.

[3] Ntuli (2006) includes domestic workers in her sample. It is difficult to ascertain how Ntuli generated a sample of earnings for the formally employed using the October Household Survey, since it is not possible from the information provided in the survey to identify formality of earnings employment.

male and female earnings equations using quantile regressions, and then employed a decomposition technique to determine whether a 'glass ceiling' or 'sticky floor' exists.

Ntuli (2006) found that the magnitude of the earnings gap decreased at the upper tails of the earnings distributions. She further found that the counterfactual earnings gaps generally declined as one proceeds from the bottom to the upper tails of the earnings distribution. However, the counterfactual earnings gaps did not show a declining tendency between 1995 and 2004 though a slight decline was evident at the 10th, 25th and 90th percentiles. Ntuli (2006) thus concluded that if discrimination is the main factor that drives these pay gaps, then female workers in the upper quantiles became more disadvantaged with time.

By contrast, other gender pay gap studies in South Africa looked primarily at the mean of the earnings distribution. Hinks (2002) for instance looked at gender earnings differentials and discrimination in South Africa using the 1995 OHS. He found a relatively small gap between male and female earnings for the African and Coloured population, while gender earnings differentials were largest for Indians and Whites. The decomposition analysis showed that the discrimination component was much larger for the latter two population groups. In fact, total discrimination was smallest for the African cohort. Hinks (2002) explained the relatively small earnings differentials among the African and Coloured population groups by the fact that earnings themselves were so low for these two cohorts. It therefore followed that consequent gender differentials and discrimination would also be quite low.

While Ntuli (2006) and Hinks (2002) analysed just the earnings gap and discrimination among those with employment, Rospabé (2001) and Gruen (2004) both analysed discrimination in access to employment as well as in earnings through the use of the Oaxaca (1973) decomposition technique. Using the 1999 OHS, Rospabé (2001) found that women were confined to a large extent to the bottom end of the skills categories, and a great part of the disparity in employment by occupation could thus be explained by discrimination in access to employment. She further found that large earnings inequalities prevailed between men and women workers, especially among African and White workers, and once more that much of the gap could be attributed to earnings discrimination. This was particularly true for the White population group.

It must be noted that Rospabé's (2001) findings for the African cohort of her sample are quite different to Hinks (2002) findings – that is, Hinks (2002) found a small earnings differential between African men and women, while Rospabé (2001) found a large differential. The difference in findings for the two studies may be as a result of the surveys used in the analyses, that is, the 1995 OHS used by Hinks (2002) may not be a representative sample, particularly of Africans. Hinks (2002) suspects that African domestic workers are considerably undersampled in the 1995 OHS, and Muller (2008) corroborates this suggestion. Alternatively gender discrimination among Africans may have increased in the period between 1995 and 1999, but this seems unlikely in light of equity legislation in the post-apartheid period.

Gruen (2004), like Rospabé (2001), looked at discrimination in both employment and earnings using the Oaxaca (1973) decomposition methodology. Unlike Rospabé (2001) and Hinks (2002) however, Gruen (2004) corrected for selection into the labour market and employment when estimating her earnings equations. This typically reduces the magnitude of the gender pay gap in comparison to uncorrected estimations, since uncorrected estimations pick up the effect of the explanatory variables on earnings, as well as the indirect effects of those same explanatory variables on labour force participation and employment. She found that gender based discrimination for both African and White workers was large and growing between 1995 and 1999. However, while African women primarily suffered from discrimination at the hiring stage, White women were found to be affected by more direct earnings discrimination. Nonetheless, Gruen's (2004) analysis of changes in the gender gap from 1995 to 1999 may also be compromised by the use of 1995 as the base year.

International reviews of the gender pay gap show that it is prevalent in most countries in the world, even though it is predominantly on the decline (Polachek & Xiang, 2006). However, the international literature shows large variations in the gender pay gap between countries: "Some countries like Australia, Belgium, the Czech Republic, Hungary, Italy, Poland and Sweden exhibit a gender pay gap of around 20% over 1970-2000 based on OECD data... [while] other countries such as Austria, Canada, South Korea and Japan maintain gender pay gaps as large as 40-50%" (Polachek & Xiang, 2006: 4). In addition, while some countries show a decreasing gap, in other countries it remains relatively constant. For instance, Polachek & Xiang (2006) note that across time the gender pay gap reduced significantly in countries like France, the United States and the United Kingdom, while for countries such as Belgium, Luxembourg, Spain, Sweden and Switzerland, the gap remained relatively unchanged.

More generally, Weichselbaumer & Winter-Ebmer (2005) in a *Meta-Analysis of the International Gender Wage Gap* found that the reported earnings gap has fallen over time from around 65% in the 1960s to 30% in the 1990s. However, while the total earnings differential more than halved across the time period 1963-1997, the decline is attributed almost entirely to an equalisation of productive characteristics, rather than to a decrease in discrimination: that is females have become better educated and trained, while the Oaxaca-Blinder earnings residual is practically constant over time.

South African studies on the gender pay gap typically have been single year studies (Hinks, 2002; Rospabé, 2001), with the exceptions of Ntuli (2006) who estimated the gap and discrimination for Africans over the 1995 to 2004 period, and Gruen (2004) who used the OHSs. Furthermore, none of the recent studies for South Africa has looked at the gender pay gap over time using just the LFSs. Comparability of the gender earnings gap over time is compromised by the fact that the OHSs and LFSs are not comparable. Specifically, questions relating to the individual's employment status changed in the cross-over from the OHS to the LFS.[4] Thus, the LFSs provide a far more detailed

---

[4] Household labour market information in South Africa in the post-apartheid period was initially collected through the October Household Surveys (OHS), conducted between 1995 and 1999. This survey was replaced by the Labour Force Survey (LFS) in 2000, and is now

explanation of what constitutes work, and therefore capture irregular and informal work activities more comprehensively than the OHSs (Casale, Muller & Posel, 2005). In addition, changes within the OHS series itself make the OHS data between years complicated to compare (Casale, Muller & Posel, 2005). Burger and Yu (2006) for instance in an analysis of wage trends in post-apartheid South Africa acknowledge that "it is widely accepted that that the OHS and LFS datasets cannot easily be compared, [but] these difficulties are often overlooked in the service of constructing a longer time series" (Burger and Yu, 2006: 2).

Given these comparability and compatibility issues, my analysis below uses post-2000 LFS data. This allows me to control for survey design in measuring changes over time, since the survey instruments have been largely unchanged from the introduction of the survey. I have opted to use the 2001 LFS rather than the 2000 LFS in light of the fact that the LFS 2000 seems to be an outlier in the series: Burger and Yu (2006) found that average earnings were dramatically higher in the September 2000 LFS than in the surveys directly preceding and following it (Burger and Yu, 2006).

Earnings studies in South Africa typically account for selection into the labour force and then for selection into employment before estimating earnings (see for instance Bhorat and Leibbrandt, 2001), given that South Africa has a very high unemployment rate and thus the decision to participate in the labour market does not guarantee that one will find employment. In the South African gender earnings gap studies, both Ntuli (2006) and Gruen (2004) correct for sample selection when estimating the earnings gap, but by contrast, neither Rospabé (2001) nor Hinks (2002) use selectivity-corrected earnings equations. A key difficulty in controlling for selection is independently identifying the selection equations, both into labour force participation and employment.

As far as quantile regression techniques are concerned, the issue of how to control for sample selection is unclear since the form of the bias over different percentiles of the earnings distribution is unknown. Some studies, including Ntuli (2006), have followed Buchinsky (1994) and estimate a labour force participation model which is transformed into several series expansions which are then included in the earnings equation as controls for selection bias (see Ntuli 2006). There is however "little consensus regarding the most appropriate selection procedure for selectivity bias in quantile regression models" (Ntuli, 2006: 3). Given the mixture of methods used, I provide both selectivity-corrected[5] and uncorrected estimates for the pooled regression. I provide only uncorrected estimates for the separated (male and female) estimations in light of the fact that it is not obvious how the selection term should be treated in the overall decomposition (see discussion in Section 4.4).

Labour market legislation in the post-apartheid period in South Africa has sought to ensure good working conditions, fair labour practices, the right to collective bargaining and trade union activity, as well as equal opportunities for all in the workplace. These principles are embodied in the Labour

conducted bi-annually, that is in March and September each year, with the September round being more comprehensive than the March one.

[5] For the quantile regressions with corrected estimates, I assume constant bias along the distribution.

Relations Act 66 of 1995, the Basic Conditions of Employment Act 75 of 1997 and the Employment Equity Act 55 of 1998. The Employment Equity Act in particular seeks to promote and ensure equal opportunities in the workplace for all employees with the two main objectives of the Act being to i) promote equal opportunity and fair treatment in employment through the elimination of unfair discrimination and ii) to implement affirmative action measures to redress the disadvantages in employment experienced by designated groups, in order to ensure their equitable representation in all occupational categories and levels in the workforce. The affirmative action measures are applicable to Black people (including Africans, Coloured and Indians), women and the disabled.

Given the introduction of affirmative action policies and the abolition of discrimination, one would expect to see better representation of women in the economy as well as a decreasing gender earnings gap in post-apartheid South Africa. While the feminisation of the labour force has been well established in the literature (Casale & Posel, 2002), the evolution of the gender earnings gap over time has received less attention. Part of the problem, as highlighted above, has been the lack of comparability between datasets. Therefore, in this thesis I use comparable LFS data to analyse whether the gender earnings gap has in fact narrowed in post-apartheid South Africa.

Chapter Three: Gender Earnings Inequality: A Descriptive Overview

3.1. Introduction

Chapter 3 descriptively analyses changes in the gender earnings gap between 2001 and 2005 using mean earnings. Data from the September LFSs of 2001 and 2005 are used. In Section 3.2 I discuss the construction of the earnings variable and motivate for using hourly earnings rather than monthly earnings. In the following section I describe and compare differences in average earnings by gender across a range of covariates, while in Section 3.4 I consider the earnings distribution more closely by describing log earnings percentile differentials.

3.2. The Earnings Variable

3.2.1. Data Sources

The empirical analysis in this study is based on nationally representative household surveys conducted by the central statistical agency in South Africa, Statistics South Africa. Specifically, the September LFSs of 2001 and 2005 are used. In the 2001 LFS about 26,000 households and 105,000 individuals were interviewed, while in the 2005 LFS 28,000 households and 109,000 individuals were interviewed. The sample in the analysis of gender earnings gaps comprises all those employed in the economy between the ages of 15 and 65 years of age. Both the 2001 and 2005 LFS data have been weighted using population weights based on the 2001 Census.

3.2.2. Construction of the Earnings Variable

a. Earnings in the Labour Force Surveys

The earnings question in both the 2001 and 2005 September LFSs reads as follows: *"What is ....'s total salary/pay at his / her main job?".*[6] The respondent answered the question for each member of the household who was 15 years and above and was employed in the seven days prior to the interview. He/she was required to estimate the salary/pay for each person in the household, and was expected to include overtime pay, allowances and bonuses before any tax or deductions in the estimated amount.

Respondents were first asked to provide a point estimate of the earnings. If they refused or did not know, they were then directed to a series of income brackets. If respondents provided a point estimate of earnings, they also had to specify the period in which the amount was earned, for instance R5000 *a month*. For those who chose to respond within categories, there were 14 categories

[6] This is question 4.15a in both questionnaires.

specifying amounts in weekly, monthly or annual equivalents, ranging from category 1 *'None'* for no income, to category 14 *'R30 001 or more'* a month. In addition to providing bracket or point responses, respondents were also allowed to refuse to answer (category 15) or specify that they did not know (category 16).

The earnings data in the LFSs therefore include a combination of both point and bracket responses. Casale and Posel (2005) showed that different methods of combining the point and bracket responses produce consistent summary measures. I chose to use the 'midpoint method' which entails assigning a point value equal to the midpoint of the corresponding earnings bracket for those that responded in brackets. The method is desirable because of its simplicity. Casale and Posel (2005) find that although the method is crude, there is no evidence to suggest that the midpoint method generates biased earnings estimates, or at least estimates that are more biased than other methods.

The earnings of those who responded with a point estimate but who failed to specify an income period where required were set to missing. Earnings were also set to missing for those who refused to answer the income question or did not know their income, that is, for those who responded in category 15 or 16. Those who reported zero earnings were kept in the sample.[7]

b. Hours Worked

Earnings have been analysed in the South African labour market in both monthly and hourly terms. However, closer examination of hours worked by the employed reveals significant variation by gender. Table 1 below shows mean hours worked disaggregated by gender in 2001 and 2005.[8]

Table 1: Mean Hours Worked per Week, by Gender: 2001 & 2005

	2001		2005		% Change	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**
Hours of Work	47.24	42.68	48.36	43.92	2.4%*	2.9%*
	0.17	*0.20*	*0.21*	*0.227*		

Source: *Own calculations, LFS(2001:2) & LFS(2005:2)*

Notes: ** indicates significant change at the 5% level of significance.*

In both years the difference between the male and female hours worked is significant.

The data are weighted. Standard errors are reported in italics below all mean estimates.

---

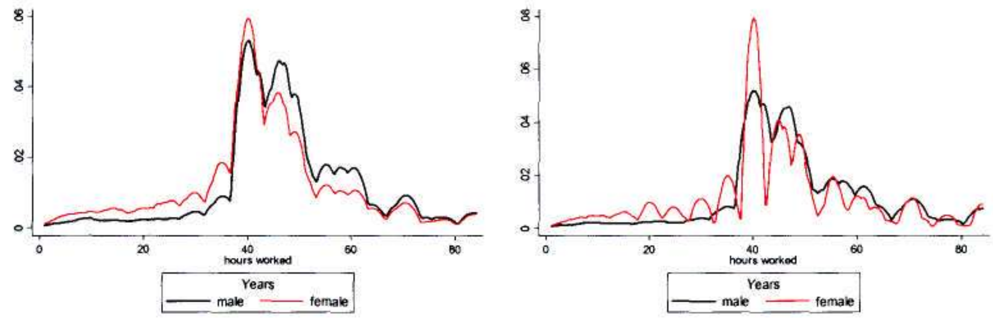
[7] Those reporting zero earnings were treated consistently in the descriptive and regression analyses. Where earnings were logged, those with earnings from zero to R1 were converted to R1 in order to maintain their earnings information.

[8] There are two hours worked question in the LFSs, that is the hours *usually* worked in the main job and in other activities, and the hours worked *in the past 7 days* in the main job and in other activities. We used the hours *usually* worked in the main job. Where respondents reported zero hours usually worked in their main job, we used hours worked in the past 7 days in the main job. Where respondents reported working more than 84 hours a week, that is more than 12 hours a day for seven days of the week, we set their hours worked to missing.

It is evident in the first instance that on average, males worked longer hours than females in 2001. Though average hours worked per week increased significantly for both men and women between 2001 and 2005, men still worked significantly longer work-weeks than women in 2005.

The kernel density plots of hours worked by males and females, as shown in Figure 1, further illustrate gender differences in hours worked. The distribution of working hours for females lies to the left of that for males and spikes earlier than that of males in 2001. In addition, in 2005 there is a clear flattening of the kernel density of hours worked for males.

Figure 1: Kernel Density Plots of Hours Worked, by Gender: 2001 & 2005



2001 ***2005***

Source: *LFS(2001:2) & LFS(2005:2).*
Notes: *The plots are based on weighted hours worked.*

Thus, an analysis of earnings inequality needs to recognise these gender differences in hours worked. In my study, I therefore generate an hourly earnings rate to explore the gender gap in earnings. Table 2 provides a brief overview of the earnings and hours worked data, as well as the sample used in the analysis.

Table 2: Overview of Earnings Data

	2001	**2005**
no. of employed (unweighted obs)	26,022	26,087
missing earnings	2,404	2,307
zero earnings	712	1,008
missing hours of work	557	583
sample	*23,148*	*23,283*
weighted sample	*9,879,899*	*10,798,943*

Source: *Own calculations, LFS(2001:2) & LFS(2005:2)*

The table shows that the number of employed people reporting zero earnings increased quite considerably between 2001 and 2005. All those reporting zero earnings were kept in the sample but those with missing earnings and hours of work data were excluded from the sample. Overall, when excluding those with missing data, I am left with 23,148 unweighted observations in 2001 and 23,283 unweighted observations in 2005. The sample was weighted up to the population using Census 2001 weights.

3.3. The Gender Earnings Gap: Mean Estimates

In this section I firstly consider the gender gap in average real earnings. I then provide a brief analysis of the earnings gap at different points in the earnings distribution.

Table 3 below describes real mean hourly earnings (in 2000 prices), disaggregated by gender.[9] The table firstly shows that while there was an increase in average hourly earnings for all the employed from 2001 to 2005, this increase was *insignificant*. Secondly, it is evident from the table that male and female earnings were significantly different in 2001 – the average earnings of females in 2001 stood at 77% of that of males. The gap, as measured by the ratio of female to male earnings (F/M), narrowed slightly between 2001 and 2005. However, although the average earnings gap decreased, male earnings remained significantly higher than female earnings in 2005.

Table 3: Real Mean Hourly Earnings, by Gender: 2001 & 2005

<table>
<tr><th></th><th colspan="3">2001</th><th colspan="3">2005</th><th colspan="2">% change</th></tr>
<tr><th></th><th>Male</th><th>Female</th><th>F/M</th><th>Male</th><th>Female</th><th>F/M</th><th>M</th><th>F</th></tr>
<tr><td rowspan="2">All employed</td><td colspan="2">13.88</td><td rowspan="2"></td><td colspan="2">14.69</td><td rowspan="2"></td><td colspan="2" rowspan="2">5.84%</td></tr>
<tr><td colspan="2">0.35</td><td colspan="2">0.46</td></tr>
<tr><td rowspan="2">All races</td><td>15.42</td><td>11.81</td><td rowspan="2">0.77</td><td>16.13</td><td>12.75</td><td rowspan="2">0.79</td><td rowspan="2">4.60%</td><td rowspan="2">7.96%</td></tr>
<tr><td>0.43</td><td>0.38</td><td>0.56</td><td>0.45</td></tr>
</table>

Source: *Own calculations, LFS(2001:2) & LFS(2005:2)*

Notes: ** indicates significant change at the 5% level of significance.*

Bolded text indicates that male and female hourly earnings are significantly different to each other at the 5% level of significance.

All earnings data have been deflated to 2000 real earnings in order to aid comparability between the two years.

The data are weighted. Standard errors are reported in italics below all mean estimates.

Having established a slight narrowing of the gender earnings gap, I disaggregate hourly earnings by a range of covariates beginning with race in Table 4, in order to identify whether the gender earnings gap has narrowed among particular groups of the employed. Race is an important marker of the gender earnings differential in South Africa, given that the African employed constitute the majority of all employed, and that repression under apartheid left African workers most vulnerable.

---

[9] All earnings data have been deflated to 2000 real earnings using the CPI deflator (released by Statistics South Africa) in order to aid comparability between the two years.

Table 4: Real Mean Hourly Earnings, by Race & Gender: 2001 & 2005

	2001			2005			% change	
	Male	Female	F/M	Male	Female	F/M	M	F
African	9.21	7.35	**0.80**	10.32	8.83	**0.86**	12.05%*	20.14%*
	0.21	*0.25*		*0.33*	*0.40*			
Coloured	12.88	10.4	**0.81**	15.76	13.29	0.84	22.36%	27.79%*
	0.58	*0.47*		*1.09*	*0.94*			
Asian	21.03	16.5	**0.78**	23.27	15.73	0.68	10.65%	-4.67%
	1.156	*1.11*		*2.22*	*1.83*			
White	39.92	28.91	**0.72**	42.92	29.91	**0.70**	7.52%	3.46%
	1.543	*1.40*		*1.96*	*1.47*			
Total	15.42	11.81	**0.77**	16.13	12.75	**0.79**	4.60%	7.96%

Source: *Own Calculations, LFS(2001:2) & LFS(2005:2)*

Notes: ** indicates significant change at the 5% level of significance*

Bolded text indicates that male and female hourly earnings are significantly different to each other at the 5% level of significance.

All earnings data have been deflated to 2000 real earnings in order to aid comparability between the two years.

The data are weighted. Standard errors are reported in italics below all mean estimates.

The race disaggregation shows that in 2001, across all race groups, females on average earned significantly less than their male counterparts, but this is particularly true for the White cohort, where the gap between male and female earnings was the largest. However, the table provides evidence of a narrowing of the gender pay gap, in particular for the African and Coloured employed. African women's earnings on average increased at a faster rate than African male earnings between 2001 and 2005, so that by 2005 the female to male ratio of average earnings for Africans had increased from 80% to 86%. However, although the gender earnings differential narrowed for Africans in the period, African women still earned significantly less than their male counterparts in 2005.

Similarly, there was a narrowing of the gender earnings differential for the Coloured cohort, such that by 2005, there was no significant difference between male and female earnings for the Coloured population. Ultimately then, the data suggests that the overall national narrowing of the average gender earnings gap was driven by the decrease in the gender earnings gap among Africans and Coloureds. It is perhaps surprising that among the employed who typically have access to greater labour market skills (Whites and Asians) the ratio of female to male average earnings increased over the period.

In Table 5 below, I describe differences in average earnings by gender across skills levels. The table highlights that a gender pay gap existed in 2001 through most of the skills distribution, but particularly for craft and trade workers, with the notable exceptions being managers, clerical workers and domestic workers. The gender gap in average earnings decreased among those employed in highly skilled occupations, particularly among managers. The narrowing of the gender earnings differential in average earnings for the highly skilled may reflect greater opportunity for better-educated women in the post-apartheid period.

Table 5: Real Mean Hourly Earnings, by Skills Levels, Occupation Group & Gender: 2001 & 2005

	2001			2005			% change	
	Male	Female	F/M	Male	Female	F/M	M	F
High-Skilled								
Manager	42.35	37.02	0.87	46.26	42.54	0.92	9.23%	14.91%
	2.19	*4.13*		*2.94*	*4.83*			
Professional	35.34	27.32	**0.77**	36.98	28.82	**0.78**	4.64%	5.49%
	1.44	*0.96*		*1.69*	*1.04*			
Semi-Skilled								
Clerical	17.93	17.07	0.95	20.27	17.93	0.88	13.05%	5.04%
	0.80	*0.73*		*1.18*	*0.83*			
Service	11.50	8.40	**0.73**	10.72	7.11	**0.66**	-6.78%	-15.36%
	0.59	*0.72*		*0.53*	*0.48*			
Skilled Agriculture & Fishing	11.51	3.41	**0.30**	10.65	2.26	**0.21**	-7.47%	-33.72%
	1.18	*1.07*		*2.97*	*0.50*			
Craft & Trade	12.03	6.23	**0.52**	10.92	5.79	**0.53**	-9.23%	-7.06%
	0.47	*0.42*		*0.48*	*0.37*			
Operator & Assembler	9.26	7.56	**0.82**	10.46	8.9	0.85	12.96%*	17.72%
	0.24	*0.48*		*0.33*	*0.94*			
Unskilled								
Elementary	5.58	4.44	**0.80**	6.72	5.08	**0.76**	20.43%*	14.41%
	0.20	*0.16*		*0.26*	*0.30*			
Domestic Worker	3.12	2.94	0.94	3.82	3.57	0.93	22.44%	21.43%*
	0.23	*0.08*		*0.25*	*0.08*			
Total	15.42	11.81	**0.77**	16.13	12.75	**0.79**	4.60%	7.96%

Source: *Own calculations, LFS(2001:2) & LFS(2005:2)*

Notes: ** indicates significant change at the 5% level of significance.*

Bolded text indicates that male and female hourly earnings are significantly different to each other at the 5% level of significance.

All earnings data have been deflated to 2000 real earnings in order to aid comparability between the two years.

The data are weighted. Standard errors are reported in italics below all mean estimates

Table 5 shows further that the gender earnings gap did not change consistently across semi-skilled occupations: the earnings differential narrowed slightly for craft and trade workers, and operators and assemblers, but it rose among clerical workers (in which White and Indian women are over-represented) and service workers, as well as among skilled agriculture and fishery workers. Among the unskilled, the gender gap in average earnings increased among workers in elementary occupations.

The differential in earnings between male and females in skilled agriculture and fishery occupations deserves some comment since it is so high. Closer examination of the data indicates that a large proportion of skilled agriculture and fishery workers are in the agricultural sector (85% in 2001).

Furthermore, a much larger proportion of women than men in this category earn zero income (72% versus 45% in 2001) and are thus likely to be subsistence farmers. The high differential between male and female earnings in this category is thus explained by the large proportion of zero income earners among women in this occupation group.

While the occupation disaggregation provides an indication of the gender earnings gap across skill levels, the table below extends this by showing the earnings gap at different education levels. The earnings literature highlights the importance of education as a key determinant of earnings. I expect to find that average earnings are monotonically linked to education levels, that is at higher levels of education mean wages will be higher. This is confirmed in Table 6 which indicates that mean hourly earnings increase at higher levels of education for both women and men. However, it is clear that at any given education level for both years, men generally earn significantly more than women, and this is particularly true as educational attainment declines.

Table 6: Real Mean Hourly Earnings, by Level of Education & Gender: 2001 & 2005

	2001			2005			% change	
	Male	**Female**	**F/M**	**Male**	**Female**	**F/M**	**M**	**F**
None	4.7	2.62	**0.56**	5.79	3.27	**0.56**	23.19%	24.81%*
	0.20	*0.12*		*0.42*	*0.18*			
Incomplete GET	6.68	4.07	**0.61**	6.84	4.3	**0.63**	2.40%	5.65%
	0.16	*0.14*		*0.21*	*0.13*			
Complete GET	11.32	7.12	**0.63**	10.16	6.68	**0.66**	-10.25%	-6.18%
	0.47	*0.26*		*0.46*	*0.34*			
Matric (grade 12)	20.11	15.67	**0.78**	19.15	14.25	**0.74**	-4.77%	-9.06%
	0.71	*0.82*		*0.78*	*0.63*			
Diploma less Matric	27.14	22.63	0.83	26.21	22.3	0.85	-3.43%	-1.46%
	2.33	*3.61*		*3.55*	*3.15*			
Diploma with Matric	31.66	24.28	**0.77**	32.14	25.58	**0.80**	1.52%	5.35%
	1.12	*0.68*		*1.60*	*1.17*			
Degree	53.63	39.43	**0.74**	61.36	47.95	**0.78**	14.41%	21.61%
	2.74	*2.06*		*3.38*	*3.40*			
Total	15.42	11.81	**0.77**	16.13	12.75	**0.79**	4.60%	7.96%

Source: *Own calculations, LFS(2001:2) & LFS(2005:2).*

Notes: ** indicates significant change at the 5% level of significance.*

Bolded text indicates that male and female hourly earnings are significantly different to each other at the 5% level of significance.

All earnings data have been deflated to 2000 real earnings in order to aid comparability between the two years.

The data are weighted. Standard errors are reported in italics below all mean estimates.

Table 6 however, also shows a narrowing of the gender earnings gap across all education categories, except for those with a matric (grade 12), where the gap increased, and no schooling where the average earnings gap was unchanged. Most striking however is the improvement in women's average earnings relative to those of men at the top end of the education distribution, that is, for those with diplomas and degrees. Specifically, women with diplomas (with matric) and degrees witnessed a narrowing of the differential. The ratio of average female to male earnings increased from 74% in 2001 to 78% in 2005 for those with degrees and from 77% to 80% for those with diplomas (with matric). The aggregate fall in the gender earnings differential therefore appears to be driven by the narrowing of the gender earnings gap particularly among those with diplomas (with matric).[10]

[10] It must be noted that the proportion of women with diplomas (with matric) in all those employed with diplomas (with matric) was roughly the same in the two years, that is, 51% in 2001 and 50% in 2005.

While occupation groups are easily dissected into skills levels which aid in understanding of the gender earnings gap, the disaggregation by industry, which represents all skills levels within that industry, is more difficult to interpret. Unsurprisingly therefore, the disaggregation in Table 7 below shows less consistent evidence of a gender gap in average earnings. There are however some interesting results.

Table 7: Real Mean Hourly Earnings, by Industry & Gender: 2001 & 2005

	2001			2005			% change	
	Male	**Female**	**F/M**	**Male**	**Female**	**F/M**	**M**	**F**
Agriculture	5.38	3.44	**0.64**	6.54	3.48	**0.53**	21.56%	1.16%
	0.39	*0.47*		*0.75*	*0.32*			
Mining	13.55	13.04	0.96	15.88	19.97	1.26	17.20%	53.14%
	0.88	*1.53*		*2.07*	*3.55*			
Manufacturing	18.11	12.18	**0.67**	17.89	11.79	**0.66**	-1.21%	-3.20%
	0.98	*0.84*		*0.99*	*1.07*			
Electricity	21.35	19.07	0.89	21.02	31.8	1.51	-1.55%	66.75%
	2.52	*2.53*		*1.90*	*12.69*			
Construction	10.15	6.47	**0.64**	10.76	13.42	1.25	6.01%	107.42%
	0.57	*0.97*		*0.91*	*4.36*			
Craft and Trade	12.26	7.33	**0.60**	12.62	6.88	**0.54**	2.94%	-6.14%
	0.62	*0.38*		*0.79*	*0.33*			
Transport	18.37	21.05	1.15	15.89	21.73	1.37	-13.50%	3.23%
	1.45	*1.91*		*1.42*	*3.36*			
Financial Services	29.26	25.29	0.86	27.47	24.68	0.90	-6.12%	-2.41%
	2.15	*2.04*		*2.39*	*1.76*			
Community Services	22.56	20.59	0.91	24.45	22.74	0.93	8.38%	10.44%
	0.72	*0.67*		*1.04*	*0.94*			
Private Households	3.12	2.94	0.94	3.82	3.57	0.93	22.44%	21.43%*
	0.23	*0.08*		*0.25*	*0.08*			
Total	15.42	11.81	**0.77**	16.13	12.75	**0.79**	4.60%	7.96%

Source: *Own calculations, LFS(2001:2) & LFS(2005:2).*

Notes: ** indicates significant change at the 5% level of significance.*

Bolded text indicates that male and female hourly earnings are significantly different to each other at the 5% level of significance.

All earnings data have been deflated to 2000 real earnings in order to aid comparability between the two years.

The data are weighted. Standard errors are reported in italics below all mean estimates.

Women are over-represented in the craft and trade sector and this sector displayed the highest gender earnings differential in 2001. Furthermore, the average gender pay differential for women in the craft and trade sector worsened over the period so that by 2005 women in this sector earned just

54% of average male earnings. The table further illustrates that women in both the agriculture and manufacturing sectors also earned significantly less than their male counterparts and furthermore witnessed a worsening of the gender pay differential between 2001 and 2005. The only industry to display a narrowing of the gender earnings differential was construction, with the female to male ratio of earnings in construction increasing from 64% to 125%. However, it must be noted that although the construction sector showed a large narrowing of the gender pay differential, less than 10% of those employed in construction in both years were female. Overall then there is no conclusive evidence of the narrowing of the gender earnings differential by industry, with the average differential actually increasing in some of the primary industries.

The descriptive analysis of real mean hourly earnings presented in this section thus indicates an overall narrowing of the gender earnings differential which has been driven by changes among Africans and Coloureds, and among skilled women (professionals) with diplomas and degrees. I explore the interaction between race and human capital further in Tables 8 and 9 where I disaggregate the earnings of the African/Coloured employed first by skills level and then by level of education. This is done in order to determine if the aggregate narrowing of the earnings differential was driven by highly skilled and educated African/Coloured women.

Table 8 describes male and female earnings by skills level for the African/Coloured employed.[11] The table indicates that there was a significant difference in male and female earnings across the three skills classes in 2001, although the gender earnings gap was lowest for the highly-skilled and highest for the low-skilled.

Table 8: Real Mean Hourly Earnings, Africans/Coloureds, by Skills Level: 2001 & 2005

	2001			2005			% change	
	Male	**Female**	**F/M**	**Male**	**Female**	**F/M**	**Male**	**Female**
High Skilled	25.75	23.10	**0.90**	30.19	28.65	0.95	17.24%*	24.03%*
	0.94	*0.71*		*1.675*	*1.664*			
Intermediate Skilled	8.84	7.41	**0.84**	9.41	7.88	**0.84**	6.45%	6.34%
	0.17	*0.25*		*0.23*	*0.34*			
Low-Skilled	4.91	3.63	**0.74**	5.64	4.18	**0.74**	14.87%*	15.15%*
	0.14	*0.08*		*0.17*	*0.10*			

Source: *Own calculations, LFS(2001:2) & LFS(2005:2)*

Notes: ** Indicates a significant change at the 10% level of significance.*

Bolded text indicates that male and female hourly earnings are significantly different to each other at the 5% level of significance.

All earnings data have been deflated to 2000 real earnings in order to aid comparability between the two years.

The data are weighted. Standard errors are reported in italics below all mean estimates.

[11] Dissecting the gender earnings gap for Africans alone did not provide significant results.

By 2005, there was a clear narrowing of the gender earnings differential for the highly skilled driven by a significantly larger increase in African/Coloured female earnings at the top end compared to male earnings. Thus, by 2005, there was no significant difference between male and female earnings for the highly skilled African/Coloured cohort. It would appear then, that among highly skilled men and women, the narrowing of the gender earnings differential was driven particularly by changes among Africans/Coloureds.

Table 9 below corroborates this finding: Over the period, the average earnings gap narrowed across all levels of education except for those with no education, but particularly among diploma-holders with matric. In 2005, the earnings of African/Coloured women with diplomas and degrees were not significantly different from their male counterparts.

Table 9: Real Mean Hourly Earnings, Africans & Coloureds, by Education Level: 2001 & 2005

	2001			2005			% change	
	Male	Female	F/M	Male	Female		Male	Female
None	**4.66**	**2.62**	**0.56**	**5.82**	**3.22**	**0.55**	24.89%*	22.90%*
	0.20	*0.12*		*0.42*	*0.17*			
Incomplete GET	**6.45**	**3.93**	**0.61**	**6.51**	**4.2**	**0.64**	0.93%	6.87%
	0.16	*0.12*		*0.16*	*0.13*			
Complete GET	**8.69**	**5.7**	**0.66**	**8**	**5.31**	**0.66**	-7.94%	-6.84%
	0.26	*0.19*		*0.28*	*0.17*			
Matric	**12.92**	**9.8**	**0.76**	**12.96**	**10.14**	**0.78**	0.31%	3.47%
	0.49	*0.37*		*0.48*	*0.50*			
Diploma less Matric	18.41	18.88	1.03	19.52	19.25	0.99	6.03%	1.96%
	1.72	*1.87*		*2.37*	*3.04*			
Diploma with Matric	**23.76**	**21.03**	**0.89**	25.2	23.99	0.95	6.06%	14.08%*
	0.84	*0.54*		*1.029*	*1.04*			
Degree	37.45	32.81	0.88	54.95	49.67	0.90	46.73%*	51.39%*
	2.44	*1.98*		*4.32*	*5.08*			

Source: *Own calculations, LFS(2001:2) & LFS(2005:2).*

Notes: ** Indicates a significant change at the 10% level of significance.*

Bolded text indicates that male and female hourly earnings are significantly different to each other at the 5% level of significance.

All earnings data have been deflated to 2000 real earnings in order to aid comparability between the two years.

The data are weighted. Standard errors are reported in italics below all mean estimates.

In summary, the evidence suggests that the national narrowing of the gender earnings gap was driven by African/Coloured women in highly skilled positions with diplomas.[12] In addition, I find both a continued existence of gender earnings gaps, particularly for those with low levels of education, and a worsening of the gender earnings differential for some semi- and unskilled workers.

3.4. Gender Earnings Inequality: Distributional Estimates

The above analysis described differences in male and female *average* earnings. In this section, I now investigate how gender differences in earnings changed along the earnings distribution. Table 10 calculates the log of earnings differentials by gender for all the employed.[13]

Table 10: Log Hourly Earnings Differentials, by Gender: 2001 & 2005

	2001	2005	% change
90th percentile (90m-90f)	0.21	0.17	-19.05%
50th percentile	0.41	0.39	-4.88%
10th percentile	0.47	0.69	46.81%

Source: *Own calculations, LFS(2001:2) & LFS(2005:2).*

Notes: *All earnings data have been deflated to 2000 real earnings in order to aid comparability between the two years. The data are weighted.*

The table shows firstly that invariant to year, males earned more than females all along the earnings distribution and secondly, that the gender earnings gap was higher at lower levels of earnings. That is, male and female earnings were closer together at the top of the distribution represented by the 90th percentile, and further apart at the bottom of the distribution represented by the 10th percentile. This corroborates the descriptive statistics presented in the tables above which indicate that the gender pay differential was higher at lower levels of education.

The table further indicates a strong narrowing of the earnings gap from 2001 to 2005 for those in the 90th percentile where the gap decreased by 19.05%. In addition to the narrowing of the earnings gap at the top end of the distribution, there was a slight narrowing of the earnings gap in the middle of the distribution. By contrast the earnings gap appears to have increased quite significantly (47%) for those in the 10th percentile. Thus, evidence suggests that the overall national narrowing of the earnings gap between 2001 and 2005 was driven by a fall in the earnings gap over the top half of the earnings distribution, and particularly at the upper tail of the distribution represented by the highly skilled, while the gender earnings gap increased at the very bottom of the distribution. The increase

[12] The proportion of highly skilled African/Coloured women in total highly skilled employed was 48.6% in 2001 and 47% in 2005, while the proportion of these women with diplomas (including matric) in total employed with diplomas (including matric) was 56% and 54% in 2001 and 2005 respectively.

[13] The differentials are calculated by firstly finding the mean hourly earnings at each of the percentiles being analysed for each of the genders. The mean earnings at each percentile are then logged for each of the genders, and the differential calculated. This follows the methodology used by Bhorat (2000) in estimating "Wage Premia and Wage Differentials in the South African Labour Market".

in the earnings gap at the bottom of the distribution is consistent with findings presented in the previous section, which showed a worsening of the earnings gap for elementary workers.

Table 11 further disaggregates the gender pay gap along the distribution for the African and Coloured employed. The table shows that in 2001, the gender earnings gap for both the African and Coloured cohorts was highest in the middle of the distribution. In contrast, while there was almost no difference in earnings for Africans at the top of the distribution (90^{th} percentile), there was a large difference in earnings for Coloured men and women at the top end of the distribution. By 2005, the earnings gap for both Africans and Coloured at the top of the distribution had disappeared, while the gap at the 50^{th} percentile decreased for both race groups, but at a faster rate for the Coloured cohort compared to the African cohort (48% versus 28%). Thus, there was clearly a decrease in the gender earnings gap for both Africans and Coloureds in the top half of the distribution, but this is particularly true of the Coloured cohort.

Table 11: Log Hourly Earnings Differentials, by Race and Gender: 2001 & 2005

	2001		2005		% change	
	African	Coloured	African	Coloured	African	Coloured
90th percentile (90m-90f)	0.04	0.22	0.00	0.00	-100.00%	-100.00%
50th percentile	0.61	0.25	0.44	0.13	-27.87%	-48.00%
10th percentile	0.47	0.15	0.79	0.13	68.09%	-13.33%

Source: *Own calculations, LFS(2001:2) & LFS(2005:2).*

Notes: *All earnings data have been deflated to 2000 real earnings in order to aid comparability between the two years. The data are weighted.*

At the bottom of the distribution, as expected, the gender earnings gap increased for the African cohort, while the Coloured cohort saw a decrease in the gender earnings differential. Consequently, by 2005 the gender earnings gap for Africans in the middle of the distribution was lower than at the bottom of the earnings distribution. The increase in the gender earnings gap for Africans at the bottom of the earnings distribution can be seen in the context of the widening earnings gap for the lower skilled and less educated, where Africans are disproportionately represented.

3.5. Conclusion

Using hourly earnings, the analysis in this chapter indicates a slight narrowing of the gender earnings differential between 2001 and 2005. Male earnings however remained significantly higher than female earnings in 2005. The race disaggregation indicates a narrowing of the gender earnings differential for the African and Coloured cohorts while the skills disaggregation shows a declining earnings differential for professionals as well as for some semi-skilled workers. The gender earnings gap did however increase for unskilled workers and for some semi-skilled workers. In addition, it is clear that a significant gender earnings gap still exists all along the skills and education distributions, and that the gap is particularly high for low-skilled workers and workers with no education.

In trying to locate the narrowing of the differential I disaggregated earnings by skills level for Africans/Coloureds specifically, and found a significant narrowing of the earnings differential for the highly skilled as well as for those with diplomas. Thus, it would appear that the aggregate narrowing of the gender earnings differential derives from changes among better educated and highly skilled Africans/Coloureds. The distribution data corroborate these findings. I find a large fall in the gender earnings differential at the 90th percentile for the aggregate sample as well as for Africans and Coloureds. The gap also narrowed, albeit more marginally, in the middle of the distribution, but it increased for those in the 10th percentile, who are likely to be elementary workers and less educated workers.

Chapter Four: Gender Earnings Inequality: A Multivariate Analysis

4.1. Introduction

The previous chapter descriptively analysed the gender earnings gap firstly by considering the change in average earnings for the 2001 to 2005 period over a range of covariates, and secondly by investigating how the gender earnings gap changed along the length of the earnings distribution. Descriptive statistics suggest a small closing of the gender earnings gap that is associated with highly-skilled and educated African/Coloured women. Significant gender differences in earnings remain however, particularly at the lower tail of the distribution where the average gender earnings gap increased over the period. In this section, I investigate the gender earnings gap more comprehensively by controlling for productivity-related characteristics of men and women.

In Section 4.3 I first use Ordinary Least Squares (OLS) to estimate a pooled earnings equation in a multivariate context. OLS regressions estimate the mean effect of the various explanatory variables (characteristics of men and women) on the dependent variable (log hourly earnings). To investigate how the gender earnings gap changes along the distribution of earnings, I then estimate earnings using quantile regression analysis. This approach allows me to test whether the gender earnings gap is higher at the top or bottom end of the earnings distribution and further, whether the gender earnings differential has changed from 2001 to 2005.

The pooled regression assumes that the returns to observable characteristics are the same for men and women. In Section 4.4 I relax this assumption by estimating separate male and female earnings equations. The separation of the models further allows me to analyse the extent of 'discrimination' through use of the Oaxaca (1973) decomposition technique.

4.2. Specifying the model

4.2.1. *OLS Mean Estimation*

Earnings are estimated using a simple Mincerian model in which earnings are a function of acquired human capital which is a function of both education and experience. The earnings function therefore takes the generic form:

$$\boxed{Y_i = \alpha + \beta X_i + u_i}$$

where Y refers to the log of hourly earnings of worker *i* and *X* is a vector of relevant independent variables including education and experience. β refers to the estimated parameters of the earnings function. The error term and constant are captured by u and α respectively.

Biases in the estimation of earnings can arise from several sources. Potential sources of bias include measurement errors as well as omitted variable bias. Specific to the South African situation is its unique problem of sample selection bias. South Africa has a very high unemployment rate and thus the decision to participate in the labour market does not guarantee that one will find employment. Estimating an earnings equation simply on the sub-sample of earners could thus bias estimates as the sub-sample may not be a representative sample. Therefore, I model earnings in three sequential phases in order to control for sample selection bias: firstly, the decision to participate in the labour market[14]; secondly, the probability of finding employment; and thirdly, the earnings of the employed (Bhorat and Leibbrandt, 2001: 107-129).

I begin with a full sample of the working-age population and model the probability of participation using variables that would impact on an individual's decision to enter into the labour market.[15] This is necessary since the sample of labour market participants is highly unlikely to be a random sample of the working-age population, that is, people choose whether or not to participate in the labour market. To correct for sample selection bias it is necessary to include in the labour force participation model variables that both influence the probability that an individual participates in the labour market and that do not also determine market earnings. Therefore, household variables which impact on the probability of participation are included. These are the number of children under 7 years old in the household, the number of 8 to 15 year olds in the household, the number of over 65 year olds in the household, as well as the number of working-age men and women in the household.

Once the probability of participation is modelled, I then model the probability of finding employment, conditional on labour force participation. Here I include race, age, education and province, as well as two variables which independently identify the employment probability equation: whether the person is the head of the household and the number of employed people in the household. The final stage involves modelling the earnings of those who found employment. (See Bhorat & Leibbrandt, 2001: 112,113; Oosthuizen, 2006: 53). Practically therefore, I use a 3-step approach, by including the inverse of the Mills ratio (lambda) from the participation probit into the employment probit, and the Mills ratio from this equation is then included as a regressor in the earnings equation. The earnings function is therefore modelled on the characteristics of earners conditional on the fact that these earners are a subsample of all the employed which is in turn a subsample of potential participants. I compare the selectivity-corrected earnings estimation results to uncorrected results for both the pooled and quantile regressions in Section 4.3. When separating male and female equations in Section 4.4 I do not take sample selection into account.

---

[14] Labour force participants include both searching and non-searching unemployed.

[15] Following Bhorat & Leibbrandt (2001) I remove all people who are in education from the sample.

4.2.2. Quantile Regression Estimation

While the OLS approach allows one to understand the impact of different covariates, and specifically gender, on the mean earnings distribution, it is also interesting to consider what happens to the earnings gap at different points on the conditional earnings distribution. This can be achieved through quantile regression techniques which were first proposed by Koenker and Basset (1978). This method refers to the generalised case of the least absolute deviations (LAD) estimator, where the quantile regression estimation procedure minimises the absolute sum of errors from a particular quantile of the log earnings distribution, in contrast to the least squares estimates which minimises the sum of squared errors. The estimation for the regression quantile thus minimises the equation below, of which one particular case is the median regression obtained by setting θ=0.5.

$$\boxed{\underset{\beta \in \Re^k}{Min} \left[\sum_{i \in \{i: y_i \geq X_{i\beta}\}} \theta |Y_i - X_i\beta| + \sum_{i \in \{i: y_i < X_{i\beta}\}} \left(1 - \theta |Y_i - X_i\beta|\right) \right]}$$

4.3. Understanding the Gender Pay Gap: Pooled Regression

To estimate the gender earnings gap in a multivariate context, I first pool the sample of men and women. I therefore include a dummy variable equal to one if the employed individual is female and zero if male. This is the crudest approach used in measuring gender earnings gaps, since the underlying assumption is that female and male earnings differ by a fixed amount (the shift parameter), and that human capital characteristics and other characteristics have the same impact on the earnings of men and women (Beblo et al, 2003).

As discussed above, the earnings function is estimated in a three-part modelling procedure that takes into account selection into labour force participation as well as employment. Given that I am primarily interested in the male-female gender earnings gap, I only show the results of the earnings estimation here. Details of the labour force participation probit and employment probit for both years can be found in Appendix 1 and 2.

The sample for the earnings estimations comprises all the working-age employed who provided earnings information. The dependent variable in both the pooled set of estimates and the separate equations is the log of hourly earnings.[16] In the human capital model developed by Mincer, earnings are a function of acquired human capital which in turn is expressed as a function mainly of education and experience. Thus, both education and a proxy for experience have been included as regressors

[16] *Those with earnings from zero to R1 an hour were converted to R1 an hour, in order to maintain their earnings information.*

in the model, and the model has further been extended to include other factors which may influence earnings such as location, occupation groups, race and industry.[17]

The effect of education on the earnings is estimated by a 5-level spline, which distinguishes between the returns to eight years at the primary/lower secondary level (grade one to eight), three years at higher secondary (grade nine to eleven), one year for completion of school (grade twelve), and then diploma and degree. Linear splines are commonly used when the effect of a continuous variable on the outcome is thought not to be linear. Where non-linearities are suspected, sometimes the quadratic form is included (such as for experience below), however, this does not give a clear indication of how the effect of the variable on the outcome changes. In other words, using "years of schooling" constrains the returns to education to be constant for each year of schooling.

In contrast, the spline allows the returns to schooling to vary across the distribution of educational attainment. Another possible solution is to create dummy variables for single years of education or to create education categories to determine the effect of education on income. However, categorisations have the drawback that the observed pattern may seem erratic, especially if some education categories have very few observations. In addition, the effect of education on income may be gradual, not stepwise, as it increases in value. Spline transformations therefore provide a way to estimate a relationship that changes gradually and continuously as the explanatory variable increases in value. The effect of the explanatory variable is assumed to be piecewise linear and each coefficient represents the slope on a particular segment (Stata Technical Bulletin 15, September 1993; Stata Technical Bulletin 18, March 1994).

The South African earnings literature uses a mixture of splines and 'years of education' in estimating earnings equations. Where splines are used, studies create the splines at different intervals (see for instance Bhorat & Leibbrandt, 2001; Chamberlain & Van der Berg, 2002; Ntuli, 2006). The descriptive chapter above showed significant differences in male and female earnings at each of the education categories used, except for those with a diploma with and without a matric. I have used the educational categories identified in Chapter Three to create my splines, grouping all those with diplomas into a single spline.[18]

Different occupation groups are typically rewarded differently. To control for occupation, I included a series of nine dummy variables representing the different occupation groups. The gender dummy in this case therefore measures the differences in earnings between men and women working in the same (broad) occupations. Although occupation groups are commonly controlled for in the gender

---

[17] I initially included both domestic workers (occupation group category) and private households (industry category). However, all domestic workers worked in private households in both 2001 and 2005, and private households primarily employed domestic workers. Therefore there appeared to be a problem of almost perfect collinearity, making it impossible to capture both an occupation and industry effect. To deal with the problem I classified all those working in private households as domestic workers, and then excluded the private household dummy from the estimation, leaving only the domestic worker dummy to capture the occupation effect.

[18] *Stata's* test command provides a very convenient way of assessing whether adjacent slopes on the splines are significantly different from each other, which I found to be the case with my specification.

earnings gap literature, a critique of including these as regressors in the earnings equation is that the occupation group in which men and women find themselves may be correlated to unobserved heterogeneity components of earnings, that is the occupation variables may not be exogenous. It is in fact observed that men and women move into different fields, in other words there is a degree of self-selection into occupations groups (Kunze, 2006: 2). Additionally, men and women may gain access to different jobs associated with different earnings, and thus estimating earnings discrimination without taking discrimination at the occupation level into account may underestimate the extent of discrimination. Rospabé (2001) for instance finds significant occupational discrimination and argues that gender disparities in occupational distribution are likely to strongly affect gender earnings differences.

Also included in the set of explanatory variables is a dummy for being in the private or public sector as well as a dummy variable that captures whether the person is married or not. A person is defined to be in the public sector if they answered the question *"Is the business or enterprise/branch where ..... works"* as 'national', 'provincial' or 'local government'. In all other cases the person was deemed to be working in the private sector. While it is not traditional practice in the South African literature to include 'marriage' in the earnings equation, I have followed Rospabé (2001) and included it here as a proxy for unobserved factors such as stability, discipline and motivation. Weichselbaumer & Winter-Ebmer (2003) for instance indicate that the marital status of an individual can be interpreted as a productivity indicator, and may have a different impact on each of the sexes. While household responsibilities may make females less productive at work, males may benefit from their wives' reproductive work and thus earn a marriage premium through being more productive. Casale and Posel (2007) conducted a study into the marriage premium for men using the LFSs in South Africa, and found a significant earnings premium, though the premium reduced significantly when they controlled for fixed effects using panel data.

Also included in the explanatory variables is a proxy for experience since more experience is generally associated with higher earnings, although a turning point may be reached. The LFSs do not ask about actual years of experience, and thus experience is commonly proxied by taking a person's age, minus their years of schooling plus six (to account for early childhood). The proxy has been criticised on the grounds that people do not work continuously after having completed school, and this is particularly true in South Africa which has a high unemployment rate (Woolard, 2002). The international gender earnings literature further highlights that women often have more interrupted work histories than their male counterparts due to family responsibilities (Kunze, 2006: 1). In the absence of data on actual experience I use a quadratic in age to proxy for this covariate. However it must be noted that if work experience itself is gendered, then the implication is that a measured gender earnings gap with such a proxy may be overestimated (or biased upwards).

Also included in the explanatory variables are dummy variables indicating whether the person is in the formal or informal sector, as well as whether the person is self-employed or wage-employed.[19] Finally, lambda captures the selection effect.

The estimates for the regressions for the two years are to be found in Table 12 below. The referent variables for the categorical variables are: Africans for race, Gauteng for province, elementary workers for occupation groups, and manufacturing industry for industries. The estimates allow us to consider firstly whether men earn more than women in any single year, and secondly whether the earnings gap has increased over the period. In the first instance it is clear from the regressions that when accounting for sample selection and controlling for other observable characteristics, women on average earned less than men in both 2001 and 2005. In addition, the estimates indicate that the gender earnings gap increased over time, though marginally: women earned 18% less than their male counterparts in 2001, but by 2005 this had increased to 20%.[20]

The lambdas are negative and significant in both the employment probit and earnings estimation for 2001, and for the earnings estimation for 2005. The lambdas represent the inverse mills ratios and are a measure of selectivity bias in the sample. The significant results suggest that sampling bias did exist, that is that a) labour market participants are not a random sample of the economically active population, and that b) the sample of employed (earners) is not a random selection of people drawn from the pool of participants. The significance of the lambda thus vindicates the use of the selection procedure.

Lambda functions as a proxy for omitted variables such as innate ability. The common sense expectation is that the co-efficient on the lambda will be positive since unmeasured factors such as ambition and intelligence that raise the earnings received in employment, also raise the probability that the person is employed. However, South African studies have typically found negative sample selection terms (see for instance Bhorat & Leibbrandt, 2001; Chamberlain & Van der Berg, 2002).

The remainder of the earnings equation looks as expected, with Asians and Whites earning substantially more than Coloured and Africans.[21] The education splines indicate that the returns to education increased as educational attainment increased. The provincial dummies capture richer versus poorer provinces.

---

[19] Individuals are classified as working in the formal sector according to how they answered the question "Is the organisation/business/enterprise where ..... works" in the formal sector, informal sector or don't know. This is question 4.22 in the LFS 2005:2 and question 4.18 in the LFS 2001:2.

Individuals are classified as self-employed if they answered to working on their own (option 3 or 4) to the question "In ......'s main work was he/she..." working for someone for pay, working in a private household, working on his/her own, or working without pay. This is question 4.3 in both questionnaires.

[20] To calculate the percentage change in earnings from the coefficient on a dummy variable in a semilogarithmic model, I used the following conversion: 100.{exp(y) – 1} (see Halvorsen and Palmquist 1980).

[21] Though the remainder of the earnings functions do not directly suggest anything about the gender pay gap, I dwell on the coefficients briefly, to illustrate that the earnings function provides the results expected.

Table 12: Earnings Equation Corrected for Sample Selection, Pooled Regression: 2001 & 2005

	2001	2005
female	-0.2037***	-0.2235***
coloured	0.3029***	0.2445***
asian	0.3803***	0.4015***
white	0.6495***	0.5200***
none_grade8	0.0420***	0.0282***
grade9_grade11	0.0748***	0.0712***
grade12	0.2152***	0.2408***
diploma	0.3023***	0.2460***
degree	0.1461***	0.1970***
western cape	-0.0445**	0.0214
eastern cape	-0.3905***	-0.2542***
northern cape	-0.3058***	-0.2870***
free state	-0.3939***	-0.3335***
kwazulu natal	-0.1884***	-0.3404***
north west	-0.2385***	-0.2288***
mpumalanga	-0.2236***	-0.2854***
limpo	-0.3994***	-0.3993***
married	0.0940***	0.0637***
experience (proxy)	0.0329***	0.0276***
experience-squared (proxy)	-0.0004***	-0.0003***
public sector	0.3903***	0.4743***
formal sector	0.5174***	0.5234***
wage-employed	-0.1195***	-0.0867***
managers	0.6711***	0.7898***
professionals	0.4945***	0.5922***
clerks	0.3368***	0.4017***
service	0.0296	-0.0098
skilled agric&fishing	0.2587***	-0.1269***
craft&trade	0.2301***	0.1585***
operators&assemblers	0.1186***	0.1779***
domestic	-0.1449***	0.0511*
agriculture	-0.7209***	-0.5218***
mining	0.2470***	0.3822***
electricity	0.1680***	0.1208**
construction	-0.1196***	-0.1223***
craft and trade	-0.2168***	-0.2581***
transport	0.0794***	0.0216
financial services	-0.0098	-0.0159
community services	-0.0966***	-0.1624***
lambda	-0.0982***	-0.1302***
constant	0.6377***	0.8169***
number of observations	22617	22862
F-statistic	1036	818
R-squared	0.6585	0.6009

Source: *Own calculations, LFS(2001:2) & LFS(2005:2).*

Notes: ****significant at the 1% level; **significant at the 5% level; *significant at the 10% level*

The data are weighted. The unspecified categories for occupation groups and industries are excluded from the table.

As expected, those working in provinces other than Gauteng earned on average less than those working in Gauteng when controlling for other characteristics. The earnings of those working in Limpopo and the Free State in particular, were lower than those working in Gauteng. Interestingly, by 2005 those working in the Western Cape did not earn significant differently from those working in Gauteng.

As expected, managers earned the highest earnings when compared to elementary workers and this was more pronounced in 2005 than in 2001. They were followed by professionals and clerks. At the bottom of the occupation distribution, domestic workers earned 13% less than elementary workers in 2001, but by 2005, they earned 5% more than elementary workers (at the 10% level of significance). The public sector dummy shows that the premium on working in the public sector increased between 2001 and 2005. Interestingly, being married earns about a 10% (7%) premium in 2001 (2005) and the significance of this variable justifies its inclusion in the regression.

The coefficient on the experience variable indicates that being in the labour market for a longer period of time is associated with a positive return on earnings, while the negative and significant co-efficient on the experience-squared variable shows that that after a certain point more experience is associated with decreasing returns. I find, as expected, that formal sector workers earned more than informal sector workers, with the earnings premium for formal sector workers standing at around 68% in both years. Interestingly, the earnings estimation for both years shows that the wage-employed earned significantly less than the self-employed.

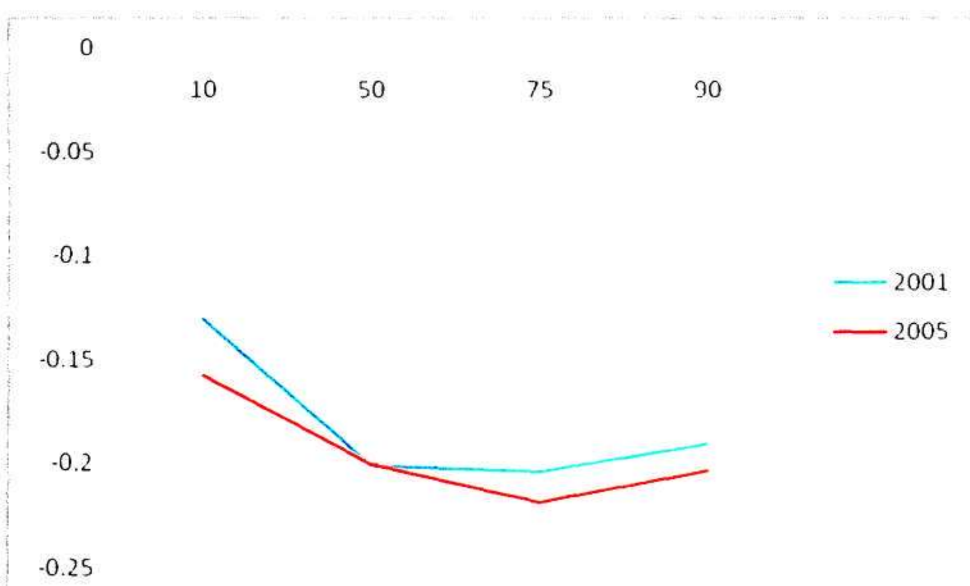
Having estimated the earnings function, using the sample selection procedure above, I then estimated the earnings function without correcting for sample selection. I find the same general result (see Appendix 3), that is, that the gender earnings gap increased over time. In addition, the coefficients all have the same signs as in the corrected estimated. However, the coefficients are larger for the uncorrected estimation, but this is to be expected since the uncorrected estimation picks up the direct effects of the explanatory variables on the earnings, as well as the indirect effects of those same explanatory variables on labour force participation and employment. Specifically, the coefficients on the female dummies indicate that the women earned even less than their male counterparts in the uncorrected estimation. This is true in both 2001 and 2005.

While the OLS regression estimates above capture the earnings differential between males and females when controlling for sample selection as well as other observable characteristics at the *mean* earnings distribution, the effect might however be different at different points of the earnings distribution. The use of quantile regressions as discussed above measures the gender earnings gap effect at different points of the conditional earnings distribution. Thus, I estimated the above earnings equation at four quantiles of the earnings distribution, that is, the 10^{th}, 50^{th} (median), 75^{th} and 90^{th} percentile.

Quantile regression techniques were first employed by Buchinsky (1994) to look at earnings disparities among women in the United States, but are now widely employed in the international literature that estimates gender earnings differentials. It is however difficult to address the problem of selection bias in quantile regressions as the form of this bias over different percentiles of the earnings distribution is unknown. Some studies have followed Buchinsky (1994) and estimate a labour force participation model which is transformed into several series expansions which are then included in the earnings equation as controls for selection bias (see Ntuli 2006). There is however "little consensus regarding the most appropriate selection procedure for selectivity bias in quantile regression models" (Ntuli, 2006: 3).

Given the lack of consensus, I estimate the quantile regressions on the mean selectivity-corrected earnings estimation reported in the table above. For the purpose of comparison, I also estimate the quantile regression estimates without taking sample selection into account. Below I show the female gender earnings coefficient as from the above regression but for four different quantiles of the earnings distribution. The figure shows firstly that in both 2001 and 2005 women earned less than their male counterparts at all points in the earnings distribution, when controlling for other observable characteristics, that is a gender earnings premium for males exists throughout the distribution. Secondly, the figure shows that in any given year, the gender earnings gap increases as one moves up the earnings distribution from the 10th to the 75th percentile and then decreases between the 75th and 90th percentile. This is true for both 2001 and 2005, except that 2001 displays a flatter gender earnings disparity between the 50th and 90th percentile. In addition, the figure shows that the gender earnings gap increased throughout the distribution, except at the 50th percentile. This provides early evidence of both a glass ceiling and a sticky floor for women, that is this evidence suggests a widening of the gender earnings differential both at the top and bottom of the earnings distribution between 2001 and 2005.

Figure 2: Quantile Regressions, Female Dummy Coefficients: 2001 & 2005



Source: *Own calculations, LFS(2001:2) & LFS(2005:2).*

Notes: *The coefficients are significant at the 1% level of significance, for all the quantiles for all years. The data are **not** weighted.*

In addition to estimating the earnings equation accounting for sample selection, I also estimated the gender earnings gap for the four quantiles of the earnings distribution without taking sample selection into account. The results are shown in Appendix 4, and are similar for the two estimations, with the exception that the estimation without sample selection shows a steeper increase in the gender earnings gap between the 10^{th} and 75^{th} percentile in a given year.

In summary then, the pooled regression estimation shows firstly that in 2001 and 2005 the gender earnings gap increased as one moves up the distribution to the 75^{th} percentile, and secondly that between 2001 and 2005 the gap increased over the quantiles of the distribution, with the exception of median income. However, the pooled regression used in the analysis of the gender earnings gap above implies that men and women have the same returns to observable characteristics. This assumption has been subjected to much debate – for instance Polachek & Xiang (2006) note that "it is widely argued that employers value similar skills differently for men and women" (Polachek & Xiang, 2006: 13). Therefore male and female earnings equations need to be estimated separately in order to obtain a more accurate reflection of the gender pay gap. In addition, estimating two separate earnings equations for males and females also allows one to understand what has happened to discrimination using the Oaxaca (1973) decomposition technique.

4.4. Understanding the Gender Pay Gap & Discrimination: Separate Specifications for Men and Women

In this section I account for the fact that males and females may have different returns to observable characteristics by estimating separate earnings equations for both genders for each year. In estimating the pooled regression, I accounted for sample selection given that the sample of employed and earners are highly unlikely to be a random sample of the working-age population. When decomposing gender earnings gaps however, it is not obvious how the selection term should be treated in the overall decomposition, that is, it is not clear whether it should be attributed to differences in endowments or included in the remuneration effect. The literature consists of several variations, but "it is not possible to point out any of these are the right one, since the appropriate procedure depends on the specific empirical problem and the data at hand" (Beblo, 2003). Given the lack of consensus on the issue I choose not to take selection into account.

The international literature commonly uses the traditional Oaxaca (1973) decomposition technique to decompose the average gender pay gap between men and women based on the OLS estimation of gender-specific earnings equations. The technique allows the overall average differential in earnings between the two groups to be decomposed into two parts: the 'explained' effect and the 'unexplained' effect, where the latter is treated as a measure of discrimination.

The technique involves first estimating separate earnings equations for males and females:

$$lnY_m = X_m b_m$$

$$lnY_f = X_f b_f$$

where Y_m and Y_f are the the earnings of males and females respectively, X_m and X_f are vectors of regressors for male and female earnings functions and b_m and b_f are vectors of parameters for male and female earnings functions.

The difference between male and female earnings, or the earnings gap can then be defined as:

$$lnY_m - lnY_f = X_m b_m - X_f b_f$$

If the two *b* vectors differ, then the same characteristics of the two groups are rewarded differently.

Using the equation above, the decomposition can be done relative to male or female earnings, depending on what one thinks about which earnings level will prevail in the absence of discrimination. In my analysis, I do the decomposition relative to male earnings, as explained further below. In this case, the equation above can be re-written as:

$$lnY_m - lnY_f = (X_m - X_f)\, b_m + X_f\, (b_m - b_f)$$

The earnings gap now consists of two parts: i) the 'explained' portion, or the actual differences in endowment levels between the two groups (the first term), and ii) the 'unexplained' portion, or the differences in the market evaluation of the same endowments (second term), where the second term is an estimate of discrimination.

The assumption that the 'unexplained' portion of the gender earnings gap is a measure only of discrimination has been subjected to much criticism. For instance, Kunze (2006) argues that unobserved skills that affect individual choices of work histories play an important role and may therefore, in part, account for the 'unexplained' portion of the differences in earnings between genders (Kunze, 2006: 9). In addition, if a human capital attribute that affects earnings is omitted, the measured 'discrimination' component would be contaminated, thereby biasing the measure of discrimination upwards. Furthermore, the 'unexplained' portion can pick up pre-labour market factors like school quality, family background and institutionalism, all of which can bias the measure of discrimination.

Table 13 shows the earnings equations for males and females. The separation of the male and female equations is justified on the basis that I reject the null hypothesis that the coefficients are the same from the male and female equations using a Chow Test.[22] This is true for both the 2001 and 2005 estimates. The coefficients shown in Table 13 below indicate that returns to observable characteristics are different for the two sexes. For instance, being white rather than African matters much more for males than for females, and this is true in both 2001 and 2005. The education variables in both years show similar returns between males and females. However, interestingly, the returns to grade 12 education increased far more for males than for females between 2001 and 2005, while the value of a diploma decreased for both sexes. Having a degree also has a higher return in 2005 than in 2001. The province variables show that the effect of being in a province other than Gauteng has a much bigger (negative) impact on female earnings rather than male earnings, but the effect seems to have dampened by 2005 compared to 2001.

---

[22] Using the error sum of squares from the separate regressions (ess_1 and ess_2 in this case), the error sum of squares from the pooled (constrained) regression (ess_c), the number of estimated parameters (k), and the number of observations in the two groups (N_1 and N_2) I used the following formula to calculate the Chow test in each year:

$$\frac{((\text{ess\_c} - (\text{ess\_1} + \text{ess\_2})) / k)}{((\text{ess\_1} + \text{ess\_2}) / (\text{N\_1} + \text{N\_2} - 2 * k))}$$

I obtained a test statistic of 242 for 2001 and 213 for 2005.

Table 13: Separate Male and Female Earnings Equations Uncorrected for Sample Selection: 2001 & 2005

	2001		2005	
	males	females	males	females
coloured	0.3162***	0.2848***	0.2321***	0.2446***
asian	0.3809***	0.3893***	0.5169***	0.2252***
white	0.7290***	0.5561***	0.6212***	0.4324***
none_grade8	0.0430***	0.0365***	0.0269***	0.0270***
grade9_grade11	0.0718***	0.0749***	0.0729***	0.0629***
grade12	0.2259***	0.2099***	0.2675***	0.2195***
diploma	0.3071***	0.3171***	0.2474***	0.2694***
degree	0.1446***	0.1524***	0.2024***	0.1984***
western cape	-0.0514**	-0.0342	0.0407	0.0136
eastern cape	-0.3381***	-0.4439***	-0.2404***	-0.2720***
northern cape	-0.2521***	-0.3733***	-0.2026***	-0.4107***
free state	-0.3415***	-0.4601***	-0.2942***	-0.3798***
kwazulu natal	-0.1598***	-0.2446***	-0.3442***	-0.3373***
north west	-0.2013***	-0.2956***	-0.2334***	-0.2353***
mpumalanga	-0.1465***	-0.3362***	-0.2649***	-0.3120***
limpopo	-0.3810***	-0.4502***	-0.3441***	-0.4893***
married	0.1598***	0.0331**	0.1114***	0.0282*
experience (proxy)	0.0375***	0.0290***	0.0305***	0.0269***
experience-squared (proxy)	-0.0005***	-0.0003***	-0.0004***	-0.0003***
public sector	0.3713***	0.4059***	0.4262***	0.5141***
formal sector	0.4888***	0.5505***	0.4978***	0.5587***
wage-employed	-0.1018***	-0.1899***	-0.0765***	-0.1015***
managers	0.6483***	0.7761***	0.7712***	0.8323***
professionals	0.4931***	0.5424***	0.6520***	0.5362***
clerks	0.3121***	0.4034***	0.4692***	0.3882***
service	0.0335	0.0722***	0.0209	-0.0191
skilled agric & fishing	0.3812***	-0.0608	-0.0454	-0.2549***
craft & trade	0.2537***	0.0673*	0.1994***	0.0019
operators & assemblers	0.1286***	0.0819*	0.2080***	0.1567***
domestic	-0.3138***	-0.0789*	-0.1026**	0.0763*
agriculture	-0.7927***	-0.5873***	-0.5459***	-0.4740***
mining	0.1910***	0.1485	0.3502***	0.3568***
electricity	0.1164**	0.1708	0.0938	0.1041
construction	-0.1485***	-0.2640***	-0.1332***	-0.1502**
craft and trade	-0.2407***	-0.2352***	-0.2560***	-0.2821***
transport	0.0335	0.1235**	-0.0292	0.1321**
financial services	-0.1102***	0.1074***	-0.0878***	0.0864**
community services	-0.1226***	-0.0826**	-0.1879***	-0.1162***
constant	0.5193***	0.5400***	0.6994***	0.6116***
number of observations	12513	10104	12375	10479
F-statistic	541.69	552.05	434.89	419.33
R-squared	0.6347	0.6815	0.5851	0.6104

Source: *Own calculations, LFS(2001:2) & LFS(2005:2.)*

Notes: ****significant at the 1% level; **significant at the 5% level; *significant at the 10% level*

The data are weighted. The unspecified categories for occupation groups and industries are excluded from the table.

The separation of the equations highlights that the marriage premium is much higher for males than for females, and this is true for both years. A large and significant male marital earnings premium is consistent with findings reported in Casale and Posel (2007). In both 2001 and 2005, experience has a larger (positive) impact on male earnings than on female earnings. However, as mentioned above, female work histories are likely to be more interrupted than those for men, particularly in developing countries. The proxy for work experience may therefore be overestimating experience, particularly for females, and this may be an explanation for the lower co-efficient on the female experience covariate compared to that for men. The co-efficient on the public sector variable indicates that being in the public sector (rather than the private sector) provides a higher return for women than for men, and this effect is more pronounced in 2005 than in 2001. The same is true for the formal sector. By comparison, being wage-employed rather than self-employed is associated with a larger earnings penalty for women than for men.

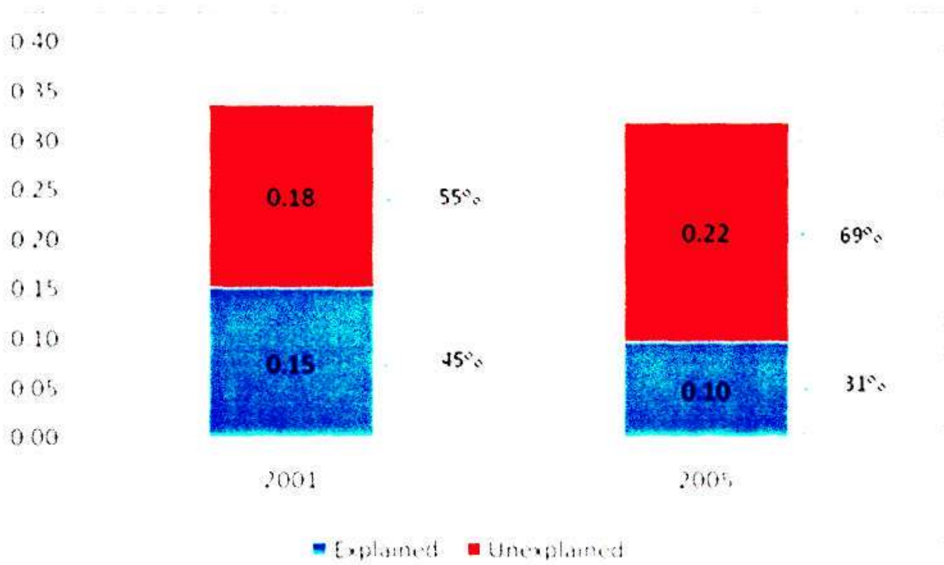
The occupation disaggregation indicates differing returns to occupation by gender. The returns to being a manager rather than an elementary worker are higher for females than males in both 2001 and 2005. The same is true for domestic workers. In contrast, being a craft and trade worker or operator and assembler rather than an elementary worker provides a lower comparative return for females than for males. The data further indicates that in 2001 it was more favourable for females to be professionals, clerks and service workers rather than elementary workers, but that by 2005 the situation was reversed. It must be noted however that the occupation categories are broad and do not identify specific occupations, and this may partly be the reason for the results generated. The same is true for the industry disaggregation, that is, the industry disaggregations represent a variety of occupation groups within each industry. Thus, though the data indicates differing returns to industry group by gender, the results are difficult to interpret, given that the disaggregation is very broad.

The separation of the equations for males and females therefore indicates that observable characteristics offer differing returns for men and women. Below I use the Oaxaca (1973) decomposition technique to decompose the gender earnings gap based on the earnings equations above. When decomposing gender earnings gaps using this technique, it is necessary to specify whether the male or female earnings estimation is the 'unbiased' measure of earnings. The choice of which of the two to use depends on the assumption on which earnings level will prevail in the absence of discrimination.

In his original article Oaxaca (1973) used both male and female earnings as referents and interpreted the two answers as upper and lower bounds of discrimination. He later indicated that this was an incorrect interpretation and introduced a neutral benchmark (that is a non-discriminatory earnings level based for example on a pooled regression of earnings) to deal with the problem and distinguish between positive and negative discrimination (Chamberlain & Van Der Berg, 2002). Many international studies use the male earnings estimation as the referent since it gives an idea of the discrimination experienced by women in comparison to what males earn (see for instance de la Rica et al, 2005). I follow this method in my estimation of discrimination.

Using the Oaxaca (1973) methodology, I firstly find from the figure below that the gender pay gap decreased marginally between 2001 and 2005. This corroborates the evidence from the descriptive statistics presented above, but goes contrary to the evidence presented for the pooled regression model. However, since I rejected the null hypothesis of equal coefficients on the covariates for the two genders, the estimated earnings gap using two separate equations holds more weight.[23]

Figure 3: Decomposition of Gender Earnings Gap: 2001 & 2005



Source: *Own calculations, LFS(2001:2) & LFS(2005:2).*

The figure above further shows the decomposition of the gender pay gap into the 'explained' and 'unexplained' parts. It shows two interesting results: Firstly, in both 2001 and 2005 the part of the gender pay gap that can be attributed to observable characteristics or endowments is less than the 'unexplained' portion. Put differently, the overall gender pay gap in both years is driven to a larger extent by the 'unexplained' component than the 'explained' component. If the 'unexplained' component is taken to be 'discrimination', then the data thus indicates that a greater proportion of the gender pay gap is explained by discrimination in earnings between the sexes than by differences in observable characteristics between the sexes.

Secondly, the 'explained' component of the gap in the overall gender pay gap decreased between 2001 and 2005, while the 'unexplained' component increased. That is, while observable characteristics accounted for 45% of the gender pay gap in 2001, by 2005 endowments accounted for only 31% of the gap. Thus, the contribution of the 'unexplained' part to the overall gender earnings gap increased from 55% to 69% between 2001 and 2005, even though the gender pay gap

[23] For instance, de la Rica et al (2005) first used the dummy variable method to estimate the gender earnings differential before separating male and female equations. They call their results from the dummy variable method 'tentative' since they reject the null hypothesis of equal coefficients on the covariates for both genders.

decreased marginally in the period. If the 'unexplained' component is taken to be discrimination, then the data indicates that discrimination in earnings increased in the period between 2001 and 2005.

In summary then, the separation of the male and female earnings equations shows differing returns to observable characteristics. Furthermore, the data indicates a slight narrowing of the gender earnings differential between 2001 and 2005. The Oaxaca (1973) decomposition shows that the overall gender earnings gap is explained to a greater extent by the 'unexplained' component rather than endowments, and this is true for both 2001 and 2005. In addition the decomposition illustrates rising discrimination in the period since the contribution of the 'unexplained' component to the overall gender earnings gap increased.

4.5. Conclusion

The multivariate analysis using a pooled regression model showed an increase in the gender earnings gap between 2001 and 2005 for the mean regression as well as over the quantiles of the distribution, except at the median. In addition, the quantile regression estimations revealed an increasing gender pay gap as one moves up the distribution in any given year. The pooled regression however assumes that the returns to observable characteristics are the same for males and females. Separating out the male and female equations for both years, I rejected the null hypothesis that the coefficients on the male and female estimations were the same, using a Chow Test.

Using separate male and female equations, I found a decreasing gender pay gap between 2001 and 2005. Decomposing the gap using the Oaxaca (1973) decomposition technique, I found that the 'unexplained' portion of the gap explains the gender earnings differential to a larger extent than endowments in both 2001 and 2005. Furthermore, the contribution of the 'unexplained' proportion increased between 2001 and 2005, even though the overall gender pay gap decreased marginally. If the 'unexplained' portion is taken to be discrimination, then the data indicates that discrimination in earnings between the sexes increased over the period. If the residual represents unobserved attributes however, then the estimations suggest a change in the selection of women and men into employment over the period.

Chapter Five: Conclusion

Unlike the race earnings gaps, the gender earnings gap has been given less attention in the South African literature. Internationally, studies across the world attempt to understand the change in the earnings gap between the sexes, and to explain why it exists. This thesis investigates the evolution of the gender earnings gap between 2001 and 2005 using the LFSs. I elected to use hourly earnings instead of monthly earnings since women on average work significantly fewer hours per week than their male counterparts.

Descriptive statistics illustrate a decrease in the overall gender earnings gap in the period under consideration, associated with relative increases in average earnings for highly skilled African/Coloured women. Male earnings however remain significantly higher than female earnings in 2005 all along the skills and education distributions. In addition, the gender earnings differential remains particularly high for low skilled workers and those with no education. In contrast to the descriptive evidence presented, the pooled regression (controlling for sample selection) indicates an increase in the gender pay gap between 2001 and 2005 for the mean regression as well as over the quantiles of the distribution, except at the median.

Given that the pooled regression assumes the same returns to observable characteristics for males and females, I then estimate separate earnings regressions for males and females. The Chow test on the separate estimations justifies the separation of the equations for the sexes. In contrast to the pooled regression, but consistent with the descriptive evidence, I find a slight decrease in the gender pay gap between 2001 and 2005 using separate male and female earnings estimations. The Oaxaca (1973) decomposition technique illustrates that the 'unexplained' part of the gap accounts for a greater proportion of the gender pay gap than endowments in both 2001 and 2005. Furthermore, the contribution of the 'unexplained' component of the gap to the overall gap increased in the period between 2001 and 2005. If the 'unexplained' component is taken to be a measure of discrimination then the data indicates an increase in discrimination in earnings for women, in the period between 2001 and 2005, even though there was a slight narrowing of the gender pay differential.

Chapter Six: Appendix

Appendix 1: Labour Force Participation Probits: 2001 & 2005

	2001		2005	
	dF/dx	x-bar	dF/dx	x-bar
female	-0.1488***	0.5285	-0.1235***	0.5270
coloured	-0.0317***	0.0984	-0.0383***	0.0972
asian	-0.1405***	0.0324	-0.1397***	0.0304
white	-0.1367***	0.1246	-0.1358***	0.1121
age_25-34	0.0573***	0.3154	0.0567***	0.3201
age_35-44	0.0393***	0.2228	0.0242***	0.2148
age_45-54	-0.0531***	0.1559	-0.0513***	0.1609
age_55-65	-0.3363***	0.1138	-0.3405***	0.1220
none_grade8	0.0112***	6.4751	0.0116***	6.7213
grade9_grade11	0.0116***	1.4133	0.0151***	1.5839
grade12	0.0487***	0.3107	0.0510***	0.3474
diploma	0.0473***	0.0899	0.0412***	0.0939
degree	0.0296***	0.0746	-0.0010	0.0753
western cape	-0.0318***	0.1097	-0.0465***	0.1139
eastern cape	-0.0620***	0.1305	-0.0958***	0.1293
northern cape	-0.0496***	0.0218	-0.0585***	0.0200
free state	0.0038	0.0693	-0.0619***	0.0645
kwazulu natal	-0.0494***	0.2076	-0.0796***	0.2001
north west	-0.0376***	0.0824	-0.0434***	0.0812
mpumalanga	-0.0132	0.0648	-0.0308***	0.0641
limpopo	-0.0209**	0.0967	-0.0535***	0.0952
no of children under 7 in hh	-0.0101***	0.9069	-0.0144***	0.7950
no of children 8-15 in hh	-0.0131***	0.8545	-0.0070***	0.7823
no of over 65 year olds in hh	-0.0386***	0.1393	-0.0322***	0.1263
no of working-age men	-0.0106***	1.5238	-0.0097***	1.4972
no of working-age women	0.0153***	1.5703	0.0131***	1.3684
observed probability		0.8044		0.8108
predicted probability (at x-bar)		0.8450		0.8529
no of observations		53560		54695
chi-squared		5552.64		4720.61
pseudo R-squared		0.188		0.1972

Source: *Own calculations, LFS(2001:2) & LFS(2005:2)*

Notes: ****significant at the 1% level; **significant at the 5% level; *significant at the 10% level. The data are weighted.*

Appendix 2: Employment Probits: 2001 & 2005

	2001		2005	
	dF/dx	x-bar	dF/dx	x-bar
female	0.0334**	0.4827	-0.0421***	0.4890
coloured	0.0450***	0.0970	0.0220	0.0952
asian	0.1287***	0.0299	0.0751**	0.0280
white	0.2459***	0.1211	0.1698***	0.1079
age_25-34	0.1820***	0.3526	0.1957***	0.3602
age_35-44	0.2583***	0.2393	0.2385***	0.2291
age_45-54	0.2499***	0.1454	0.2257***	0.1516
age_55-65	0.2837***	0.0599	0.2276***	0.0648
none_grade8	-0.0077***	6.7483	0.0010	7.0085
grade9_grade11	-0.0057	1.5396	-0.0048	1.7291
grade12	0.0367***	0.3441	0.0554***	0.3843
diploma	0.1544***	0.1023	0.1754***	0.1041
degree	0.0287*	0.0865	0.0400**	0.0829
western cape	0.0250	0.1098	-0.0076	0.1139
eastern cape	-0.0007	0.1212	-0.0048	0.1186
northern cape	0.0587***	0.0207	-0.0263	0.0193
free state	0.0217*	0.0724	0.0172	0.0642
kwazulu natal	-0.0137	0.1999	-0.0338**	0.1918
north west	0.0198	0.0818	-0.0305*	0.0816
mpumalanga	-0.0035	0.0654	-0.0227	0.0652
limpo	-0.0199	0.0957	-0.0699***	0.0932
no employed in hh	0.4288***	1.3064	0.4385***	1.3236
head of hh	0.4914***	0.4373	0.4777***	0.4530
lambda	-0.1635**	0.2768	0.0646	0.2661
observed probability		0.5919		0.6132
predicted probability (at x-bar)		0.6627		0.7020
no of observations		42841		42990
chi-squared		6842.35		5049.85
pseudo R-squared		0.4826		0.4953

Source: *Own calculations, LFS(2001:2) & LFS(2005:2)*

Notes: ****significant at the 1% level; **significant at the 5% level; *significant at the 10% level. The data are weighted.*

Appendix 3: Earnings Equation Uncorrected for Sample Selection, Pooled Regression: 2001 & 2005

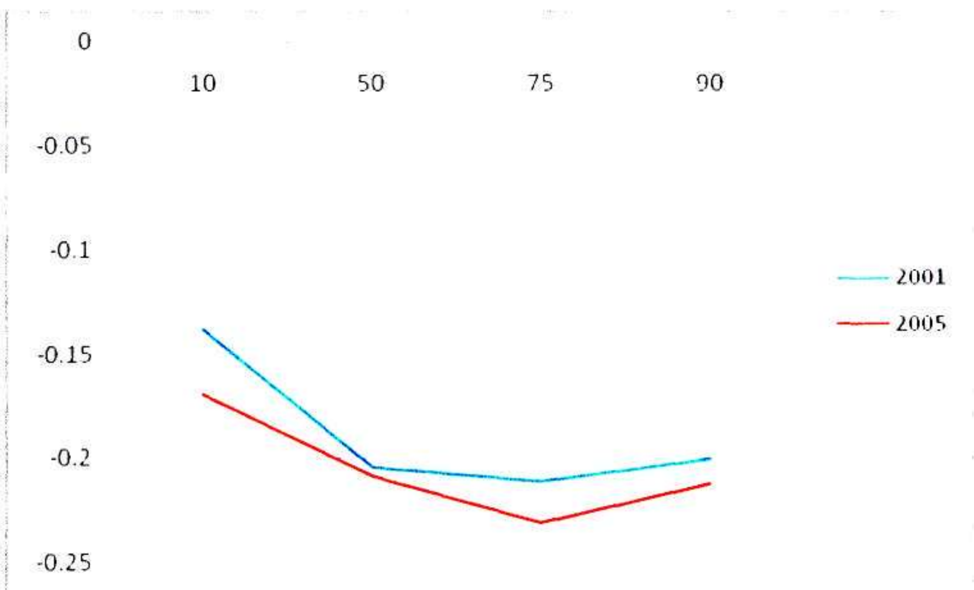
	2001	2005
female	-0.2152***	-0.23972***
coloured	0.309644***	0.247189***
asian	0.394945***	0.413618***
white	0.66863***	0.5378***
none_grade8	0.041924***	0.02855***
grade9_grade11	0.075281***	0.071142***
grade12	0.223769***	0.25314***
diploma	0.315616***	0.264452***
degree	0.148302***	0.19899***
western cape	-0.04033**	0.024512
eastern cape	-0.3947***	-0.25671***
northern cape	-0.30649***	-0.29317***
free state	-0.39425***	-0.33489***
kwazulu natal	-0.19345***	-0.34446***
north west	-0.24139***	-0.23794***
mpumalanga	-0.22505***	-0.28712***
limpopo	-0.40654***	-0.41232***
married	0.09541***	0.06555***
experience (proxy)	0.03468***	0.029878***
experience-squared (proxy)	-0.00042***	-0.00036***
public sector	0.390681***	0.472537***
formal sector	0.520429***	0.527423***
wage-employed	-0.11941***	-0.08719***
managers	0.675138***	0.796306***
professionals	0.496222***	0.598364***
clerks	0.340134***	0.407863***
service	0.031047*	-0.00711
skilled agric & fishing	0.257856***	-0.12139***
craft&trade	0.232659***	0.16023***
operators & assemblers	0.120506***	0.183002***
domestic	-0.13985***	0.058323**
agriculture	-0.71463***	-0.51761***
mining	0.247714***	0.382599***
electricity	0.168356***	0.112186*
construction	-0.11951***	-0.11996***
craft and trade	-0.21629***	-0.25899***
transport	0.080203***	0.019124
financial services	-0.00847	-0.01517
community services	-0.09579***	-0.16156***
constant	0.567384***	0.730387***
number of observations	22617	22862
F-statistic	1058.63	834.37
R-squared	0.6578	0.5999

Source: *Own calculations, LFS(2001:2) & LFS(2005:2)*

Notes: ****significant at the 1% level; **significant at the 5% level; *significant at the 10% level*

The data are weighted. The unspecified categories for occupation groups and industries are excluded from the table.

Appendix 4: Quantile Regressions, Female Dummy Coefficients, Uncorrected for Sample Selection: 2001 & 2005



Source: *Own calculations, LFS(2001:2) & LFS(2005:2)*

Notes: *The coefficients are significant at the 1% level of significance, for all the quantiles for all years.*

*The data are **not** weighted.*

Chapter Seven: References

Ajwad, M.I. & Kurukulasuriya. (2002). Ethnic and Gender Wage Disparities in Sri Lanka, World Bank Research Working Paper 2859.

Albrecht, J., Vureen, A. & Vroman, S. (2004). Decomposing the Gender Wage Gap in the Netherlands with Sample selection Adjustments, IZA Discussion Paper No. 1400.

Allanson, P. & J. P. Atkins. (2001). Labour Market Reform and the Evolution of the Racial Wage Hierarchy in Post-Apartheid South Africa, DPRU Working Paper No. 01/59. University of Cape Town.

Beblo, M., Beninger, D., Heinze, A. & Laisney, F. (2003). Measuring Selectivity-Corrected Gender Wage Gaps in the EU, Discussion Paper No. 03-74. Centre for European Economic Research.

Bhorat, H. (2000). Wage Premia and Wage Differentials in the South African Labour Market, DPRU Working Paper No. 00/43.

Bhorat, H. & Leibbrandt, M. (2001). *Modelling Vulnerability and Low Earnings in the South African Labour Market.* Chapter 3 in Bhorat, H; Leibbrandt, M; Maziya, M, Van der Berg, S; and Woolard, I. 2001. *Fighting poverty: Labour Markets and Inequality in South Africa.* Cape Town: UCT Press.

Blau, F. & Kahn, L.M. (2003). "Understanding International Differences in the Gender Pay Gap", *Journal of Labour Economics*, Vol. 21:(1).

Buchinsky, M. (1994). "Changes in the US Wage Structure 1963-1987: An Application of Quantile Regression", *Econometrica*, Vol. 62.

Burger, R. & Jafta, R. (2006). Returns to Race: Labour Market Discrimination in Post-Apartheid South Africa, Stellenbosch University Economic Working Paper 04/06.

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

Casale, D & Posel, D. (2002). "The Continued Feminisation of the Labour Force in South Africa: An Analysis of Recent Data and Trends", *The South African Journal of Economics*, Vol. 70:(1).

Casale, D., Muller, C., & Posel, D. (2004). "'Two Million Net New Jobs': A Reconsideration of the Rise in Employment in South Africa", *The South African Journal of Economics*, Vol. 72:(5).

Casale, D. & Posel, D. (2005). Who replies in brackets and what are the implications? An analysis of earnings data in South Africa, ERSA Working Paper No. 7.

Casale, D. & Posel, D. (2007). Bridewealth and the Marital Earnings Premium for Men: Evidence from South Africa, ERSA Working Paper No. 57

Chamberlain, D & Van der Berg, S. (2002). Earning Functions labour Market Discrimination and Quality of Education in South Africa, Stellenbosch University Economic Working Paper 2/2002.

de la Rica, S., Dolado J.J & Llorens, V. (2005). Ceilings and Floors, Gender Wage Gaps by Education in Spain, IZA Discussion Paper No. 1483.

Erichsen, G. & Wakeford, J. (2001). Racial Wage Discrimination in SA before and after the first Democratic Election, DPRU Working Paper No 01/49.

González, P., Santos, M.C. & Santos, L.D. (2005). The Gender Wage Gap in Portugal: Recent Evolution and Decomposition, DP 2005 – 05. Research Centre on Industrial, Labour and Managerial Economics.

Gruen, C. (2004). "Direct and Indirect Gender wage discrimination in the South Africa labour market", *International Journal of Manpower*, Vol. 25:(3-4).

Gunawardena, D. (2006). Exploring Gender Wage Gaps in Sri Lanka: A Quantile Regression Approach, Paper presented during the 5th PEP Research Network General Meeting. Addis Ababa, Ethiopia, 18-22 June 2006.

Heintz, J. & Posel, D. (2008). "Revisiting Informal Employment and Segmentation in the South African Labour Market", *South African Journal of Economics*, Vol. 76:(1)

Hinks, T. (2002). "Gender, Wage differentials and Discrimination in the new South Africa", *Applied Economics*, Vol. 34:(16)

Hyder, A. & Reilly, B. (2005). The Public Sector Pay Gap in Pakistan: A Quantile Regression Analysis, PRUS Working Paper 33.

Isemonger, A.G. & Roberts, N.J. (1999). "Post Entry Gender Discrimination in the South African Labour Market", *Journal of Studies in Economics and Econometrics*, Vol 23:(2).

Koenker, R & Basset, G. (1978). "Regression Quantiles", *Econometrica*, Vol. 46.

Kunze, A. (2006). Gender Wage Gap Studies: Consistency and Decomposition, Working Paper No 48/06. Institute for Research in Economics and Business Administration.

Maziya, M. (2001). *Contemporary Labour Market Policy and Poverty in South Africa*, Chapter 8 in Bhorat, H, Leibbrandt, M., Maziya, M., Van Der Berg, S. and Woolard, I. *Fighting Poverty: Labour Markets and Inequality in South Africa*. Cape Town: UCT Press.

Mincer, J. (1974). *Schooling Experience and Earnings*. New York: Colombia University Press for the National Bureau of Economic Research.

Montenegro, C. (2001). Wage Distribution in Chile: Does Gender Matter? A Quantile Regression Approach, Working Paper Series No. 20. World Bank Development Research Group.

Muller, C. (2008). Trends in the gender wage gap and gender discrimination among part-time and full-time workers in post-apartheid South Africa, Paper Presented at the 13th Annual Conference on Econometric Modeling in Africa. University of Pretoria, 9-11 July 2008.

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", *American Economic Review*, Vol 86:(2).

Nielsen, H.S. & Rosholm, M. (2001). "The Public-Private wage gap in Zambia in the 1990s: A quantile regression approach", *Empirical Economics*, Vol. 26:(1).

Ntuli, M. (2007). Exploring gender wage 'discrimination' in South Africa, 1995-2004: A Quantile Regression Approach, IPC Working Paper Series No. 56.

Oaxaca, R.L. (1973). "Male Female Wage Differentials in Urban Labour Markets", *International Economic Review*, Vol. 14:(3).

Oosthuizen, M. (2006). The Post-Apartheid Labour Market: 1995-2003, DPRU Working Paper No. 06/103.

Pena-Boquete, Y., Destefanis, S. & Fernandez-Grela, M. (2007). The Distribution of Gender Wage Discrimination in Italy and Spain: A Comparison Using the ECHP, Paper Presented at the XXII National Conference of Labour Economics. Napoli, 13-14 September 2007.

Pham, T.H. & Reily, B. (2006). The Gender Pay Gap in Vietnam, 1993-2002: A Quantile Regression Approach, PRUS Working Paper No. 34.

Polachek, S.W. & Xiang, J. (2006). The Gender Pay Gap: A Cross-Country Analysis, Unpublished Paper. SUNY-Binghamton, February 2006.

Republic of South Africa. (1995). *Labour Relations Act No. 66 of 1995.* Available at www.labour.gov.za

Republic of South Africa. (1997). *Basic Conditions of Employment Act No. 75 of 1997.* Available at www.labour.gov.za

Republic of South Africa. (1998). *Employment Equity Act No. 55 of 1998.* Available at www.labour.gov.za

Rospabé, S. (2001). An Empirical Evaluation of Gender Discrimination in Employment Occupation and Wage in South Africa in the late 1990s, Mimeographed University of Cape Town.

Rospabe, S. (2002). "How Did Labour Market Racial Discrimination Evolve After the End of Apartheid?" *South African Journal of Economics,* Vol. 70:(1).

Stata Technical Bulletin, STB-15. September 1993.

Stata Technical Bulletin, STB-18. March 1994.

Weichselbaumer, D. & Winter-Ebmer, R. (2005). "A Meta-Analysis of the International Gender Wage Gap", *Journal of Economic Surveys*, Vol 19(3).

Woolard, I. (2002). A Comparison of Wage Levels and Wage Inequality in the Public and Private Sectors, 1995 and 2000, DPRU Working Paper No. 02/62.

UNIVERSITY OF
KWAZULU-NATAL

RESEARCH OFFICE (GOVAN MBEKI CENTRE)
WESTVILLE CAMPUS
TELEPHONE NO.: 031 – 2603587
EMAIL : ximbap@ukzn.ac.za

06 OCTOBER 2008

MS. S GOGA (201500851)
ECONOMICS

Dear Ms. Goga

ETHICAL CLEARANCE APPROVAL NUMBER: HSS/0605/08M

I wish to confirm that ethical clearance has been approved for the following project:

"Understanding the Gender Earnings Gap in the Post-Apartheid South African Labour Market"

PLEASE NOTE: Research data should be securely stored in the school/department for a period of 5 years

Yours faithfully

....
MS. PHUMELELE XIMBA

cc. Supervisor (Prof. D Posel)
cc. Mrs. C Haddon