

A Quantile Regression Analysis of the Conditional Female Wage Distribution in Post-Apartheid South Africa



by
Safia Khan
KHNSAF003

A thesis submitted in partial fulfillment of the requirements for the degree of
Master of Commerce in Applied Economics
School of Economics
University of Cape Town

Supervisor
Prof. Murray Leibbrandt

Final Version Dated: 2014 12 08

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

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

Abstract

The democratisation of South Africa in 1994 brought about increased economic opportunity for labour market participants. The impact of democratisation on the labour market outcomes of both males and females of various races has been studied extensively in terms of mean earnings. However, because of the heterogenous distribution of South African labour force participants, a mean analysis is not sufficient. This paper investigates changes in the post-Apartheid conditional earnings distribution of women in South Africa of all races using conditional quantile regression analysis. Women are exclusively analysed because the labour market choices and constraints they face are unique. A two part model is estimated with the first part correcting for sample selection into employment in a South African context. Then, quantile regression analysis is employed to parsimoniously describe the entire conditional wage distribution aiding the understanding of female labour force dynamics in South Africa. Individual data on the South African labour market from Statistics S.A is utilised for the analysis. White women are found to earn more than black and coloured women at all quantiles of the wage distribution conditional on education, skills level and sector of employment. Inequality is found to be increasing among women in higher skilled occupations. Women in lower skilled occupations have witnessed an increase in real wages at all quantiles of the conditional wage distribution.

Contents

1	Introduction	3
2	Literature Review	4
2.1	South African Labour Market Context	5
2.2	Quantile Regression in the Context of Wage Equations	8
3	Economic Model	11
3.1	Conditional Expectation Representation of the Wage Offer	12
3.2	Conditional Quantile Regression Representation of the Wage Offer: 15	
4	Methodology	16
4.1	First Step Estimation: Probability of Employment	17
4.2	Quantile Regression Estimation: Approximating the Bias Term by a Polynomial Series Expansion	19
4.3	Heteroscedasticity Concerns	21
5	Data	21
5.1	Data Issues	22
5.2	Summary Statistics	25
6	Results	28
6.1	Employment Selection Equation Results	28
6.2	Earnings Equations: Conditional Quantile Regression Results . .	31
6.3	Discussion	45
7	Conclusion	50
	References	54
8	Appendix	57

1 Introduction

The South African Apartheid system officially fell in 1994 with the onset of a new democracy. The democratisation of the country brought to an end forty-six years of official racial discrimination of the (non-white) majority by a white led nationalist minority, and a promise of new opportunity for South Africa as a nation. In the labour market, such opportunity began with the implementation of constitutional changes to ensure equal opportunity and fair treatment to all labour market participants. Affirmative action and the Employment Equity Act of 1998 were introduced, and the importance of women in the South African labour force was acknowledged (Casale, 2004). It has been twenty years since the implementation of these policies at the judicial level, and a rich array of literature exists on the range of issues that underlie the workings of the South African labour market (Casale, 2004; Fallon and Lucas, 1998; Kingdon and Knight, 2000, 2004; Knight and McGrath, 1987). These include social issues such as inequality and the distribution of income that are closely linked to the labour market and its dynamics. Further, the history of patriarchy in conjunction with Apartheid implies that women (especially women of colour) are the most vulnerable labour market participants (Kingdon and Knight, 2004; Ntuli, 2007; Bhorat and Leibbrandt, 1999a). The role that women play in the South African labour force has been a focal question in many studies (Casale, 2004; Casale and Posel, 2002; Winter, 1999; Van der Westhuizen, Goga and Oosthuizen, 2007). A substantial increase in the labour force participation of women since the end of apartheid has been documented, and a feminisation of the labour force in post-apartheid South Africa has become apparent (Casale, 2004).

This paper thus focuses on the labour market outcomes of women in South Africa. The aim is to contribute to understanding the dynamics of the female labour market by providing a full characterization of the evolution of the female labour force in South Africa. This includes a full description of each stage of the labour market process, from employment to earnings, taking into account the labour force participation choice along the way. The emphasis is on earnings of women of differing race, where the earnings outcome is evaluated using quantile regression analysis of the conditional wage distribution.

More explicitly, the dynamics of the labour force and the outcomes for women

in the labour market are considered with respect to: (1) The labour force participation decisions of women, (2) the employment/unemployment labour market outcome and (3) realised earnings . Upon modeling entry into the labour force simultaneously with the employment/unemployment outcome the conditional distribution of earnings of females (across quantiles) in the post-apartheid era is evaluated.

The contribution made here uses recent developments in quantile regression as a tool for analyzing the conditional earnings distribution at different points of the distribution for women. The aim is to make statements about what factors are driving the variation at different points of the conditional earnings distribution of women in South Africa. Using quantile regression assists in explaining the inherent heterogeneity of earnings outcomes for women across different points of the distribution over different variables. This is important for understanding the dynamics of female labour force behaviour and outcomes, especially across variables like race, education, and skills level.

In order to achieve this aim, the next section presents a brief review of the multifaceted South African labour market, and a review on quantile regression in the context of earnings models. This is followed by the economic model which carefully lays out the theoretical underpinnings at each stage of the labour market process for women of all races in South Africa. This begins with modeling the employment outcome and ends with the conditional (quantile) earnings outcome. The methodology section discusses how each stage of the economic model is estimated, it is at this stage that the participation choice of women into the labour market is accounted for. The data used for the analysis is then introduced, after which the results are presented. There is a brief discussion of these results, and a conclusion which draws the paper to a close.

2 Literature Review

The literature on both, the South African labour market and quantile regression in the context of earnings models, is broad and multi-faceted. The literature on the South African labour market spans the issues of selection, employment, earnings and discrimination; while the literature on quantile regressions is technical and delves into the realm of econometric theory. For this reason the literature

review is split into two sections, one discussing the South African labour market context and the other discussing the literature on quantile regression as applied to earnings.

2.1 South African Labour Market Context

The effects of pre-labour market discrimination in South Africa makes accounting for the selection bias associated with the female labour force complicated (Heckman, 1979). Differential provision of education, training and access to public services, and the influences of these on the labour market outcomes of different race groups must be taken into consideration (Knight and McGrath, 1987; Moll, 1991; Case and Deaton, 1998). In conjunction with these socio-economic issues it is important to address the issue of how to define the labour force, and choose whom to omit or include in trying to create an adequate description of the labour market. The labour market in South Africa is not understudied and a rich array of literature exists on the issues that underlie it. A concise review follows highlighting the key results and themes in the literature concerning female employment, and more importantly, earnings function estimation in post-apartheid South Africa.

To begin, Kingdon and Knight (2000) address the issue concerning how the labour force should be defined. They investigate whether the non-searching unemployed state (broad unemployment rate) is distinct from the searching unemployed state (narrow unemployment rate). The non-searching unemployed are found to be more deprived than the searching unemployed, where search is hampered by cost of search, poverty, and distance from urban centers. The authors find no evidence of behavioural distinctions between employment states, implying that non-search is due to constraints faced by individuals as opposed to a “taste” or preference for unemployment. The use of the broad unemployment rate in the South African context is suggested. Most of the literature that follows conduct analyses based on the broad and narrow definitions of unemployment. This study uses the broad unemployment rate.

At each stage of the labour market description process – (1) labour force participation status, (2) the incidence of employment, and (3) the earnings outcome – the knock-on effects of the apartheid regime’s pre-labour market discrimination is visible. Labour force participation of females has been on the uptake in

the years following the collapse of the apartheid government (Casale and Posel, 2002). However, Casale (2004) observes that this increase translated into an increase in unemployment rates for the country. She finds that there has been some opportunity for advancement of females in the labour force, but that white women seem to have been the main beneficiaries. Ntuli and Wittenberg (2013) use decomposition analysis to explain the observed increase in labour force participation of black women during 1995–2004. They show that black women are more likely to enter the labour force if they have higher education and assume urban residency, but that labour force participation does not guarantee employment. Borat and Leibbrandt (1999b) find that, for black women, personal, household, and regional variables are significant determinants of labour force participation in South Africa.

Kingdon and Knight (2004) observe that unemployment rates for different groups reveal great disparity in the incidence of unemployment by race, age, education, region and importantly gender. This finding has been confirmed by Borat and Leibbrandt (1999a) and makes the question of understanding female labour force dynamics an interesting one.

As far as earnings are concerned, as early as 1980 the racial wage hierarchy in South Africa was being studied upon binary comparisons of the wage position of whites relative to other race groups (Knight and McGrath, 1987; Moll, 1991). Allanson, Atkins and Hinks (2002) made use of the first waves of nationally representative labour market related data to describe the evolution of the racial wage hierarchy over time. The authors find that there was a substantial wage gap (for males) between non-whites and whites in the mid 1990's, but that this gap had eroded when compared in gross terms to the wage gap prevalent in the 1980's. A comparison to the behaviour of female wages would be interesting to look at.

Casale (2004); Van der Westhuizen et al. (2007) and Winter (1999) found that as far as returns to employment (earnings) are concerned, that discrimination by gender and race are prevalent. African women who do find employment are found to earn considerably less than their white counterparts, with large differences between races, especially at lower skills levels. African women are thus found to be the most vulnerable participants in the labour force (Bhorat and Leibbrandt, 1999b; Van der Westhuizen et al., 2007). Understanding how

exactly the difference in earnings differs for women in South Africa across race and skills level is thus necessary.

Few of the papers written on labour force participation explicitly take into account the unique context of the South African labour market when conceptualising earnings models and the associated selection biases (Bhorat and Leibbrandt, 1999a). Some South African studies have attempted to model selection into the labour market by establishing such an interface between the estimated earnings equations and sample selection equations (Schultz and Mwabu, 1997; Fallon and Lucas, 1998; Winter, 1999; Hofmeyr, 1998). All of these studies consider selection into the labour market based on all potential labour market participants. Winter (1999), in her characterization of the gender bias inherent in earnings for women omits the sample selection term in her earnings function. Fallon and Lucas (1998) move from the employment decision straight to earnings, ignoring the choice to participate. Schultz and Mwabu (1997) attempt to correct for selection into the labour market, selection into employment and participation in the informal sector using one participation equation identified by agricultural variables. Their selection term is insignificant.

As a remedy, Bhorat and Leibbrandt (1999b) address the issue of the participation decision being separate from the employment outcome by building a three step participation, employment and earnings model. This model captures the transition between labour force participation and the selection into employment. Van der Westhuizen et al. (2007) extend the work done by Casale (2004) and Bhorat and Leibbrandt (1999b) by employing an extensive descriptive analysis then estimating a participation-employment-earnings model for females between 1995-2005. The authors find that the post-apartheid feminisation of the labour force has been driven by more women entering the labour force, with majority of women entering finding jobs that are unskilled and low-paid. The inability of the labour market to absorb new entrants resulted in an over-representation of young women, and African women among the unemployed. Bhorat and Leibbrandt (1999b) using 1995 South African labour market data, tell a similar story with black women being the group most affected by labour market discrimination.

Finally, the above studies (Bhorat and Leibbrandt (1999b), Van der Westhuizen et al. (2007), Allanson *et al.* (2002), *inter alia*) have investigated the

participation choice, the employment outcome, the earnings function and wage gaps exclusively at the means of their distributions. The standard earnings regression fits models based only on the conditional expectation of earnings on the covariates considered. Since the wage distribution is skewed to the left, with a long right tail, the conditional expectation of income estimated is a misrepresentative measure of location. These models implicitly standardize the size of the wage gap at every point of the conditional distribution and ignore the potential for there to be a distribution of wages that corresponds to the change in the conditional wage structure over time (Machado and Mata, 2005; Ntuli, 2007). To remedy this, the wage structure can be estimated by focusing on different points in the wage distribution using quantile regression, the justification for which lies in extreme value theory (Yu, Lu and Stander, 2003; Buchinsky, 1998a). This is discussed in the next section.

2.2 Quantile Regression in the Context of Wage Equations

Conditional quantile regression analysis is used to characterize the entire distribution of wages given a set of regressors. Since the objective function is a (weighted) sum of absolute deviations, it renders a more robust measure of location than simple mean regression (Buchinsky, 1998a). This is useful in the case of wages since it implies that the estimated coefficient vector is not sensitive to outlier observations. A short review of the literature on quantile regression as applied to wages, education and the labour market follows, with emphasis on the empirical findings.

First, Buchinsky (1994) in his seminal work, studies the general distribution of wages in the USA across different education and skills levels. He uses quantile regression to parsimoniously describe the conditional wage distribution. Buchinsky analyses wage inequality – defined as the difference between two conditional quantiles – via narrowly defined education-experience groups. He finds that when using a single quantile regression, it appears that at different quantiles the returns to experience and education are different, but that the changes over time have followed a similar pattern. By breaking skills into separate groups he finds that the returns to education at different quantiles do not all follow the same pattern of change for all skills groups. In his paper he estimates the model using a linear programming algorithm; he derives a new

estimator for the asymptotic covariance matrix as well as a new algorithm for estimating the censored quantile regression.

Relevant to the present context, Buchinsky (1998b) studies the dynamics of the female wage structure in the USA between 1968 and 1990. He analyses changes in the conditional log-wage quantiles across nine age-groups for women between the ages of 20 and 64, focusing on returns to education. He corrects for sample selection (Mroz, 1987) using a non-parametric correction procedure similar to the one suggested by Heckman (1979). In the first step participation equation he finds that traditional family and household duties played a less significant role in determining labour force participation status in the late 1980s than in the late 1960s, with the significant variables being number of children and family size. Buchinsky (1998b) rejects the parametric probit specification of the sample selection equation against the non-parametric alternative.

For the second stage conditional log-wage quantile regression estimation, Buchinsky (1998b) finds pronounced returns to college education among the upper quantiles. High school graduates exhibit a less systematic upward trend in their earnings over the period of concern. Buchinsky (1998b) notes that the conditional mean return for all levels of education is consistently below the conditional median, which once again makes a case for the use of quantile regression to analyse the distribution of wages over time (Yu et al., 2003).

Machado and Mata (2005) identify sources of increased wage inequality in Portugal between 1986 and 1995 by estimating marginal wage distributions consistent with a conditional distribution estimated by quantile regression. The authors decompose changes in the wage distribution over time and distinguish between changes in the characteristics of the working population and the returns to these characteristics. They find that increases in education levels contribute to wage inequality. The authors importantly note that the heterogeneity between types of workers (high and low skilled) is substantial. The quantile regression results point out that wages are more spread out in higher skills groups, and an increase in the level of schooling reduces the weight of the low spread group, contributing to increased wage dispersion and overall inequality. These results are in line with Buchinsky (1998a) and Fitzenberger and Kurz (1997). The authors discover that a higher proportion of more educated individuals is inequality increasing. Individual characteristics (age, education, etc) and re-

turns to these characteristics contributed in the same direction to increases in wage inequality.

In a South African context, Ntuli (2007) applies quantile regression decomposition analysis to the description of the conditional wage distribution. Her analysis focuses on the gender wage-gap, with emphasis on the vulnerable in the labour force. She applies a “Heckman like” selection correction and then obtains coefficients of gender wage discrimination for Africans at different points of the wage distribution. However, like Borat and Leibbrandt (1999b) she models the participation – employment – earnings paradigm by first independently estimating the inverse Mill’s ratio to correct for selection into the labour force for females. The problem with this approach will be discussed in the methodology section. Ntuli’s estimation of the correction terms are fully parametric. Since her analysis is centered around counter-factual decompositions, Ntuli does not delve into the dynamics of the female wage distribution or the findings from her selection equations. She concludes that a significant gender wage gap is prevalent in South Africa.

Schultz and Mwabu (1997) offer varying explanations for returns to education in South Africa by race, through the analysis of conditional quantiles of the wage function. The authors posit that education and ability are complementary and that the return to education (controlling for race) should be higher at higher quantiles of the wage distribution. They find that this is the case among highly educated white males using South African data for 1993. The authors also find that the wage returns to schooling for males were twice as high for blacks than for whites, and that returns to schooling were higher at more advanced levels of schooling. These latter results are consistent with the ones by Machado and Mata (2005) : increasing education, is expected to increase wage inequality *ceteris paribus*. Whether these findings hold for women is an interesting question and one that will be answered in the following sections.

A common thread in Buchinsky (1994, 1998b,a); Machado and Mata (2005); Ntuli (2007); Schultz and Mwabu (1997), *inter alia*, is the importance of the estimation of the variance-covariance matrix when applying quantile regression estimation. If the assumption of a constant conditional error variance between quantiles is not met, then the errors across quantiles may be heteroscedastic, implying a bias in the estimates of the quantile regression variances. Further,

the assumptions made on the quantile regression specification can complicate the asymptotic properties of the coefficient estimates in the presence of sample selection (Huber and Melly, 2011). These issues surrounding heteroscedasticity can be dealt with in a range of ways and will be discussed in the methodology section.

The model presented in the next section follows the papers by Ntuli (2007) and Buchinsky (1998a) closely, by accounting for the complex structure involved in selection into the South African labour market (Bhorat and Leibbrandt, 1999b). The model is designed to aid identification at each part of the participation – employment – earnings paradigm by making explicit linkages between the employment selection process and the earnings equation. It will be interesting to observe whether the results at each stage of the female labour market process are in accordance with the female labour market literature presented earlier (Casale, 2004; Winter, 1999; Van der Westhuizen et al., 2007), whether the quantile regression findings posit an increase in earnings inequality over time (Schultz and Mwabu, 1997; Machado and Mata, 2005; Buchinsky, 1998b,a); and whether these results are robust across quantiles, race and time.

3 Economic Model

The economic model introduced considers selection into the labour force in the context of quantile regression (Buchinsky, 1998b). The conditional quantiles of the earnings distribution explain how the percentiles (τ) of the earnings distribution change given a set of regressors. The conditional quantile of the observed wage depends on the bias term from the employment outcome equation. That is, the bias associated with observing a non-random sample of female workers (Heckman, 1979).

For the linear form of the conditional quantile regression wage offer equation, the bias term in its unknown form can not be corrected for using Heckman’s simple parametric correction for sample selection (Heckman, 1976; Buchinsky, 1998b). Buchinsky (1998b) estimates the correction term non-parametrically. In this paper, the correction term is estimated fully parametrically, then linearised using a polynomial expansion before being included in the wage offer equation to correct for the selectivity bias (Buchinsky, 1998b). The linearisa-

tion is intended to correct the functional form of the selection correction so that it is compatible with each respective quantile across each point of the conditional earnings distribution.

The economic model is introduced below, followed by the methodology. The parametric estimation technique used to obtain the selection correction term is outlined in the methodology.

3.1 Conditional Expectation Representation of the Wage Offer

The theoretical underpinning of the employment equation lies in “seek and hire theory” (Mohanty, 2010). The employer’s hiring decision determines whether a woman will be employed or not and is assumed to be a linear function of the characteristics of each woman x_1 and an idiosyncratic error term ε_1 (the i subscript has been dropped for notational ease):

$$y_1 = x_1' \beta_1 + \varepsilon_1. \tag{1}$$

The vector of individual characteristics x_1 consists of each woman’s age, education, marital status, household size, the number of children in the household and a range of demographic variables. The idiosyncratic error term is normally distributed; $\varepsilon_1 \sim N(0, 1)$. It is assumed that $E(\varepsilon_1 | x_1) = 0$.

The binary outcome employment equation can be expressed as the employment index:

$$I_1 = I(x_1' \beta_1 + \varepsilon_1 > 0)$$

where $I(\cdot)$ is the usual indicator function.

The wage offer equation takes the form of the conventional latent index

y_2^* (Mincer and Polachek, 1974). It is assumed to depend linearly on a set of labour market and individual characteristics of each woman x_2 (industry, occupation, hours worked, among others) and has a normal zero mean idiosyncratic error ε_2 :

$$\begin{aligned} y_2^* &= x_2' \beta_2 + \varepsilon_2, \\ \varepsilon_2 &\sim N(0, 1). \end{aligned} \tag{2}$$

The wage outcome is observed if the employer hires the woman ($I_1 = 1$) at a wage y_2^* . The observed wage y_2 outcome is given by

$$y_2 = y_2^* \times I_1. \tag{3}$$

That is, the wage offer is only observed if the employer hires the woman. The condition $I_1 = 1$ is sufficient to observe the wage outcome.

The observed wage equation (3) can be written in terms of the regular expectation based formulation:

$$\begin{aligned} E(y_2 | x_2) &= E(y_2^* | x_2, I_1 = 1) \\ &= E(x_2' \beta_2 + \varepsilon_2 | x_2, -x_1' \beta_1 + \varepsilon_1 > 0) \\ &= x_2' \beta_2 + E(\varepsilon_2 | \varepsilon_1 > -x_1' \beta_1). \end{aligned} \tag{4}$$

It is clear that the conditional expectation of the error of the observed wage equation $E(\varepsilon_2 | \cdot)$ is dependent on the unobservables of the employment offer (ε_1). The errors of the employment offer (ε_1) and the latent earnings equation

(ε_2) are jointly normally distributed,

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \sim N \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Heckman's (1979) famous result constitutes correcting the wage offer equation (3) for the selection bias associated with the employment outcome where the employment outcome is characterised by a single index model (I_1). For the classic Heckman correction (Heckman, 1979) $E(\varepsilon_2 | \cdot)$ in (4) takes the form:

$$\begin{aligned} E(\varepsilon_2 | \varepsilon_1 > -x_1' \beta_1) &= E(\sigma_{12}\varepsilon_1 + v | \varepsilon_1 > -x_1' \beta_1) \\ &= \sigma_{12} E(\varepsilon_1 | \varepsilon_1 > -x_1' \beta_1), \end{aligned}$$

where $\varepsilon_2 = \sigma_{12}\varepsilon_1 + v$ and $v \sim N(0, 1)$. The coefficient σ_{12} captures the correlation between ε_1 and ε_2 . Heckman's (1979) simplification of $E(\varepsilon_2 | \cdot)$ under the joint normality assumption of ε_1 and ε_2 implies that the conditional error term of the employment index $E(\varepsilon_1 | \cdot)$ can be expressed in terms of the inverse Mill's ratio evaluated at $x_1' \beta_1$:

$$\begin{aligned} E(\varepsilon_2 | \varepsilon_1 > -x_1' \beta_1) &= \sigma_{12} \frac{\phi(x_1' \beta_1)}{\Phi(x_1' \beta_1)} \\ &= \sigma_{12} \lambda(x_1' \beta_1) \end{aligned} \tag{5}$$

The expression in (5) is the Heckman correction accounting for the selection bias in the wage equation from observing a non-random sample of employed women (Mroz, 1987).

In the case of the observed wage conditional quantile regression, the expectation based sample selection correction term in (5) can not be used to correct for

selection bias. The expression in (5) is evaluated at the mean and the correction term for the mean is not suitable at different points of the conditional earnings distribution. As a result an appropriate sample selection correction term must be estimated for each respective quantile. The estimation of the quantile specific sample selection correction is outlined in the methodology section. First, however, the conditional quantile regression representation of the wage offer is presented.

3.2 Conditional Quantile Regression Representation of the Wage Offer:

The latent wage offer equation can be written in quantile regression form following Koenker and Bassett (1978) as:

$$y_2^* = x_2' \beta_\tau + \varepsilon_\tau, \quad (6)$$

where $Q_\tau(\varepsilon_\tau | x_2) = 0$ assuming that $Q_\tau(y_2^* | x_2') = x_2' \beta_\tau$ and $0 < \tau < 1$. It is clear that $Q_\tau(\cdot)$ is the τ th quantile of y_2^* .

A woman will be employed conditional on being offered a job by the employer ($I_1 = 1$). This implies that the conditional quantile of the observed wage in the presence of the selection bias from the employment equation is:

$$\begin{aligned} Q_\tau(y_2 | x_2' \beta_\tau) &= Q_\tau(y_2^* | x_2, I_1 = 1) \\ &= Q_\tau(x_2' \beta_\tau + \varepsilon_2 | x_2, I_1 = 1) \\ &= x_2' \beta_\tau + Q_\tau(\varepsilon_\tau | I_1 = 1), \end{aligned} \quad (7)$$

and in general: $Q_\tau(\varepsilon_\tau | I_1 = 1) \neq 0$. The conditional quantile of the observed wage is equal to the conditional quantile of the latent wage offer equation given that the employment index takes on the value one.

The similarity between the conditional quantile representation of the observed wage (7) and the conditional expectation representation of the observed wage (4) in the presence of sample selection is apparent.

Assume that $Q_\tau(\varepsilon_\tau | I_1 = 1)$ is only a function of the index $x_1' \beta_1$, which holds under homoscedasticity. The observed wage equation (7) then becomes:

$$y_2 = x_2' \beta_\tau + h_\tau(x_1' \beta_1) + \tilde{\varepsilon}_\tau, \quad (8)$$

where $Q_\tau(\tilde{\varepsilon}_\tau) = 0$. The function $h_\tau(\cdot)$ is a quantile specific selection correction for the employment selection process (I_1) and depends on $x_1' \beta_1$. The function $h_\tau(x_1' \beta_1)$ is defined as:

$$\begin{aligned} h_\tau(x_1' \beta_1) &= Q_\tau(\varepsilon_\tau | I_1 = 1) \\ &= Q_\tau(\varepsilon_\tau | \varepsilon_1 > -x_1' \beta_1), \end{aligned} \quad (9)$$

and the exact functional form of $h_\tau(\cdot)$ is unknown. To make the selection correction $h_\tau(\cdot)$ compatible with each quantile regression performed at the τ th quantile (8), $h_\tau(\cdot)$ is approximated using a polynomial series expansion that is fully parametric. This expansion is outlined in the methodology section below.

4 Methodology

The economic model presented above illustrates that first the employment choice is modeled from which the selection effect $h_\tau(\cdot)$ is obtained. The quantile regressions are modeled second using this selection effect to correct for the associated bias.

Given the model, the estimation procedure is conducted in two steps. The first step estimates the probability of employment as outlined in the first part of the model, but takes special consideration of the South African labour market

context and the choice of women to participate in the labour force Borhat and Leibbrandt (1999b). The second step uses the estimates from the probability of employment to obtain an estimate of $h_\tau(\cdot)$ after which the conditional earnings quantile regressions are estimated.

4.1 First Step Estimation: Probability of Employment

As was discussed in the literature review, the nature of the South African labour market is complex. The labour market consists of a sequential dependent three step process made up of labour force participation, employment and earnings- Borhat and Leibbrandt (1999b). The employment of a woman in South Africa according to Borhat and Leibbrandt (1999b) depends on her choice to participate before the employment offer is made by the employer. After the participation choice is made, the employment outcome is observed for a non-random sample of women and it is this selection effect from the exogenous employment offer that is used to correct the earnings equation (Heckman, 1979; Meng and Schmidt, 1985).

In the case of the single employment index (I_1) outlined in Section 3.1 the probability of a woman working takes the form:

$$\begin{aligned} P(I_1 = 1 | x_1, \beta_1) &= P(\varepsilon_1 > -x_1' \beta_1) \\ &= \Phi(x_1' \beta_1). \end{aligned} \tag{10}$$

This corresponds with a probit specification. However because we know that participation and employment are closely related processes in South Africa (the first being a choice, the second being exogenously determined), it will be inconsistent to estimate the selection bias by the probit specification outlined in (10) (Meng and Schmidt, 1985). The single outcome probit ignores the choice a woman makes to participate in the labour force (Bhorat and Leibbrandt, 1999b).

Instead the probability of employment can be modeled first by introducing a temporary index, I_P , a linear participation index dependent on a set of re-

gressors x_p that determine the labour force participation decision of a woman. This index takes the form:

$$I_p = (x_p' \beta_p + \varepsilon_p > 0), \quad (11)$$

where $\varepsilon_p \sim N(0, 1)$.

The probability of a woman working is then dependent on her participation choice (I_P) and the employment outcome (I_1) from the economic model:

$$\begin{aligned} P(I_P = 1, I_1 = 1) &= P(\varepsilon_p > -x_p' \beta_p, \varepsilon_1 > -x_1' \beta_1) \\ &= \Phi(x_p' \beta_p, x_1' \beta_1, \rho). \end{aligned} \quad (12)$$

The correlation between the unobservables of the employment index (I_1) and the participation index (I_P) is captured by ρ . The above specification is a bivariate probit model with partial observability. There is partial observability because if a woman chooses not to participate in the labour force ($I_P = 0$) the employment outcome (I_1) is censored (not observed). The specification in (12) is used to obtain the estimates of $x_1' \beta_1$ that are needed for the sample selection correction function $h_\tau(\cdot)$ outlined in Section 3.2. The bivariate probit specification that accounts for the participation decision ensures that the estimates of $x_1' \beta_1$ are consistent (Meng and Schmidt, 1985). A key assumption used here is that ρ is uncorrelated with the unobservables of the wage offer equation (ε_1).

Further, in estimating the bivariate probit model, and the quantile regressions that follow special attention is paid to exclusion restrictions. At each stage of the selection process variables are selected and then carefully omitted to aid identification.

4.2 Quantile Regression Estimation: Approximating the Bias Term by a Polynomial Series Expansion

Once the estimates for $x_1' \beta_1$ are obtained from the bivariate probit model outlined above an estimate of the sample selection correction function $h_\tau(\cdot)$ can be obtained.

The observed wage equation in quantile regression from is given by (8). The exact functional form of the sample selection correction $h_\tau(\cdot)$ is unknown. In the case of the conditional expectation model the parametric normality assumption on the errors of the employment equation (1) and the wage outcome (2) imply that sample selection bias can be corrected for using the inverse Mill's ratio (5). This normality assumption does not hold in the case of quantile regression, because the conditional outcome at the τ th quantile is not evaluated upon the conditional expectation¹. Since the inverse Mill's ratio is a correction derived at this conditional expectation it would be incorrect to use this (mean based) correction as a regressor across quantiles to correct for sample selection bias. It is, however, used as a starting point to obtain the correct functional form of the selection correction bias for each of the τ quantiles.

Following Heckman's (1979) result derived in equation (5) the choice of polynomial expansion² of the correction term $h_\tau(x_1' \beta_1)$ outlined in (8) takes the form:

$$\begin{aligned} h_\tau(\lambda(x_1' \beta_1)) &= \text{Poly}[\delta_\tau' (\lambda(x_1' \beta_1))] \\ &= \sum_{j=1}^J a_j [\delta_\tau' \widehat{\lambda(x_1' \beta_1)}]^{j-1}. \end{aligned} \quad (13)$$

The polynomial expansion corrects the functional form of the parametric inverse Mill's ratio estimate so that it can be included as a regressor in each respective quantile regression to correct for the sample selection bias associated

¹The quantile regression objective function is estimated by minimal distance and is a weighted sum of absolute deviations (Buchinsky, 1998a)

²Note that the mathematical operator Poly[.] used below is a non-standard operator defined for the sole purpose of this paper and for simplicity of notation and exposition of the estimated polynomial correction term $h_\tau(\cdot)$.

with women in the labour force.

Other polynomial expansion forms such as:

$$\text{Poly}(x_1' \beta_1) = \sum_{j=1}^J a_j \left(\frac{\hat{f}(x_1' \beta_1)}{\hat{F}(x_1' \beta_1)} \right)^{j-1}, \quad (14)$$

where $\hat{f}(\cdot)$ and $\hat{F}(\cdot)$ are the nonparametric densities and CDF's of the error terms from the employment selection index could have been used. Nevertheless, the polynomial expansion of $h_\tau(\cdot)$ using the Heckman correction when studying the female labour force (Mroz, 1987) is chosen.

Using the polynomial expansion of the selection correction term (13) in the observed wage equation (8) implies that the observed wage equation at the τ th quantile becomes:

$$y_2 = x_2' \beta_\tau + \sum_{j=1}^J a_j [\delta_\tau' \lambda(x_1' \hat{\beta}_1)]^{j-1} + \tilde{\varepsilon}_\tau. \quad (15)$$

The expression $\lambda(x_1' \hat{\beta}_1)$ is the estimate of the inverse Mill's ratio obtained from the bivariate probit model (12). The polynomial expansion in (13) is included here as a regressor. The j products of a_j and δ_τ can not be disentangled, and are hence jointly estimated as j single coefficients in each of the quantile regression models.

It is worth mentioning at this point that the conventional method used to obtain standard errors in the case of quantile regression (Koenker and Bassett, 1978) is unproblematic in the case of homoscedasticity. However, if the homoscedasticity assumption fails then the method of Koenker and Bassett (1978) will understate the standard errors (Rogers, 1992). Exploring heteroscedastic errors is one of the major benefits of using quantile regression in the context of the earnings equation. To remedy this, the method of Gould (1992) is used to obtain estimates of the quantile regression standard errors by using bootstrap resampling.

4.3 Heteroscedasticity Concerns

If the variance of the error terms are not fully independent of the covariates in the employment selection equation, then the coefficient estimates in the employment selection equation will be inconsistent. This inconsistency may then lead to quantile regression coefficient estimates that are inconsistent (Melly and Huber, 2007). In order to avoid quantile regression estimates that are inconsistent it is important to test the selection equations for heteroscedasticity³, and to adjust for initial heteroscedasticity before estimating the respective earnings quantile regressions. To do so univariate heteroscedastic probits are estimated for the employment index⁴, evidence of heteroscedasticity in the selection equation is found and corrected for throughout the rest of the study. The next section presents a discussion of the data used. After which the results are presented following the sequence in the methodology discussed above.

5 Data

The dataset used for this study is the Post Apartheid Labour Market Series⁵ (PALMS). The PALMS dataset is a stacked cross-sectional dataset consisting of data from thirty-nine surveys conducted by Statistics South Africa (Stats S.A) between 1994 and 2012. These include the October Household Surveys (OHS) from 1994-1999, the bi-annual Labour Force Surveys (LFS) from 2000-2007 and the Quarterly Labour Force Surveys (QLFS) that were conducted from 2008 onwards. The data is presented at the individual level, but no attempt has been made to link individuals or households across waves, despite there being a partial panel element to the earlier rounds of the LFS.

The OHS's are independent cross-sectional annual surveys undertaken by Stats S.A from 1993-1999. The OHS was designed with the particular aim of making data available on the population, particulars of dwellings, the labour market and public services in South Africa. The LFS, on the other hand, is semi-annual rotating panel household survey, specifically designed to measure

³Differences in the variance of unobservables of women at different stages of the labour force process may be correlated with the observed variables in the selection equations.

⁴These results are not presented here, but all results that follow correct for heteroscedasticity

⁵PALMS v2.1 (1994-2012)

the dynamics of employment and unemployment in South Africa. For the purpose of this study, to keep the data at an annual level of analysis, the second installment of each LFS (LFS_2) is used. This second installment (LFS_2) was conducted each September from 2000-2007. Finally, the QLFS are an extension to the LFS, the objective of which is to collect quarterly information about persons in the labour market (employed, unemployed and not economically active). For consistency the third wave of the LFS is used, to keep the data at an annual level and to ensure comparability with the OHS and the LFS_2 waves.

PALMS Version 2.1 contains variables which are defined across all thirty-nine waves of the OHS, LFS and QLFS⁶. For these data sets, samples were explicitly stratified by province, magisterial district, urban/rural status and population group. All subsequent analyses take into consideration sampling weights to ensure that the estimated results are population representative. The PALMS dataset contains weights created using entropy estimation, which results in consistent demographic and geographic trends over time for both the individual and the household. These weights are applied at the individual level.

For this study the PALMS dataset is pooled⁷ into five relatively homogenous sub-periods; 1994-1995, 1997-1998, 2000-2002, 2005-2007 and 2010-2011. Each of these periods is treated as a cross-section and no panel structure is assumed. The first two periods consist of data from the OHS. The periods 2000-2002 and 2005-2007 consist of data from the second installment of the LFS (LFS_2), and the period 2010-2011 consists of the third quarter of data from the QLFS.

5.1 Data Issues

The sub-sample of interest consists of working age females (16-65) from the black, white and coloured race groups. The Indian/Asian population group was omitted from the analysis from the outset due to small sample sizes across periods⁸. A woman is defined as employed if she worked in the week prior to the interview, or did not, but had a paid job to return to and reported a positive income. Working women are considered from across all occupations and sectors

⁶Special dwellings such as prisons, hospitals, boarding houses, hotels, guest houses, schools and churches were excluded from the entire PALMS sample.

⁷The reason the data are pooled is that following the modeling procedure outlined in Section 4, at each stage of the estimation process observations are dropped. In order to maintain an adequate sample size for estimation pooling across periods was necessary.

⁸Less than 5% of total observations in the data belonged to this group.

of the economy, bar the informal sector.

A woman is defined as officially unemployed⁹ if she has not held a paid job for at least seven days prior to the interview but has taken active steps to find work in the four weeks preceding the interview, and would be willing and able to work within two weeks after the interview. A woman is defined as a discouraged worker if she has not actively searched for a job in the last four weeks, but is willing and able to take up employment within two weeks after the interview. The broad definition of the labour force includes discouraged workers among the unemployed. For this study the broad definition of unemployment is used following Kingdon and Knight (2000).

A woman is defined as not economically active (out of the labour force) if she is not able to work. This includes the case of full time home-makers, scholars and students, those who are retired, and those who are unwilling or unable to work.

Due to data constraints it is not possible to identify the early retirees or the disabled, implying that only scholars and students have been removed from the not economically active population. The not economically active thus includes early retirees and the disabled¹⁰. It should be borne in mind that the selection effect later estimated to account for women who are involved in child rearing activities is conflated by the presence of early retirees and the disabled¹¹.

A key data issue relates to the way income has been reported over time. The initial waves of PALMS which correspond to the OHS, report data on both, income from a main job and income from self-employment. The later waves of PALMS, which correspond to the LFS and QLFS, contain data only on income from an individuals main job (which could in turn be self-employment). For consistency, the dependent variable (log of monthly wages) used in the conditional earnings quantile regressions has been defined using only income from an individual's main job, converted to a monthly period. The earnings variable is

⁹According to the narrow definition of the labour force.

¹⁰Assuming disability occurs randomly across the population implies that the disabled among the not economically active can be ignored.

¹¹The inclusion of retirees and disabled individuals is still expected to lead to a bias in the estimate of the selection effect.

in real terms¹².

As is the case with collecting data on earnings, some individuals reported an earnings amount and others provided bracketed responses in each time period. The extent of bracketed responses varies between surveys (the OHS, LFS and QLFS). Ignoring bracketed responses, and taking into account only the reported earnings amounts is likely to misrepresent the earnings distribution. This is because women with higher incomes are less likely to provide an exact earning amount and more likely to respond by choosing an income bracket. Further, the quality of the responses to income questions varied greatly between datasets with the 2011 income data (from the QLFS) being of notably poor quality.

To remedy the issue surrounding differences in quality of reported incomes, and to adjust for bracketed responses multiply imputation analysis is used. The real earnings variable is the one that was imputed to account for bracketed responses. Income values were not imputed for women who reported zero incomes. The imputation method used was predictive mean matching. The (weighted) kernel densities of real-earnings and imputed real-earnings are presented for the periods 1994-1995, 2000-2002 and 2010-2011 in the Appendix (Figure A.1). The kernel densities illustrate that for 1994-1995 and 2000-2002 there seem to have been many bracket responses among the upper ends of the income distribution, for this reason the imputed real-earnings density is placed slightly to the right of the ordinary real-earnings density. For the period 2010-2011, the imputed density and the real-earnings density coincide. This illustrates that for this period, there were barely any bracket responses in the data. Imputed real-earnings is used as the outcome variable in the rest of this study¹³.

Another problem concerning the data is that PALMS does not include data on actual work experience. This could be computed using potential experience (age - education - 6) as a proxy, but this does not make for a good proxy for womens' actual work experience because of the decisions associated with home-making, child-rearing, and large scale unemployment. The lack of a control for actual work experience should be noted from the outset.

¹²All nominal earnings has been deflated to the base year 2000 using the Consumer Price Index

¹³This corresponds to multiply imputed real income variable 1 (imputed_real_v1) in PALMS.

The final data limitation is that there is no identifier linking each woman in the dataset to her own children. It is possible to identify whether there are children present in a woman's household, but not how many of these children belong to a woman biologically.

All variables (described in detail below) are defined for all five sub-periods, with the exception of the household characteristics variables, which have been excluded from the original data-sets for the periods 2005-2007 and 2010-2011.

5.2 Summary Statistics

Before analysing the empirical results, the summary statistics of the explanatory variables used at each stage of the estimation procedure are briefly discussed. Tables A.1, A.2, A.3, A.4 and A.5 in the Appendix present the means and standard deviations of these variables. The variables are split into different sub-groups; industry, occupation, province, individual characteristics and household characteristics. The summary statistics are presented for each respective race group (black, coloured and white) for each of the five periods considered. While the full tables of summary statistics are presented in the Appendix, the discussion below highlights some of the key themes presented in these statistics.

Following the discussion in the literature review, the skills levels that women possess are important determinants of earnings. Sector of employment is strongly correlated with skills level. Looking at sector of employment over time, it can be seen that the manufacturing, trade and service sectors were the densest ones for women in all three race groups. For instance in the period 1994-1995, 28% of black, 27% of coloured and 37% of white women were employed in the services sector. Over time participation in the finance sector was more prominent for white women, but the proportion of non-white women participating in the finance sector grew steadily between the mid 1990's and the late 2000's. Coloured women have participated quite substantially in the agricultural sector over time (15% of coloured women worked in agriculture in 1994-1995, 11% worked in agriculture in 2000-2002 and 6% worked in agriculture in 2010-2011). For black females, participation in the agricultural sector increased between 1994-1995 and 2000-2002, from 9% to 14%, and declined again by 2010-2011 to 4%. The mining, construction and utilities sectors account for between zero and three

percent of the employment of women in each race group over all five periods of consideration. These tend to be male dominated sectors historically.

Occupation is also tightly related to skills level. For occupation, it is clear that the higher and medium skilled occupations (professional, technician and clerks) exhibit a higher participation rate for white women, in line with the *a priori* expectation regarding the pre-labour market discrimination associated with apartheid. Over time white women increasingly took up managerial positions, increasing from 7% in 1994-1995 to 18% by 2010-2011 with slight fluctuations in between. It is worth noting that a moderate proportion of non-white women are employed in medium skilled occupations (technicians and clerks), with coloured women taking up an increasing proportion of positions as clerks between the periods 1994-1995 (16%) and 2010-2011 (22%). For black and coloured women the low-skilled occupations, elementary and domestic work, are the most common forms of occupation. For these non-white women there is not much evidence of a change in the proportion of women working as domestic workers and elementary workers between 1997-1998 and 2010-2011. This reflects low inter-occupation mobility of women in lower skilled occupations, and emphasises potential skill disparities by race.

With respect to provincial location, the rural provinces of Limpopo, the North-West and the Eastern-Cape were collectively home to 44% of black women in the period in the period 1994-1995, and this proportion decreased to 37% by 2010-2011. This reflects that over time black women may be migrating out of rural provinces. Overall, the distribution of each race across provinces still reflects the large scale geographical segregation by race that is prevalent in South Africa, between each period there have been slight fluctuations around the proportion of each race group living in each province but the patterns have remained relatively constant over time. For instance, 40% of white women and 14% of black women lived in the Gauteng in 1994, while 39% of white women and 20% black women lived in Gauteng by 2010-2011.

The average household size for black women is higher than that of coloured and white women over all five periods of consideration. The average household size is roughly constant for each race group over time, with the average household size of black women being just over 6 individuals, for coloured women it is just about 6 individuals, and for white women the mean household size is about

3 individuals per household. On average it is also clear that over all five periods white women are older at a mean age of 41 in the period 2010-2011. The lower mean ages of black and coloured women (34 and 36 in 2010-2011) reflects both higher fertility and higher mortality among these groups.

Marital status is an important variable of consideration, especially in relation to the labour force participation decision of women. As far as marital status is concerned, in the period 1994-1995 26% of black women, 32% of coloured women and 51% of white women were married or cohabiting. These proportions have remained relatively constant over time, with the number of black women who are married decreasing to 22% by the period 2010-2011. An insight at this point is that marital status does not seem to be positively correlated to household size.

Years of education¹⁴ will be an important explanatory variable in the earnings equation. Immediately, it is noticeable that white women have a consistently higher mean level of education over all time periods when compared to blacks and coloureds. The mean level of education of black women has steadily increased since the end of apartheid. It was reported at 5.15 (grade 5) years in 1994-1995, increased to 5.56 years by 2000-2002 and increased further to 6.59 (grade 7) years by 2010-2011. The same is true of coloured women who reported average years of education at 5.34 (grade 5) in 1994-1995 which increased to 7.34 (grade 9) by 2010-2011.

Finally, as far as access to amenities is concerned, household characteristics and urban/rural status are reported for the periods 1994-1995, 1997-1998 and 2000-2002. Here the legacy of apartheid regarding the discrepancy in service delivery across race is apparent. In the period 1994-1995, 65% of black women lived in rural areas whereas only 17% of coloured women and 8% of white women lived in rural areas. In this period, 98% of white households had piped water, whereas 75% of coloured women reported having piped water and only 26% of black women had access to piped water. All white homes used a flush toilet where only 40% of black homes and 83% of coloured homes used a flush toilet. Interestingly, 57% of black women lived in a brick house, 76% of coloured

¹⁴Years of education is a variable that takes into account the years of schooling of a woman. The range of the variable is 0-16. A value of 13 corresponds to matriculation plus a diploma. A value of 14 corresponds to matric and a technical certificate of a two year duration. A value of 15 corresponds to a bachelors degree. A value of 16 corresponds to a master's or doctoral degree.

women lived in a brick house and 88% of white women lived in a brick house. No white or coloured women lived in a traditional dwelling where 26% of black individuals lived in a traditional home. From these statistics important residential differences as well as the gap in service delivery and access to services across race groups at the end of apartheid is visible (Leibbrandt, Woolard, Finn and Argent, 2010). The knock-on effects of these on the labour market outcomes of women in the South African labour force are considered in the following section.

6 Results

6.1 Employment Selection Equation Results

The employment selection equation results consist of the results from the bi-probit selection equation outlined in (12)¹⁵. The dependent variable of the main outcome equation is whether or not a woman is employed according to the broad definition of employment (discouraged workers are counted as unemployed), this outcome is estimated simultaneously with the labour force participation choice of a woman (I_p) under the joint normality assumption of errors.

The independent variables used for the employment outcome include: race, province, age, age-squared, a marital dummy, years of education, number of children in the household under the age of six and number of children in the household between the ages of seven and fifteen. The participation choice of a woman is simultaneously estimated using the same independent variables as those used in the employment equation but for the first three periods (1994-1995; 1997-1998; 2000-2002) additional regressors capturing household characteristics and urban/rural status are included. These variables are related with socioeconomic standing and may influence the labour force participation decision of a woman. For the periods 2005-2007, 2010-2011 these household characteristics variables are not included in the data and are hence omitted.

The results of the bivariate probit model (Meng and Schmidt, 1985) are presented in Table 1¹⁶. The coefficients of the variables in the employment part

¹⁵As per the discussion in the Methodology section, the estimates of the simple employment probit (10) render inconsistent estimates of the coefficient estimates $(x_1' \beta_1)$. Nevertheless, the simple probit estimates are presented in the Appendix (Table A.7).

¹⁶Selected controls are presented here. The full table with all controls is presented in the Appendix (Table A.6)

Table 1: Bi-Probit Results: Broad Definition of Employment

	1994-1995	1997-1998	2000-2002	2005-2007	2010-2011
Dependent Variable: Employment Outcome					
Household Size	-0.097*** (0.011)	-0.022*** (0.004)	-0.059*** (0.003)	-0.051*** (0.004)	-0.025*** (0.005)
Num Child (0-6)	0.078** (0.028)	-0.029*** (0.008)	-0.007** (0.003)	-0.017*** (0.003)	-0.011*** (0.003)
Num Child (7-15)	0.048* (0.023)	-0.023** (0.007)	-0.001 (0.002)	-0.001 (0.002)	-0.017*** (0.003)
Age	0.127*** (0.020)	0.209*** (0.005)	0.219*** (0.004)	0.199*** (0.004)	0.276*** (0.004)
Age-Squared	-0.001* (0.000)	-0.003*** (0.000)	-0.003*** (0.000)	-0.002*** (0.000)	-0.003*** (0.000)
Years of Educ	0.028*** (0.008)	0.056*** (0.002)	0.047*** (0.002)	0.058*** (0.002)	0.082*** (0.002)
Married	-0.011 (0.060)	-0.255*** (0.016)	-0.228*** (0.016)	-0.175*** (0.016)	-0.150*** (0.016)
Black	-2.062*** (0.230)	-0.314*** (0.024)	-0.201*** (0.024)	-0.090** (0.033)	-0.185*** (0.027)
Coloured	-1.119*** (0.227)	0.082** (0.031)	0.038 (0.033)	0.033 (0.041)	-0.024 (0.035)
Northern Cape	-0.514*** (0.126)	-0.204*** (0.036)	-0.171*** (0.036)	-0.233*** (0.033)	-0.115*** (0.034)
North-West	-0.113 (0.084)	-0.138*** (0.027)	-0.229*** (0.027)	-0.261*** (0.030)	-0.261*** (0.029)
Limpopo	-0.279** (0.100)	-0.201*** (0.029)	-0.140*** (0.029)	-0.352*** (0.029)	-0.235*** (0.027)
Dependent Variable: Labour Force Participation					
Household Size	0.005 (0.010)	0.006 (0.004)	-0.017*** (0.003)	-0.018*** (0.004)	-0.003 (0.004)
Num Child (0-6)	-0.127*** (0.022)	-0.053*** (0.008)	-0.011*** (0.003)	-0.015*** (0.003)	-0.011*** (0.003)
Num Child (7-15)	-0.038* (0.017)	-0.032*** (0.007)	-0.003 (0.002)	-0.003 (0.002)	-0.016*** (0.002)
Age	0.079*** (0.011)	0.173*** (0.004)	0.149*** (0.004)	0.137*** (0.005)	0.283*** (0.004)
Age-Squared	-0.001*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.003*** (0.000)
Years of Educ	0.079*** (0.004)	0.059*** (0.002)	0.062*** (0.002)	0.072*** (0.002)	0.087*** (0.002)
Married	-0.378*** (0.033)	-0.386*** (0.016)	-0.353*** (0.018)	-0.252*** (0.017)	-0.300*** (0.017)
Black	0.646*** (0.058)	0.290*** (0.027)	0.329*** (0.030)	0.395*** (0.038)	0.141*** (0.028)
Coloured	0.308*** (0.047)	0.232*** (0.033)	0.234*** (0.037)	0.256*** (0.046)	0.098** (0.037)
Piped Water	0.018 (0.046)	0.009 (0.019)	0.110*** (0.021)		
Public Tap	-0.404*** (0.063)	0.057* (0.025)	0.092** (0.032)		
Borehole	-0.257*** (0.065)	0.001 (0.035)	0.059 (0.042)		
Natural Source	-0.515*** (0.062)	-0.034 (0.029)	0.015 (0.035)		
Traditional Dwelling	-0.287*** (0.052)	0.070** (0.022)	-0.032 (0.027)		
Flat/Cluster	0.054 (0.043)	0.038 (0.023)	0.095** (0.032)		
Informal	-0.042 (0.070)	0.126*** (0.024)	0.068* (0.028)		
Chemical Toilet	-0.205 (0.152)	-0.043 (0.074)	0.477** (0.182)		
Pit Toilet	0.027 (0.055)	0.006 (0.019)	-0.002 (0.029)		
Bucket	-0.113 (0.061)	-0.024 (0.030)	-0.030 (0.038)		
Rural	0.082* (0.041)	-0.103*** (0.019)	-0.111*** (0.025)		
Rho	0.081	0.994	0.980	0.988	0.994
N	29395	49519	48812	87890	55818

Standard errors in parentheses. All results are estimated using sampling weights.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

of the model are of interest here. First, the tetrachoric correlation coefficients (ρ) are around .99 for the periods 1997-1998, 2000-2002, 2005-2007 and 2010-2011. This illustrates that the participation choice and the hiring decision of the employer are dependent processes, and that the simple probit estimation of the employment outcome (Appendix Table A.7) presents inconsistent estimates. The correlation coefficient for the first period considered (1994-1995) is only 8% indicating that the two processes of participation and employment are uncorrelated, this is unlikely and brings into question the quality of data in this period.

The coefficients from the employment equation illustrate that household size has a negative and significant effect on the probability of employment of a white woman. From the second period onwards, the number of children below the age of six in the household have a negative effect on the probability of employment of white women holding all else constant.

As far as race is concerned, Table 1 shows that for all periods considered black women have a lower probability of employment than white women *ceteris paribus*. Coloured women display a lower and significant probability of employment than white women for the period 1994-1995. The years of education and age variables have a positive and significant effect on the probability of employment.

Relative to white women, black and coloured woman have a higher probability of labour force participation *ceteris paribus*. This illustrates the racial differences within the labour force in line with the findings of Casale (2004). It is important to note at this point that the findings from the bivariate probit illustrate that although women of colour face a higher probability of participation this does not ensure positive employment probabilities. This is precisely what was found by Ntuli and Wittenberg (2013).

The effects of other covariates on the probability of labour force participation are similar to the effect of the covariates on the probability of employment. Household size, the number of children per household and marital status all have negative and significant effects on the probability of labour force participation, holding all else constant. Years of education and age have a positive and significant effect on the probability of participation. The direction of these effects are the same for the probability of employment.

Upon comparison of the results in Table 1 presented above to the results of the ordinary (inconsistent) employment probit presented in Appendix Table A.7 it is clear that the impact on the covariates in the employment equation on the probability of employment (simple probit) are the same as in the bivariate probit model. This finding is reassuring.

Overall the results from the bivariate employment selection equation clearly illustrate that in the South African context the labour force participation choice and the employment offer are correlated processes. The important point is that the zero earners evaluated in this part of the model are an integral part of the labour market in South Africa and play an important role in the employment outcome, which in turn affects the earnings part of the model. The selection effect from the employment part of the model is obtained using the method outlined in Section 4.2. The coefficients from the employment part of the selection equation (biprobit) are obtained, after which the inverse Mill's ratio is calculated, this selection effect is then linearised to make it compatible with each quantile. The linearised correction term is included as a control in the conditional log-wage quantile regressions, the results for which are presented below.

6.2 Earnings Equations: Conditional Quantile Regression Results

The results of the conditional quantile earnings model (15) are presented below. The wage equation contains only employed women, and the focus is on the characterisation of changes at various points of the conditional log-wage distribution. The findings from the summary statistics provided evidence in support of differences across race regarding socio-economic characteristics. The first step estimation results illustrated that these differences affect the conditional probability that a woman will (1) participate in the labour force and (2) find employment. Given these findings, the earnings outcome is evaluated for women who are employed using conditional quantile regression with emphasis on differences in skills levels of women, as captured by industry and occupation, across race.

The dependent variable used in the quantile regression equations is log of

monthly real earnings. The independent variables used include race, industry, occupation, log of hours worked in the last week, age, age-squared, a marital dummy, number of children in the household under the age of six and between seven and fifteen¹⁷. Interaction terms between race and secondary schooling were omitted because they were reported insignificant. The sample selection correction terms estimated via the polynomial expansion from the first step estimation are also included as independent variables in the quantile regressions¹⁸.

A variable for tenure or experience is replaced by the weak proxy of age and age-squared in the earnings equations, so there may be some omitted variables bias. The number of children in the household are not expected to affect the earnings outcome directly but are expected to be negatively related to labour market experience. These variables are thus included in the wage equation. Following Buchinsky (1994) the definition of inequality is seen as the gap between two conditional log-wage quantiles.

The quantile regressions were estimated at five selected quantiles: 0.1, 0.25, 0.5 (the median), 0.75 and 0.9. The results presented in Tables 2, 3, 4, 5 and 6 consist of the quantile regressions for each period in the dataset.

For the first period (1994-1995) following the end of apartheid, the conditional quantile regression output is presented in Table 2¹⁹. Concerning race, black and coloured women earn significantly less than their white counterparts at all quantiles of the conditional wage distribution. For a black woman at the bottom 10% of the conditional wage distribution, earnings are 20% less than a white woman, holding all else constant. As quantile increases, the gap between

¹⁷The base race is white, the base industry is manufacturing, the base occupation is plant operator and the base province is Gauteng. Controls for province are also included in the model.

¹⁸The polynomial correction function uses the logic of the Taylor Series approximation. The true value of the correction term function $h_{\tau}(\cdot)$ is unknown, and is seen to be approximated by an expansion that is based on an infinite sum of the terms calculated from the functions derivative at a single point. The choice of j is based on the notion that if $j = \sqrt{n}$ then the results of the linearisation would result in a reliable estimate of $h_{\tau}(\cdot)$. However, \sqrt{n} in this sample is quite large. For the period 2005-2007, alone, $j = \sqrt{n} = 181$, this would mean that 181 additional variables would have to be added to the quantile regression leading to obvious identification problems. Further, the value of the correction terms are small, and approach zero rapidly as j increases. Various values for j were tested and it was decided that $j = 5$ represented a choice for j that was high enough to achieve the desired linearised convergence, but would avoid the problem of almost perfect collinearity that would arise from a higher value of j such as \sqrt{n} .

¹⁹The complete table with all coefficients is presented in Table A.8 (Appendix)

white and black earnings increases. A black woman at the median earns 22% less than a white woman at the median of the income distribution, *ceteris paribus*. At the top end of the income distribution a black woman earns 28% less than a white woman, all else being equal. This may be interpreted as evidence of racial discrimination against black women, and the extent of discrimination increases as the conditional earnings quantile increases. This is an issue that is picked up in the discussion section (6.3).

For coloured women, the magnitude of the gap between white and coloured wages is not as big as the gap between the wages of black and white women. For a coloured woman in the bottom decile, earnings are 9% less than that of a white woman in the base occupation and province. The wage gap between coloured and white women also increases as conditional log-wage quantile increases. At the median coloured women earn 11% less than white women, and at the 90th percentile coloured women earn 20% less than white women, *ceteris paribus*. In comparing these results to those from the mean-based regression, one can see how much information is lost.

The returns to occupation by quantile tie in perfectly with the observations made about occupation in the summary statistics section. Relative to the base occupation of plant operator, it can be seen that for the more highly skilled occupations (manager, professional and technician) there is a wage premium for the top half of the distribution (the median, the 75th percentile and the 90th percentile). For clerks, agricultural and artisanal workers, at the median, 25th percentile and 10th percentile, earnings are significantly less than for the base profession. Working as an elementary worker or domestic service worker at all conditional log-wage quantiles implies that a woman in these occupations will earn significantly less than a plant operator, *ceteris paribus*.

For sector of occupation, all sectors aside from agriculture exhibit a positive coefficient for all quantiles. This means that the mining, utilities, construction, trade, transport, finance and service sectors all pay a wage premium relative to manufacturing (the base sector). For the construction, trade, transport, finance and services sectors the wage premium relative to the manufacturing sector is decreasing as quantile increases. For instance, a woman at the bottom decile of construction earns monthly wage 39% higher than a (white) woman in manufacturing, but a woman in the 90th percentile of construction earns a wage

Table 2: Imputed Quantile Regression-With Heckman Correction: 1994-1995

	.10	.25	.5	.75	.9	Mean Reg
Log Hrs Worked	0.451*** (0.029)	0.340*** (0.028)	0.269*** (0.018)	0.190*** (0.020)	0.142*** (0.021)	0.308*** (0.062)
Num Child (0-6)	-0.022** (0.008)	-0.001 (0.007)	0.004 (0.005)	0.006 (0.007)	0.005 (0.007)	0.013 (0.025)
Num Child (7-15)	-0.026** (0.008)	-0.021*** (0.005)	-0.015** (0.005)	-0.007 (0.004)	-0.006 (0.006)	-0.006 (0.021)
Age	0.040*** (0.007)	0.042*** (0.005)	0.043*** (0.004)	0.040*** (0.004)	0.045*** (0.005)	0.020 (0.019)
Age-Squared	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000 (0.000)
Years of Educ	0.097*** (0.003)	0.101*** (0.002)	0.105*** (0.003)	0.099*** (0.002)	0.087*** (0.002)	0.076*** (0.008)
Married	-0.045** (0.016)	-0.043*** (0.013)	-0.029** (0.010)	-0.013 (0.013)	-0.007 (0.011)	-0.043 (0.048)
Black	-0.198** (0.067)	-0.239*** (0.039)	-0.222*** (0.038)	-0.218*** (0.029)	-0.284*** (0.059)	0.022 (0.172)
Coloured	-0.087* (0.040)	-0.116*** (0.026)	-0.109*** (0.023)	-0.142*** (0.024)	-0.197*** (0.033)	0.245* (0.117)
Manager	-0.280** (0.092)	-0.174** (0.055)	0.076* (0.033)	0.405*** (0.042)	0.865*** (0.090)	-0.205 (0.193)
Professional	0.038 (0.043)	0.020 (0.041)	0.104*** (0.030)	0.224*** (0.022)	0.351*** (0.053)	-0.444* (0.174)
Technician	0.071 (0.041)	0.078* (0.033)	0.173*** (0.023)	0.312*** (0.020)	0.421*** (0.040)	0.231 (0.142)
Clerk	-0.185*** (0.036)	-0.209*** (0.036)	-0.148*** (0.016)	-0.045 (0.025)	0.023 (0.035)	-0.340* (0.135)
Agricultural	-0.906*** (0.090)	-0.844*** (0.080)	-0.479*** (0.093)	-0.026 (0.194)	0.550*** (0.132)	0.304 (0.325)
Artisan	-0.536*** (0.056)	-0.356*** (0.042)	-0.153*** (0.027)	-0.038 (0.037)	0.067 (0.049)	-0.225 (0.161)
Elementary	-1.048*** (0.046)	-1.021*** (0.035)	-0.795*** (0.029)	-0.493*** (0.025)	-0.255*** (0.034)	-0.369** (0.134)
Domestic Service	-0.728*** (0.031)	-0.699*** (0.034)	-0.526*** (0.019)	-0.321*** (0.023)	-0.147*** (0.039)	-0.445*** (0.121)
Agri (Sector)	0.354*** (0.043)	0.266*** (0.024)	0.087** (0.029)	-0.105*** (0.024)	-0.074 (0.042)	-0.248* (0.102)
Mining	0.609*** (0.102)	0.450*** (0.068)	0.345*** (0.081)	0.414*** (0.078)	0.451*** (0.109)	0.088 (0.311)
Utilities	0.308* (0.122)	0.415*** (0.081)	0.339*** (0.071)	0.290*** (0.074)	0.188* (0.083)	0.500 (0.422)
Construction	0.392*** (0.062)	0.346*** (0.050)	0.290*** (0.055)	0.194* (0.083)	0.197** (0.065)	0.315 (0.253)
Trade	0.383*** (0.036)	0.371*** (0.021)	0.289*** (0.014)	0.189*** (0.015)	0.156*** (0.017)	0.167* (0.075)
Transport	0.611*** (0.078)	0.558*** (0.039)	0.469*** (0.026)	0.357*** (0.043)	0.298*** (0.038)	0.250 (0.189)
Finance	0.541*** (0.043)	0.454*** (0.025)	0.352*** (0.010)	0.236*** (0.017)	0.201*** (0.021)	0.210 (0.119)
Services	0.471*** (0.032)	0.418*** (0.023)	0.313*** (0.016)	0.174*** (0.014)	0.113*** (0.023)	0.222** (0.071)
λ_{mean}						0.763 (1.326)
N	25930	25930	25930	25930	25930	16006

Standard errors in parentheses. All results estimated using sampling weights.

Dependent variable: Natural Logarithm of Imputed Real Earnings (2000 prices).

All standard errors obtained via bootstrap resampling.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

that is 19% higher than a woman in manufacturing, *ceteris paribus*. The fact that there are differential returns to working in different sectors and differential returns to race implies that using the mean regression to characterise the conditional log-wage distribution would lead to misrepresentative conclusions.

Other explanatory variables that are interesting determinants of earnings at each conditional quantile include log of hours worked, age, marital status and years of education. The impact of these variables on the conditional earnings outcome is discussed briefly for the initial period, but for the sake of brevity for each subsequent time period the discussion focuses only on returns to occupation and sector at different quantiles highlighting other variables where necessary.

At the turn of the new democracy log of hours worked in the last week is a significant determinant of the logged monthly wage of (white) women in the manufacturing sector, holding all else constant. The elasticity of wages with respect to hours worked is quite high. For the bottom decile, a 1% increase in hours worked is expected to lead to a 45% increase in monthly wages, holding all else constant. The magnitude of the elasticity of wages with respect to hours worked is decreasing by quantile. For the 90th percentile a 1% increase in hours worked in the last week leads to a 14% increase in monthly earnings, *ceteris paribus*.

Age is seen to have a linear significant impact on the earnings of all quantiles; the squared age variable has a significant zero effect on earnings at all quantiles. The effect of an increase in age by one year leads to an increase in monthly real earnings of between 4% and 5% for all quantiles. There appears to be no differential return to age at different ends of the income distribution in this period.

Years of education has a significant positive impact on log monthly earnings at all points of the conditional income distribution. For the bottom four quantiles a 1 year increase in education is expected to lead to a 10% increase in earnings, holding all else constant. For the 90th percentile, an increase in education by one year leads to an increase in earnings of 9%.

Interestingly marital status and number of children in the household under the age of fifteen report significant coefficients for the median and below. A married woman at the bottom decile earns 5% less than a non-married woman

at the bottom decile, *ceteris paribus*. A woman at the bottom decile with one child between the ages of seven and fifteen earns 3% less than a woman with no school-going children, *ceteris paribus*. For the 25th percentile and above the number of children below the ages of six in the household have no effect on the earnings of women *ceteris paribus*.

As far as the selection effect from observing a non-random part of the employed female population is concerned, the linearised sample selection correction terms²⁰ are at large not significant for the period 1994-1995. The selection effect for the mean regression is also not significant.

The results from later sub-periods are now presented. While the results are interesting their own right, changes over time relative to the base period are highlighted. For the second period (1997-1998) the conditional quantile regression estimates are presented in Table 3²¹. As far as race is concerned, for 1997-1998, black women exhibit negative coefficients over all five quantiles. The conditional log-wage gap between black and white women now decreases up to the 75th percentile then increases slightly at the 90th percentile, in the previous period the gap increased with quantile. For coloured women, the gap between coloured and white conditional log-wage is smaller than the gap between coloured and black conditional log-wages. For instance a coloured woman at the median earns 22% less than a white woman at the median, where a black woman at the median earns 40% less than a white woman at the median *ceteris paribus*. The gap between coloured and white wages decreases between the first two quantiles then increases from the median on as quantile increases.

In terms of returns to working in certain occupations; artisan, elementary and domestic services exhibit negative significant coefficients for all quantiles, this is inline with what was found for the period 1994-1995. The gap between the base occupation and domestic services occupation is now larger (more negative) for all quantiles. Clerks earn significantly less than plant operators for the median and 75th percentile. Agricultural workers earn significantly less than plant operators for the median and below this is opposite to the results from the preceding period. At the top half of the conditional quantile distribution technicians earn more than their plant operating counterparts. Professionals

²⁰These are presented in Table A.8 in the Appendix.

²¹The complete table with all coefficients and controls is presented in Table A.9 (Appendix)

Table 3: Imputed Quantile Regression-With Heckman Correction: 1997-1998

	.10	.25	.5	.75	.9	Mean Reg
Log Hrs Worked	0.172*** (0.040)	0.062*** (0.011)	0.020 (0.013)	0.002 (0.006)	0.007 (0.007)	0.041*** (0.012)
Num Child (0-6)	-0.061** (0.019)	-0.027** (0.009)	-0.011 (0.011)	-0.000 (0.012)	0.007 (0.015)	-0.044** (0.016)
Num Child (7-15)	-0.057** (0.021)	-0.042*** (0.008)	-0.020* (0.010)	-0.004 (0.011)	-0.002 (0.013)	-0.045*** (0.013)
Age	0.180** (0.070)	0.125*** (0.024)	0.055 (0.039)	0.013 (0.043)	-0.004 (0.044)	0.179*** (0.046)
Age-Squared	-0.002* (0.001)	-0.001*** (0.000)	-0.001 (0.000)	-0.000 (0.001)	0.000 (0.001)	-0.002*** (0.001)
Years of Educ	0.125*** (0.020)	0.109*** (0.008)	0.088*** (0.010)	0.071*** (0.011)	0.060*** (0.012)	0.116*** (0.013)
Married	-0.223* (0.087)	-0.104*** (0.031)	0.010 (0.047)	0.054 (0.054)	0.079 (0.049)	-0.200*** (0.055)
Black	-0.492*** (0.118)	-0.464*** (0.044)	-0.397*** (0.064)	-0.371*** (0.072)	-0.399*** (0.085)	-0.603*** (0.075)
Coloured	-0.238*** (0.061)	-0.206*** (0.035)	-0.218*** (0.024)	-0.303*** (0.030)	-0.384*** (0.043)	-0.261*** (0.055)
Manager	-0.256** (0.081)	-0.045 (0.068)	0.078 (0.041)	0.246*** (0.043)	0.377*** (0.090)	-0.009 (0.092)
Professional	0.170 (0.088)	0.178** (0.055)	0.175*** (0.036)	0.190*** (0.037)	0.196*** (0.057)	0.172* (0.082)
Technician	0.002 (0.053)	0.048 (0.063)	0.078* (0.033)	0.120*** (0.036)	0.105* (0.046)	-0.025 (0.079)
Clerk	-0.020 (0.049)	-0.088 (0.062)	-0.118*** (0.031)	-0.080* (0.040)	-0.086 (0.053)	-0.161* (0.076)
Agricultural	-0.392** (0.140)	-0.440*** (0.075)	-0.328*** (0.062)	-0.216 (0.111)	0.027 (0.169)	-0.395*** (0.110)
Artisan	-0.544*** (0.064)	-0.388*** (0.064)	-0.213*** (0.036)	-0.139*** (0.041)	-0.118* (0.047)	-0.399*** (0.078)
Elementary	-0.610*** (0.042)	-0.517*** (0.048)	-0.380*** (0.024)	-0.279*** (0.033)	-0.267*** (0.050)	-0.509*** (0.072)
Domestic Service	-0.715*** (0.059)	-0.743*** (0.044)	-0.655*** (0.022)	-0.508*** (0.027)	-0.381*** (0.047)	-0.688*** (0.068)
Agri (Sector)	-0.437*** (0.091)	-0.358*** (0.024)	-0.407*** (0.031)	-0.452*** (0.033)	-0.319*** (0.081)	-0.345*** (0.049)
Mining	0.533** (0.163)	0.519*** (0.078)	0.499*** (0.056)	0.408*** (0.059)	0.370** (0.113)	0.579*** (0.144)
Utilities	0.440*** (0.066)	0.251** (0.085)	0.406* (0.162)	0.383*** (0.096)	0.510 (0.343)	-0.029 (0.205)
Construction	0.348* (0.137)	0.286*** (0.087)	0.129 (0.070)	0.161 (0.085)	0.125 (0.097)	0.341** (0.122)
Trade	-0.108** (0.041)	0.053* (0.026)	0.115*** (0.020)	0.142*** (0.022)	0.192*** (0.027)	0.003 (0.033)
Transport	0.339*** (0.084)	0.419*** (0.053)	0.471*** (0.055)	0.489*** (0.043)	0.376*** (0.064)	0.436*** (0.094)
Finance	0.370*** (0.045)	0.346*** (0.032)	0.356*** (0.028)	0.337*** (0.031)	0.318*** (0.039)	0.423*** (0.060)
Services	0.325*** (0.041)	0.427*** (0.025)	0.457*** (0.029)	0.407*** (0.024)	0.363*** (0.030)	0.463*** (0.040)
λ_{mean}						0.863** (0.299)
<i>N</i>	16453	16453	16453	16453	16453	11096

Standard errors in parentheses. All results estimated using sampling weights.

Dependent variable: Natural Logarithm of Imputed Real Earnings (2000 prices).

All standard errors obtained via bootstrap resampling.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

earn more than plant operators for the 25th percentile and above, and their earnings are between 18% and 20% higher than that of plant operators, *ceteris paribus*.

Sector of employment exhibits some different results for the period 1997-1998 relative to 1994-1995, agriculture now exhibits a negative coefficient for all quantiles, and this effect is significant at the 1% level. Mining, utilities and construction exhibit positive and significant coefficients. Mining will pay a sectoral premium relative to manufacturing for all five conditional log-wage quantiles. The transport, finance and service sectors report positive and significant coefficients for all five quantiles, but the pattern of returns to working in these sectors has changed since 1994-1995. For instance, instead of the wage gap between these sectors and the base sector decreasing by quantile, the gap for services and transport increases after the first quantile (.10) and then decreases again by the last (.90) quantile.

Marital status is reported to have a significant negative effect on the conditional monthly wage of a woman in the bottom two quantiles of -22% (at 0.10) and -10% (at 0.25), respectively. For this period, age displays a quadratic effect. The elasticity of wage with respect to hours worked is only significant for the quantiles below the median. Years of education exhibits a positive effect for each quantile. The effect of years of education on conditional log-wages decreases as quantile increases.

Regarding the selection effect for this period, none of the polynomial selection correction terms are reported significant in these quantile regressions²². The lack of significance may be evidence of multicollinearity between the linearised quantile regression correction terms since the mean regression sample selection coefficient is significant.

Overall for the 1990's it has been shown that the results from the quantile regressions differ from the results presented by the mean regressions. Across quantiles, black and coloured women earn less than white women. Conditional earnings also differ based on quantile, occupation and sector. Individuals in lower skilled jobs and sectors earn less than women in higher skilled sectors

²²These are presented in Table A.9 in the Appendix

across all quantiles for the 1990's. Increasing years of education is expected to lead to an increase in wages for all quantiles, although the size of the increase may vary.

For the 2000's the results are presented in Tables 4, 5 and 6²³. The coefficients for these three periods are interpreted jointly, for simplicity and brevity. For black women, in 2000-2002, the conditional log-wage gap between black and white wages increased between the period 1995-1997 and 2000-2002 for the bottom decile. In 2000-2002 the conditional log-wage gap no longer increased by quantile, instead for this period it decreased to the median, then increased to the .90 quantile. This u-shaped pattern is also prevalent for the period 2005-2007. For the period 2010-2011, the wage gap decreases from 64% at the .10 quantile to 36% at the median then increases by a percent thereafter. Within each period considered, the wage gap seems to follow a rough u-shape pattern between the bottom decile and top quantile. However, the dynamics of the wage gap between periods warrants further investigation. The discussion in Section 6.3 systematically investigates the change in wages by race across quantiles over time.

For coloured women, for all three periods in the 2000's, the conditional log-wage gap between white and coloured wages exhibits a u-shape across quantiles. It decreases to the median than increases. The size of the gap is larger than it was in the mid 1990's for all quantiles. The gap between coloured and white womens' wages is still smaller than the black-white conditional wage gap.

As far as occupation is concerned, in the first period (1994-1995) it was seen that for some professions (manager, clerk, agricultural workers) there were differential returns relative to being a plant operator, by quantile. That is, for the lower quantiles, a woman in the aforementioned profession would earn less than a plant operator, and for higher quantiles, a woman in the aforementioned professions would earn more than a plant operator (*ceteris paribus*). By 1997-1998, this pattern had changed slightly and by the 2000's it can be seen that systematic differential effects between quantiles for each occupation have disappeared. In other words there is no longer a pattern of negative coefficients for lower quantiles and positive coefficients for higher quantiles across various

²³The complete tables with all coefficients are presented in Table A.10, A.11 and A.12 in the Appendix.

Table 4: Imputed Quantile Regression-With Heckman Correction: 2000-2002

	.10	.25	.5	.75	.9	Mean Reg
Log Hrs Worked	0.203*** (0.020)	0.111*** (0.013)	0.033*** (0.006)	0.018*** (0.002)	0.012* (0.005)	0.054*** (0.004)
Num Child (0-6)	-0.009* (0.004)	-0.009*** (0.002)	-0.005* (0.002)	-0.005* (0.002)	-0.008* (0.004)	-0.006** (0.002)
Num Child (7-15)	-0.007* (0.003)	-0.008*** (0.001)	-0.010*** (0.002)	-0.006*** (0.002)	-0.006** (0.002)	-0.010*** (0.002)
Age	0.028 (0.017)	0.041*** (0.010)	0.054*** (0.009)	0.061*** (0.010)	0.081*** (0.012)	0.070*** (0.010)
Age-Squared	-0.000 (0.000)	-0.000** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Years of Educ	0.067*** (0.004)	0.071*** (0.004)	0.073*** (0.003)	0.077*** (0.002)	0.085*** (0.002)	0.073*** (0.003)
Married	0.022 (0.018)	0.007 (0.016)	-0.004 (0.013)	-0.000 (0.015)	0.000 (0.022)	-0.022 (0.013)
Black	-0.564*** (0.032)	-0.470*** (0.024)	-0.452*** (0.020)	-0.496*** (0.020)	-0.543*** (0.030)	-0.523*** (0.021)
Coloured	-0.392*** (0.039)	-0.258*** (0.027)	-0.240*** (0.027)	-0.285*** (0.021)	-0.328*** (0.030)	-0.310*** (0.025)
Manager	0.455*** (0.090)	0.536*** (0.046)	0.506*** (0.032)	0.539*** (0.030)	0.554*** (0.044)	0.499*** (0.048)
Professional	0.543*** (0.082)	0.489*** (0.056)	0.388*** (0.026)	0.331*** (0.035)	0.346*** (0.046)	0.407*** (0.043)
Technician	0.317*** (0.055)	0.323*** (0.035)	0.274*** (0.032)	0.246*** (0.029)	0.219*** (0.040)	0.250*** (0.034)
Clerk	0.174** (0.057)	0.086* (0.037)	0.022 (0.033)	0.014 (0.029)	0.031 (0.034)	0.026 (0.032)
Agricultural	-1.025*** (0.156)	-0.693*** (0.069)	-0.614*** (0.045)	-0.279** (0.103)	0.047 (0.059)	-0.655*** (0.051)
Artisan	-0.517*** (0.085)	-0.382*** (0.049)	-0.372*** (0.032)	-0.269*** (0.039)	-0.198*** (0.057)	-0.348*** (0.035)
Elementary	-0.529*** (0.050)	-0.590*** (0.035)	-0.567*** (0.027)	-0.460*** (0.028)	-0.371*** (0.037)	-0.529*** (0.030)
Domestic Service	-0.563*** (0.044)	-0.663*** (0.035)	-0.698*** (0.022)	-0.625*** (0.022)	-0.547*** (0.035)	-0.643*** (0.028)
Agri (Sector)	0.130*** (0.025)	0.043 (0.028)	-0.055* (0.022)	-0.152*** (0.024)	-0.202*** (0.035)	-0.030 (0.023)
Mining	0.643*** (0.100)	0.669*** (0.079)	0.698*** (0.059)	0.742*** (0.069)	0.592*** (0.067)	0.725*** (0.073)
Utilities	0.577* (0.262)	0.764*** (0.100)	0.808*** (0.056)	0.727*** (0.085)	0.847*** (0.091)	0.718*** (0.094)
Construction	0.127 (0.102)	0.122* (0.055)	0.191*** (0.041)	0.233** (0.074)	0.568*** (0.116)	0.157** (0.056)
Trade	-0.140*** (0.034)	-0.013 (0.022)	0.141*** (0.015)	0.210*** (0.014)	0.269*** (0.022)	0.107*** (0.015)
Transport	0.229*** (0.044)	0.436*** (0.050)	0.554*** (0.044)	0.549*** (0.042)	0.553*** (0.052)	0.534*** (0.046)
Finance	0.343*** (0.043)	0.445*** (0.031)	0.482*** (0.026)	0.473*** (0.016)	0.451*** (0.022)	0.508*** (0.026)
Services	0.388*** (0.037)	0.528*** (0.025)	0.630*** (0.032)	0.643*** (0.021)	0.589*** (0.019)	0.636*** (0.019)
λ_{mean}						-0.013 (0.058)
N	32006	32006	32006	32006	32006	24831

Standard errors in parentheses. All results estimated using sampling weights.

Dependent variable: Natural Logarithm of Imputed Real Earnings (2000 prices).

All standard errors obtained via bootstrap resampling.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Imputed Quantile Regression-With Heckman Correction: 2005-2007

	.10	.25	.5	.75	.9	Mean Reg
Log Hrs Worked	0.218*** (0.029)	0.118*** (0.021)	0.044*** (0.007)	0.022*** (0.004)	0.018*** (0.004)	0.056*** (0.004)
Num Child (0-6)	-0.003 (0.005)	-0.006 (0.004)	-0.004 (0.003)	-0.003 (0.003)	-0.004 (0.004)	-0.011*** (0.003)
Num Child (7-15)	-0.012*** (0.003)	-0.014*** (0.004)	-0.013*** (0.002)	-0.010*** (0.002)	-0.007* (0.003)	-0.015*** (0.002)
Age	-0.007 (0.017)	0.033* (0.016)	0.033** (0.011)	0.033** (0.012)	0.028 (0.015)	0.056*** (0.011)
Age-Squared	0.000 (0.000)	-0.000 (0.000)	-0.000* (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.001*** (0.000)
Years of Educ	0.063*** (0.007)	0.077*** (0.006)	0.074*** (0.003)	0.070*** (0.004)	0.076*** (0.004)	0.076*** (0.003)
Married	0.013 (0.023)	0.028 (0.019)	0.030* (0.013)	0.039* (0.015)	0.077*** (0.019)	-0.002 (0.013)
Black	-0.548*** (0.032)	-0.479*** (0.024)	-0.460*** (0.017)	-0.509*** (0.015)	-0.620*** (0.021)	-0.543*** (0.021)
Coloured	-0.317*** (0.039)	-0.267*** (0.029)	-0.270*** (0.018)	-0.349*** (0.017)	-0.457*** (0.023)	-0.330*** (0.025)
Manager	0.317*** (0.077)	0.475*** (0.049)	0.644*** (0.037)	0.690*** (0.028)	0.677*** (0.073)	0.532*** (0.042)
Professional	0.622*** (0.045)	0.669*** (0.045)	0.630*** (0.029)	0.473*** (0.033)	0.293*** (0.057)	0.570*** (0.039)
Technician	0.147*** (0.042)	0.328*** (0.061)	0.442*** (0.032)	0.343*** (0.027)	0.197*** (0.054)	0.313*** (0.033)
Clerk	0.201*** (0.048)	0.208*** (0.048)	0.218*** (0.024)	0.144*** (0.027)	0.028 (0.047)	0.165*** (0.032)
Agricultural	-1.041*** (0.081)	-0.761*** (0.114)	-0.405*** (0.045)	-0.115* (0.057)	-0.009 (0.065)	-0.583*** (0.053)
Artisan	-0.474*** (0.035)	-0.287*** (0.037)	-0.219*** (0.032)	-0.177*** (0.031)	-0.181** (0.064)	-0.262*** (0.033)
Elementary	-0.516*** (0.032)	-0.473*** (0.034)	-0.405*** (0.027)	-0.386*** (0.029)	-0.401*** (0.051)	-0.439*** (0.030)
Domestic Service	-0.473*** (0.028)	-0.454*** (0.031)	-0.420*** (0.024)	-0.444*** (0.025)	-0.459*** (0.049)	-0.445*** (0.028)
Agri (Sector)	0.386*** (0.037)	0.329*** (0.021)	0.204*** (0.020)	0.058* (0.026)	0.001 (0.032)	0.219*** (0.022)
Mining	0.673*** (0.081)	0.667*** (0.051)	0.747*** (0.070)	0.774*** (0.068)	0.724*** (0.072)	0.796*** (0.066)
Utilities	0.605*** (0.150)	0.745*** (0.151)	0.874*** (0.093)	0.988*** (0.067)	1.025*** (0.075)	0.802*** (0.093)
Construction	0.266*** (0.045)	0.133*** (0.032)	-0.000 (0.034)	0.096*** (0.025)	0.134 (0.071)	0.117*** (0.035)
Trade	-0.259*** (0.033)	-0.092*** (0.019)	0.042** (0.013)	0.122*** (0.016)	0.169*** (0.020)	0.011 (0.014)
Transport	0.196*** (0.042)	0.305*** (0.071)	0.462*** (0.023)	0.546*** (0.063)	0.687*** (0.057)	0.378*** (0.046)
Finance	0.209*** (0.043)	0.385*** (0.038)	0.378*** (0.013)	0.429*** (0.020)	0.471*** (0.027)	0.415*** (0.024)
Services	0.242*** (0.032)	0.421*** (0.021)	0.578*** (0.026)	0.678*** (0.022)	0.681*** (0.024)	0.544*** (0.017)
λ_{mean}						0.078 (0.072)
N	32975	32975	32975	32975	32975	24759

Standard errors in parentheses. All results estimated using sampling weights.

Dependent variable: Natural Logarithm of Imputed Real Earnings (2000 prices).

All standard errors obtained via bootstrap resampling.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Imputed Quantile Regression-With Heckman Correction: 2010-2011

	.10	.25	.5	.75	.9	Mean Reg
Log Hrs Worked	0.038*** (0.008)	0.024** (0.008)	0.015*** (0.004)	0.007 (0.004)	0.006* (0.003)	0.024*** (0.004)
Num Child (0-6)	-0.008 (0.006)	-0.006 (0.005)	-0.008* (0.004)	-0.004 (0.003)	0.004 (0.004)	-0.018*** (0.004)
Num Child (7-15)	-0.006 (0.004)	-0.007 (0.005)	-0.003 (0.004)	0.003 (0.005)	0.010 (0.007)	-0.012** (0.004)
Age	-0.065 (0.052)	-0.048 (0.049)	-0.047 (0.045)	-0.040 (0.034)	-0.118 (0.067)	0.077* (0.031)
Age-Squared	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.000)	0.001 (0.001)	-0.001* (0.000)
Years of Educ	0.050** (0.018)	0.061*** (0.015)	0.063*** (0.013)	0.059*** (0.011)	0.036 (0.019)	0.103*** (0.009)
Married	0.011 (0.032)	0.032 (0.036)	0.066* (0.030)	0.054** (0.020)	0.114** (0.039)	-0.023 (0.022)
Black	-0.637*** (0.054)	-0.522*** (0.045)	-0.363*** (0.048)	-0.370*** (0.026)	-0.366*** (0.061)	-0.455*** (0.030)
Coloured	-0.471*** (0.036)	-0.320*** (0.038)	-0.307*** (0.029)	-0.314*** (0.030)	-0.334*** (0.047)	-0.289*** (0.033)
Manager	0.466*** (0.087)	0.537*** (0.064)	0.764*** (0.044)	0.833*** (0.059)	0.807*** (0.073)	0.651*** (0.060)
Professional	0.535*** (0.076)	0.600*** (0.067)	0.798*** (0.040)	0.735*** (0.058)	0.704*** (0.060)	0.665*** (0.060)
Technician	0.310*** (0.070)	0.420*** (0.059)	0.610*** (0.051)	0.501*** (0.047)	0.305*** (0.048)	0.469*** (0.053)
Clerk	0.248*** (0.066)	0.186*** (0.046)	0.278*** (0.047)	0.259*** (0.044)	0.179*** (0.040)	0.281*** (0.052)
Agricultural	-0.396 (0.256)	-0.219 (0.160)	-0.116 (0.119)	-0.026 (0.103)	0.278* (0.118)	-0.120 (0.139)
Artisan	-0.096 (0.090)	-0.046 (0.051)	-0.013 (0.057)	0.028 (0.073)	0.087 (0.096)	-0.042 (0.062)
Elementary	-0.341*** (0.078)	-0.338*** (0.052)	-0.341*** (0.045)	-0.377*** (0.048)	-0.406*** (0.047)	-0.329*** (0.052)
Domestic Service	-0.175** (0.063)	-0.297*** (0.044)	-0.317*** (0.044)	-0.387*** (0.039)	-0.406*** (0.049)	-0.272*** (0.048)
Agri (Sector)	0.431*** (0.053)	0.254*** (0.044)	0.160*** (0.021)	-0.025 (0.040)	-0.168*** (0.035)	0.188*** (0.047)
Mining	0.663*** (0.099)	0.665*** (0.109)	0.815*** (0.079)	0.805*** (0.087)	0.742*** (0.064)	0.789*** (0.090)
Utilities	0.370** (0.117)	0.525* (0.266)	0.711*** (0.110)	0.840*** (0.147)	0.886*** (0.161)	0.739*** (0.143)
Construction	0.180* (0.073)	-0.006 (0.057)	-0.046 (0.054)	0.185* (0.086)	0.347*** (0.074)	0.137* (0.060)
Trade	0.046 (0.035)	0.147*** (0.023)	0.201*** (0.018)	0.251*** (0.027)	0.433*** (0.029)	0.191*** (0.025)
Transport	0.162 (0.106)	0.183*** (0.055)	0.336*** (0.069)	0.453*** (0.058)	0.591*** (0.067)	0.335*** (0.053)
Finance	0.275*** (0.042)	0.322*** (0.032)	0.345*** (0.015)	0.347*** (0.026)	0.428*** (0.040)	0.350*** (0.032)
Services	0.185*** (0.023)	0.273*** (0.025)	0.442*** (0.024)	0.590*** (0.028)	0.674*** (0.038)	0.422*** (0.025)
λ_{mean}						0.323* (0.164)
N	18826	18826	18826	18826	18826	18817

Standard errors in parentheses. All results estimated using sampling weights.
 Dependent variable: Natural Logarithm of Imputed Real Earnings (2000 prices).
 All standard errors obtained via bootstrap resampling.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

occupations.

For 2000-2002, highly skilled workers (managers, professionals, and technicians) all report significant positive coefficients across all five quantiles. For clerks, the bottom two quantiles report positive and significant coefficients. A white woman who is a clerk at the .10 quantile earns 17% more than a woman who is a plant operator holding all else constant. By 2005-2007 the coefficients on clerks are significant up to the 75th percentile, and by 2010-2011 the coefficients on clerks are positive and significant for all quantiles. Clerks have increased in importance over this period in terms of relative earnings potential. For 2005-2007 and 2010-2011 managers, professionals, and technicians, all report significant positive coefficients across all five quantiles.

In 2000-2002, all significant coefficients for lower skilled occupations (artisans, elementary workers and domestic service workers) are negative for all quantiles. For agricultural workers, these coefficients are negative for the 75th percentile and below. By 2005-2007, agricultural workers earn less than plant operators for the median and below. For artisans, elementary workers and domestic service workers all quantiles earn less than plant operators in 2005-2007. By 2010-2011, agricultural workers earn less than plant operators only for the bottom quantile, artisans do not report any significant coefficients and elementary and domestic service workers earn significantly less than plant operators at all quantiles. The pattern that agricultural workers exhibit illustrates that there may be decreasing importance of agricultural work on the earnings outcome of women in the labour force holding all else constant. The magnitude of the returns to working in different occupations differ by quantile over time and are investigated in Section 6.3.

The patterns for sector of occupation following the periods in the 2000's are interesting. First, the returns to working in different sectors exhibit different behaviour across quantiles when compared to the mid 1990's. For 1994-1995 the returns to working in the construction, trade, transport, finance and services sectors all exhibited decreasing returns to earnings as quantile increased. This meant that in this period, moving up the earnings distribution decreased the gap between the manufacturing sector and the above mentioned sectors. By 2000-2002 however for the construction, trade, transport, finance and services sectors the gap between the manufacturing sector and these sectors is shown

to be increasing by quantile. This illustrates that returns in the manufacturing sector seem to have decreased by quantile over time, the manufacturing sector also seems to have decreased in importance overall (it has a lower impact on earnings).

The agricultural sector displays decreasing returns by quantile for all three periods in the 2000's. For the period 2000-2002, this effect is positive for the first two quantiles then negative for the median and above. For the period 2005-2007 the coefficient on agriculture is positive and significant up to the 75th quantile, for 2010-2011, this effect is positive and significant for the .10, .25, .50 and .90 quantiles. The mining and utilities sectors exhibit positive coefficients relative to manufacturing for all three periods in the 2000's. For these two sectors the effect of being employed in one of them on wage is not increasing or decreasing by quantile, but exhibits different behaviour across the distribution.

The coefficients on years of education are significant for all quantiles for the period 2000-2002, 2005-2007, and for four quantiles in 2010-2011. The returns to education seem to have changed over the years. While the return to education is still positive over all time periods and quantiles, the magnitude of the effect has decreasing slightly over all quantiles between the 1990's and the 2000's.

Finally, concerning the sample selection effects for the 2000's²⁴, it is clear that as was the case for both periods in the 1990's at large almost all of the sample selection correction effects are not significant. One exception to this is the 90th percentile for the period 2000-2002. For the period 2000-2002 and 2005-2007 the mean sample selection correction terms are also not reported as significant. However, before concluding that there is no bias inherent in the earnings function it is worth mentioning that a few issues could be causing this insignificance; the first is potential multicollinearity of the orders of the polynomial expansion. These variables are all highly correlated which may decrease reported significance. The other reason is an incorrect specification of the polynomial selection correction in the Methodology (Section 4.2). The correction of these terms may be a possible area for improvement of this study.

The results from the conditional quantile regressions illustrated that us-

²⁴Presented in Tables, A.10, A.11 and A.12 in the Appendix.

ing mean regression to characterise the female earnings distribution provides a misrepresentative measure of the impact of determinants of earnings. The quantile regressions presented above not only provide insight into the determinants of earnings at different points of the distribution, they also provide a look at the evolution of earnings for women in the South African labour force over time. Racial differences in earnings are evident, so are differentials in returns to education, occupation and sector of employment. To what extent skills levels and observed individual characteristics are driving the differences in wages of women of different races across quantiles is investigated graphically in the next section.

6.3 Discussion

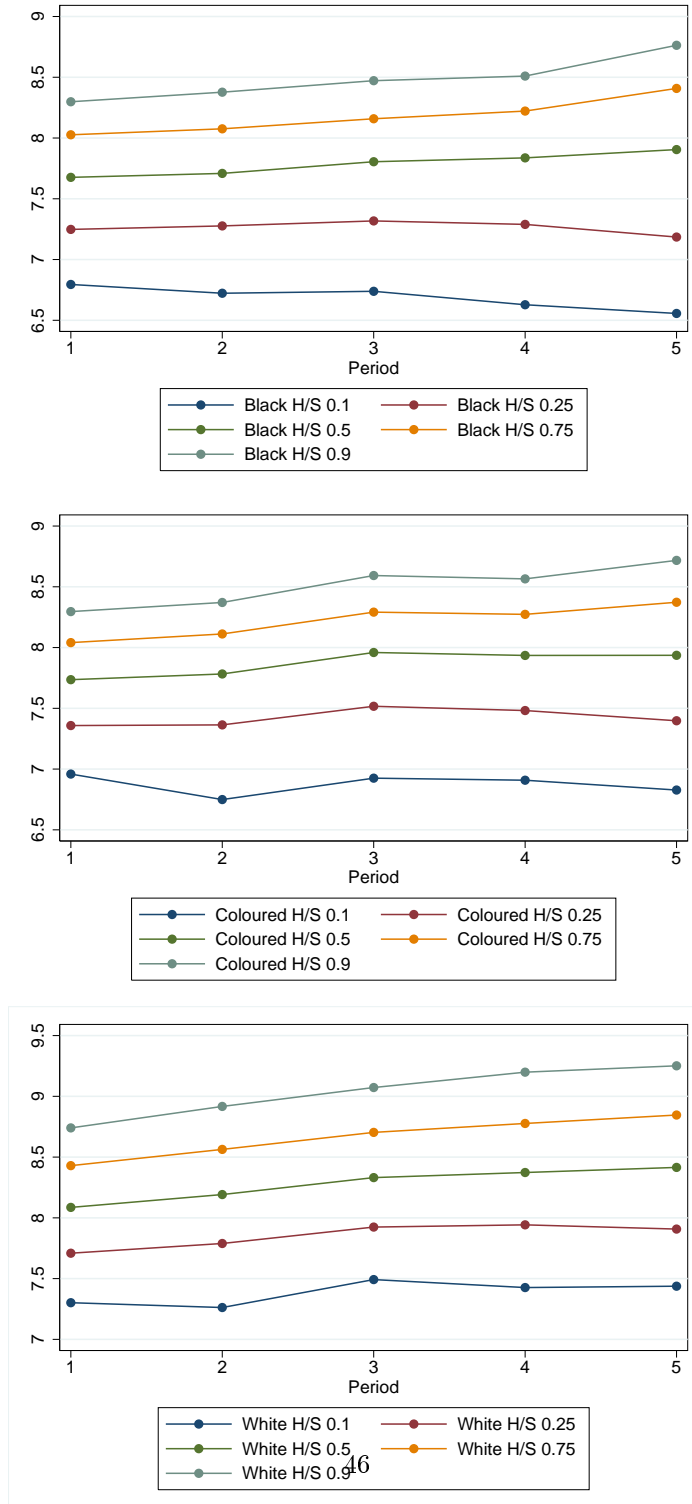
To summarise the results presented in the above section, the graphs of predicted (log) wage by race are presented. The sample was split into a high skilled occupation group and a low skilled occupation group and earnings per quantile were predicted for mean values of the explanatory variables of each respective race group, conditional on a woman being high skilled or low skilled. High skilled occupations consist of professionals, managers, technicians and plant operators. Low skilled occupations include clerks, domestic service workers, elementary workers and artisans. The choice to split the sample by high and low skilled workers stems from the large variation in coefficients on occupation between quantiles in Section 6.2. For the graphs presented below, predicted log-earnings appear on the y-axis and the time period²⁵ on the x-axis.

Figure 1 plots predicted income of high-skilled workers by quantile for all three races on separate graphs. The top most graph predicts earnings for black high-skilled women, the second graph predicts earnings for coloured women in high-skilled professions and the third graph presents predicted earnings for white women who hold high skills. The first thing to note is that all predicted quantile curves do not cross. The range of earnings of women in high skilled professions is between R812.41 ($\ln(6.7)$) at the 10th percentile and R6002.91 p.m ($\ln(8.7)$) at the 90th percentile.

For black women in high-skilled professions over all quantiles there has been

²⁵Period 1 corresponds to 1994-1995, period 2 to 1997-998, period 3 to 2000-2002, period 4 to 2005-2007 and period 5 to 2010-2011.

Figure 1: Predicted Log of Monthly Wage By Race (High Skilled Women)



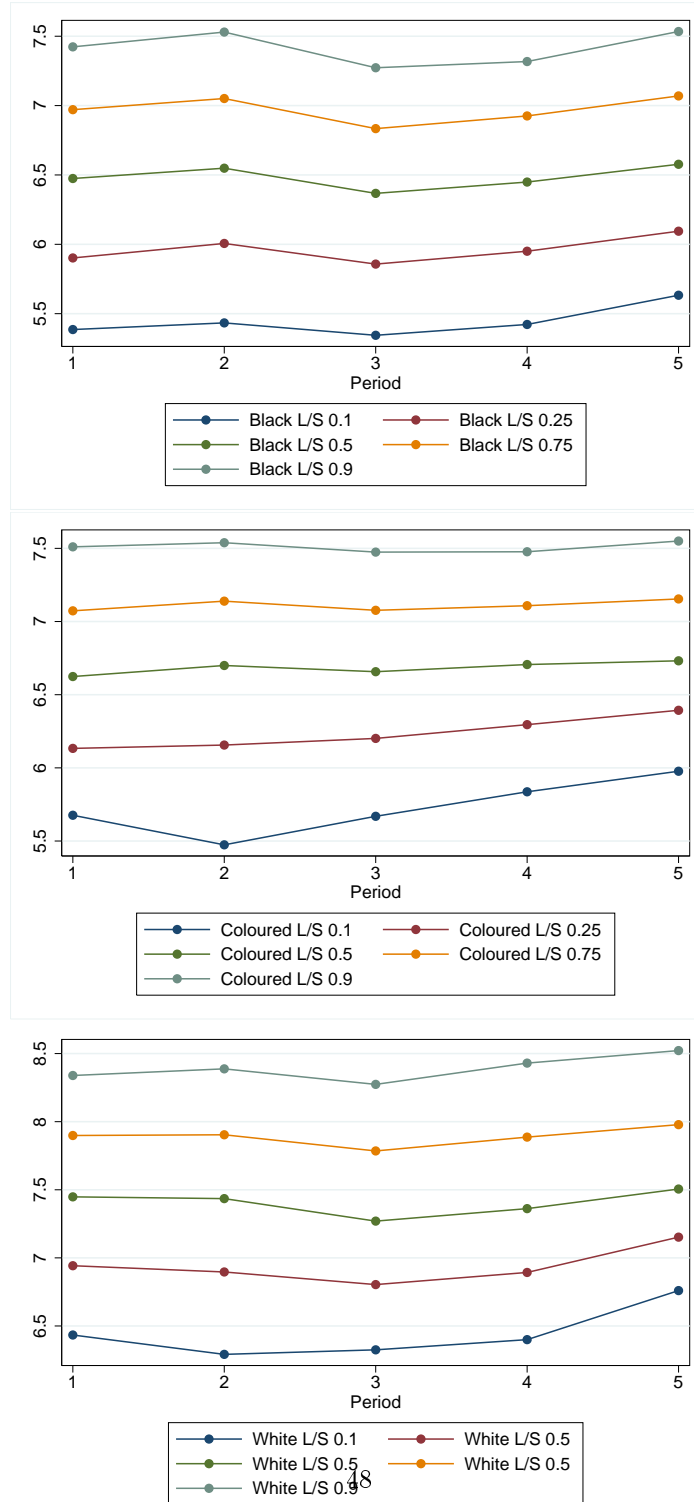
a change in predicted earnings over time. For the median and above there has been an increase in predicted earnings since the end of apartheid. For the bottom two quantiles of the conditional wage distribution there has been a decrease in predicted earnings between the mid 1990's and late 2000's. Visual inspection of the graph illustrates that the dispersion of predicted wages between quantiles has increased over time. The 90-10 and 75-25 measures of inequality²⁶ have increased over time. That is, the difference in predicted earnings of black women in the top 25% of the conditional earnings distribution and the bottom 25% of women in the conditional earnings distribution has increased between the first period and the last period.

This general pattern of change is the same for coloured and white women. For both coloured and white women in high-skilled occupations the predicted average wage for skilled women at the median and above increased between the mid 1990's and late 2000's. The increase in predicted earnings of white women at the median and above is smoother than for coloured and black women. For the bottom two quantiles of high-skilled coloured women it can be seen that there has been a decrease in predicted monthly wage by quantile over time. White highly-skilled women are the only race group that observe an increase in predicted wage at the bottom two quantiles between period 1 and period 5. In all graphs the dispersion of predicted monthly earnings, holding all else constant, has increased over time. The 75-25 and 90-10 measures of inequality have increased over all five periods of consideration irrespective of race. This corresponds with the findings of Machado and Mata (2005) who showed that wages tend to be more dispersed in higher skills groups.

Figure 2 presents the conditional quantile predicted earnings of women in lower skilled occupations. The first graph corresponds to predicted earnings of black women in low-skilled occupations, the second to earnings of coloured women in low-skilled occupations and the third to earnings of white women in low skilled occupations, respectively. The graphs on predicted earnings for women in lower skilled occupations show that the trends followed by the predicted earnings curves are not as systematic as they were for women in higher skilled occupations. This coincides with the observation made by Buchinsky (1994) that predicted earnings across quantiles do not follow the same pattern

²⁶Defined as the gap between the 90th percentile and 10th percentile, and the gap between the 75th percentile and 25th percentile.

Figure 2: Predicted Log of Monthly Wage By Race (Low Skilled Women)



of change for high and low skilled groups. The range of predicted earnings for low-skilled women is much lower than the range of earnings of women in high skilled occupations. The minimum log of monthly earnings for all race groups is 6.5 which corresponds to a monthly figure of R665.14. The maximum monthly earnings of women in low skilled occupations is natural logarithm 8.5 which corresponds to a monthly wage of R4914.77.

For all women in low-skilled occupations conditional predicted earnings between period 1 and period 5 have increased, but the pattern of change has been different across race between periods 1 and 5. For black women at all quantiles, predicted earnings increased by period two, decreased during periods 3 and 4 and increased again by period 5. For low-skilled coloured women at the top three quantiles predicted earnings remained relatively unchanged over time and the bottom two quantiles saw an increase between periods 1 and 5. For white women in low skilled occupations between period 1 and 3 there was a dip in predicted earnings and an increase again by period 5.

The 90-10 and 75-25 measure of earnings inequality for women in low skilled occupations illustrate that there has been a decrease in inequality of the earnings of women in these occupations between quantiles. That is, for each race group, the gap between the top 10% and bottom 10% of women in low-skilled occupations has decreased. This is a reassuring finding as it illustrates that for the lower end of income earners there has been a convergence in predicted (conditional) earnings. Thus, for these low skilled workers, within race groups, inequality is decreasing. However, between race groups it can be seen that at each quantile white women earn more than coloured women and black women, and that coloured women earn more than black women, holding all else constant. For women in these lower skilled occupations this racial discrepancy in predicted wages is visible over time. This may allude to differences in earnings across race but can not solely be attributed to racial discrimination since the mean characteristics of each race group differ as well.

Figures 1 and 2 jointly illustrate some important findings that bring to fruition the complexities involved in understanding the dynamics of the female labour force and earnings in South Africa. By analysing earnings differences by skills levels, it is found that the monetary range of monthly earnings is much higher for women in high-skilled jobs than women in low skilled ones. Condi-

tioning on the mean characteristics of each respective race group illustrates that even though apartheid is over, the knock on effects of its policies are still visible as can be seen by earnings of women. At higher and lower skills levels white women earn more than black and coloured women at each quantile. However, in the case of lower skilled women the gaps between black and white earnings per quantile is larger. This directly corresponds to the results of Van der Westhuizen et al. (2007) who show that black women earn less than their white counterparts, especially in lower skilled jobs.

Predicted wages for highly skilled working women of all races has become more dispersed over time. That is, for each racial group with high skills earnings inequality is increasing. The distribution of predicted conditional quantile earnings of women in high skilled occupations is more dispersed at the extremes of the income distribution (90-10 measure of earnings inequality) and less dispersed for the intermediate quantiles (75-25 measure of earnings inequality).

For women in lower skilled occupations, conditional on race, earnings inequality is decreasing. There is an increase in real wages and a convergence of wages at all quantiles of lower skilled jobs. The decrease in the inter-quantile gap of the real wages of women in lower skilled occupations may be explained by the findings of Borat and Leibbrandt (1999b) who show that the feminisation of the South African labour force is characterised by black women finding low skilled jobs. The increase in competition for low skilled jobs might be driving the wages between quantiles closer together for women in these occupations. This however does not explain why there is still inequality between races at the lower end of the income distribution and points directly to the question of discrimination. Nonetheless, the above results shed important light on the drivers of female earnings, provide key insights into the dynamics of earnings across the distribution as well the evolution of earnings inequality that female labour force participants with various skills sets are exposed to.

7 Conclusion

This study employs conditional quantile regression analysis to characterise the conditional female earnings distribution in post-apartheid South Africa. Standard mean regression analysis of the conditional earnings distribution is not an

adequate tool to analyse a distribution that is inherently heterogenous. This paper remedies this issue by analysing the conditional quantiles of earnings of women in South Africa. Specifically, returns to women in occupations of different skills levels and races are evaluated by focusing on the .10, .25, .50, .75 and .90 quantiles. Prior to modeling the conditional earnings distribution, the employment outcome of women is evaluated simultaneously with the labour force participation choice that women make. This accounts for the unique context of the South African labour market. The conditional quantile earnings equation corrects for the bias associated with the employment offer, after which measures of inter-quantile inequality are evaluated.

The econometric model has been carefully derived to aid the understanding of the dynamics of the female labour force by accounting for the binary participation choice that women make, and the simultaneous employment outcome that is exogenously determined. The conditional earnings quantile regressions take into account this multi-layered process while accounting for the bias associated with observing a non-random part of the female population.

Importantly, the participation - employment selection part of the model illustrates that black women and coloured women faced a positive probability of labour force participation in each period following the end of apartheid. Over time, education also played an important role in the labour force participation probability of individuals. As far as the employment outcome is concerned, increases in education affect the probability of employment positively. However, racial differences vary, and positive probabilities of participation do not translate to increases in employment probabilities for black women. Black women face a lower employment probability than white women over all time periods. Geographical location is another important determinant of the employment outcome, since demand for labour varies across province. These zero-earner dynamics are a crucial part of the earnings distribution.

With respect to the employment outcome the results presented are consistent with those reported in the literature that use both mean-regression analysis and quantile regression analysis. The results illustrate key stylised facts surrounding the evolution of earnings of women in the South African labour market. Black and coloured women earn a lower wage than white women for each quantile of the wage distribution. Education affects earnings positively, by quantile. This ef-

fect holds over time. Returns to being employed in various sectors have evolved, and so have returns to holding various occupations. For instance, agriculture has decreased in importance, whereas returns to holding a managerial position have increased.

By splitting the sample into lower skilled and higher skilled occupations it is seen that earnings in high skilled groups are more dispersed and inequality is increasing. Interestingly, earnings in lower skilled groups are less dispersed and inequality is decreasing conditional on race along with an increase in real earnings for women in these occupations. This could be because of increases in participation numbers of women over time, changes in minimum wage legislation and women finding jobs in lower skilled professions.

More importantly the quantile regressions illustrate that white women earn more than women of colour in both high and low skilled occupations. The gap between white women and non-white women is larger in lower skilled occupations. This may be interpreted as evidence of racial discrimination, but the characteristics of women of different races at different quantiles are different. The history of apartheid and its influence on the socio-economic and individual characteristics of women of varying race groups has its part to play in the inter-race wage gap. In order to detect if the gap between wages of women of different races within the same skills group is indeed a manifestation of racial discrimination a counter-factual analysis employing Oaxaca-Blinder decompositions could be used.

Overall, it can be concluded that the distribution of earnings for women in the South African labour market is a complex but approachable issue. Much light has been shed on the drivers of earnings at different points of the earnings distribution. The interplay between access to education and services and the choice of women to participate coupled with the legacy of apartheid make the alleviation of poverty and reduction in inequality a precarious issue for policy makers. The increase in inequality among women in higher skilled occupations is concerning. The consensus in the literature that increases in education will increase the inequality gap is equally as disconcerting. Despite this, one reassuring result is that women in lower skilled occupations have observed an increase real earnings.

In light of these findings there remains a question surrounding how the pattern of increasing inequality can be remedied in the context of women in the South African labour market. The first policy recommendation would look to increasing the skills levels of women at the bottom end of the income distribution, the second recommendation would be to create more viable education opportunities. The geographic segmentation and isolation of lower skilled groups also contributes to the pattern of inequality, and finding ways to create job opportunities for women in these remote locations could assist with poverty alleviation in the long term. The honest answer however, is that inequality is not going to decrease among the more highly skilled in South Africa, and any policy designed should be aimed at uplifting those women with lower skills sets and less opportunity via redistribution.

References

- Allanson, P., Atkins, J. P., and Hinks, T. (2002). No end to the racial wage hierarchy in South Africa? *Review of Development Economics*, 6(3):442–459.
- Bhorat, H. and Leibbrandt, M. (1999a). Correlates of vulnerability in the South African labour market. Working Papers 99027, University of Cape Town, Development Policy Research Unit.
- Bhorat, H. and Leibbrandt, M. (1999b). Modelling vulnerability and low earnings in the South African labour market. Working Papers 99032, University of Cape Town, Development Policy Research Unit.
- Buchinsky, M. (1994). Changes in the U.S Wage Structure 1963-1987: Application of Quantile Regression. *Econometrica: Journal of the Econometric Society*, pages 405–458.
- Buchinsky, M. (1998a). Recent Advances in quantile regression models: A practical guideline for empirical research. *Journal of human resources*, pages 88–126.
- Buchinsky, M. (1998b). The dynamics of changes in the female wage distribution in the U.S.A: a quantile regression approach. *Journal of applied econometrics*, 13(1):1–30.
- Casale, D. (2004). What has the feminisation of the labour market" bought" women in South Africa?: Trends in Labour Force Participation, Employment and Earnings, 1995-2001. *Journal of Interdisciplinary Economics*, 15(3/4):251–276.
- Casale, D. and Posel, D. (2002). The continued feminisation of the labour force in South Africa. *South African Journal of Economics*, 70(1):156–184.
- Case, A. and Deaton, A. (1998). School quality and educational outcomes in South Africa. Technical report.
- Fallon, P. and Lucas, R. (1998). South African Labor Markets: Adjustment and Inequalities. *World Bank Informal Discussion Papers on Aspects of the Economy of South Africa*, 12.
- Fitzenberger, B. and Kurz, C. (1997). New insights on earnings trends across skill groups in West Germany.

- Gould, W. (1992). Quantile regression with bootstrapped standard errors. *Stata Technical Bulletin*, 9:19–21.
- Heckman, J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. In *Annals of Economic and Social Measurement, Volume 5, number 4*, pages 475–492. NBER.
- Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):pp. 153–161.
- Hofmeyr, J. (1998). *Segmentation in the South African labour market*. Citeseer.
- Huber, M. and Melly, B. (2011). *Quantile regression in the presence of sample selection*. School of Economics and Political Science, Department of Economics, University of St. Gallen.
- Kingdon, G. and Knight, J. (2000). *Are Searching and non-searching unemployment distinct states when unemployment is high?: The case of South Africa*. Centre for the Study of African Economies, Institute of Economics and Statistics, University of Oxford.
- Kingdon, G. G. and Knight, J. (2004). Race and the incidence of unemployment in South Africa. *Review of Development Economics*, 8(2):198–222.
- Knight, J. B. and McGrath, M. D. (1987). *The erosion of apartheid in the South African labour market: measures and mechanisms*. University of Oxford Institute of Economics and Statistics.
- Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50.
- Leibbrandt, M., Woolard, I., Finn, A. and Argent, J. (2010). Trends in South African Income Distribution and Poverty since the Fall of Apartheid. *OECD Social, Employment and Migration Working Papers*, No. 101.
- Machado, J. A. and Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of applied Econometrics*, 20(4):445–465.
- Melly, B. and Huber, M. (2007). Sample selection, heteroscedasticity, and quantile regression. Technical report, Mimeo.

- Meng, C.-L. and Schmidt, P. (1985). On the cost of partial observability in the bivariate probit model. *International Economic Review*, 26(1):71–85.
- Mincer, J. and Polachek, S. (1974). Family investments in human capital: Earnings of women. In *Marriage, Family, Human Capital, and Fertility*, pages 76–110. NBER.
- Mohanty, M. S. (2010). Effects of Positive Attitude and Optimism on Employment: Evidence from the US data. *Journal of Socio-Economics*, 39(2):258–270.
- Moll, P. G. (1991). The decline of discrimination against colored people in South Africa, 1970 to 1980. *Journal of Development Economics*, 37(1):289–307.
- Mroz, T. A. (1987). The sensitivity of an empirical model of married women’s hours of work to economic and statistical assumptions. *Econometrica: Journal of the Econometric Society*, pages 765–799.
- Ntuli, M. (2007). Exploring gender wage discrimination in South Africa, 1995-2004: A quantile regression approach. *International Policy Center, Gerald R. Ford School of Public Policy, University of Michigan*, IPC Working Paper Series Number 56.
- Ntuli, M. and Wittenberg, M. (2013). Determinants of black women’s labour force participation in post-Apartheid South Africa. *Journal of African Economies*, 22(3):347–374.
- Rogers, W. H. (1992). Quantile regression standard errors. *Stata Technical Bulletin*, 9:16–19.
- Schultz, T. P. and Mwabu, G. (1997). Labor unions and the distribution of wages and employment in South Africa. *Indus. & Lab. Rel. Rev.*, 51:680.
- Van der Westhuizen, C., Goga, S., and Oosthuizen, M. (2007). Women in the South African Labour Market, 1995 - 2005. Working Papers 07118, University of Cape Town, Development Policy Research Unit.
- Winter, C. (1999). *Women workers in South Africa: Participation, pay and prejudice in the formal labour market*. ERIC Clearinghouse.
- Yu, K., Lu, Z., and Stander, J. (2003). Quantile regression: Applications and current research areas. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 52(3):331–350.

8 Appendix

Figure A.1: Imputed Versus Non-imputed Earnings Densities

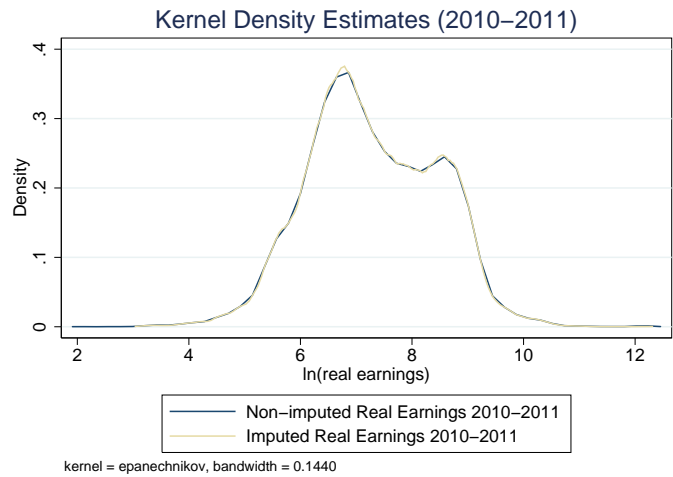
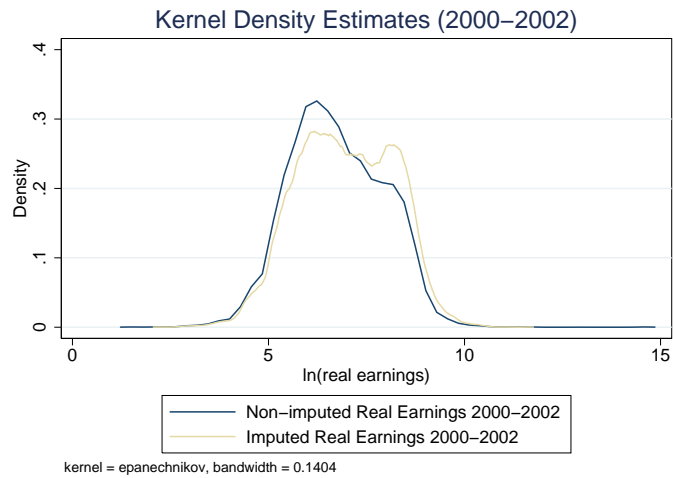
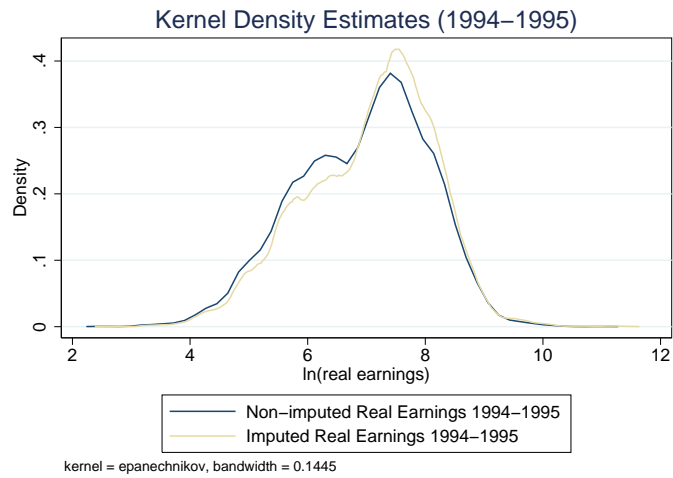


Table A.1: Summary Stats 1994-1995

	Black	Coloured	White
Industry			
Manufacturing	0.12 (0.322)	0.21 (0.406)	0.11 (0.312)
Agriculture	0.09 (0.286)	0.15 (0.356)	0.02 (0.123)
Mining	0.00 (0.066)	0.00 (0.026)	0.01 (0.087)
Utilities	0.00 (0.045)	0.00 (0.050)	0.01 (0.089)
Construction	0.01 (0.076)	0.01 (0.080)	0.02 (0.134)
Trade	0.19 (0.392)	0.21 (0.407)	0.22 (0.412)
Transport	0.01 (0.089)	0.01 (0.103)	0.05 (0.218)
Finance	0.02 (0.147)	0.05 (0.209)	0.21 (0.405)
Services	0.42 (0.494)	0.27 (0.445)	0.37 (0.482)
Occupation			
Plant/Machine Operator	0.04 (0.199)	0.10 (0.294)	0.01 (0.087)
Manager	0.02 (0.148)	0.01 (0.110)	0.07 (0.256)
Professional	0.05 (0.210)	0.03 (0.163)	0.10 (0.299)
Technician	0.12 (0.323)	0.07 (0.259)	0.19 (0.390)
Clerk	0.09 (0.282)	0.16 (0.370)	0.45 (0.498)
Agriculture	0.00 (0.061)	0.01 (0.082)	0.01 (0.082)
Artisan	0.05 (0.213)	0.04 (0.197)	0.03 (0.176)
Elementary Worker	0.30 (0.456)	0.27 (0.446)	0.03 (0.158)
Domestic Worker	0.34 (0.473)	0.31 (0.462)	0.12 (0.322)
Province			
Gauteng	0.14 (0.351)	0.08 (0.279)	0.40 (0.490)
Western Cape	0.02 (0.154)	0.60 (0.489)	0.19 (0.389)
Eastern Cape	0.18 (0.386)	0.13 (0.331)	0.07 (0.259)
Northern Cape	0.01 (0.088)	0.11 (0.314)	0.02 (0.153)
Freestate	0.07 (0.260)	0.02 (0.147)	0.07 (0.261)
Kwa-Zulu Natal	0.23 (0.423)	0.03 (0.176)	0.12 (0.321)
Northwest	0.10 (0.294)	0.01 (0.116)	0.05 (0.215)
Mpumalanga	0.08 (0.275)	0.01 (0.075)	0.05 (0.226)
Limpopo	0.16 (0.364)	0.00 (0.046)	0.02 (0.155)
Individual Characteristics			
Household Size	6.61 (2.819)	5.94 (2.660)	3.69 (1.511)
Number Children (0-6)	1.38 (1.243)	1.09 (1.077)	0.49 (0.777)
Number Children (7-15)	1.65 (1.390)	1.35 (1.252)	0.67 (0.939)
Age	33.19 (13.069)	34.11 (12.733)	37.72 (13.580)
Age-Squared	1272.51 (998.132)	1325.93 (976.681)	1606.98 (1083.396)
Years of Education	5.15 (4.291)	5.80 (4.101)	9.31 (4.377)
Married (or Cohabiting)	0.26 (0.440)	0.32 (0.468)	0.51 (0.500)
Household Characteristics			
Piped Water in Dwelling	0.26 (0.436)	0.75 (0.433)	0.98 (0.151)
Piped Water on Site	0.22 (0.416)	0.19 (0.389)	0.00 (0.045)
Public Tap (Main source of water)	0.16 (0.368)	0.03 (0.170)	0.00 (0.018)
Borehole (on or off site)	0.12 (0.325)	0.01 (0.081)	0.02 (0.124)
Water Tanker, Dam, Pool, etc	0.24 (0.428)	0.03 (0.163)	0.01 (0.075)
Brick House	0.57 (0.495)	0.76 (0.426)	0.88 (0.321)
Traditional Dwelling	0.26 (0.440)	0.00 (0.067)	0.00 (0.015)
Flat/Apartment	0.02 (0.136)	0.18 (0.388)	0.11 (0.311)
Informal Shack	0.10 (0.297)	0.03 (0.174)	0.00 (0.019)
Flush-toilet (on or off site)	0.40 (0.490)	0.83 (0.374)	1.00 (0.028)
Chemical Toilet (on or off site)	0.01 (0.116)	0.00 (0.055)	0.00 (0.014)
Pit Latrine	0.52 (0.500)	0.05 (0.227)	0.00 (0.022)
Bucket Toilet	0.07 (0.250)	0.11 (0.313)	0.00 (0.011)
Rural	0.65 (0.478)	0.17 (0.374)	0.08 (0.275)
Observations	90916	22025	18512

All summary statistics are computed using sampling weights.

Table A.2: Summary Stats 1997-1998

	Black	Coloured	White
Industry			
Manufacturing	0.11 (0.317)	0.20 (0.402)	0.11 (0.316)
Agriculture	0.08 (0.267)	0.13 (0.337)	0.01 (0.109)
Mining	0.00 (0.059)	0.00 (0.023)	0.01 (0.104)
Utilities	0.00 (0.057)	0.00 (0.041)	0.01 (0.108)
Construction	0.01 (0.091)	0.00 (0.064)	0.01 (0.120)
Trade	0.22 (0.414)	0.20 (0.400)	0.18 (0.386)
Transport	0.01 (0.115)	0.02 (0.146)	0.06 (0.236)
Finance	0.03 (0.181)	0.08 (0.265)	0.22 (0.416)
Services	0.24 (0.428)	0.21 (0.407)	0.36 (0.480)
Occupation			
Plant/Machine Operator	0.03 (0.164)	0.07 (0.248)	0.01 (0.106)
Manager	0.03 (0.159)	0.04 (0.202)	0.12 (0.320)
Professional	0.07 (0.247)	0.06 (0.228)	0.16 (0.369)
Technician	0.10 (0.299)	0.10 (0.304)	0.21 (0.407)
Clerk	0.08 (0.271)	0.17 (0.374)	0.35 (0.476)
Agriculture	0.02 (0.142)	0.01 (0.084)	0.00 (0.068)
Artisan	0.06 (0.246)	0.06 (0.234)	0.03 (0.156)
Elementary Worker	0.22 (0.411)	0.23 (0.424)	0.02 (0.125)
Domestic Worker	0.40 (0.490)	0.27 (0.442)	0.11 (0.309)
Province			
Gauteng	0.16 (0.365)	0.07 (0.260)	0.38 (0.485)
Western Cape	0.03 (0.165)	0.61 (0.488)	0.19 (0.391)
Eastern Cape	0.17 (0.379)	0.13 (0.338)	0.08 (0.266)
Northern Cape	0.01 (0.089)	0.11 (0.318)	0.02 (0.153)
Freestate	0.07 (0.252)	0.02 (0.142)	0.07 (0.256)
Kwa-Zulu Natal	0.23 (0.423)	0.03 (0.175)	0.13 (0.335)
Northwest	0.09 (0.290)	0.01 (0.108)	0.05 (0.216)
Mpumalanga	0.08 (0.274)	0.01 (0.079)	0.06 (0.230)
Limpopo	0.16 (0.362)	0.00 (0.056)	0.03 (0.159)
Individual Characteristics			
Household Size	6.61 (3.159)	5.83 (2.497)	3.91 (1.778)
Number Children (0-6)	1.40 (1.289)	1.06 (1.065)	0.55 (0.865)
Number Children (7-15)	1.69 (1.520)	1.39 (1.306)	0.76 (1.052)
Age	33.23 (12.903)	34.52 (12.772)	38.17 (13.626)
Age-Squared	1270.74 (981.200)	1354.60 (979.379)	1642.49 (1087.298)
Years of Education	5.27 (4.409)	6.16 (4.234)	9.50 (4.398)
Married (or Cohabiting)	0.26 (0.436)	0.33 (0.469)	0.51 (0.500)
Household Characteristics			
Piped Water in Dwelling	0.21 (0.408)	0.78 (0.413)	0.99 (0.118)
Piped Water on Site	0.26 (0.441)	0.16 (0.366)	0.01 (0.089)
Public Tap (Main source of water)	0.24 (0.430)	0.04 (0.191)	0.00 (0.031)
Borehole (on or off site)	0.06 (0.245)	0.00 (0.062)	0.00 (0.063)
Water Tanker, Dam, Pool, etc	0.22 (0.411)	0.02 (0.131)	0.00 (0.032)
Brick House	0.60 (0.490)	0.72 (0.447)	0.83 (0.375)
Traditional Dwelling	0.21 (0.408)	0.02 (0.127)	0.00 (0.063)
Flat/Apartment	0.02 (0.149)	0.20 (0.402)	0.13 (0.336)
Informal Shack	0.13 (0.334)	0.04 (0.201)	0.00 (0.018)
Flush-toilet (on or off site)	0.46 (0.498)	0.90 (0.306)	0.99 (0.072)
Chemical Toilet (on or off site)	0.01 (0.082)	0.00 (0.055)	0.00 (0.027)
Pit Latrine	0.49 (0.500)	0.04 (0.207)	0.00 (0.052)
Bucket Toilet	0.05 (0.208)	0.06 (0.231)	0.00 (0.042)
Rural	0.57 (0.495)	0.16 (0.364)	0.05 (0.220)
Observations	94775	13780	8204

All summary statistics are computed using sampling weights.

Table A.3: Summary Stats 2000-2002

	Black	Coloured	White
Industry			
Manufacturing	0.10 (0.298)	0.20 (0.401)	0.11 (0.315)
Agriculture	0.14 (0.351)	0.11 (0.317)	0.02 (0.155)
Mining	0.00 (0.049)	0.00 (0.045)	0.01 (0.100)
Utilities	0.00 (0.055)	0.00 (0.057)	0.00 (0.060)
Construction	0.01 (0.109)	0.01 (0.071)	0.01 (0.109)
Trade	0.25 (0.435)	0.20 (0.402)	0.19 (0.393)
Transport	0.01 (0.108)	0.02 (0.125)	0.05 (0.208)
Finance	0.04 (0.195)	0.09 (0.286)	0.24 (0.426)
Services	0.20 (0.399)	0.20 (0.402)	0.36 (0.480)
Occupation			
Plant/Machine Operator	0.03 (0.169)	0.08 (0.266)	0.01 (0.075)
Manager	0.01 (0.103)	0.02 (0.131)	0.12 (0.325)
Professional	0.03 (0.168)	0.03 (0.159)	0.13 (0.339)
Technician	0.10 (0.299)	0.11 (0.315)	0.23 (0.422)
Clerk	0.07 (0.257)	0.19 (0.392)	0.36 (0.480)
Agriculture	0.09 (0.287)	0.00 (0.064)	0.01 (0.108)
Artisan	0.05 (0.226)	0.04 (0.204)	0.02 (0.132)
Elementary Worker	0.26 (0.436)	0.24 (0.429)	0.02 (0.126)
Domestic Worker	0.36 (0.480)	0.29 (0.454)	0.11 (0.307)
Province			
Gauteng	0.17 (0.379)	0.09 (0.287)	0.37 (0.483)
Western Cape	0.03 (0.171)	0.63 (0.483)	0.20 (0.402)
Eastern Cape	0.17 (0.374)	0.11 (0.317)	0.08 (0.264)
Northern Cape	0.01 (0.089)	0.10 (0.300)	0.03 (0.163)
Freestate	0.06 (0.245)	0.02 (0.132)	0.08 (0.272)
Kwa-Zulu Natal	0.23 (0.421)	0.03 (0.173)	0.12 (0.328)
Northwest	0.09 (0.288)	0.01 (0.111)	0.05 (0.214)
Mpumalanga	0.08 (0.273)	0.00 (0.065)	0.05 (0.221)
Limpopo	0.15 (0.361)	0.00 (0.046)	0.02 (0.151)
Individual Characteristics			
Household Size	7.01 (3.419)	6.12 (2.799)	3.79 (1.670)
Number Children (0-6)	3.68 (3.769)	2.71 (3.068)	1.04 (2.015)
Number Children (7-15)	5.04 (5.076)	3.77 (4.034)	1.54 (2.574)
Age	33.34 (12.894)	34.82 (12.930)	38.84 (13.726)
Age-Squared	1277.80 (977.456)	1379.89 (988.790)	1696.80 (1099.689)
Years of Education	5.56 (4.498)	6.40 (4.322)	9.94 (4.371)
Married (or Cohabiting)	0.23 (0.421)	0.32 (0.468)	0.50 (0.500)
Household Characteristics			
Piped Water in Dwelling	0.22 (0.415)	0.75 (0.432)	0.98 (0.151)
Piped Water on Site	0.31 (0.464)	0.18 (0.387)	0.01 (0.105)
Public Tap (Main source of water)	0.20 (0.402)	0.05 (0.211)	0.00 (0.017)
Borehole (on or off site)	0.05 (0.227)	0.00 (0.068)	0.01 (0.102)
Water Tanker, Dam, Pool, etc	0.21 (0.406)	0.01 (0.118)	0.00 (0.038)
Brick House	0.60 (0.490)	0.72 (0.449)	0.81 (0.390)
Traditional Dwelling	0.19 (0.395)	0.01 (0.096)	0.00 (0.048)
Flat/Apartment	0.02 (0.152)	0.18 (0.382)	0.15 (0.361)
Informal Shack	0.14 (0.343)	0.06 (0.235)	0.00 (0.033)
Flush-toilet (on or off site)	0.43 (0.494)	0.90 (0.304)	1.00 (0.043)
Chemical Toilet (on or off site)	0.00 (0.062)	0.00 (0.045)	0.00 (0.000)
Pit Latrine	0.54 (0.498)	0.05 (0.216)	0.00 (0.043)
Bucket Toilet	0.03 (0.174)	0.05 (0.221)	0.00 (0.000)
Rural	0.55 (0.497)	0.17 (0.377)	0.08 (0.268)
Observations	130873	18924	11396

All summary statistics are computed using sampling weights.

Table A.4: Summary Stats 2005-2007

	Black	Coloured	White
Industry			
Manufacturing	0.09 (0.289)	0.17 (0.372)	0.10 (0.297)
Agriculture	0.08 (0.270)	0.08 (0.264)	0.03 (0.159)
Mining	0.00 (0.065)	0.00 (0.041)	0.01 (0.094)
Utilities	0.00 (0.069)	0.00 (0.068)	0.01 (0.073)
Construction	0.02 (0.147)	0.01 (0.112)	0.02 (0.137)
Trade	0.29 (0.454)	0.24 (0.427)	0.19 (0.395)
Transport	0.02 (0.135)	0.03 (0.171)	0.04 (0.194)
Finance	0.06 (0.235)	0.12 (0.322)	0.27 (0.442)
Services	0.21 (0.411)	0.23 (0.418)	0.34 (0.473)
Occupation			
Plant/Machine Operator	0.03 (0.166)	0.06 (0.236)	0.01 (0.090)
Manager	0.03 (0.164)	0.03 (0.181)	0.15 (0.359)
Professional	0.05 (0.214)	0.05 (0.208)	0.13 (0.337)
Technician	0.09 (0.288)	0.12 (0.322)	0.26 (0.437)
Clerk	0.09 (0.288)	0.22 (0.413)	0.32 (0.466)
Agriculture	0.04 (0.199)	0.01 (0.074)	0.01 (0.098)
Artisan	0.06 (0.241)	0.05 (0.215)	0.02 (0.151)
Elementary Worker	0.26 (0.438)	0.21 (0.405)	0.01 (0.119)
Domestic Worker	0.35 (0.477)	0.26 (0.441)	0.09 (0.281)
Province			
Gauteng	0.19 (0.391)	0.06 (0.234)	0.37 (0.483)
Western Cape	0.03 (0.180)	0.62 (0.486)	0.23 (0.424)
Eastern Cape	0.16 (0.363)	0.11 (0.314)	0.07 (0.263)
Northern Cape	0.01 (0.100)	0.11 (0.319)	0.02 (0.139)
Freestate	0.06 (0.243)	0.03 (0.180)	0.06 (0.242)
Kwa-Zulu Natal	0.22 (0.417)	0.04 (0.203)	0.11 (0.315)
Northwest	0.09 (0.280)	0.02 (0.137)	0.05 (0.227)
Mpumalanga	0.09 (0.282)	0.00 (0.056)	0.04 (0.201)
Limpopo	0.15 (0.360)	0.00 (0.037)	0.03 (0.170)
Individual Characteristics			
Household Size	6.39 (3.112)	5.89 (2.971)	3.75 (1.574)
Number Children (0-6)	3.20 (3.198)	2.21 (2.554)	0.78 (1.505)
Number Children (7-15)	4.39 (4.201)	2.87 (3.250)	1.11 (1.871)
Age	33.85 (12.999)	35.56 (13.205)	39.82 (13.997)
Age-Squared	1314.63 (987.534)	1438.66 (1017.389)	1781.23 (1128.581)
Years of Education	6.07 (4.616)	6.85 (4.376)	10.31 (4.193)
Married (or Cohabiting)	0.23 (0.418)	0.32 (0.466)	0.51 (0.500)
Observations	134742	22712	9404

All summary statistics are computed using sampling weights.

Table A.5: Summary Stats 2010-2011

	Black	Coloured	White
Industry			
Manufacturing	0.09 (0.279)	0.16 (0.363)	0.11 (0.309)
Agriculture	0.04 (0.189)	0.06 (0.228)	0.02 (0.131)
Mining	0.01 (0.077)	0.00 (0.028)	0.01 (0.107)
Utilities	0.00 (0.051)	0.00 (0.068)	0.01 (0.077)
Construction	0.02 (0.140)	0.01 (0.111)	0.03 (0.163)
Trade	0.26 (0.438)	0.24 (0.430)	0.18 (0.385)
Transport	0.02 (0.149)	0.02 (0.154)	0.04 (0.206)
Finance	0.08 (0.273)	0.13 (0.341)	0.25 (0.435)
Services	0.28 (0.448)	0.26 (0.438)	0.35 (0.478)
Occupation			
Plant/Machine Operator	0.03 (0.169)	0.05 (0.217)	0.00 (0.069)
Manager	0.03 (0.170)	0.06 (0.230)	0.18 (0.383)
Professional	0.04 (0.188)	0.05 (0.227)	0.14 (0.347)
Technician	0.12 (0.321)	0.13 (0.336)	0.26 (0.436)
Clerk	0.12 (0.329)	0.22 (0.417)	0.30 (0.459)
Agriculture	0.00 (0.059)	0.00 (0.044)	0.01 (0.080)
Artisan	0.04 (0.186)	0.04 (0.191)	0.01 (0.098)
Elementary Worker	0.26 (0.439)	0.19 (0.395)	0.02 (0.135)
Domestic Worker	0.36 (0.481)	0.25 (0.435)	0.09 (0.279)
Province			
Gauteng	0.20 (0.401)	0.08 (0.265)	0.39 (0.488)
Western Cape	0.04 (0.194)	0.65 (0.478)	0.19 (0.390)
Eastern Cape	0.15 (0.354)	0.12 (0.324)	0.09 (0.287)
Northern Cape	0.02 (0.126)	0.09 (0.291)	0.02 (0.147)
Freestate	0.06 (0.241)	0.02 (0.131)	0.07 (0.253)
Kwa-Zulu Natal	0.23 (0.418)	0.03 (0.164)	0.11 (0.309)
Northwest	0.08 (0.267)	0.01 (0.108)	0.06 (0.236)
Mpumalanga	0.09 (0.281)	0.01 (0.080)	0.06 (0.236)
Limpopo	0.14 (0.352)	0.00 (0.054)	0.02 (0.130)
Individual Characteristics			
Household Size	6.08 (3.306)	5.59 (2.652)	3.38 (1.642)
Number Children (0-6)	3.78 (4.144)	2.76 (3.281)	1.06 (2.175)
Number Children (7-15)	4.50 (4.578)	3.29 (3.441)	1.38 (2.488)
Age	34.40 (13.200)	36.27 (13.365)	40.88 (13.928)
Age-Squared	1357.28 (1010.326)	1494.12 (1037.018)	1865.07 (1143.217)
Years of Education	6.59 (4.656)	7.34 (4.424)	10.60 (4.201)
Married (or Cohabiting)	0.22 (0.413)	0.32 (0.466)	0.51 (0.500)
Observations	73276	9601	5945

All summary statistics are computed using sampling weights.

Table A.6: Bi-Probit Results: Broad Definition of Employment

	1994-1995	1997-1998	2000-2002	2005-2007	2010-2011
Dependent Variable: Employment Outcome					
Household Size	-0.097*** (0.011)	-0.022*** (0.004)	-0.059*** (0.003)	-0.051*** (0.004)	-0.025*** (0.005)
Num Child (0-6)	0.078** (0.028)	-0.029*** (0.008)	-0.007** (0.003)	-0.017*** (0.003)	-0.011*** (0.003)
Num Child (7-15)	0.048* (0.023)	-0.023** (0.007)	-0.001 (0.002)	-0.001 (0.002)	-0.017*** (0.003)
Age	0.127*** (0.020)	0.209*** (0.005)	0.219*** (0.004)	0.199*** (0.004)	0.276*** (0.004)
Age-Squared	-0.001* (0.000)	-0.003*** (0.000)	-0.003*** (0.000)	-0.002*** (0.000)	-0.003*** (0.000)
Years of Educ	0.028*** (0.008)	0.056*** (0.002)	0.047*** (0.002)	0.058*** (0.002)	0.082*** (0.002)
Married	-0.011 (0.060)	-0.255*** (0.016)	-0.228*** (0.016)	-0.175*** (0.016)	-0.150*** (0.016)
Black	-2.062*** (0.230)	-0.314*** (0.024)	-0.201*** (0.024)	-0.090** (0.033)	-0.185*** (0.027)
Coloured	-1.119*** (0.227)	0.082** (0.031)	0.038 (0.033)	0.033 (0.041)	-0.024 (0.035)
Western Cape	0.310** (0.109)	-0.135*** (0.029)	0.033 (0.030)	0.086* (0.038)	0.099** (0.032)
Eastern Cape	-0.074 (0.083)	-0.316*** (0.027)	-0.038 (0.026)	-0.145*** (0.026)	-0.194*** (0.027)
Northern Cape	-0.514*** (0.126)	-0.204*** (0.036)	-0.171*** (0.036)	-0.233*** (0.033)	-0.115*** (0.034)
Freestate	0.039 (0.081)	-0.013 (0.029)	-0.025 (0.028)	-0.184*** (0.029)	-0.119*** (0.028)
Kwa-Zulu Natal	0.137 (0.073)	-0.165*** (0.024)	-0.042 (0.024)	-0.074** (0.026)	-0.061* (0.024)
North-West	-0.113 (0.084)	-0.138*** (0.027)	-0.229*** (0.027)	-0.261*** (0.030)	-0.261*** (0.029)
Mpumalanga	-0.336*** (0.081)	-0.062* (0.029)	0.012 (0.028)	-0.083** (0.030)	-0.025 (0.029)
Limpopo	-0.279** (0.100)	-0.201*** (0.029)	-0.140*** (0.029)	-0.352*** (0.029)	-0.235*** (0.027)
Dependent Variable: Labour Force Participation					
Household Size	0.005 (0.010)	0.006 (0.004)	-0.017*** (0.003)	-0.018*** (0.004)	-0.003 (0.004)
Num Child (0-6)	-0.127*** (0.022)	-0.053*** (0.008)	-0.011*** (0.003)	-0.015*** (0.003)	-0.011*** (0.003)
Num Child (7-15)	-0.038* (0.017)	-0.032*** (0.007)	-0.003 (0.002)	-0.003 (0.002)	-0.016*** (0.002)
Age	0.079*** (0.011)	0.173*** (0.004)	0.149*** (0.004)	0.137*** (0.005)	0.283*** (0.004)
Age-Squared	-0.001*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.003*** (0.000)
Years of Educ	0.079*** (0.004)	0.059*** (0.002)	0.062*** (0.002)	0.072*** (0.002)	0.087*** (0.002)
Married	-0.378*** (0.033)	-0.386*** (0.016)	-0.353*** (0.018)	-0.252*** (0.017)	-0.300*** (0.017)
Black	0.646*** (0.047)	0.290*** (0.033)	0.329*** (0.037)	0.395*** (0.046)	0.141*** (0.037)
Coloured	0.308*** (0.049)	0.232*** (0.031)	0.234*** (0.036)	0.256*** (0.043)	0.098** (0.033)
Western Cape	-0.148** (0.051)	-0.296*** (0.028)	-0.173*** (0.031)	-0.063 (0.029)	-0.060 (0.027)
Eastern Cape	-0.293*** (0.074)	-0.392*** (0.037)	-0.205*** (0.040)	-0.294*** (0.037)	-0.363*** (0.034)
Northern Cape	-0.276*** (0.057)	-0.293*** (0.031)	-0.259*** (0.033)	-0.216*** (0.033)	-0.216*** (0.028)
Freestate	-0.066 (0.046)	-0.065* (0.026)	-0.067* (0.029)	-0.201*** (0.029)	-0.172*** (0.024)
Kwa-Zulu Natal	-0.280*** (0.062)	-0.214*** (0.029)	-0.156*** (0.032)	-0.253*** (0.033)	-0.312*** (0.029)
North-West	-0.323*** (0.065)	-0.150*** (0.030)	-0.265*** (0.033)	-0.281*** (0.033)	-0.378*** (0.028)
Mpumalanga	-0.163* (0.074)	-0.136*** (0.031)	-0.106** (0.036)	-0.076* (0.033)	-0.021 (0.027)
Limpopo	-0.297*** (0.046)	-0.238*** (0.019)	0.065 (0.021)	-0.105** (0.021)	-0.413*** (0.027)
Piped Water	0.018 (0.046)	0.009 (0.019)	0.110*** (0.021)		
public tap	-0.404*** (0.063)	0.057* (0.025)	0.092** (0.032)		
Borehole	-0.257*** (0.065)	0.001 (0.035)	0.059 (0.042)		
Natural Source	-0.515*** (0.062)	-0.034 (0.029)	0.015 (0.035)		
Traditional Dwelling	-0.287*** (0.052)	0.070** (0.022)	-0.032 (0.027)		
Flat/Cluster	0.054 (0.043)	0.038 (0.023)	0.095** (0.032)		
Informal	-0.042 (0.070)	0.126*** (0.024)	0.068* (0.028)		
Chemical Toilet	-0.205 (0.152)	-0.043 (0.074)	0.477** (0.182)		
Pit Toilet	0.027 (0.055)	0.006 (0.019)	-0.002 (0.029)		
Bucket	-0.113 (0.061)	-0.024 (0.030)	-0.030 (0.038)		
Rural	0.082* (0.041)	-0.103*** (0.019)	-0.111*** (0.025)		
Rho	0.081	0.994	0.980	0.988	0.994
N	29395	49519	48812	87890	55818

Standard errors in parentheses. All results are estimated using sampling weights.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.7: Employment Probit Results

	1994-1995	1997-1998	2000-2002	2005-2007	2010-2011
Household Size	-0.042*** (0.004)	-0.029*** (0.004)	-0.051*** (0.003)	-0.056*** (0.004)	-0.029*** (0.005)
Num Child (0-6)	-0.043*** (0.007)	-0.025** (0.008)	-0.014*** (0.002)	-0.019*** (0.003)	-0.014*** (0.003)
Num Child (7-15)	-0.019** (0.007)	-0.019** (0.007)	-0.004* (0.002)	-0.001 (0.002)	-0.017*** (0.003)
Age	0.168*** (0.004)	0.189*** (0.004)	0.202*** (0.003)	0.195*** (0.004)	0.254*** (0.004)
Age-Squared	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.003*** (0.000)
Years of Educ	0.059*** (0.002)	0.053*** (0.002)	0.039*** (0.002)	0.056*** (0.002)	0.081*** (0.003)
Married	-0.243*** (0.014)	-0.249*** (0.015)	-0.200*** (0.012)	-0.166*** (0.016)	-0.152*** (0.017)
Black	-0.136*** (0.018)	-0.317*** (0.023)	-0.193*** (0.020)	-0.127*** (0.031)	-0.197*** (0.026)
Coloured	0.129*** (0.022)	0.060 (0.031)	0.003 (0.027)	-0.007 (0.040)	-0.058 (0.035)
Western Cape	-0.036 (0.024)	-0.110*** (0.029)	0.085*** (0.026)	0.090* (0.038)	0.101** (0.032)
Eastern Cape	-0.350*** (0.022)	-0.395*** (0.027)	-0.070*** (0.020)	-0.148*** (0.027)	-0.192*** (0.028)
Northern Cape	-0.375*** (0.030)	-0.175*** (0.036)	-0.138*** (0.030)	-0.252*** (0.034)	-0.134*** (0.036)
Freestate	0.083** (0.027)	0.025 (0.028)	-0.013 (0.023)	-0.196*** (0.029)	-0.117*** (0.029)
Kwa-Zulu Natal	-0.224*** (0.021)	-0.155*** (0.024)	0.021 (0.019)	-0.078** (0.026)	-0.057* (0.024)
North-West	-0.245*** (0.027)	-0.152*** (0.027)	-0.227*** (0.022)	-0.264*** (0.031)	-0.281*** (0.031)
Mpumalanga	-0.270*** (0.028)	-0.050 (0.028)	0.026 (0.024)	-0.090** (0.031)	-0.029 (0.030)
Limpopo	-0.413*** (0.028)	-0.211*** (0.028)	-0.147*** (0.022)	-0.349*** (0.029)	-0.211*** (0.028)
N	70341	56246	83916	87890	55818

Standard errors in parentheses. All results estimated using sampling weights.

Dependent variable: Employment Status.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.8: Imputed Quantile Regression-With Heckman Correction: 1994-1995

	.10	.25	.5	.75	.9	Mean Reg
Log Hrs Worked	0.451*** (0.029)	0.340*** (0.028)	0.269*** (0.018)	0.190*** (0.020)	0.142*** (0.021)	0.308*** (0.062)
Num Child (0-6)	-0.022** (0.008)	-0.001 (0.007)	0.004 (0.005)	0.006 (0.007)	0.005 (0.007)	0.013 (0.025)
Num Child (7-15)	-0.026** (0.008)	-0.021*** (0.005)	-0.015** (0.005)	-0.007 (0.004)	-0.006 (0.006)	-0.006 (0.021)
λ^1	-0.939* (0.457)	-0.599 (0.343)	-0.540 (0.312)	-0.770** (0.243)	-0.725 (0.539)	-
λ^2	2.672* (1.354)	1.804 (1.060)	1.507 (0.943)	2.382** (0.778)	2.560 (1.697)	-
λ^3	-3.477 (1.778)	-2.571 (1.401)	-1.994 (1.231)	-3.280** (1.115)	-3.729 (2.263)	-
λ^4	1.846 (1.004)	1.466 (0.788)	1.069 (0.686)	1.855** (0.676)	2.183 (1.265)	-
λ^5	-0.333 (0.199)	-0.280 (0.159)	-0.190 (0.135)	-0.357* (0.141)	-0.433 (0.247)	-
Age	0.040*** (0.007)	0.042*** (0.005)	0.043*** (0.004)	0.040*** (0.004)	0.045*** (0.005)	0.020 (0.019)
Age-Squared	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000 (0.000)
Years of Educ	0.097*** (0.003)	0.101*** (0.002)	0.105*** (0.003)	0.099*** (0.002)	0.087*** (0.002)	0.076*** (0.008)
Married	-0.045** (0.016)	-0.043*** (0.013)	-0.029** (0.010)	-0.013 (0.013)	-0.007 (0.011)	-0.043 (0.048)
Black	-0.198** (0.067)	-0.239*** (0.039)	-0.222*** (0.038)	-0.218*** (0.029)	-0.284*** (0.059)	0.022 (0.172)
Coloured	-0.087* (0.040)	-0.116*** (0.026)	-0.109*** (0.023)	-0.142*** (0.024)	-0.197*** (0.033)	0.245* (0.117)
Western Cape	-0.093*** (0.019)	-0.131*** (0.019)	-0.138*** (0.019)	-0.123*** (0.017)	-0.127*** (0.021)	-0.746*** (0.092)
Eastern Cape	-0.357*** (0.030)	-0.305*** (0.028)	-0.275*** (0.016)	-0.227*** (0.016)	-0.219*** (0.023)	-0.303*** (0.082)
Northern Cape	-0.454*** (0.037)	-0.443*** (0.030)	-0.404*** (0.022)	-0.382*** (0.026)	-0.425*** (0.028)	-1.112*** (0.117)
Freestate	-0.801*** (0.034)	-0.720*** (0.031)	-0.548*** (0.027)	-0.385*** (0.019)	-0.391*** (0.018)	-0.720*** (0.092)
Kwa-Zulu Natal	-0.221*** (0.016)	-0.204*** (0.018)	-0.192*** (0.014)	-0.161*** (0.016)	-0.131*** (0.019)	-0.437*** (0.079)
North-West	-0.310*** (0.027)	-0.301*** (0.027)	-0.246*** (0.023)	-0.222*** (0.019)	-0.186*** (0.034)	-0.383*** (0.101)
Mpumalanga	-0.183*** (0.031)	-0.195*** (0.028)	-0.148*** (0.020)	-0.079*** (0.018)	-0.066* (0.031)	-0.316** (0.113)
Limpopo	-0.194*** (0.041)	-0.164*** (0.030)	-0.140*** (0.024)	-0.043 (0.032)	-0.004 (0.019)	-0.472*** (0.116)
Manager	-0.280** (0.092)	-0.174** (0.055)	0.076* (0.033)	0.405*** (0.042)	0.865*** (0.090)	-0.205 (0.193)
Professional	0.038 (0.043)	0.020 (0.041)	0.104*** (0.030)	0.224*** (0.022)	0.351*** (0.053)	-0.444* (0.174)
Technician	0.071 (0.041)	0.078* (0.033)	0.173*** (0.023)	0.312*** (0.020)	0.421*** (0.040)	0.231 (0.142)
Clerk	-0.185*** (0.036)	-0.209*** (0.036)	-0.148*** (0.016)	-0.045 (0.025)	0.023 (0.035)	-0.340* (0.135)
Agricultural	-0.906*** (0.090)	-0.844*** (0.080)	-0.479*** (0.093)	-0.026 (0.104)	0.550*** (0.132)	0.304 (0.325)
Artisan	-0.536*** (0.056)	-0.356*** (0.042)	-0.153 (0.027)	-0.038 (0.037)	0.067 (0.049)	-0.225 (0.161)
Elementary	-1.048*** (0.046)	-1.021*** (0.035)	-0.795** (0.029)	-0.493*** (0.025)	-0.255** (0.034)	-0.369** (0.134)
Domestic Service	-0.728*** (0.031)	-0.699*** (0.034)	-0.526*** (0.019)	-0.321*** (0.023)	-0.147*** (0.039)	-0.445*** (0.121)
Agri (Sector)	0.354*** (0.043)	0.266*** (0.024)	0.087** (0.029)	-0.105*** (0.024)	-0.074 (0.042)	-0.248* (0.102)
Mining	0.609*** (0.102)	0.450*** (0.068)	0.345*** (0.081)	0.414*** (0.078)	0.451*** (0.109)	0.088 (0.311)
Utilities	0.308 (0.122)	0.415*** (0.081)	0.339*** (0.071)	0.290*** (0.074)	0.188* (0.083)	0.500 (0.422)
Construction	0.392*** (0.062)	0.346*** (0.050)	0.290*** (0.055)	0.194* (0.083)	0.197** (0.065)	0.315 (0.253)
Trade	0.383*** (0.036)	0.371*** (0.021)	0.289*** (0.014)	0.189*** (0.015)	0.156*** (0.017)	0.167* (0.075)
Transport	0.611*** (0.078)	0.558*** (0.039)	0.469*** (0.026)	0.357*** (0.043)	0.298*** (0.038)	0.250 (0.189)
Finance	0.541*** (0.043)	0.454*** (0.025)	0.352*** (0.010)	0.236*** (0.017)	0.201*** (0.021)	0.210 (0.119)
Services	0.471*** (0.032)	0.418*** (0.023)	0.313*** (0.016)	0.174*** (0.014)	0.113*** (0.023)	0.222** (0.071)
λ_{mean}	-	-	-	-	-	0.763 (1.326)
N	25930	25930	25930	25930	25930	16006

Standard errors in parentheses. All results estimated using sampling weights.
 Dependent variable: Natural Logarithm of Imputed Real Earnings (2000 prices).
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.9: Imputed Quantile Regression-With Heckman Correction: 1997-1998

	.10	.25	.5	.75	.9	Mean Reg
Log Hrs Worked	0.172*** (0.040)	0.062*** (0.011)	0.020 (0.013)	0.002 (0.006)	0.007 (0.007)	0.041*** (0.012)
Num Child (0-6)	-0.061** (0.019)	-0.027** (0.009)	-0.011 (0.011)	-0.000 (0.012)	0.007 (0.015)	-0.044** (0.016)
Num Child (7-15)	-0.057** (0.021)	-0.042*** (0.008)	-0.020* (0.010)	-0.004 (0.011)	-0.002 (0.013)	-0.045*** (0.013)
λ^1	6.025 (5.382)	0.465 (2.033)	1.413 (1.693)	2.593 (3.254)	-3.943 (3.314)	-
λ^2	-8.356 (8.942)	0.523 (3.282)	-1.447 (2.869)	-4.649 (5.478)	5.486 (5.985)	-
λ^3	5.960 (7.313)	-1.094 (2.605)	0.177 (2.485)	3.171 (4.632)	-4.215 (4.966)	-
λ^4	-1.908 (2.835)	0.695 (0.979)	0.339 (1.027)	-0.894 (1.849)	1.622 (1.930)	-
λ^5	0.225 (0.416)	-0.136 (0.139)	-0.104 (0.161)	0.080 (0.279)	-0.238 (0.283)	-
Age	0.180** (0.070)	0.125*** (0.024)	0.055 (0.039)	0.013 (0.043)	-0.004 (0.044)	0.179*** (0.046)
Age-Squared	-0.002* (0.001)	-0.001*** (0.000)	-0.001 (0.000)	-0.000 (0.001)	0.000 (0.001)	-0.002*** (0.001)
Years of Educ	0.125*** (0.020)	0.109*** (0.008)	0.088*** (0.010)	0.071*** (0.011)	0.060*** (0.012)	0.116*** (0.013)
Married	-0.223* (0.087)	-0.104*** (0.031)	0.010 (0.047)	0.054 (0.054)	0.079 (0.049)	-0.200*** (0.055)
Black	-0.492*** (0.118)	-0.464*** (0.044)	-0.397*** (0.064)	-0.371*** (0.072)	-0.399*** (0.085)	-0.603*** (0.075)
Coloured	-0.238*** (0.061)	-0.206*** (0.035)	-0.218*** (0.024)	-0.303*** (0.030)	-0.384*** (0.043)	-0.261*** (0.055)
Western Cape	-0.225*** (0.067)	-0.142*** (0.039)	-0.130*** (0.035)	-0.067 (0.048)	-0.130* (0.053)	-0.168** (0.055)
Eastern Cape	-0.657*** (0.118)	-0.486*** (0.040)	-0.324*** (0.074)	-0.169* (0.079)	-0.165* (0.083)	-0.599*** (0.078)
Northern Cape	-0.634*** (0.090)	-0.522*** (0.045)	-0.445*** (0.046)	-0.319*** (0.051)	-0.246*** (0.062)	-0.511*** (0.067)
Freestate	-0.735*** (0.061)	-0.566*** (0.031)	-0.471*** (0.033)	-0.304*** (0.038)	-0.234*** (0.044)	-0.463*** (0.044)
Kwa-Zulu Natal	-0.397*** (0.062)	-0.298*** (0.027)	-0.249*** (0.043)	-0.180*** (0.040)	-0.163** (0.054)	-0.332*** (0.054)
North-West	-0.475*** (0.063)	-0.327*** (0.043)	-0.225*** (0.038)	-0.147*** (0.042)	-0.164*** (0.044)	-0.323*** (0.053)
Mpumalanga	-0.338*** (0.046)	-0.269*** (0.034)	-0.235*** (0.024)	-0.174*** (0.028)	-0.203*** (0.035)	-0.238*** (0.046)
Limpopo	-0.515*** (0.081)	-0.349*** (0.033)	-0.288*** (0.045)	-0.129* (0.052)	-0.144** (0.050)	-0.520*** (0.061)
Manager	-0.256** (0.081)	-0.045 (0.068)	0.078 (0.041)	0.246*** (0.043)	0.377*** (0.090)	-0.009 (0.092)
Professional	0.170 (0.088)	0.178** (0.055)	0.175*** (0.036)	0.190*** (0.037)	0.196*** (0.057)	0.172* (0.082)
Technician	0.002 (0.053)	0.048 (0.063)	0.078* (0.033)	0.120*** (0.036)	0.105* (0.046)	-0.025 (0.079)
Clerk	-0.020 (0.049)	-0.088 (0.062)	-0.118*** (0.031)	-0.080* (0.040)	-0.086 (0.053)	-0.161* (0.076)
Agricultural	-0.392** (0.140)	-0.440*** (0.075)	-0.328*** (0.062)	-0.216 (0.111)	0.027 (0.169)	-0.395*** (0.110)
Artisan	-0.544*** (0.064)	-0.388*** (0.064)	-0.219*** (0.036)	-0.139*** (0.041)	-0.118* (0.047)	-0.399*** (0.078)
Elementary	-0.610*** (0.042)	-0.517*** (0.048)	-0.380*** (0.024)	-0.270*** (0.033)	-0.267*** (0.050)	-0.509*** (0.072)
Domestic Service	-0.715*** (0.059)	-0.743*** (0.044)	-0.655*** (0.022)	-0.508*** (0.027)	-0.381*** (0.047)	-0.688*** (0.068)
Agri (Sector)	-0.437*** (0.091)	-0.358*** (0.024)	-0.407*** (0.031)	-0.452*** (0.033)	-0.319*** (0.081)	-0.345*** (0.049)
Mining	0.533** (0.163)	0.519*** (0.078)	0.499*** (0.056)	0.408*** (0.059)	0.370** (0.113)	0.579*** (0.144)
Utilities	0.440*** (0.066)	0.251** (0.085)	0.406* (0.162)	0.383*** (0.096)	0.510 (0.343)	-0.029 (0.205)
Construction	0.348 (0.137)	0.286*** (0.087)	0.129 (0.070)	0.161 (0.085)	0.125 (0.097)	0.341** (0.122)
Trade	-0.108** (0.041)	0.053 (0.026)	0.115*** (0.020)	0.142*** (0.022)	0.192*** (0.027)	0.003 (0.033)
Transport	0.339*** (0.084)	0.419*** (0.053)	0.471*** (0.055)	0.489*** (0.043)	0.376*** (0.064)	0.436*** (0.094)
Finance	0.370*** (0.045)	0.346*** (0.032)	0.356*** (0.028)	0.337*** (0.031)	0.318*** (0.039)	0.423*** (0.060)
Services	0.325*** (0.041)	0.427*** (0.025)	0.457*** (0.029)	0.407*** (0.024)	0.363*** (0.030)	0.463*** (0.040)
λ_{mean}	-	-	-	-	-	0.863** (0.299)
N	16453	16453	16453	16453	16453	11096

Standard errors in parentheses. All results estimated using sampling weights.
 Dependent variable: Natural Logarithm of Imputed Real Earnings (2000 prices).
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.10: Imputed Quantile Regression-With Heckman Correction: 2000-2002

	.10	.25	.5	.75	.9	Mean Reg
Log Hrs Worked	0.203*** (0.020)	0.111*** (0.013)	0.033*** (0.006)	0.018*** (0.002)	0.012* (0.005)	0.054*** (0.004)
Num Child (0-6)	-0.009* (0.004)	-0.009*** (0.002)	-0.005* (0.002)	-0.005* (0.002)	-0.008* (0.004)	-0.006** (0.002)
Num Child (7-15)	-0.007* (0.003)	-0.008*** (0.001)	-0.010*** (0.002)	-0.006*** (0.002)	-0.006** (0.002)	-0.010*** (0.002)
λ^1	2.873 (2.337)	1.469 (1.442)	1.988 (1.155)	-0.273 (1.612)	-4.134*** (0.937)	-
λ^2	-6.377 (4.316)	-3.227 (2.816)	-3.780 (2.275)	0.200 (2.853)	7.139*** (1.689)	-
λ^3	5.636 (3.657)	2.804 (2.574)	3.103 (2.081)	0.024 (2.388)	-5.532*** (1.513)	-
λ^4	-2.238 (1.430)	-1.101 (1.094)	-1.156 (0.888)	-0.084 (0.932)	1.992** (0.636)	-
λ^5	0.327 (0.209)	0.160 (0.173)	0.159 (0.143)	0.027 (0.136)	-0.264** (0.098)	-
Age	0.028 (0.017)	0.041*** (0.010)	0.054*** (0.009)	0.061*** (0.010)	0.081*** (0.012)	0.070*** (0.010)
Age-Squared	-0.000 (0.000)	-0.000** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Years of Educ	0.067*** (0.004)	0.071*** (0.004)	0.073*** (0.003)	0.077*** (0.002)	0.085*** (0.002)	0.073*** (0.003)
Married	0.022 (0.018)	0.007 (0.016)	-0.004 (0.013)	-0.000 (0.015)	0.000 (0.022)	-0.022 (0.013)
Black	-0.564*** (0.032)	-0.470*** (0.024)	-0.452*** (0.020)	-0.496*** (0.020)	-0.543*** (0.030)	-0.523*** (0.021)
Coloured	-0.392*** (0.039)	-0.258*** (0.027)	-0.240*** (0.027)	-0.285*** (0.021)	-0.328*** (0.030)	-0.310*** (0.025)
Western Cape	-0.048 (0.042)	-0.009 (0.025)	-0.034 (0.019)	-0.032* (0.015)	-0.082*** (0.022)	-0.006 (0.023)
Eastern Cape	-0.512*** (0.030)	-0.466*** (0.024)	-0.387*** (0.017)	-0.323*** (0.015)	-0.304*** (0.026)	-0.442*** (0.020)
Northern Cape	-0.448*** (0.038)	-0.366*** (0.030)	-0.332*** (0.028)	-0.351*** (0.022)	-0.357*** (0.044)	-0.414*** (0.028)
Freestate	-0.626*** (0.035)	-0.594*** (0.030)	-0.542*** (0.020)	-0.448*** (0.021)	-0.432*** (0.024)	-0.556*** (0.021)
Kwa-Zulu Natal	-0.210*** (0.033)	-0.212*** (0.022)	-0.231*** (0.015)	-0.253*** (0.019)	-0.282*** (0.024)	-0.221*** (0.018)
North-West	-0.253*** (0.032)	-0.230*** (0.023)	-0.242*** (0.026)	-0.197*** (0.020)	-0.239*** (0.032)	-0.230*** (0.023)
Mpumalanga	-0.319*** (0.033)	-0.267*** (0.032)	-0.240*** (0.021)	-0.236*** (0.018)	-0.260*** (0.024)	-0.289*** (0.022)
Limpopo	-0.308*** (0.028)	-0.319*** (0.026)	-0.320*** (0.018)	-0.273*** (0.018)	-0.290*** (0.028)	-0.337*** (0.022)
Manager	0.455*** (0.090)	0.536*** (0.046)	0.506*** (0.032)	0.539*** (0.030)	0.554*** (0.044)	0.499*** (0.048)
Professional	0.543*** (0.082)	0.489*** (0.056)	0.388*** (0.026)	0.331*** (0.035)	0.346*** (0.046)	0.407*** (0.043)
Technician	0.317*** (0.055)	0.323*** (0.035)	0.274*** (0.032)	0.246*** (0.029)	0.219*** (0.040)	0.250*** (0.034)
Clerk	0.174** (0.057)	0.086* (0.037)	0.022 (0.033)	0.014 (0.029)	0.031 (0.034)	0.026 (0.032)
Agricultural	-1.025*** (0.156)	-0.693*** (0.069)	-0.614*** (0.045)	-0.279** (0.103)	0.047 (0.059)	-0.655*** (0.051)
Artisan	-0.517*** (0.085)	-0.382*** (0.049)	-0.372*** (0.032)	-0.269*** (0.039)	-0.198*** (0.057)	-0.348*** (0.035)
Elementary	-0.529*** (0.050)	-0.590*** (0.035)	-0.567*** (0.027)	-0.460*** (0.028)	-0.371*** (0.037)	-0.529*** (0.030)
Domestic Service	-0.563*** (0.044)	-0.663*** (0.035)	-0.698*** (0.022)	-0.625*** (0.022)	-0.547*** (0.035)	-0.643*** (0.028)
Agri (Sector)	0.130*** (0.025)	0.043 (0.028)	-0.055* (0.022)	-0.152*** (0.024)	-0.202*** (0.035)	-0.030 (0.023)
Mining	0.643*** (0.100)	0.669*** (0.079)	0.698*** (0.059)	0.742*** (0.069)	0.592*** (0.067)	0.725*** (0.073)
Utilities	0.577* (0.262)	0.764*** (0.100)	0.808*** (0.056)	0.727*** (0.085)	0.847*** (0.091)	0.718*** (0.094)
Construction	0.127 (0.102)	0.122* (0.055)	0.191*** (0.041)	0.233** (0.074)	0.568*** (0.116)	0.157** (0.056)
Trade	-0.140*** (0.034)	-0.013 (0.022)	0.141*** (0.015)	0.210*** (0.014)	0.269*** (0.022)	0.107*** (0.015)
Transport	0.229*** (0.044)	0.436*** (0.050)	0.554*** (0.044)	0.549*** (0.042)	0.553*** (0.052)	0.534*** (0.046)
Finance	0.343*** (0.043)	0.445*** (0.031)	0.482*** (0.026)	0.473*** (0.016)	0.451*** (0.022)	0.508*** (0.026)
Services	0.388*** (0.037)	0.528*** (0.025)	0.630*** (0.032)	0.643*** (0.021)	0.589*** (0.019)	0.636*** (0.019)
λ_{mean}	-	-	-	-	-	-0.013 (0.058)
N	32006	32006	32006	32006	32006	24831

Standard errors in parentheses. All results estimated using sampling weights.
 Dependent variable: Natural Logarithm of Imputed Real Earnings (2000 prices).
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.11: Imputed Quantile Regression-With Heckman Correction: 2005-2007

	.10	.25	.5	.75	.9	Mean Reg
Log Hrs Worked	0.218*** (0.029)	0.118*** (0.021)	0.044*** (0.007)	0.022*** (0.004)	0.018*** (0.004)	0.056*** (0.004)
Num Child (0-6)	-0.003 (0.005)	-0.006 (0.004)	-0.004 (0.003)	-0.003 (0.003)	-0.004 (0.004)	-0.011*** (0.003)
Num Child (7-15)	-0.012*** (0.003)	-0.014*** (0.004)	-0.013*** (0.002)	-0.010*** (0.002)	-0.007* (0.003)	-0.015*** (0.002)
λ^1	-4.102* (2.080)	-1.187 (2.243)	-1.465 (1.261)	-0.188 (1.413)	-3.924* (1.818)	-
λ^2	4.662 (3.975)	0.620 (4.353)	2.222 (2.585)	-0.328 (2.770)	6.114 (3.492)	-
λ^3	-2.666 (3.609)	0.254 (4.041)	-1.816 (2.457)	0.561 (2.554)	-4.448 (3.202)	-
λ^4	0.688 (1.518)	-0.301 (1.763)	0.756 (1.090)	-0.255 (1.100)	1.519 (1.382)	-
λ^5	-0.060 (0.237)	0.062 (0.289)	-0.122 (0.182)	0.036 (0.179)	-0.195 (0.224)	-
Age	-0.007 (0.017)	0.033* (0.016)	0.033** (0.011)	0.033** (0.012)	0.028 (0.015)	0.056*** (0.011)
Age-Squared	0.000 (0.000)	-0.000 (0.000)	-0.000* (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.001*** (0.000)
Years of Educ	0.063*** (0.007)	0.077*** (0.006)	0.074*** (0.003)	0.070*** (0.004)	0.076*** (0.004)	0.076*** (0.003)
Married	0.013 (0.023)	0.028 (0.019)	0.030* (0.013)	0.039* (0.015)	0.077*** (0.019)	-0.002 (0.013)
Black	-0.548*** (0.032)	-0.479*** (0.024)	-0.460*** (0.017)	-0.509*** (0.015)	-0.620*** (0.021)	-0.543*** (0.021)
Coloured	-0.317*** (0.039)	-0.267*** (0.029)	-0.270*** (0.018)	-0.349*** (0.017)	-0.457*** (0.023)	-0.330*** (0.025)
Western Cape	-0.114* (0.046)	-0.068* (0.029)	-0.073*** (0.020)	-0.084*** (0.018)	-0.121*** (0.027)	-0.064* (0.023)
Eastern Cape	-0.370*** (0.035)	-0.334*** (0.030)	-0.277*** (0.024)	-0.285*** (0.021)	-0.291*** (0.031)	-0.375*** (0.020)
Northern Cape	-0.204*** (0.058)	-0.250*** (0.033)	-0.234*** (0.035)	-0.283*** (0.026)	-0.279*** (0.040)	-0.300*** (0.027)
Freestate	-0.402*** (0.045)	-0.379*** (0.041)	-0.309*** (0.026)	-0.307*** (0.034)	-0.289*** (0.037)	-0.408*** (0.023)
Kwa-Zulu Natal	-0.205*** (0.026)	-0.228*** (0.019)	-0.230*** (0.012)	-0.239*** (0.020)	-0.244*** (0.027)	-0.258*** (0.018)
North-West	-0.091 (0.049)	-0.184*** (0.033)	-0.161*** (0.021)	-0.186*** (0.029)	-0.211*** (0.035)	-0.207*** (0.025)
Mpumalanga	-0.269*** (0.042)	-0.195*** (0.027)	-0.152*** (0.025)	-0.179*** (0.023)	-0.172*** (0.028)	-0.221*** (0.022)
Limpopo	-0.338*** (0.050)	-0.340*** (0.035)	-0.249*** (0.029)	-0.254*** (0.037)	-0.268*** (0.047)	-0.357*** (0.028)
Manager	0.317*** (0.077)	0.475*** (0.049)	0.644*** (0.037)	0.690*** (0.028)	0.677*** (0.073)	0.532*** (0.042)
Professional	0.622*** (0.045)	0.669*** (0.045)	0.630*** (0.029)	0.473*** (0.033)	0.293*** (0.057)	0.570*** (0.039)
Technician	0.147*** (0.042)	0.328*** (0.061)	0.442*** (0.032)	0.343*** (0.027)	0.197*** (0.054)	0.313*** (0.033)
Clerk	0.201*** (0.048)	0.208*** (0.048)	0.218*** (0.024)	0.144*** (0.027)	0.028 (0.047)	0.165*** (0.032)
Agricultural	-1.041*** (0.081)	-0.761*** (0.114)	-0.405*** (0.045)	-0.115* (0.057)	-0.009 (0.065)	-0.583*** (0.053)
Artisan	-0.474*** (0.035)	-0.287*** (0.037)	-0.219*** (0.032)	-0.177*** (0.031)	-0.181** (0.064)	-0.262*** (0.033)
Elementary	-0.516*** (0.032)	-0.473*** (0.034)	-0.405*** (0.027)	-0.386*** (0.029)	-0.401*** (0.051)	-0.439*** (0.030)
Domestic Service	-0.473*** (0.028)	-0.454*** (0.031)	-0.420*** (0.024)	-0.444*** (0.025)	-0.459*** (0.049)	-0.445*** (0.028)
Agri (Sector)	0.386*** (0.037)	0.329*** (0.021)	0.204*** (0.020)	0.058* (0.026)	0.001 (0.032)	0.219*** (0.022)
Mining	0.673*** (0.081)	0.667*** (0.051)	0.747*** (0.070)	0.774*** (0.068)	0.724*** (0.072)	0.796*** (0.066)
Utilities	0.605*** (0.150)	0.745*** (0.151)	0.874*** (0.093)	0.988*** (0.067)	1.025*** (0.075)	0.802*** (0.093)
Construction	0.266*** (0.045)	0.133*** (0.032)	-0.000 (0.034)	0.096*** (0.025)	0.134 (0.071)	0.117*** (0.035)
Trade	-0.259*** (0.033)	-0.099*** (0.019)	0.049** (0.013)	0.122*** (0.016)	0.169*** (0.020)	0.011 (0.014)
Transport	0.196*** (0.042)	0.305*** (0.071)	0.462*** (0.023)	0.546*** (0.063)	0.687*** (0.057)	0.378*** (0.046)
Finance	0.209*** (0.043)	0.385*** (0.038)	0.378*** (0.013)	0.429*** (0.020)	0.471*** (0.027)	0.415*** (0.024)
Services	0.242*** (0.032)	0.421*** (0.021)	0.578*** (0.026)	0.678*** (0.022)	0.681*** (0.024)	0.544*** (0.017)
λ_{mean}	-	-	-	-	-	0.078 (0.072)
N	32975	32975	32975	32975	32975	24759

Standard errors in parentheses. All results estimated using sampling weights.
 Dependent variable: Natural Logarithm of Imputed Real Earnings (2000 prices).
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.12: Imputed Quantile Regression-With Heckman Correction: 2010-2011

	.10	.25	.5	.75	.9	Mean Reg
Log Hrs Worked	0.038*** (0.008)	0.024** (0.008)	0.015*** (0.004)	0.007 (0.004)	0.006* (0.003)	0.024*** (0.004)
Num Child (0-6)	-0.008 (0.006)	-0.006 (0.005)	-0.008* (0.004)	-0.004 (0.003)	0.004 (0.004)	-0.018*** (0.004)
Num Child (7-15)	-0.006 (0.004)	-0.007 (0.005)	-0.003 (0.004)	0.003 (0.005)	0.010 (0.007)	-0.012** (0.004)
λ^1	-3.256* (1.434)	-1.663 (1.380)	1.267 (1.146)	1.395 (1.191)	-2.166 (1.626)	-
λ^2	3.986 (2.710)	1.233 (2.417)	-3.634 (1.927)	-3.149 (2.135)	2.307 (2.897)	-
λ^3	-2.704 (2.250)	-0.543 (2.093)	3.204* (1.567)	2.423 (1.975)	-1.916 (2.586)	-
λ^4	0.894 (0.839)	0.148 (0.827)	-1.182* (0.594)	-0.792 (0.848)	0.816 (1.043)	-
λ^5	-0.113 (0.114)	-0.020 (0.119)	0.156 (0.085)	0.092 (0.135)	-0.130 (0.153)	-
Age	-0.065 (0.052)	-0.048 (0.049)	-0.047 (0.045)	-0.040 (0.034)	-0.118 (0.067)	0.077* (0.031)
Age-Squared	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.000)	0.001 (0.001)	-0.001* (0.000)
Years of Educ	0.050** (0.018)	0.061*** (0.015)	0.063*** (0.013)	0.059*** (0.011)	0.036 (0.019)	0.103*** (0.009)
Married	0.011 (0.032)	0.032 (0.036)	0.066* (0.030)	0.054** (0.020)	0.114** (0.039)	-0.023 (0.022)
Black	-0.637*** (0.054)	-0.522*** (0.045)	-0.363*** (0.048)	-0.370*** (0.026)	-0.366*** (0.061)	-0.455*** (0.030)
Coloured	-0.471*** (0.036)	-0.320*** (0.038)	-0.307*** (0.029)	-0.314*** (0.030)	-0.334*** (0.047)	-0.289*** (0.033)
Western Cape	-0.049 (0.040)	-0.156** (0.048)	-0.068* (0.032)	-0.124*** (0.028)	-0.201*** (0.050)	-0.038 (0.032)
Eastern Cape	-0.298*** (0.051)	-0.281*** (0.038)	-0.164*** (0.048)	-0.184*** (0.027)	-0.093* (0.047)	-0.215*** (0.035)
Northern Cape	-0.301*** (0.073)	-0.251*** (0.047)	-0.082 (0.047)	-0.099*** (0.028)	-0.085* (0.037)	-0.149*** (0.041)
Freestate	-0.381*** (0.046)	-0.368*** (0.038)	-0.276*** (0.035)	-0.237*** (0.035)	-0.141*** (0.041)	-0.290*** (0.033)
Kwa-Zulu Natal	-0.232*** (0.029)	-0.257*** (0.032)	-0.166*** (0.029)	-0.180*** (0.018)	-0.142*** (0.030)	-0.152*** (0.027)
North-West	-0.084 (0.073)	-0.090 (0.049)	-0.075 (0.052)	-0.121** (0.037)	-0.086 (0.051)	-0.197*** (0.043)
Mpumalanga	-0.229*** (0.047)	-0.262*** (0.035)	-0.166*** (0.021)	-0.144*** (0.021)	-0.102*** (0.029)	-0.171*** (0.031)
Limpopo	-0.327*** (0.052)	-0.363*** (0.048)	-0.264*** (0.047)	-0.183*** (0.043)	-0.063 (0.059)	-0.311*** (0.040)
Manager	0.466*** (0.087)	0.537*** (0.064)	0.764*** (0.044)	0.833*** (0.059)	0.807*** (0.073)	0.651*** (0.060)
Professional	0.535*** (0.076)	0.600*** (0.067)	0.798*** (0.040)	0.735*** (0.058)	0.704*** (0.060)	0.665*** (0.060)
Technician	0.310*** (0.070)	0.420*** (0.059)	0.610*** (0.051)	0.501*** (0.047)	0.305*** (0.048)	0.469*** (0.053)
Clerk	0.248*** (0.066)	0.186*** (0.046)	0.278*** (0.047)	0.259*** (0.044)	0.179*** (0.040)	0.281*** (0.052)
Agricultural	-0.396 (0.256)	-0.219 (0.160)	-0.116 (0.119)	-0.026 (0.103)	0.278* (0.118)	-0.120 (0.139)
Artisan	-0.096 (0.090)	-0.046 (0.051)	-0.013 (0.057)	0.028 (0.073)	0.087 (0.096)	-0.042 (0.062)
Elementary	-0.341*** (0.078)	-0.338*** (0.052)	-0.341*** (0.045)	-0.377*** (0.048)	-0.406*** (0.047)	-0.329*** (0.052)
Domestic Service	-0.175** (0.063)	-0.297*** (0.044)	-0.317*** (0.044)	-0.387*** (0.039)	-0.406*** (0.049)	-0.272*** (0.048)
Agri (Sector)	0.431*** (0.053)	0.254*** (0.044)	0.160*** (0.021)	-0.025 (0.040)	-0.168*** (0.035)	0.188*** (0.047)
Mining	0.663*** (0.099)	0.665*** (0.109)	0.815*** (0.079)	0.805*** (0.087)	0.742*** (0.064)	0.789*** (0.090)
Utilities	0.370** (0.117)	0.525* (0.266)	0.711*** (0.110)	0.840*** (0.147)	0.886*** (0.161)	0.739*** (0.143)
Construction	0.180* (0.073)	-0.006 (0.057)	-0.046 (0.054)	0.185* (0.086)	0.347*** (0.074)	0.137* (0.060)
Trade	0.046 (0.035)	0.147*** (0.023)	0.201*** (0.018)	0.251*** (0.027)	0.433*** (0.029)	0.191*** (0.025)
Transport	0.162 (0.106)	0.183*** (0.055)	0.336*** (0.069)	0.453*** (0.058)	0.591*** (0.067)	0.335*** (0.053)
Finance	0.275*** (0.042)	0.322*** (0.032)	0.345*** (0.015)	0.347*** (0.026)	0.428*** (0.040)	0.350*** (0.032)
Services	0.185*** (0.023)	0.273*** (0.025)	0.442*** (0.024)	0.590*** (0.028)	0.674*** (0.038)	0.422*** (0.025)
λ_{mean}	-	-	-	-	-	0.323* (0.164)
<i>N</i>	18826	18826	18826	18826	18826	18817

Standard errors in parentheses. All results estimated using sampling weights.
 Dependent variable: Natural Logarithm of Imputed Real Earnings (2000 prices).
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$