Racial Wage Discrimination in South Africa

Before and After the First Democratic Election

Gaute Erichsen
School of Economics
University of Cape Town

Foreword

This paper is written as part of the Master in Economics Program at the University of Cape Town. The present work is based on a forthcoming paper written for the Development Policy Research Unit. In that regard I would like to thank USAID for financial support and my supervisor Jeremy Wakeford for more help than anyone could expect.

Estimations have been performed using STATA 6.0.

Abstract

Apartheid in South Africa was formally discarded by the first free election in 1994. Prior to 1994 discrimination in the labour market was embodied in a number of policies (pass laws, occupational colour barring etc.). While such polices may be eliminated by the ANC government, it is apparent that the elimination of racial wage discrimination altogether will be a lengthy process.

In the present paper, racial wage discrimination is treated via a multilateral wage decomposition technique. Each observed wage differential is broken down into a productivity component and a discrimination component so that the extent of racial wage discrimination can be estimated. Using data collected just before (1993) and just after (1995) the first democratic election, it can be concluded that previous findings of long-term declining discrimination are reversed in the post apartheid era.

0 Executive Summary

Several studies focusing on racial discrimination in the South African labour market have been conducted. Different approaches have been used, but the major findings have been similar. Knight and McGrath (1987) found that discrimination against Africans (blacks) decreased by 50 per cent between 1976 and 1985. Work done by Moll (1998) and Hinks (1999) using data from 1980 and 1993, and 1980 and 1994, respectively, further led to the conclusion that racial wage discrimination was declining in South Africa. However, discrimination studies using post apartheid data cannot conclude that the previous decline in discrimination is continuing. Based on data from 1995 and 1997 Hinks and Allanson (2000) find discriminatory patterns to vary among racial groups.

The present paper focuses on wage discrimination for full-time employees. Such discrimination is estimated by decomposing differences in mean wages of South Africa's four racial groups, into a productivity component and a discrimination component. Following this approach, discrimination is defined as unequal pay for similar work. Other discriminatory factors, such as occupational barriers and differing access to education and quality of education, are at best controlled for. Some of these factors are difficult to quantify and are not my primary concern, since I primarily focus on discrimination within the labour market.

Section 2 of the paper provides a review of major developments in the methodology of estimating discrimination based on wage differentials. The point of departure is Oaxaca's (1973A) decomposition of a wage differential into a productivity component and a discrimination component. Possible estimation methods of different wage decompositions are also discussed. Next, following Neumark (1988), it is shown how different wage decompositions link to the theory of employer discrimination.

Section 3 describes the specification of the earnings functions, which enable estimation of discrimination coefficients. Using ordinary least squares, the natural logarithm of earnings is estimated as a function of a range of individual characteristics, *viz*: years of education; potential work experience; rural versus urban residency; union membership versus non-membership; industry; occupation; and province of residence.

The data used in the analysis are drawn from the 1993 Southern Africa Labour and Development Research Unit (SALDRU) Household Survey (SALDRU93) and the 1995 October Household Survey (OHS95), conducted by Statistics South Africa (then the Central Statistical Services). Using data collected just before and just after the formal demise of apartheid, it is hoped that additional insight into changes in the nature and magnitude of discrimination may be provided. The methodology is applied to males and females separately, so as not to introduce gender bias into the estimates.

A number of important assumptions and restrictions underlie the results. Firstly, the focus is on discrimination among full time wage earning employees. Other forms of discrimination, labelled pre-labour market discrimination, have at best been controlled for (e.g. dummy variables for occupational categories) and at worst not controlled for owing to lack of data (e.g. quality of education). Secondly, the decomposition method estimates discrimination as the residual difference between observed wage differentials and the wage component explained by productivity characteristics. Misspecification of the earnings functions will thus induce bias into

the discrimination estimates. Thirdly, a more fundamental assumption is that the presence of discrimination has only distributional effects. In other words, the wage bill is regarded as constant whether discrimination is present or not. Fourthly, to make discrimination estimates operational, one needs a notion of what the wage structure would be like in the absence of discrimination. This paper uses average characteristics based on a pooled sample containing all racial groups as a proxy for the non-discriminatory wage structure. Lastly, certain ad hoc assumptions were unavoidable when attempting to make data from 1993 congruent with data from 1995.

The major findings can be summarised as follows. African males were subject to discrimination in the form of underpayment in both 1993 and 1995. However, the largest share of the observed wage differential between this group's average wage and the average wage of all workers was due to below average productivity characteristics. For Coloured and Asian males, substantial wage differentials were mostly explained by above or below average productivity characteristics (again relative to the average male worker). An exception was found for Asian males in 1995, where above average productivity characteristics accounted for some 78 per cent of the observed wage differential. Based on discrimination coefficients, it was found that discrimination was most remarkable towards white males. Whites are overpaid, illustrated by the fact that in 1993 and 1995, productivity characteristics could only explain approximately 65 per cent and 54 per cent of the observed wage differential, respectively.

Among females, discrimination exhibits similar patterns as for males. Differing productivity compared to the average female worker explains by far the largest proportion of the observed wage differentials for Africans, Coloureds and Asians. But productivity as the explanatory force, is not as predominant among Africans as among Coloureds and Asians. The results for Asian females in 1995 are an exception. In 1995 productivity accounted for only 62 per cent of the (average) wage differential between Asian females and all females. Again, the data suggest that the most striking discriminatory effect is defined by White overpayment. For White females, above average productivity characteristics account for 71 per cent and 66 per cent of the wage differential in 1993 and 1995, respectively.

In sum, using data from 1993 and 1995, variation in discrimination trends among racial groups seems to be the rule. At the least, I cannot conclude that discrimination is declining as an overall trend for all racial groups. But what do the differences when comparing the 1993 results with the 1995 results really reflect? Two possible explanations seem plausible: the differences reflect either changes over time or data inconsistencies. If the time span of two years is too short to pick up considerable changes in the underlying wage structure, then variation in discriminatory trends probably reflects data inconsistencies. A more likely scenario is that changing discrimination coefficients incorporate both changes over time and problems with the data. Unfortunately, it is not clear how to separate these effects. Despite these problems, one finding stands out, namely that Whites are overpaid relative to their measured productivity. However, some of this effect could be due to superior quality of education, which is not controlled for.

Contents

0 Executive Summary ... *iii*

0.1 List of Figures ... vi

0.2 List of Tables ... vi

1 Introduction ... *1*

2 Review of the Wage Differential Methodology ... *6*

2.1 Decomposition theory ... 6

2.2 Estimation of discrimination coefficients ... 10

2.3 Decomposition's foundation in discrimination theory ... 13

3 Data Used and the Determination of Earnings Functions ... *18*

3.1 Data ... 18

3.2 Sample Definition ... 18

3.3 Earnings Functions ... 18

4 Results ... *23*

4.1 Males ... 23

4.2 Females ... 30

5 Conclusions ... *36*

6 References ... *38*

7 Appendix ... *40*

0.1 List of Figures

Figure Number	Title	Page
Figure 1	Welfare Costs of Discrimination	2
Figure 2	Wage-Experience Profile	21
Figure 3	Male Wage-Education Profile	26
Figure 4	Male Wage-Experience Profile	26
Figure 5	Female Wage-Education Profile	33
Figure 6	Female Wage-Experience Profile	33
Figure A1	1993 Male Wage-Education Profile	43
Figure A2	1995 Male Wage-Education Profile	43
Figure A3	1993 Male Wage-Experience Profile	44
Figure A4	1995 Male Wage-Experience Profile	44
Figure A5	1993 Female Wage-Education Profile	49
Figure A6	1995 Female Wage-Education Profile	49
Figure A7	1993 Female Wage-Experience Profile	50
Figure A8	1995 Female Wage-Experience Profile	50

0.2 List of Tables

Table Number	Title	Page
Table 1	Weighted Mean Values for the Variables Included in the Male Earnings Function	24
Table 2	Estimated Coefficients from the All Male Earnings Function	25
Table 3	Decomposition of Male Earnings Differentials	28
Table 4	Weighted Mean Values for the Variables Included in the Female Earnings Function	31
Table 5	Estimated Coefficients from the All Female Earnings Function	32
Table 6	Decomposition of Female Earnings Differentials	34
Table A1	Estimated Coefficients from the African Male Earnings Function	39
Table A2	Estimated Coefficients from the Coloured Male Earnings Function	40
Table A3	Estimated Coefficients from the Asian Male Earnings Function	41
Table A4	Estimated Coefficients from the White Male Earnings Function	42
Table A5	Estimated Coefficients from the African Female Earnings Function	45
Table A6	Estimated Coefficients from the Coloured Female Earnings Function	46
Table A7	Estimated Coefficients from the Asian Female Earnings Function	47
Table A8	Estimated Coefficients from the White Female Earnings Function	48

1 Introduction

In 1980 the geometric mean wage of full time working South African males was R245 per month (Hinks 1999). In 1995 the geometric mean monthly wage had increased to R1305 per month. However, mean wages varied significantly among South Africa's racial groups. As an illustration, the mean African wage was 48 per cent below the all-male average in 1980 and 37 per cent below the all-male average in 1995. Whites, on the other hand, received mean wages 203 per cent above the all-male average in 1980 and 296 per cent more than the all-male average in 1995. Large wage differences between races are thus clearly predominant in South Africa, both in 1980 and 1995.

How can such wage differences be explained?
According to economic theory workers should be paid according to their productivity. Differing productivity characteristics among racial groups could therefore serve as an explanation for observed wage differences. However, can productivity characteristics fully explain observed wage differences?A number of studies (Oaxaca 1973A, Oaxaca and Ransom 1988, Cotton 1988, Neumark 1988, Hinks 1999) conclude that productivity accounts for parts of observed wage differences only, leaving discrimination as a possible explanation for the remaining parts.Labour market discrimination may thus be defined as present if workers having identical productivity characteristics in a physical or material sense are paid differently owing to some exterior characteristic (e.g. race, gender, union membership etc).

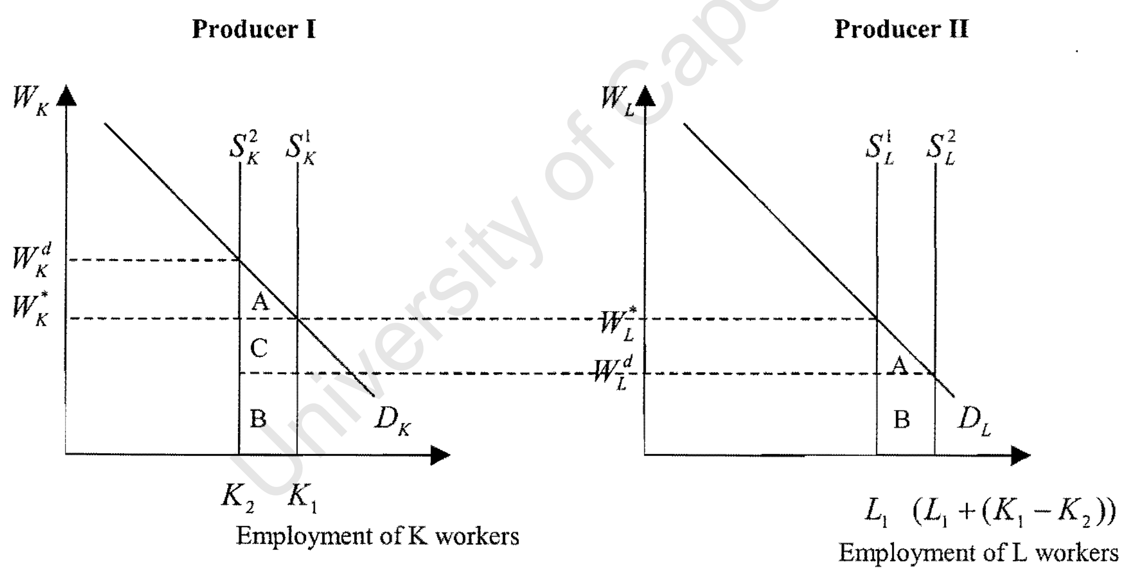
This paper focuses on wage discrimination among races in the South African labour market for full time employees. Several other studies concerning the same topic have been conducted. These studies have differed in their approach to discrimination, but the major findings have been similar. Knight and McGrath (1987) found that discrimination against Africans (blacks) decreased by 50 per cent between 1976 and 1985. Work done by Moll (1998) and Hinks (1999) using data from 1980 and 1993, and 1980 and 1994, respectively, concluded with the finding that racial wage discrimination was declining in South Africa. However, discrimination studies using post apartheid data cannot conclude that the previous decline in discrimination is continuing. Based on data from 1995 and 1997 Hinks and Allanson (2000) find discriminatory patterns to vary among racial groups.

But why this interest in discrimination?
Firstly, most people agree that workers having identical productivity characteristics and doing the same job, but being paid differently, is unjustifiable. The slogan "equal pay for equal work" seems fair to most people. In an economic sense discrimination attracts interest beyond the fairness perspective. Economic theories of discrimination predict welfare losses associated with any form of discrimination in the labour market (Sapsford and Tzannatos 1993). The implication of welfare losses from discrimination is illustrated in the following example (ibid.).

Assume that the labour market consists of two producers (I and II) and two labour groups (K and L). The two producers have identical technology, produce the same good and use labour as the only input in production. The two groups of workers, K and L, are equally productive. Further, it is imposed that initially the labour market is completely segregated, meaning that producer I employs only K workers, while producer II employs only L workers. Lastly, the labour supply is completely inelastic and identical for both groups of employees. Producers pay wages equalling the value of the marginal product of labour. Since labour demand and supply are, *de facto*, identical in the two segments of the labour market, group K and L workers receive identical wages. In Figure 1 equal wages for both K and L workers are illustrated by $W_K^* = W_L^*$. To analyse the effects of discrimination, an arbitrary higher wage for K workers ($W_K^d - W_K^*$) is imposed. The implication of discrimination is thus to increase the wage for a certain group of workers. Higher wages, in the labour market for producer I, imply that fewer workers are employed. To maintain full employment, the displaced K workers have to seek employment from producer II. The increased supply of labour facing producer II decreases wages for all L workers and the displaced K workers. K workers who are still employed by producer I are therefore better off with discrimination, while L workers and displaced K workers are worse off.

In addition to the redistributional effects of discrimination, Figure 1 may be used to examine welfare effects. Reduced employment by producer I gives a net reduction in welfare, illustrated by the trapezoid A+C+B. However, increased employment by producer II gives an increase in welfare, illustrated by the area A+B. Area C is thus a net welfare loss induced by discrimination. Although the same number of workers are employed, inefficiencies are created since displaced K workers are less productive when employed by producer II in addition to the original L workers.

FIGURE 1 Welfare Costs of Discrimination



Having argued that discrimination creates welfare losses it is not surprising that economists have been preoccupied with analyses of discrimination. As far back as 1922 Edgeworth analysed discrimination between genders. However, Becker presented a more comprehensive treatment of discrimination in 1957. According to Becker (1971), the cause of discrimination is individuals having a "taste for discrimination." Furthermore, discrimination occurs only when individuals act according to their tastes, e.g. in the form of direct payments or reduced income, in order to be associated with workers from certain groups. Becker defines three types of discrimination rooted in "tastes", namely employer discrimination, employee discrimination and consumer discrimination.

When employer discrimination is present, the employer's decisions involve a tradeoff between profits and the ratio of workers from different groups. If an employer dislikes group L workers, his utility increases as the ratio of L workers decreases. Employers are thus willing to forfeit profits in order to employ a low ratio of employees from disliked groups. In other words, employer discrimination implies that preferred groups receive higher wages than in a situation with no discrimination, while disliked groups are paid less.[1]

The same conclusion can be reached based on slightly different assumptions. Instead of employers having tastes for discrimination one can define employees to have tastes for different groups of co-workers. If a certain type of employee, e.g. foremen, have preferences over co-workers from different groups (race, gender, union etc), then foremen's decisions to take employment is influenced by a trade-off between the wage and the ratio of co-workers from different groups. A rational firm takes their foremen's preferences into account and profits by offering higher wages to workers who are preferred by foremen, since the cost of hiring foremen thereby is reduced. Again, employees from disliked groups are paid less than in a situation with no discrimination, while workers from groups preferred by foremen are better paid.

Finally, consumer discrimination is present if consumers are willing to pay higher prices to purchase goods produced by specific groups of workers.

Although tastes for discrimination, within a neoclassical framework, lead to plausible explanations for discriminatory behaviour, some major objections persist. In the case of employer discrimination, a long run argument for discrimination is hard to maintain. In the long run capital is flexible and will be attracted by the most profitable firms. The most discriminatory employers, who are willing to forfeit the most profit in order to keep the relative number of workers from certain groups low, will thus be the least profitable and have the smallest market shares. Moreover, if non-discriminating firms exist, then these are the only firms that will survive in the long run. Neoclassical models for employer discrimination therefore predict the end of discrimination in the long run. Following Arrow (1973B), the model of employer discrimination predicts the absence (in the long run) of the phenomenon it was designed to explain. However, by introducing costs of adjusting the ratio of workers (the cost of hiring and firing) from different groups, it might be profitable to keep the ratio of workers fixed given a marginal change in wages. In other words, it can be profitable not to change the ratio of employees from different groups, and discriminatory firms might survive even in the long run.

Apart from the long term implications of models based on tastes for discrimination, it can be argued that such models are of little value since the source of discrimination, namely the tastes, are unexplained. The tastes for discrimination are exogenous in the models and are therefore not analysed. Arrow (1973A) offers an alternative source of discrimination, namely the "perception of reality." If employers believe that group K workers are more productive than L workers, the latter will only be employed if $W_L < W_K$. If Arrow's notion of "perceptions" is to add additional insight, an explanation of how the "perceptions" are formed has to be given. One possible explanation is that if individuals act in a discriminatory fashion, they tend to develop beliefs which justify their actions (Marshall 1974). However, beliefs based on the perception of reality will not persist, if proven wrong in reality.

[1]Only relative discrimination matters. If an employer dislikes L workers, workers from other groups are being preferred relatively speaking. Relative discrimination applies to wages as well. If L workers are underpaid due to discrimination, other groups are relatively overpaid.

There are a number of other approaches to the theory of discrimination. One example is Marshall's (ibid.) bargaining model *"which incorporates features of the neoclassical models but specifies the motives of the various actors and the contexts within which they operate on the basis of empirical evidence rather than a priori deductive reasoning."* A problem with models based on empirical evidence is that they tend to become quite complex and fail to produce general results. Hence neoclassical models, as described above, have an advantage in their general form and straightforward predictions. That might be the reason why such models are a key reference in much work addressing discrimination, including this paper.

Given the existence of discrimination, how can it be measured?
If observed wage differentials are decomposed into a productivity component and a discrimination component, one way of measuring discrimination is as the residual between the wage differential and the estimated productivity component.

This paper addresses how decomposition techniques can be utilised in order to measure discrimination. Once the approach is established, it will be applied to full time employees in the South African labour market. Since discrimination is analysed only for full time employees and is measured by a residual approach, some initial remarks concerning the interpretation of discrimination are in order.

Using decomposition techniques, discrimination is defined as unequal pay for similar work. Having said this, wage discrimination should not be interpreted as workers from different groups having the same occupation, identical jobs and working for the same employer, but being paid differently. Discrimination rather reflects different pay to workers falling under the same occupational categories, working for different employers, etc. Moreover, only discrimination within the labour market is addressed, meaning that other discriminatory factors such as the quality of and access to education are not controlled for. The occupational distribution is controlled for by using dummy variables. Access to education is defined as pre-labour market discrimination, meaning that once within the workforce employees are not paid differently due to original differing access to education. Accounting for the quality of education is more complicated since the quality of education directly affects employees' productivity. Unfortunately no information potentially describing the quality of education was available and the impact of the quality of education is thus omitted. Since discrimination is estimated as a residual, it is crucial that all factors that have an impact on productivity are accounted for. However, in empirical work, accounting for all productivity factors is at best extremely costly and difficult, and at worst impossible. This is important to keep in mind since the discrimination coefficient, measured by the residual between the wage differential and the productivity component, captures both the effects of discrimination and unobserved group differences (e.g. the quality of education) in productivity and tastes (Altonji and Blank 1999).

The empirical work is based on data from 1993 and 1995. Using data collected just before and just after the fall of apartheid, it is hoped that additional insight into changes in and the magnitude of discrimination may be provided.

Section 2 of the paper provides a review of major developments in the methodology of estimating discrimination based on wage differentials. The point of departure is Oaxaca's (1973A) decomposition of a wage differential into a productivity component and a discrimination component. Possible estimation methods of different wage decompositions are also discussed.Next, following Neumark (1988), it is shown how different wage decompositions link to the theory of employer discrimination.

Section 3 describes the specification of the earnings functions, which enable estimation of discrimination coefficients. Using ordinary least squares, the natural logarithm of earnings is estimated as a function of a range of individual characteristics, *viz*: years of education; potential work experience; rural versus urban residency; union membership versus non-membership; industry; occupation; and province of residence.

Finally, results are presented and conclusions are drawn.

2 Review of the Wage Differential Methodology

2.1 Decomposition theory

Economic work concerning wage discrimination has focused on decomposing wage differentials into two components, namely a productivity component and a discrimination component. Since Oaxaca's (1973A) pioneering study of gender discrimination in the United States labour market, the methodology concerning discrimination based on wage differences has evolved considerably. A review of the main developments on how to handle decomposition between two groups and also how researchers deal with multilateral comparisons is given below. In the literature explanations of how to derive the desired decomposition equation are generally rather brief. This review will be more thorough in an attempt to make it easier to follow the various steps of the mathematical derivations. The original contributors to the methodology included Oaxaca (1973A), Neumark (1988), Cotton (1988), Oaxaca and Ransom (1988 and 1994) and Hinks (1999).

The gross (unadjusted) wage differential

The observed or actual wage differential G is interchangeably defined as

$$G_{kl} = \frac{\overline{W}_k - \overline{W}_l}{\overline{W}_l} \Leftrightarrow G_{kl} + 1 = \frac{\overline{W}_k}{\overline{W}_l}\,^{2}, \tag{1}$$

where k and l represent groups or categories within a chosen qualitative variable (e.g. racial groups, genders, union members versus non-members, etc.), $\overline{W}_{k,l}$ is the geometric mean wage of the various groups, and G gives the relative difference between the mean wage of groups k and l.

Applying a logarithmic transformation to (1) one obtains

$$\ln(G_{kl} + 1) = \ln\left(\frac{\overline{W}_k}{\overline{W}_l}\right) = \ln\left(\overline{W}_k\right) - \ln\left(\overline{W}_l\right). \tag{2}$$

This format is particularly useful given the standard semi-logarithmic form of the earnings function regression equation.

The productivity differential

Assuming that labour would be paid their marginal productivity in the absence of discrimination we have:

$$\overline{W}_k^{\,0} = \overline{MP}_k \quad \text{and} \tag{3a}$$

$$\overline{W}_l^{\,0} = \overline{MP}_l \tag{3b}$$

where the superscript 0 refers to the absence of discrimination. This allows us to define the productivity differential interchangeably as

---

[2] Mean wages are computed as geometric means defined as: $\overline{W} = e^{\left(\frac{1}{n}\sum_{i=1}^{n} \ln W_i\right)} = e^{\overline{\ln W_i}}$.

$$Q_{kl} = \frac{\overline{MP}_k - \overline{MP}_l}{\overline{MP}_l} \quad \Leftrightarrow \quad Q_{kl} + 1 = \left(\frac{\overline{W}_k^{\,0}}{\overline{W}_l^{\,0}} \right). \tag{4}$$

Again, subscripts refer to groups, while Q gives the relative difference between mean wages of different groups if employees were paid the non-discriminatory wage. In other words, Q represents the earnings differential attributable solely to differences in (mean) productivity between the groups k and l.

In logarithmic form (4) becomes

$$\ln(Q_{kl} + 1) = \ln\left(\frac{\overline{W}_k^{\,0}}{\overline{W}_l^{\,0}} \right) = \ln(\overline{W}_k^{\,0}) - \ln(\overline{W}_l^{\,0}). \tag{5}$$

The market discrimination coefficient

Given the gross wage differential and the productivity differential the market discrimination coefficient is defined as the relative difference between the actual mean wage difference and the mean wage difference in the absence of discrimination:

$$D_{kl} = \left(\frac{\overline{W}_k / \overline{W}_l - \overline{W}_k^{\,0} / \overline{W}_l^{\,0}}{\overline{W}_k^{\,0} / \overline{W}_l^{\,0}} \right) \quad \Leftrightarrow \quad D_{kl} + 1 = \left(\frac{\overline{W}_k / \overline{W}_l}{\overline{W}_k^{\,0} / \overline{W}_l^{\,0}} \right). \tag{6}$$

In logarithmic form (6) becomes

$$\ln(D_{kl} + 1) = \ln\left(\frac{\overline{W}_k / \overline{W}_l}{\overline{W}_k^{\,0} / \overline{W}_l^{\,0}} \right) = \ln\left(\frac{\overline{W}_k}{\overline{W}_l} \right) - \ln\left(\frac{\overline{W}_k^{\,0}}{\overline{W}_l^{\,0}} \right). \tag{7}$$

Inserting (2) and (5) into (7) and rearranging we obtain

$$\ln(G_{kl} + 1) = \ln(Q_{kl} + 1) + \ln(D_{kl} + 1). \tag{8}$$

Equation (8) gives the traditional decomposition of the logarithmic mean wage differential into a productivity component ($\ln(Q_{kl} + 1)$) and a discrimination component ($\ln(D_{kl} + 1)$) (Oaxaca and Ransom 1994). According to this decomposition, discrimination takes the form of one group being either over- or underpaid. Assume group k is overpaid. Group l now has nothing to gain from discrimination vanishing, since the absence of discrimination only implies reduced wages for group k. Assume instead that discrimination materialises in group l being underpaid. In the absence of discrimination group l wages would increase. Group k wages would remain unchanged, since discrimination defined by equation (8) restricts discrimination to either over- or underpayment (Cotton 1998). In both scenarios only one group is affected if discrimination should disappear. This argument is valid only if total actual output or the total wage bill remains unaltered in the absence of discrimination. Equation (8) is thus implicitly based on an assumption stating that a hypothetical elimination of discrimination will have only redistributional effects

(ibid.). Discrimination defined solely as over- or underpayment, originally used by Oaxaca (1973A), has more recently been recognised as overly restrictive.
Arguing that both groups would be affected by the removal of discrimination leads to a further decomposition of equation (8). Discrimination can now be attributed to both over- and underpayments, relative to a non-discriminatory wage. One group receives above the non-discriminatory wage, benefiting from the other group being paid less than the non-discriminatory wage.

From (7) we have

$$\ln(D_{kl}+1)=\ln\left(\frac{\overline{W}_k}{\overline{W}_l}\right)-\ln\left(\frac{\overline{W}_k^{\,0}}{\overline{W}_l^{\,0}}\right)$$

$$=\left(\ln\overline{W}_k-\ln\overline{W}_l\right)-\left(\ln\overline{W}_k^{\,0}-\ln\overline{W}_l^{\,0}\right)$$

$$=\ln\left(\frac{\overline{W}_k}{\overline{W}_k^{\,0}}\right)-\ln\left(\frac{\overline{W}_l}{\overline{W}_l^{\,0}}\right).$$

Therefore we can write

$$\ln(D_{kl}+1)=\ln(\delta_k+1)-\ln(\delta_l+1), \qquad (9)$$

where

$$\delta_i=\frac{\overline{W}_i}{\overline{W}_i^{\,0}}-1 \qquad \text{i = k, l.} \qquad (10)$$

δ_i is the differential between the current mean wage of group i, and the mean wage of the same group in the absence of discrimination. If $\delta_i \geq 0$ then $\ln(\delta_i+1)$ can be interpreted as the group i wage advantage. If $\delta_i \leq 0$ then $\ln(\delta_i+1)$ is interpreted as the group i wage disadvantage. From (9) we see that discrimination can be separated into a wage advantage ($\ln(\delta_i+1)$, when $\delta_i \geq 0$) and a wage disadvantage ($\ln(\delta_i+1)$, when $\delta_i \leq 0$).[3]

Inserting (9) into (8) yields

$$\ln(G_{kl}+1)=\ln(Q_{kl}+1)+\left[\ln(\delta_k+1)-\ln(\delta_l+1)\right]. \qquad (11)$$

Equation (11) gives an expanded decomposition where the discrimination component is broken down into overpayment $(\ln(\delta_k+1))$ and underpayment $(\ln(\delta_j+1))$. Estimations according to (11) have dominated the literature, whether the focus has been on unions (Oaxaca and Ransom 1988), gender (Neumark 1988) or race (Cotton 1988, Oaxaca and Ransom 1994).

[3]If $\delta_i = 0$ for both k and l, then no group has a wage advantage/disadvantage.

A limitation of the methodology described thus far is that it handles comparisons between only two groups at the time. Hinks (1999) describes a method that allows comparisons of several groups against a pooled non-discriminatory wage structure. Following this approach it can be determined whether each of the different groups is favoured or discriminated against.

Using (2) and defining $\overline{W}$ as the mean wage in the labour market as a whole, we can write

$$\ln(G_{kl}+1) = \ln\left(\frac{\overline{W}_k}{\overline{W}_l}\right) = \ln(\overline{W}_k) - \ln(\overline{W}_l) + \ln\overline{W} - \ln\overline{W}$$

$$= \ln\left(\frac{\overline{W}_k}{\overline{W}}\right) - \ln\left(\frac{\overline{W}_l}{\overline{W}}\right)$$

$$= \ln(\gamma_k + 1) - \ln(\gamma_l + 1), \qquad (12)$$

where

$$\gamma_i = \frac{\overline{W}_i}{\overline{W}} - 1 \text{ , i = k, l.} \qquad (13)$$

γ_i is the differential between the mean wage of group i relative to the mean wage of all employees (i.e. all groups combined).

In similar fashion, we define $\overline{W}^0$ as the non-discriminatory mean wage for the whole labour force and rewrite (5) as

$$\ln(Q_{kl}+1) = \ln\left(\frac{\overline{W}_k^0}{\overline{W}_l^0}\right) = \ln(\overline{W}_k^0) - \ln(\overline{W}_l^0) + \ln(\overline{W}^0) - \ln(\overline{W}^0)$$

$$= \ln\left(\frac{\overline{W}_k^0}{\overline{W}^0}\right) - \ln\left(\frac{\overline{W}_l^0}{\overline{W}^0}\right)$$

$$= \ln(\theta_k + 1) - \ln(\theta_l + 1), \qquad (14)$$

where

$$\theta_i = \frac{\overline{W}_i^0}{\overline{W}^0} - 1 \text{ , i = k, l.} \qquad (15)$$

θ_i is the differential between the mean wage of group i in the absence of discrimination and the non-discriminatory wage of the entire workforce.

Substituting (12) and (14) into (11) we obtain

$$[\ln(\gamma_k+1) - \ln(\gamma_l+1)] = [\ln(\theta_k+1) - \ln(\theta_l+1)] + [\ln(\delta_k+1) - \ln(\delta_l+1)]. \qquad (16)$$

The decomposition given by (16) has the advantage that each component is defined relative to the entire workforce and is thus independent of a comparison between two specific groups of workers (e.g Whites versus Blacks).

If (16) is manipulated slightly this point becomes clearer. Let k = g, where g is equal to the different groups of the population (e.g. African, Coloured, Asian and White) and let l = H, where H represents the entire workforce (i.e. all four groups combined).

Then we have

$$\ln(\gamma_H + 1) = \ln\left(\frac{\overline{W}_H}{\overline{W}}\right) = \ln\overline{W} - \ln\overline{W} = 0,$$

$$\ln(\theta_H + 1) = \ln\left(\frac{\overline{W}_H^0}{\overline{W}^0}\right) = \ln\overline{W}^0 - \ln\overline{W}^0 = 0 \text{ ,and}$$

$$\ln(\delta_H + 1) = \ln\left(\frac{\overline{W}_H}{\overline{W}_H^0}\right) = \ln\overline{W}_H - \ln\overline{W}_H = 0 \text{ if } \overline{W}_H = \overline{W}_H^0.$$

Therefore, assuming that the current mean wage of the population equals the mean wage in the absence of discrimination (i.e. $\overline{W}_H = \overline{W}_H^0$), we see that equation (16) reduces to

$$\ln(\gamma_g + 1) = \ln(\theta_g + 1) + \ln(\delta_g + 1). \tag{17}$$

From (17) it is obvious that Hinks' decomposition compares the individual group g to the entire workforce. In other words decompositions following Hinks' approach define discrimination of a group g relative to the mean wage of the whole workforce.

2.2 Estimation of discrimination coefficients

To make any of the decompositions suggested by (8), (11) or (17) operational one needs to make certain assumptions concerning what the wage structure would be in the absence of discrimination (which can also be interpreted as the competitive wage structure).

Common to all approaches is that estimations are based on earnings functions of the form

$$\ln W_{gi} = X_{gi}^{'}\beta_g + u_{gi}, \tag{18}$$

where g refers to group g and i to individual i in group g.
Equation (18) relates the natural logarithm of the wage for a specific group to a vector of individual characteristics (both quantitative and qualitative), X_{gi}, and a disturbance term, u (which is assumed to have Gauss-Markov properties). β_g is a vector of coefficients. (In Hinks' estimations g may also be interpreted as the entire sample H.)

Equation (18) can be estimated by

$$(\ln \hat{W}_{gi}) = X'_{gi}\hat{\beta}_g \,. \qquad (19)$$

Further, the estimate of the natural logarithm of a group's mean wage is given by

$$\overline{\ln \hat{W}_g} = \overline{X}'_g \hat{\beta}_g = \ln(e^{\overline{\ln \hat{W}_g}}) = \ln(\hat{\overline{W}}_g)\,. \qquad (20)$$

[4]

$\hat{\overline{W}}_g$ is the estimated geometric mean wage of group g. $\overline{X}'_g$ is a vector of mean characteristics for group g, and $\hat{\beta}_g$ a vector consisting of the estimated coefficients from the earnings function given by (19).

Using (20), equation (2) can then be estimated by

$$\ln(\hat{G}_{kl} + 1) = \overline{X}'_k \hat{\beta}_k - \overline{X}'_l \hat{\beta}_l \,. \qquad (21)$$

If one assumes that discrimination takes the form of one group being under- or overpaid while the other group receives the non-discriminatory wage, then equation (8) can be estimated by applying (22a) or (22b) which follow.

Firstly, assume group k is overpaid. In the absence of discrimination this group would thus receive the same wage as group l. Hence the group l wage is interpreted as the non-discriminatory wage in this case.

$$\begin{aligned}\ln(\hat{G}_{kl} + 1) &= \overline{X}'_k \hat{\beta}_k - \overline{X}'_l \hat{\beta}_l = \overline{X}'_k \hat{\beta}_k - \overline{X}'_l \hat{\beta}_l + \overline{X}'_k \hat{\beta}_l - \overline{X}'_k \hat{\beta}_l \\ &= (\overline{X}'_k - \overline{X}'_l)\hat{\beta}_l + \overline{X}'_k(\hat{\beta}_k - \hat{\beta}_l) \end{aligned} \qquad (22a)$$

Related to equation (8), the productivity component of the wage differential is reflected by $(\overline{X}'_k - \overline{X}'_l)\hat{\beta}_l$ and the discrimination component by $\overline{X}'_k(\hat{\beta}_k - \hat{\beta}_l)$.
When estimating discrimination the observed differences in sample mean wages are normally used rather than trying to estimate wage differences, i.e. based on (22a). Using sample mean wages imply that only the productivity component is estimated. The discrimination component is then derived taking the difference between the log of the wage differential and the estimated productivity component. This is illustrated in the equation below.

$\ln(\hat{D}_{kl} + 1) = \ln(G_{kl} + 1) - (\overline{X}'_k - \overline{X}'_l)\hat{\beta}_l$, where G_{kl} is the sample wage differential.

Assume instead that group l is underpaid. Now discrimination implies that group l would be paid the same as group k in the absence of discrimination.

[4] Since the regression line passes through the mean values of the variables.

$$\ln(G_{kl}+1) = \bar{X}_k^{'}\hat{\beta}_k - \bar{X}_l^{'}\hat{\beta}_l = \bar{X}_k^{'}\hat{\beta}_k - \bar{X}_l^{'}\hat{\beta}_l + \bar{X}_l^{'}\hat{\beta}_k - \bar{X}_l^{'}\hat{\beta}_k \qquad (22b)$$
$$= (\bar{X}_k^{'} - \bar{X}_l^{'})\hat{\beta}_k + \bar{X}_l(\hat{\beta}_k - \hat{\beta}_l)$$

Again, discrimination estimates are derived from an equation as

$$\ln(\hat{D}_{kl}+1) = \ln(G_{kl}+1) - (\bar{X}_k' - \bar{X}_l')\hat{\beta}_k \ .$$

If one prefers the notion that discrimination results in one group being overpaid by subsidisation from an underpaid group, then estimation along the line of equation (11) can be expressed as

$$\ln(G_{kl}+1) = (\bar{X}_k^{'} - \bar{X}_l^{'})\hat{\beta}^* + \left[\bar{X}_k^{'}(\hat{\beta}_k - \hat{\beta}^*) - \bar{X}_l^{'}(\hat{\beta}_l - \hat{\beta}^*)\right]. \qquad (23)$$

$\hat{\beta}^*$ refers to a vector of coefficients describing the non-discriminatory wage structure.
$\hat{\beta}^*$ may be defined as

$$\hat{\beta}^* = \phi\hat{\beta}_k + (1-\phi)\hat{\beta}_l \, . \qquad (24)$$

The coefficients describing the non-discriminatory wage structure of the whole workforce are therefore some (linear) combination of the coefficients for the different groups of workers. Since, $\hat{\beta}_i$ (i =k, l), is estimated from the earnings functions, an estimation of β^* becomes a question of how to determine ϕ.

If ϕ takes either the value 0 or 1, we see that (23) reduces itself to (22a) or (22b). So the assumption of one group being either over- or underpaid, as made by Oaxaca (1973A), is a special case of the more general specification given by (23). Cotton (1988) suggests that the fraction of the entire sample that each group constitutes should be used as estimators for ϕ and $(1-\phi)$. This weighting is based on a number of assumptions. First, one group is paid more than what the wage would be in the absence of discrimination and the other group less. Second, β^* is a strict linear combination of the groups' wage structures (β_k, β_l). These assumptions imply that $\phi \in (0,1)$, and that the non-discriminatory wage structure will be closer to the current majority group's wage structure.[5] Mathematically,

$$\hat{\phi} = \frac{K}{K+L}, \qquad (25)$$

where K and L are the sample sizes of the respective groups.
Another, more general, weighting scheme is suggested by Oaxaca and Ransom (1994). Here a row vector of coefficients describing the non-discriminatory wage structure is defined as

[5]Cotton does an analysis based on race in the US (for males only). In this case white workers are favoured by discrimination and whites are also the majority group in the workforce. Thus the non-discriminatory wage structure is closest to the white wage structure. But if the group disfavoured by discrimination constitutes the majority of the workforce, then the non-discriminatory wage structure will be closest to the disfavoured group.

$$\hat{\beta}^{**} = \Phi\hat{\beta}_k + (I - \Phi)\hat{\beta}_l, \tag{26}$$

where Φ is defined as

$$\Phi = \left[X_p' X_p\right]^{-1} X_k' X_k = \left[X_k' X_k + X_l' X_l\right]^{-1} X_k' X_k. \tag{27}$$

X_p refers to a vector of mean characteristics for a pooled sample including all individual groups.

Neumark (1988) shows how $\hat{\beta}^{**}$ can be estimated as the set of coefficients from a least squares regression based on a pooled sample containing all groups. By substituting $\hat{\beta}^{**}$ for $\hat{\beta}^{*}$, equation (23) still holds, but now with a different interpretation of the coefficients describing the non-discriminatory wage structure.

If a multilateral decomposition is desired, estimates of the different components of equation (17) are given by (Hinks 1999)

$$\ln(\hat{\gamma}_g + 1) = \ln\hat{\bar{W}}_g - \ln\hat{\bar{W}} = \bar{X}_g'\hat{\beta}_g - \bar{X}'\hat{\beta}, \tag{28}$$

$$\ln(\hat{\theta}_g + 1) = \ln\hat{\bar{W}}_g^0 - \ln\hat{\bar{W}}^0 = (\bar{X}_g - \bar{X})'\hat{\beta} \text{ and} \tag{29}$$

$$\ln(\hat{\delta}_g + 1) = \ln\hat{\bar{W}}_g - \ln\hat{\bar{W}}_g^0 = \bar{X}_g'(\hat{\beta}_g - \hat{\beta}). \tag{30}$$

Again, g represents the particular group in question, and $\hat{\beta}$ is a vector of coefficients describing the non-discriminatory wage structure. $\hat{\beta}$ is estimated by a least squares regression on a (pooled) sample containing all groups.[6]

If sample mean wages are utilised rather than estimating mean wages, the discrimination component in equation (17) can be estimated following

$\ln(\hat{\delta}_g + 1) = \ln(\gamma_g + 1) - (\bar{X}_g - \bar{X})'\hat{\beta}$, where γ_g is the sample mean wage differential.

2.3 Decomposition's foundation in discrimination theory

As shown in the previous section, to render any decomposition operational certain assumptions concerning the non-discriminatory wage structure have to be made. Since any estimation results hinge on these assumptions it is natural to explore how well they are rooted in theory. The following section is based on Neumark (1988), which again is based on Becker and Arrow.

[6] Again we can show that $\ln(\hat{\gamma}_k + 1) - \ln(\hat{\gamma}_l + 1) = \ln(\hat{\theta}_k + 1) - \ln(\hat{\theta}_l + 1) + \ln(\hat{\delta}_k + 1) - \ln(\hat{\delta}_l + 1)$ is reduced to $\ln(\hat{\gamma}_g + 1) = \ln(\hat{\theta}_g + 1) + \ln(\hat{\delta}_g + 1)$ by letting k = g where g is a sample group, letting l = H where H denotes the pooled sample, and assuming that $\hat{\bar{W}} = \hat{\bar{W}}^0$. The last assumption implies that the estimated mean wage of the full sample equals the estimated non-discriminatory mean wage.

Using a model of employer discrimination the link between employers' discriminatory tastes and the definition of the competitive wage (non-discriminatory wage) is established.

In the following model the economy consists of identical firms. The firms produce Y with labour as the only input. There are two types of labour, A and B (e.g. skilled and unskilled). Within each type of labour there are two groups of workers, K and L (e.g. gender, race, union membership etc). The two groups of workers are perfect substitutes in production. Without discrimination workers would be paid a wage equal to their marginal product of labour.

$W_{KA}^{0} = MP_{KA} = MP_{LA} = W_{LA}^{0}$ [7] for type A workers and

$W_{KB}^{0} = MP_{KB} = MP_{LB} = W_{LB}^{0}$ for type B workers.

Productivity can thus differ between types of workers, but is assumed identical for a certain type of worker (e.g. skilled). Further we assume a strictly concave production function and fixed supply of labour (full employment).
Production can then be described as:

$$A = K_A + L_A \quad (31)$$
$$B = K_B + L_B \quad (32)$$
$$Y = f(A,B) = f[(K_A + L_A),(K_B + L_B)] \quad (33)$$
$$\Pi = f[(K_A + L_A),(K_B + L_B)] - W_{KA}K_A - W_{LA}L_A - W_{KB}K_B - W_{LB}L_B \quad (34)$$

Equations (31) and (32) indicate that the supply of different types of labour is completely inelastic, while (33) defines the production function and (34) expresses the firm's profit when the product price is numeraire.

Since we allow employers to have preferences over groups of labour they maximise not profits alone, but a strictly concave utility function where profit is an argument together with two types of employees (A, B) belonging to two different groups (K, L).

$$U = U(\Pi, K_A, L_A, K_B, L_B) \quad (35)$$

where $\frac{\partial U}{\partial \Pi} \geq 0$, $\frac{\partial U}{\partial K_A} \geq 0$ [8], $\frac{\partial U}{\partial K_B} \geq 0$, $\frac{\partial U}{\partial L_A} \leq 0$, $\frac{\partial U}{\partial L_B} \leq 0$

When the marginal utility of a group of workers is positive they are subject to nepotism, and when it is negative the group of workers is disfavoured. In the model it is imposed that group K workers are subject to nepotism and that group L workers are disfavoured.

[7] W_{ij}^{0} is the competitive wage of type j workers from group i.

MP_{ij} is the marginal product of type j workers from group i.
j=A, B
i=K, L

[8] Relatively speaking, discrimination is present as long as either the marginal utility of both types of K workers or L workers is different from zero. When restrictions are placed on all the marginal utilities, this is to enable the links to be made from discrimination theory to the various wage decompositions mentioned in Section 2.2.

Optimal behaviour of the employer is given by

$$\max_{\Pi,KA,LA,KB,LB}\{U(\Pi,K_A,L_A,K_B,L_B\}$$

The first order conditions are then given by[9]

$$\frac{\partial U}{\partial \Pi}\frac{\partial \Pi}{\partial K_A}+\frac{\partial U}{\partial K_A}=0 \quad \Leftrightarrow \quad \frac{\partial U}{\partial \Pi}\left(\frac{\partial f(A,B)}{\partial K_A}-W_{KA}\right)+\frac{\partial U}{\partial K_A}=0 \quad \text{and} \tag{36a}$$

$$\frac{\partial U}{\partial \Pi}\frac{\partial \Pi}{\partial L_A}+\frac{\partial U}{\partial L_A}=0 \quad \Leftrightarrow \quad \frac{\partial U}{\partial \Pi}\left(\frac{\partial f(A,B)}{\partial L_A}-W_{LA}\right)+\frac{\partial U}{\partial L_A}=0. \tag{36b}$$

Equivalent conditions hold for type B workers (unskilled).

Using rearrangements of (36a) and (36b) we can define

$$d_{KA}=-\frac{\partial U/\partial K_A}{\partial U/\partial \Pi} \quad \text{and} \quad d_{LA}=-\frac{\partial U/\partial L_A}{\partial U/\partial \Pi}. \tag{37}$$

d_{KA}, d_{LA}, can be interpreted as discrimination coefficients (or the difference between marginal productivity and the wage rate). If $d_{Kj} < 0$ (marginal utility of group K workers is positive) group K workers are subject to nepotism. If $d_{Lj} > 0$ group L workers are disfavoured. Substituting from equation (37) into (36a) and (36b) gives

For type A workers

$$W_{KA}=\frac{\partial f(A,B)}{\partial K_A}-d_{KA} \quad \text{, where} \quad d_{KA}\le 0 \quad \text{since} \quad \frac{\partial U}{\partial K_A}\ge 0 \tag{38a}$$

$$W_{LA}=\frac{\partial f(A,B)}{\partial L_A}-d_{LA} \quad \text{, where} \quad d_{LA}\ge 0 \quad \text{since} \quad \frac{\partial U}{\partial L_A}\le 0 \tag{38b}$$

For type B workers

$$W_{KB}=\frac{\partial f(A,B)}{\partial K_B}-d_{KB} \quad \text{, where} \quad d_{KB}\le 0 \quad \text{since} \quad \frac{\partial U}{\partial K_B}\ge 0 \tag{38c}$$

$$W_{LB}=\frac{\partial f(A,B)}{\partial L_B}-d_{LB} \quad \text{, where} \quad d_{LB}\ge 0 \quad \text{since} \quad \frac{\partial U}{\partial L_B}\le 0 \tag{38d}$$

---

[9] To make the first order conditions complete we should include $\frac{\partial U}{\partial \Pi}=0$, which is equivalent to the first order condition of profit maximisation. However, this term does not influence the following discussion and is therefore disregarded.

From (38a)-(38d) we see that the favoured group (K) of workers is paid at least their marginal product while the disfavoured group (L) gets paid at most their marginal product.

For type A workers (since the marginal productivities of K_A and L_A are assumed equal)

$$W_{KA} \geq \frac{\partial f(A,B)}{\partial A} \geq W_{LA} \tag{39a}$$

For type B workers (since the marginal productivities of K_B and L_B are assumed equal)

$$W_{KB} \geq \frac{\partial f(A,B)}{\partial B} \geq W_{LB} \tag{39b}$$

The model of employer discrimination has established a framework that enables linking the theory of discrimination to various ways of decomposing wage differences.

Let us first assume that there is no nepotism. Given the model this implies that $d_{KA} = d_{KB} = 0$ and $d_{LA}, d_{LB} \geq 0$. (At least one of the discrimination coefficients for group L workers must be strictly greater than zero. If not there is no discrimination.) From (38a) and (38c) we see that group K workers will be paid their marginal product, while L workers will be paid less than their marginal product. The discriminatory wage structure therefore pays group K workers according to their productivity, while group L workers are paid less then they are worth. In the absence of discrimination group K will retain their discriminatory wage while group L workers will receive a higher wage than with discrimination present. This situation maps to the wage decomposition where discrimination takes the form of one group being underpaid (see Equation (22b)).

Let employer discrimination be characterised by nepotism towards group K only. Then $d_{LA} = d_{LB} = 0$ and $d_{KA}, d_{KB} \leq 0$. Now group L workers will be paid their marginal product, while K workers will receive a wage in excess of their marginal productivity. In the absence of discrimination wages for group L workers would be unaltered while group K workers would be paid less. This scenario maps to the wage decomposition where discrimination is characterised by one group being overpaid (see equation (22a)).

When employer discrimination takes the form of either nepotism or disfavourism, a hypothetical elimination of discrimination, and the introduction of a competitive wage, will alter the received wage for only one group.

Perhaps the most interesting case is when we allow employer discrimination to take the form of both nepotism and disfavourism simultaneously. In other words, given the presence of both nepotism and disfavourism, one might ask what the wage structure would be like if discrimination should disappear? For further discussion to be meaningful an additional restriction on the utility function is imposed. It is assumed that the utility function is homogenous of degree zero within each type of labour.[10] The intuition behind this restriction is that employers are concerned only about the ratio of worker groups within each type of labour; the absolute number of workers from different groups is of no concern to the employer.

[10] $U(\Pi, t_1 K_A, t_1 L_A, t_2 K_B, t_2 L_B) = U$

Since the utility function is homogenous of degree zero it follows from Euler's rule (Sydsæter, Strøm and Berck 1999) that

$$\frac{\partial U}{\partial K_A}K_A + \frac{\partial U}{\partial L_A}L_A = 0 \qquad \text{and} \tag{40a}$$

$$\frac{\partial U}{\partial K_B}K_B + \frac{\partial U}{\partial L_B}L_B = 0. \tag{40b}$$

Dividing through by $-\frac{\partial U}{\partial \Pi}$ and using (38a-d) we obtain

$$\left(\frac{\partial f(A,B)}{\partial A} - W_{KA}\right)K_A + \left(\frac{\partial f(A,B)}{\partial A} - W_{LA}\right)L_A = 0 \Leftrightarrow \frac{\partial f(A,B)}{\partial A} = \frac{W_{KA}K_A + W_{LA}L_A}{K_A + L_A} = W_A^0 \tag{41a}$$

$$\left(\frac{\partial f(A,B)}{\partial B} - W_{KB}\right)K_B + \left(\frac{\partial f(A,B)}{\partial B} - W_{LB}\right)L_B = 0 \Leftrightarrow \frac{\partial f(A,B)}{\partial B} = \frac{W_{KB}K_B + W_{LB}L_B}{K_B + L_B} = W_B^0 \tag{41b}$$

Again the notation is simplified by noting that the marginal productivities of K_A and L_A are assumed to be equal (and thus A can represent either), and similiarly for K_B and L_B.The wage received in the absence of discrimination is a weighted average of the wages received with employer discrimination present. This holds for both types of labour. The competitive wage of the group facing nepotism would thus be lower in the absence of employer discrimination and the wage of the disfavoured group would be higher. This interpretation of discrimination maps to wage decompositions where the current wage structure is described by both over- and under-payments (see equation (23)).

3 Data Used and the Determination of Earnings Functions

3.1 Data

The data for the empirical analysis are derived from the 1993 Southern Africa Labour and Development Research Unit (SALDRU) Household Survey (SALDRU93) and the 1995 October Household Survey (OHS95).

SALDRU93 was a comprehensive household survey of the entire South African population (former TBVC states were also included). The survey was of a two-stage self-weighting design. First 360 clusters (enumerated areas) were determined and next 25 households from each cluster were surveyed. Thus some 9000 households were questioned and with an average of 5 persons per household the sample consists of approximately 45 000 individuals.

The October Household Survey has been conducted annually by Statistics South Africa since 1993, but 1994 was the first year in which a sample representative of the whole population was defined. The OHS95 survey was stratified based on provinces. Then 3000 enumerated areas were defined and next 10 households were questioned from each enumerated area. Thus some 30 000 households were surveyed in OHS95. Assuming an average of 5 persons per household, approximately 150 000 individuals were surveyed in the OHS95.

3.2 Sample Definition

Discrimination is estimated as the residual between the observed wage differential and the wage component explained by productivity characteristics. Since discrimination is analysed among full time employees, factors not addressing different pay for equal work are excluded. This implies that factors such as quality of education, the nature of experience and occupational barriers are treated as pre-labour market discrimination. This is not to insinuate that such factors are unimportant, but they are very difficult to quantify. The issue of discrimination in access to occupations is addressed through the use of dummy variables, while I have not controlled for quality of education. Females are treated separately so as not to induce gender bias into the discrimination estimates. Further, the focus is on regular wage employment. To capture the labour force an age restriction ranging from 15 to 64 is imposed. Moreover, self employed, casual workers, not currently employed workers and workers with non-positive wages are omitted from the sample. Employees with non-positive figures reported for hours worked last week are also excluded. Finally, observations with missing data on any variable included in the earnings functions, were disregarded. The samples then consist of regular, wage earning employees between 15 and 64 years of age. Given the described restrictions the sample size was reduced to 2223 male and 1570 female observations in 1993 and 11 727 male and 5822 female observations in 1995.

3.3 Earnings Functions

Utilising the residual as the basis for the discrimination estimates implies that correct specifications of the earning functions are crucial so as not to induce omitted variable bias into the estimates. Obviously, discrimination will seem smaller, the more of the wage differential that can be explained by productivity differences. Thus one should be careful concerning which variables to include in the model specification. Below, the choice of variables and the reasoning why they should be included in the earning functions are described.

Dependent Variable

The dependent variable is the log of the worker's gross wage. The wage variable includes the normal wage and overtime payment. Gross wages are used as the dependent variable, since deductions of a redistributional nature will lead to underestimation of the wage differential. SALDRU93 data on regular wage employees include self-employed professionals. Cross tabulation of occupation and employment status showed that no professionals reported their income from self-employment in OHS95. Some respondents defined themselves as semi self-employed and for this group the dependent variable is defined as gross wage, plus gross income from own account activities. (Regarding income from own account activities, actual expenses were deducted.)

Independent Variables

All discrimination estimates are based on estimated earnings functions. The earnings functions include two types of independent variable: data set variables, which are related to the design of the survey and should be included to get reliable point estimates and variance estimates; and individual variables, both quantitative and qualitative, which contain individual-specific information.

Data Set Variables

Weights

Household surveys are often designed to be self-weighting so that the sample reflects the population that is surveyed. But often the obtained sample fails to be self-weighting. SALDRU93 failed to represent the whole population because whites were over-represented as non-respondents and because some areas could not be surveyed due to the crime situation. In situations where the population characteristics are not reflected by the available sample, probability weights can be used as a means of correction. Weights are defined according to the inverse of the probability of being in the sample. Observations that are more likely to be in the sample than in the population will be weighted down and vice versa. As an example, white observations were weighted up in SALDRU93 since white observations were underrepresented in the sample.

If a sample does not represent the population it is taken from, failing to include weights will result in biased point estimates. Probability weights are thus included when estimating mean values. Using weights in regressions is a controversial issue, but following Deaton's (1997) advice weights are also included in the earnings function regressions. [11]

Clusters

Both SALDRU93 and OHS95 were clustered surveys. When data are collected from clusters information is "lost" since neighbouring households are more likely to produce similarities than two random households (income, crime rate, number of children, etc.). Observations within clusters are therefore not truly independent. Imagine an extreme case where data are gathered from an agricultural based cluster in a rural area in a developing country. If so, the marginal

[11] " If regressions are primarily descriptive, exploring association by looking at the mean of one variable conditional on others, the answer is straight forward: use the weight and correct the standard errors for design" (Deaton 1997, p 71).

household is not likely to give much additional information (Deaton 1997). Failing to take clusters into account will produce variances of the estimated variables that are too small.

Individual Characteristics

Education

To account for education a continuous variable measuring years of education is used. The education variable is obtained by transforming information on educational categories into number of years of education. An obvious problem using this variable is that access to and quality of education are unspecified. Different access to education may be regarded as pre-labour market discrimination. However, the quality of education is undoubtedly correlated with a worker's productivity, and is therefore likely to be reflected in wage differentials. But measuring quality of education is complicated, and the available data sets did not contain potential indicators regarding the quality of education (e.g. mean school results in nation-wide, standardised examinations). The influence of the quality of education on wages is thus omitted.

Education Squared

A squared term for years of education is included, allowing for the possibility that education's influence on wages is nonlinear.

Experience

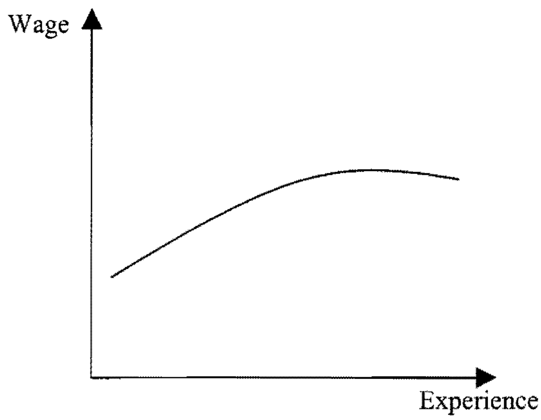
Potential experience is included as another independent variable. Actual experience is hard to measure and potential experience is used as a proxy. Potential experience is defined as worker's age minus the number of years of schooling and the age when first attending school (Mincer 1993).[12] While the accuracy of this measure can be questioned, it is arguable that simply using age itself is the greater of two evils.

Experience Squared

A quadratic term for experience is also included in the regression specification. This is to allow for increasing and then decreasing returns to experience over the life cycle, as suggested by human capital theory. The empirically observed wage-experience profile is illustrated in Figure 1 (Oaxaca 1973B). Based on this pattern one would expect the coefficient of potential experience to be positive and the coefficient of the squared term to be negative.

[12] Potential Experience = Age – Years of Education - 6

FIGURE 2 Wage-Experience Profile



Using potential experience as a proxy for actual experience the quantity of experience is likely to be overestimated, since temporary retreats from the labour market are not accounted for. Thus the peak of the wage-experience profile is unlikely to be accurate. Potential gains from experience also depend on other labour market characteristics such as occupational categories. As an illustration, one would imagine that a newly employed secretary has less to gain from experience compared to an inexperienced trainee manager (Bergmann 1973). Furthermore, any gains from experience that are linked, for example, to the quality of education, are unaccounted for.

For females a more reliable proxy for actual experience could be obtained by adjusting for number of children. But SALDRU93 reports number of children only for females younger than 49 years old. OHS95 reports number of children for females younger than 55 years. I was therefore deprived of the option of adjusting for number of children unless I wanted to restrict the sample size significantly.

Another variable that might influence wage earnings is tenure. But information on tenure is not available for both data sets and thus this variable has to be excluded.

Location

A dummy variable for location is included to test the hypothesis that workers in urban areas earn more than those in rural areas. SALDRU93 reports location in three categories: Metropolitan, Urban and Rural. Location in a metropolitan area indicates residence in one of the major South African cities. In this analysis metropolitan and urban categories are collapsed into one urban category.

Union membership

A dummy for union membership is included to enable a hypothesis test on whether unionised workers earn higher wages than non-unionised employees. SALDRU93 does not include specific information on union membership, but paid up union members are recorded and this figure is used as a proxy for union membership. This union dummy variable is assumed to be exogenous.

Industry

Dummy variables for nine one-digit industrial categories are included. Making industrial categories comparable across the two data sets did not involve any major difficulties.

Occupation

The earnings functions include dummies for eight one-digit occupational categories, since wages are observed to vary substantially among occupations. Occupation may also serve as an indication of workers' abilities other than quantifiable variables such as education. Given the historical background of South Africa, the occupational distribution reflects past discrimination in the way that certain groups were excluded from high earning occupations. The scope of the estimations is to determine discrimination in the form of unequal pay for equal work. Occupational barriers for certain groups are thus defined as pre-labour market discrimination.

The occupational breakdown in SALDRU93 is more limited than in the OHS95 survey. Hence adjustments to make OHS95 categories congruent with SALDRU93 categories had to be made. In 1993 'Professionals, Assosiate-Professionals and Technicians' were treated as one occupational category. Thus the more complete breakdown of the OHS95 survey had to be collapsed to fit the 1993 data. Similarly, 'Clerical and Sales' occupations were defined as one category in 1993 while these occupations were reported separately in 1995. Again 1995 data were adjusted to fit 1993 data. A further obstacle was that 'Farming and related' occupations from 1993 had changed to 'Skilled Agriculture and Fishery' workers in 1995. No attempt was made to adjust for this change and a reduction in the number of workers in agricultural occupations from 1993 to 1995 is expected. Further reductions in agricultural occupations imply increased numbers in unskilled occupations. In OHS95 army employees were not reported in separate occupations. Army workers therefore had to be excluded from the OHS95 sample.

Province

Dummies for South Africa's nine provinces are included. Provinces were defined differently in 1993 and 1995, but based on information about regions, data from 1993 can be transformed to match the provincial definition from 1994 onwards. Ideally, information on province of work would be preferred to information on province of residence. But data on the former were not available for SALDRU93. Province of residence is thus used as a proxy for province of work for both data sets. Since data are collected just before and just after the fall of the apartheid regime, significant movements between provinces may be expected.

After running ordinary least squares regressions according to the earnings funtion specifications described above, Wald adjusted F-tests for the joint significance of respectively education, experience, industry, occupation and province variables were performed. All F-values turned out to be significantly different from zero at the five per cent level.

4 Results

4.1 Males

Table 1 shows the weighted estimated mean values for males based on the 1993 and 1995 samples. It can be seen that the overall nominal mean wage has increased from 1993 to 1995. Coloureds and Whites contradict the overall trend by receiving lower nominal mean wages in 1995 compared to 1993, although the decrease is very small, particularly for Whites.

The mean number of years of education has increased when measured for all males. The increase in years of education is most apparent for Africans, while Coloureds have less education in 1995 than in 1993, on average. The definition of the potential experience variable would normally imply less potential experience when years of education increase. Still, Table 1 suggests increasing potential experience alongside increasing educational levels. A possible explanation for this anomaly is differing age distributions for the two years. If the age distribution in 1995 is more skewed toward the right than in 1993, the estimated mean values are consistent.

The urban-rural distribution suggests that approximately 12 per cent more male employees are located in urban areas in 1995 compared with 1993. Urbanisation is most noticeable among Africans. Coloureds and Asians show increasing concentrations in rural areas. Increasing numbers of male workers are organised in unions in 1995 (although the increase is only 3 per cent). Coloureds and Asians exhibit contradictive trends vis-à-vis the overall population.

The industrial distribution does not change dramatically from 1993 to 1995. A smaller proportion of employees in the mining and transport industries is noted, while the wholesale and service industries are expanding in terms of their share of wage employment. Mean values for occupational categories are more or less constant across the two data sets. As expected, agricultural occupations are declining, but the decline is a result of differing occupational definitions for the two surveys. In SALDRU93 agricultural occupations include all agricultural occupations, while OHS95 includes only skilled agricultural workers. Matching the decline in agricultural occupations, an increase in unskilled occupations is observed.

Provincial dummies reflect province of residence. Since data are collected the years before and after the first democratic election, the substantial movements among provinces which are observed, may be plausible. Especially prominent are movements out of Mpumalanga and North West Province and into Gauteng and Free State – presumably reflecting the geographical distribution of jobs. However, the overall means conceal large provincial changes for certain racial groups. For example, it seems unlikely that 18 per cent of Whites moved *out* of the Western Cape within a two-year periods, as indicated by the data. Part of the explanation may lie in the conversion of the SALDRU93 data from the four old provinces to the nine new provinces.

TABLE 1 Weighted Mean Values for the Variables Included in the Male Earning Functions

$$\delta_g(\theta_g + 1)$$

TABLE 2 Estimated Coefficients from the All Male Earning Functions

	1993 Male Earnings Regression					1995 Male Earnings Regression				
Variable	Coef.	Std. Err.	t	P>|t|	Dum Int	Coef.	Std. Err.	t	P>|t|	Dum Int
yrsed	-0.004	0.014	-0.285	0.776		-0.005	0.008	-0.566	0.576	
yrsedsq	0.007	0.001	7.407	0.000		0.007	0.001	13.355	0.000	
exp	0.051	0.005	11.372	0.000		0.037	0.003	14.061	0.000	
expsq	-0.001	0.000	-8.660	0.000		0.000	0.000	-9.093	0.000	
Rural	-0.299	0.077	-3.905	0.000	-0.259	-0.213	0.032	-6.567	0.000	-0.192
Urban	0.000	Base Case			0.000	0.000	Base Case			0.000
Member	0.031	0.057	0.552	0.581	0.032	0.108	0.014	7.458	0.000	0.114
Nonmem	0.000	Base Case			0.000	0.000	Base Case			0.000
Agriculture	-0.447	0.152	-2.944	0.004	-0.361	-0.734	0.045	-16.362	0.000	-0.520
Mining	0.081	0.132	0.613	0.541	0.084	-0.019	0.052	-0.359	0.722	-0.018
Manufact	0.000	Base Case			0.000	0.000	Base Case			0.000
Electricity	0.052	0.092	0.563	0.574	0.053	0.164	0.048	3.427	0.002	0.178
Construct	-0.121	0.073	-1.669	0.097	-0.114	-0.223	0.038	-5.869	0.000	-0.200
Wholesale	-0.174	0.053	-3.282	0.001	-0.160	-0.194	0.019	-10.018	0.000	-0.176
Transport	0.082	0.060	1.351	0.178	0.085	0.097	0.027	3.657	0.001	0.102
Finance	0.142	0.108	1.312	0.191	0.153	-0.023	0.036	-0.650	0.520	-0.023
Service	-0.125	0.059	-2.116	0.035	-0.117	-0.095	0.025	-3.722	0.001	-0.090
Professiona	0.824	0.090	9.163	0.000	1.280	0.679	0.028	24.085	0.000	0.972
Manager	1.086	0.100	10.821	0.000	1.963	1.046	0.047	22.080	0.000	1.845
Clerk	0.551	0.067	8.205	0.000	0.735	0.337	0.023	14.560	0.000	0.400
Service	0.482	0.082	5.849	0.000	0.620	0.387	0.036	10.867	0.000	0.473
Agriculture	-0.428	0.196	-2.186	0.030	-0.348	0.792	0.086	9.174	0.000	1.209
Trade	0.520	0.072	7.236	0.000	0.683	0.482	0.034	14.204	0.000	0.620
Skilled	0.313	0.059	5.304	0.000	0.368	0.241	0.018	13.536	0.000	0.273
Unskilled	0.000	Base Case			0.000	0.000	Base Case			0.000
W Cape	0.097	0.077	1.257	0.210	0.101	0.041	0.031	1.312	0.199	0.042
N Cape	-0.174	0.237	-0.735	0.463	-0.160	-0.126	0.058	-2.190	0.036	-0.119
E Cape	-0.090	0.115	-0.782	0.435	-0.086	-0.170	0.022	-7.748	0.000	-0.156
KZN	0.000	Base Case			0.000	0.000	Base Case			0.000
FS	-0.238	0.083	-2.884	0.004	-0.212	-0.374	0.031	-12.168	0.000	-0.312
Mpumal	0.104	0.102	1.021	0.308	0.109	0.006	0.043	0.137	0.892	0.006
N Province	-0.018	0.077	-0.233	0.816	-0.018	0.038	0.045	0.840	0.407	0.038
North W	-0.109	0.122	-0.895	0.372	-0.103	-0.151	0.061	-2.484	0.018	-0.140
Gauteng	0.142	0.101	1.408	0.161	0.153	0.086	0.037	2.352	0.025	0.090
_cons	5.682	0.116	48.962	0.000		5.995	0.069	86.994	0.000	

1993: F = 72, R-squared = 0.69, N = 2223

1995: F = 2238, R-squared = 0.70, N = 11,727

FIGURE 3 Male Wage-Education Profile

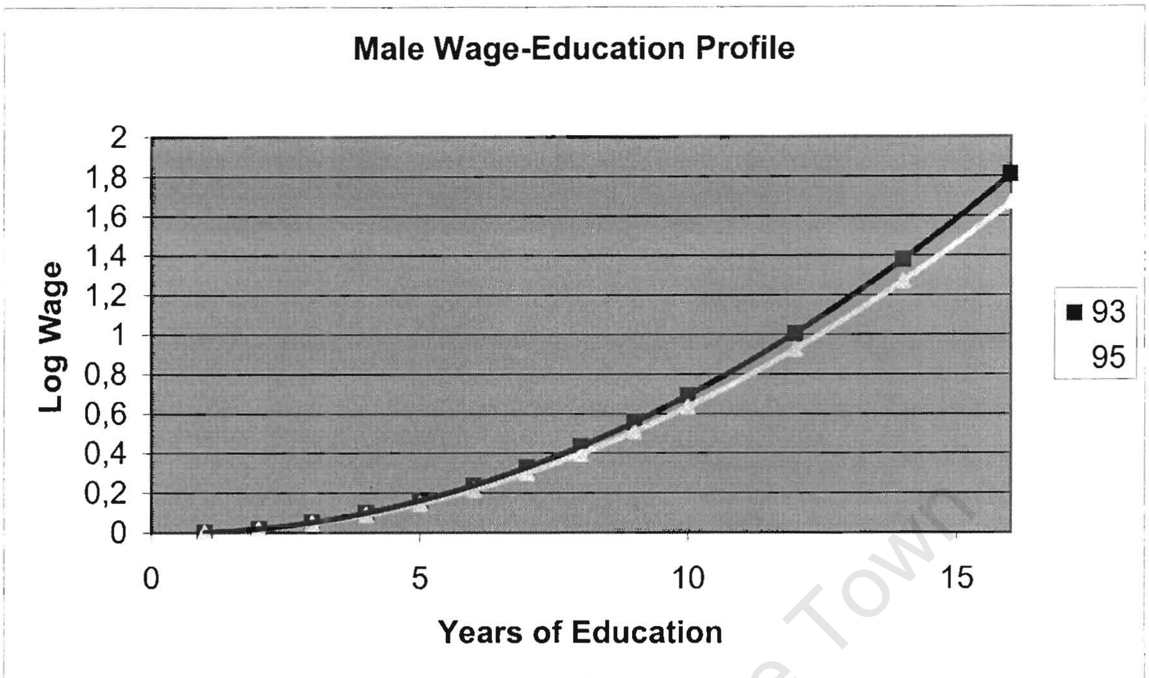
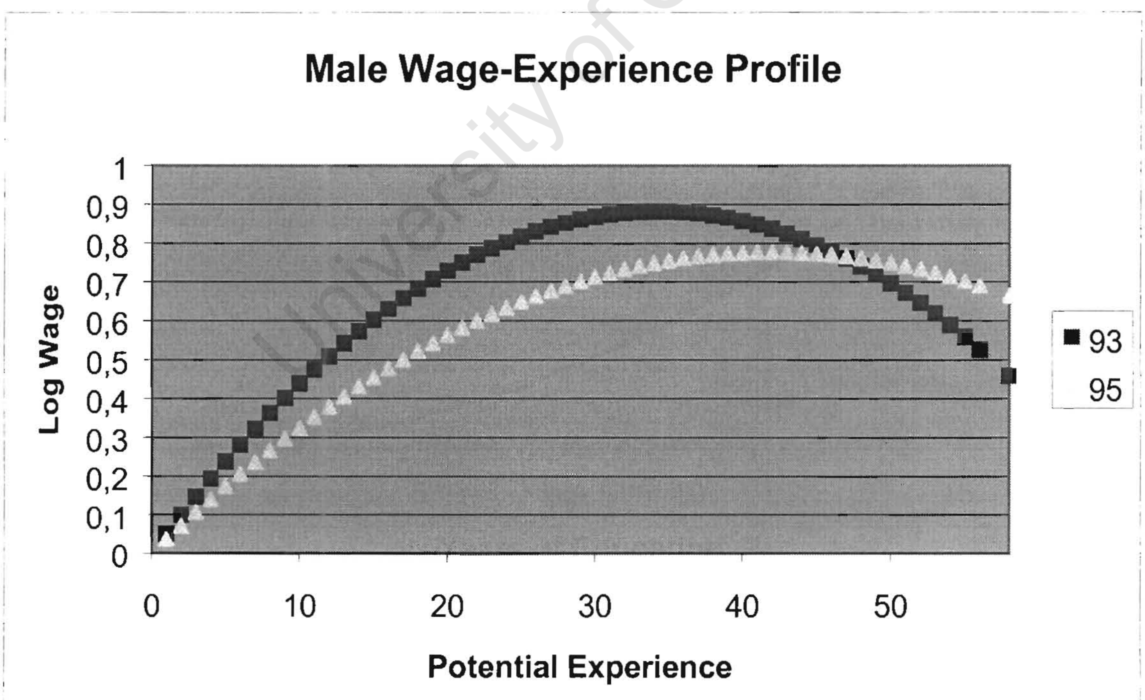


FIGURE 4 Male Wage-Experience Profile[13]



[13] It should be noted that the Wage-Experience profiles reflect only the estimated coefficients for the experience and experience squared variables in the earnings functions. The profiles in Figure 4 (and all the other Wage-Experience profiles) therefore do not illustrate the returns from potential experience over individual employees' life span, since the data do not follow workers over their life cycle.

Table 2 displays the results of the all-males weighted earnings regression. Given the semi-logarithmic form of the earning function, the interpretations of the coefficients are especially convenient. The interpretation of the coefficients of the continuous variables is that they measure the constant proportional or relative change in the dependent variable for a given absolute change in the value of an independent variable[14] (Gujarati 95).

The estimated coefficients of the dummy variables are also easily interpreted if Halvorsen and Palmquist's (1980) rule is applied: "Take the antilog of the estimated dummy coefficient and subtract one from it." Applying this rule, the coefficients of the dummy variables express the relative change in the mean value of the dependent variable when the value of a dummy variable changes from zero to one. Transformed values for the dummy coefficients according to Halvorsen and Palmquist's rule are reported in Table 2 in the column labelled "Dum Int."

The influence of education and potential experience on wages is illustrated in Figures 3 and 4. These figures are created by imposing a second order polynomial relation using the estimated coefficients for education and education squared, and experience and experience squared, respectively, as reported in Table 2. The values of all other variables are kept fixed. Figure 3 shows that education has positive and increasing marginal returns on wages. The impact of education on wages seems to be more or less constant across the two periods, 1993 and 1995. Figure 4 relates potential experience to wages in a manner consistent with human capital theory.

The results reported in Table 2, enable the rejection of the hypothesis that urban and rural locations yield identical wages. Rural location results in reduced earnings but the negative impact of rural location on earnings is reduced in 1995 compared to 1993. Based on the regression results, it can be argued that union members earn more than the non-unionised. The coefficient for union members is insignificant (5 per cent level) for the 1993 data, but based on 1995 data the coefficient is significant and indicates that union members earn 11 per cent more than non-unionised employees.

Using the manufacturing industry as the base case, it is observed that employees in the agriculture, wholesale and service industries earn significantly less than employees in the manufacturing industry both in 1993 and 1995. Furthermore, wages are lower in the construction industry, while employees in electricity and transport industries earn higher wages than in the manufacturing industry; however, these three results are statistically significant only in 1995. The coefficient for the mining industry dummy changes sign from 1993 (positive) to 1995 (negative), but it is insignificant for both years. Both finance industry coefficients are positive but insignificant.

Unskilled occupations were used as the base case and as expected all other occupations earn higher incomes, except for agricultural occupations in 1993. The coefficient for the agricultural occupational dummy further change sign moving from 1993 to 1995 data. The changing sign can again be explained by the new definition of agricultural occupations in 1995. As mentioned before agricultural occupations in 1995 included only skilled agricultural workers. It is

[14] An example using a single regression model illustrates this point.

$\ln(W) = \alpha + \beta X \Leftrightarrow W = e^{\alpha} e^{\beta X}$ The derivative of ln (W) with respect to X equals β.

The derivative of W with respect to X is given by: $\partial W / \partial X = e^{\alpha} e^{\beta X} \beta = W\beta$

Thus $\partial \ln(W) / \partial X = \beta = W^{-1}\, \partial W / \partial X$

interesting to note further that in both years the mangerial category returns the highest income – ceteris paribus – even higher than the professional category. This could be explained by the fact that the professional category includes associate professionals and technicians.

Most of the coefficients for individual provincial dummies are insignificant. It can be noted that employees with residential status in the Free State earn less than workers in KwaZulu Natal. The 1995 regression indicates that Gauteng workers earn 9 per cent more than employees in KwaZulu Natal. However, given the potential problems with this variable discussed earlier, not much reliance is placed on these results.

From Table 1 the estimated[15] observed wage differentials between the individual racial groups and the overall male mean wage can be computed. Furthermore, the estimated productivity component of the wage differential can be derived, based on the coefficients from the all-male earnings regression reported in Table 2 and the difference between the overall mean characteristics and the mean characteristics of the various races (see equation (29)). These results are reported in Table 3 below.

TABLE 3 Decomposition of Male Wage Differentials

Year	1993				1995			
Race	African	Coloured	Asian	White	African	Coloured	Asian	White
$\hat{\gamma}_g$	-0.369	-0.020	0.846	2.960	-0.273	-0.211	0.858	2.390
$\hat{\theta}_g$	-0.325	0.084	0.940	1.913	-0.187	-0.221	0.669	1.297
$\hat{\delta}_g(\hat{\theta}_g+1)$ [16]	-0.045	-0.105	-0.094	1.047	-0.086	0.011	0.189	1.093
$\hat{\delta}_g$	-0.066	-0.097	-0.049	0.360	-0.106	0.014	0.113	0.476

$\hat{\gamma}_g = \dfrac{\hat{\bar{W}}_g - \hat{\bar{W}}}{\hat{\bar{W}}}$ and expresses group g's mean wage relative to the mean wage of the whole sample (all males). Gamma thus defines the gross wage differential.

$\hat{\theta}_g = \dfrac{\hat{\bar{W}}_g^0 - \hat{\bar{W}}^0}{\hat{\bar{W}}^0}$ and expresses group g's non-discriminatory mean wage relative to the non-discriminatory mean wage of the whole sample. Theta defines the productivity differential.

$\hat{\delta}_g(\hat{\theta}_g+1) = \dfrac{\hat{\bar{W}}_g - \hat{\bar{W}}_g^0}{\hat{\bar{W}}_g^0}\left(\dfrac{\hat{\bar{W}}_g^0}{\hat{\bar{W}}^0}\right) = \dfrac{\hat{\bar{W}}_g - \hat{\bar{W}}_g^0}{\hat{\bar{W}}^0}$ and shows the difference between group g's mean wage and non-discriminatory wage relative to the non-discriminatory wage of the whole sample. $\delta_g(\theta_g+1)$ therefore illustrates the discrimination differential. Comparing $\delta_g(\theta_g+1)$ with δ_g as

[15] The wage differential is estimated in the sense that mean values are computed for the whole sample and individual racial groups.

[16] $\theta_g + \delta_g(\theta_g+1) = \dfrac{\bar{W}_g - \bar{W}^0}{\bar{W}^0} = \gamma_g$, assuming $\bar{W}^0 = \bar{W}$. $\hat{\delta}_g(\hat{\theta}_g+1)$ is thus computed as the difference between $\hat{\gamma}_g$ and $\hat{\theta}_g$.

defined by equation (10), we see that the former relates the difference between the mean wage and the mean wage in the absence of discrimination to the pooled mean wage, while δ_g relates the same difference to the mean wage of its own group.

Table 3 illustrates South Africa's wage ranking where Africans earn the least and Whites the most. Respectively Africans, Coloureds, Asians and Whites earn 37 per cent below, 2 per cent below, 85 per cent above and 296 per cent above the 1993 mean male monthly wage of R1128[17]. Of the African gross wage differential 88 per cent can be explained by productivity differences. The remaining 12 per cent of the gross wage differential can be interpreted as discrimination in the form of underpayment. Based on productivity characteristics, Coloureds are expected to earn more than the average male wage. Thus Coloureds are also subject to underpayment. Table 3 indicates that Asians are underpaid. Although the Asian mean wage was above the average male wage, Asians would have been paid even more in the absence of discrimination, reflecting their superior productivity. Whites earn 296 per cent above the male mean wage. Of the White wage differential, 65 per cent is due to above average productivity characteristics. The remaining 35 per cent of the wage differential results from Whites being overpaid.

Corresponding numbers for 1995 indicate that the wage ranking among races remains unaltered. In 1995 Africans earn 27 per cent less than the mean wage of R1305. Of the wage differential some 68 per cent is explained by productivity differences. Thus the residual, comprising the discrimination component, has increased for Africans from 1993 to 1995. The estimations suggest that Coloureds earn less compared to the overall mean wage in 1995 than in 1993. But different from 1993 is that below average payments are more than explained by (below average) productivity characteristics. The data thus suggest that Coloureds are overpaid in 1995. Asians and Whites earn respectively 85 per cent and 239 per cent more than the mean male wage. Some 78 per cent and 54 per cent of the Asian and White wage differential can be related to above average productivity characteristics, respectively.

Comparing the discrimination coefficient δ_g for 1993 and 1995 I find that discrimination is declining only in the Coloured group. In relative terms Africans are more extensively underpaid in 1995, Whites are overpaid relatively more, and Asians change from being faced by underpayment to overpayment. The results thus contradict the argument that in South Africa the long term trend has been declining discrimination (Moll 1998, Hinks 1999). On the other hand the findings are consistent with previous work based on post apartheid data sets. Hinks and Allanson (2000) also fail to find that discrimination is declining for all racial groups utilising the OHS95 and OHS97 surveys.

[17] The male mean monthly wage equals $e^{7.028} = 1127.77$

4.2 Females

Estimates for the female sub-sample are produced in the same fashion as for males. The results can therefore be interpreted equivalently for both males and females. From Table 4 below it can be observed that the overall nominal mean wage has increased from 1993 to 1995. An increase in wages is evident for all racial groups.

The mean number of years of education has increased both for all females and for the respective racial groups. As for males, the increase in education is most substantial for Africans. In contrast to males, however, females exhibit less potential experience 1995 than in 1993. Decreased potential experience is consistent with what one might expect due to the construction of the potential experience variable. Asians violate the general trend, having more experience in 1995 than in 1993. A possible explanation is differing age distributions for Asian females in the two years. The data suggest that females in general are also subject to urbanisation and increasing union membership. Coloureds and Asians deviate from the overall pattern.

The industrial distribution of employees is more or less constant based on data from 1993 and 1995. However, in terms of occupational categories a female-specific trend is noticeable, namely movements from service occupations into clerk and sales occupations. As for males, a decrease in agricultural occupations and an increase in unskilled occupations are observed. A likely explanation is the differing definitions of agricultural occupations in the two surveys.

Considerable movements across provinces are also observed for females. The main tendencies are similar for both genders, in that the figures in Table 4 indicate movements out of Mpumalanga and North West Province and into Gauteng and Free State. As before, certain reservations about the reliability of these figures are pertinent.

Table 5 displays the results from the pooled female weighted earnings regression. The influence of education and potential experience on wages is illustrated in Figures 5 and 6, respectively. As for males, a second order polynomial relationship between wages and education/experience is imposed, using the estimated coefficients from the earnings functions and disregarding all other estimated values. From Figure 5 it is seen that earnings are positively related to years of education. Furthermore, marginal returns to education are increasing. The wage-experience profile exhibits similar patterns to those suggested by human capital theory.

Rural location reduces earning by 28 per cent in 1993 and 16 per cent in 1995. Union membership has a positive influence on earnings. In 1993 union membership increased monthly earnings by 29 per cent while this figure was reduced to 14 per cent in 1995.

Among industrial categories estimations suggest that workers in the primary, mining, wholesale, and service industries earn less than those in the manufacturing industry. Workers in electricity, transport and finance industries earn more than manufacturing workers. However, some of the coefficients are not statistically significant, most notably for the mining sector. The coefficient for the construction industry changes sign from 1993 to 1995, but this coefficient is insignificant for both years.

TABLE 4 Weighted Mean Values for the Variables Included in the Female Earning Functions

		1993 Mean Values					1995 Mean Values				
Variable Name	**Code**	**All**	**African**	**Coloured**	**Asian**	**White**	**All**	**African**	**Coloured**	**Asian**	**White**
Log Gross Mth Wage	lnwage	6,646	6,211	6,729	7,159	7,739	7,170	6,932	6,907	7,504	7,831
Years of Education	yrsed	8,438	7,403	8,190	10,899	11,087	10,097	9,300	9,220	11,146	12,313
Years of Ed Squared	yrsedsq	89,717	72,821	78,616	127,015	139,447	116,216	103,786	95,420	132,434	154,787
P. Work Experience	exp	20,681	22,660	19,226	14,864	17,464	19,726	21,557	17,257	15,492	17,415
P. Work Exp Squared	expsq	568,258	654,503	479,003	360,448	437,358	520,303	594,778	409,062	350,876	431,862
Rural	locdum1	0,391	0,602	0,122	0,011	0,081	0,257	0,384	0,147	0,035	0,052
Urban	locdum2	0,609	0,398	0,878	0,989	0,919	0,743	0,616	0,853	0,965	0,948
Member	uniond1	0,231	0,206	0,436	0,416	0,093	0,315	0,374	0,295	0,361	0,177
Not Member	uniond2	0,769	0,794	0,564	0,584	0,907	0,685	0,626	0,705	0,639	0,823
Agric/Fish/Forest/Hunt	inddum1	0,106	0,137	0,134	0,000	0,017	0,087	0,114	0,141	0,004	0,006
Mining/Quarrying	inddum2	0,008	0,007	0,000	0,000	0,021	0,008	0,009	0,000	0,011	0,010
Manufacturing	inddum3	0,154	0,131	0,268	0,449	0,054	0,154	0,143	0,245	0,336	0,101
Electricity	inddum4	0,004	0,001	0,000	0,011	0,016	0,005	0,003	0,003	0,000	0,011
Construction	inddum5	0,012	0,011	0,012	0,011	0,016	0,008	0,004	0,004	0,013	0,019
Wholesale	inddum6	0,162	0,131	0,219	0,135	0,213	0,224	0,217	0,267	0,205	0,220
Transort	inddum7	0,032	0,020	0,012	0,011	0,091	0,013	0,007	0,007	0,021	0,028
Finance	inddum8	0,045	0,015	0,012	0,056	0,157	0,097	0,044	0,061	0,123	0,238
Service	inddum9	0,477	0,547	0,343	0,326	0,415	0,404	0,458	0,271	0,285	0,367
Professional	occdum1	0,176	0,166	0,131	0,180	0,242	0,246	0,256	0,114	0,245	0,294
Manager	occdum2	0,047	0,005	0,008	0,034	0,204	0,017	0,010	0,011	0,018	0,038
Clerk & Sale	occdum3	0,205	0,105	0,201	0,405	0,454	0,323	0,209	0,294	0,439	0,591
Service	occdum4	0,209	0,315	0,038	0,056	0,069	0,086	0,100	0,119	0,024	0,045
Agriculture	occdum5	0,043	0,072	0,004	0,000	0,000	0,002	0,002	0,004	0,000	0,000
Craft & Trade	occdum6	0,015	0,009	0,023	0,056	0,013	0,029	0,032	0,042	0,065	0,009
Skilled Worker	occdum7	0,098	0,084	0,210	0,258	0,010	0,058	0,060	0,107	0,157	0,011
Unskilled Worker	occdum8	0,208	0,243	0,385	0,011	0,008	0,239	0,331	0,308	0,052	0,011
Western Cape	provdum1	0,207	0,045	0,745	0,000	0,305	0,135	0,027	0,661	0,024	0,121
Northern Cape	provdum2	0,013	0,004	0,067	0,000	0,000	0,012	0,005	0,039	0,000	0,016
Eastern Cape	provdum3	0,098	0,120	0,075	0,000	0,075	0,113	0,144	0,086	0,016	0,069
KwaZulu/Natal	provdum4	0,189	0,180	0,106	0,989	0,077	0,187	0,214	0,046	0,770	0,104
Free State	provdum5	0,013	0,021	0,000	0,000	0,000	0,056	0,069	0,015	0,000	0,056
Mpumalanga	provdum6	0,131	0,186	0,003	0,011	0,103	0,046	0,064	0,004	0,003	0,034
Northern Province	provdum7	0,084	0,114	0,000	0,000	0,087	0,057	0,087	0,011	0,000	0,018
North West	provdum8	0,110	0,165	0,000	0,000	0,063	0,071	0,097	0,009	0,003	0,053
Gauteng	provdum9	0,156	0,165	0,003	0,000	0,290	0,324	0,294	0,131	0,183	0,528
Sample Size		1570	957	259	89	265	5822	3272	1090	306	1154

TABLE 5 Estimated Coefficients from the All Female Earning Functions

	1993 Female Earnings Regression					1995 Female Earnings Regression				
Variable	Coef.	Std. Err.	t	P>|t|	Dum Int	Coef.	Std. Err.	t	P>|t|	Dum Int
yrsed	-0.010	0.019	-0.509	0.612		0.026	0.009	2.728	0.010	
yrsedsq	0.005	0.001	4.365	0.000		0.004	0.001	7.698	0.000	
exp	0.037	0.005	7.898	0.000		0.025	0.003	9.086	0.000	
expsq	-0.001	0.000	-5.772	0.000		0.000	0.000	-4.320	0.000	
Rural	-0.327	0.074	-4.427	0.000	-0.279	-0.177	0.022	-8.201	0.000	-0.162
Urban	0.000	Base Case			0.000	0.000	Base Case			0.000
Member	0.258	0.040	6.459	0.000	0.294	0.133	0.015	8.658	0.000	0.142
Nonmem	0.000	Base Case			0.000	0.000	Base Case			0.000
Agriculture	-0.379	0.092	-4.136	0.000	-0.315	-0.611	0.059	-10.326	0.000	-0.457
Mining	-0.016	0.179	-0.087	0.931	-0.015	-0.012	0.168	-0.070	0.945	-0.012
Manufact	0.000	Base Case			0.000	0.000	Base Case			0.000
Electricity	0.696	0.183	3.801	0.000	1.006	0.153	0.094	1.630	0.113	0.165
Construct	-0.125	0.247	-0.506	0.613	-0.118	0.058	0.137	0.424	0.675	0.060
Wholesale	-0.075	0.058	-1.308	0.192	-0.072	-0.255	0.033	-7.834	0.000	-0.225
Transport	0.235	0.088	2.665	0.008	0.265	0.039	0.087	0.446	0.659	0.040
Finance	0.165	0.075	2.195	0.029	0.179	0.122	0.032	3.878	0.000	0.130
Service	-0.224	0.055	-4.073	0.000	-0.201	-0.033	0.032	-1.036	0.308	-0.033
Professiona	1.452	0.086	16.976	0.000	3.272	0.698	0.026	27.039	0.000	1.009
Manager	1.359	0.098	13.795	0.000	2.890	0.827	0.079	10.453	0.000	1.288
Clerk	0.865	0.077	11.207	0.000	1.376	0.497	0.024	20.909	0.000	0.644
Service	0.119	0.075	1.586	0.114	0.126	0.176	0.038	4.613	0.000	0.192
Agriculture	-0.140	0.135	-1.040	0.299	-0.131	0.113	0.103	1.096	0.281	0.119
Trade	0.490	0.219	2.234	0.026	0.633	0.053	0.061	0.869	0.392	0.055
Skilled	0.410	0.069	5.980	0.000	0.507	0.159	0.029	5.444	0.000	0.172
Unskilled	0.000	Base Case			0.000	0.000	Base Case			0.000
W Cape	0.210	0.060	3.512	0.001	0.234	0.026	0.045	0.578	0.567	0.026
N Cape	-0.178	0.138	-1.292	0.198	-0.163	-0.203	0.041	-4.892	0.000	-0.184
E Cape	-0.030	0.091	-0.330	0.742	-0.029	-0.139	0.033	-4.278	0.000	-0.130
KZN	0.000	Base Case			0.000	0.000	Base Case			0.000
FS	-0.258	0.281	-0.920	0.358	-0.228	-0.377	0.056	-6.720	0.000	-0.314
Mpumal	-0.146	0.090	-1.617	0.107	-0.136	-0.025	0.049	-0.518	0.608	-0.025
N Province	-0.002	0.085	-0.029	0.977	-0.002	-0.017	0.040	-0.417	0.679	-0.017
North W	0.054	0.104	0.517	0.606	0.055	-0.085	0.046	-1.853	0.073	-0.082
Gauteng	0.211	0.079	2.675	0.008	0.235	0.176	0.032	5.537	0.000	0.193
_cons	5.389	0.123	43.756	0.000		5.804	0.066	87.478	0.000	

1993: F = 95.24, R-squared = 0.70, N = 1570

1995: F = 6396, R-squared = 0.66, N = 5822

FIGURE 5 Female Wage-Education Profile

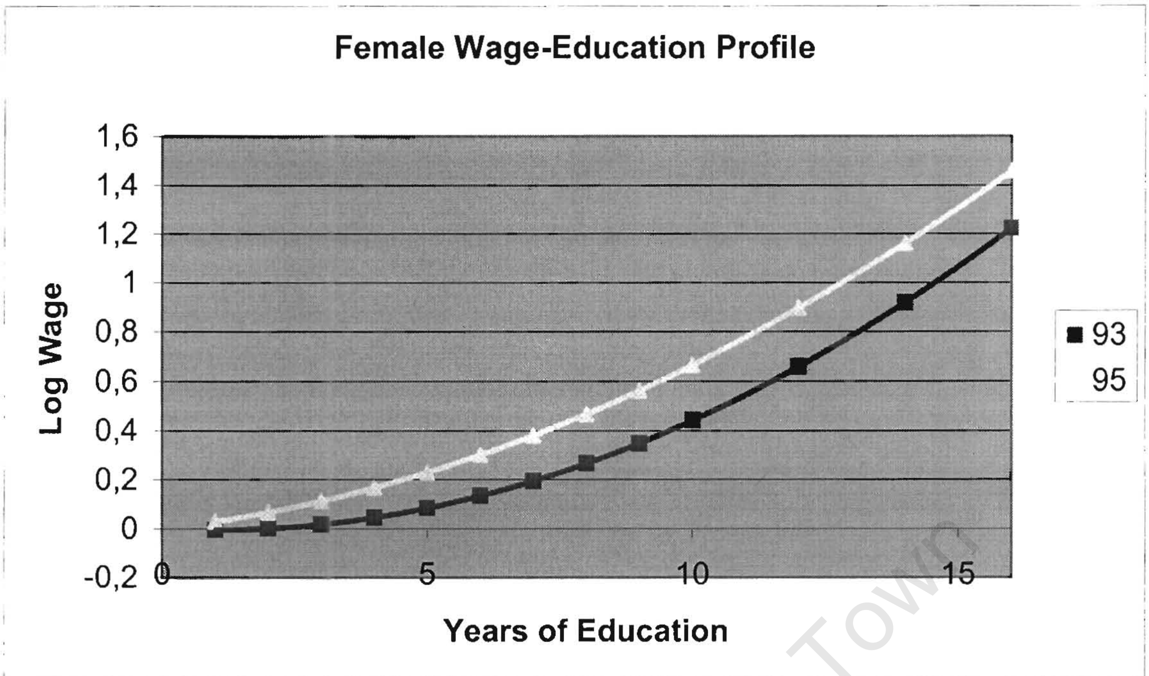
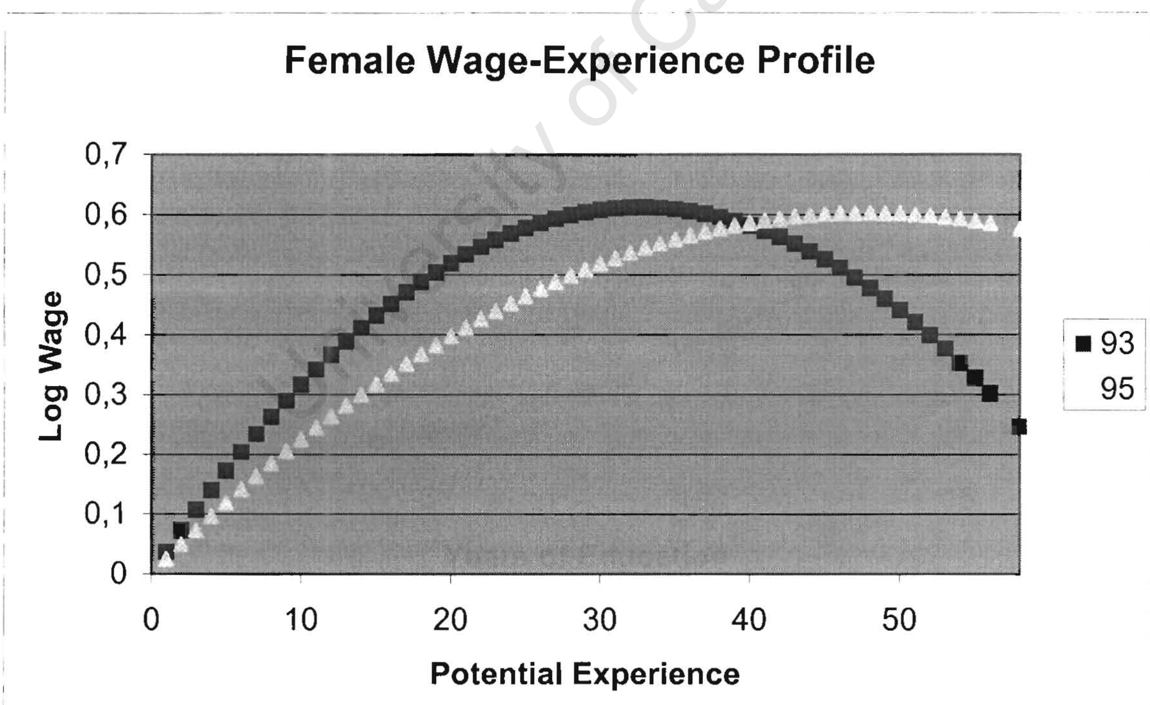


FIGURE 6 Female Wage-Experience Profile



Defining unskilled occupations as the base case, it is observed that almost all other occupations yield higher earnings. The exception is agricultural occupations in 1993. This coefficient further changes sign from 1993 (negative) to 1995 (positive), which again can be explained by the fact that agricultural occupations included only skilled workers in 1995 while all agricultural workers were included in 1995.

Most of the individual dummy variables for provinces are insignificant. Only the coefficient for Gauteng is significant for both years; the positive sign indicates that residents of Gauteng enjoy higher earnings, on average.

Based on the estimated mean values in Table 5, the observed wage differential between each racial group and the whole sample of females can be determined. Using the regression coefficients for all females and the differences in mean values between each racial group and the entire female workforce, the productivity component of the observed wage differential can be estimated. These results are presented in Table 7.

TABLE 6 Decomposition of Female Wage Differentials

Year	1993				1995			
Race	African	Coloured	Asian	White	African	Coloured	Asian	White
$\hat{\gamma}_g$	-0.352	0.086	0.670	1.984	-0.212	-0.231	0.396	0.936
$\hat{\theta}_g$	-0.304	0.087	0.659	1.410	-0.143	-0.232	0.246	0.617
$\hat{\delta}_g(\hat{\theta}_g+1)$	-0.049	-0.001	0.011	0.574	-0.069	0.002	0.150	0.319
$\hat{\delta}_g$	-0.070	-0.001	0.007	0.238	-0.080	0.002	0.121	0.197

The 1993 results support the notion of a South African wage hierarchy where Whites earn the most and Africans the least. Similar figures for 1995 indicate that the wage ranking has changed. From 1993 to 1995 Coloureds and Africans swapped places in the wage hierarchy with Coloureds now earning the least. In 1993, respectively Africans, Coloureds, Asians and Whites earned 35 per cent less, and 8.6 per cent, 67 per cent and 198 per cent above an average monthly wage of R770.[18] For Africans 86 per cent of the wage differential can be explained by below average productivity characteristics while the remaining component potentially reflects discrimination. Productivity characteristics explain almost fully the observed wage differential for Coloureds and Asians. For Whites, productivity characteristics explain roughly 71 percent of the wage differential.

Moving to 1995, Africans, Coloureds, Asians and Whites earned respectively 21 per cent less, 23 per cent less, 40 per cent more and 94 per cent more than an average monthly wage of R1300.[19] The components of the wage differentials explained by productivity characteristics were respectively for Africans, Coloureds, Asians and Whites some 67 per cent, 100 per cent, 62 per cent and 66 per cent.

[18] $e^{6.646} = 770$

[19] $e^{7.170} = 1300$

Based on the estimations, Whites are the only group where the unexplained component of the wage differential is decreasing. Thus it is only on behalf of Whites it can argued that discrimination is declining. For the other racial groups discrimination is potentially increasing. This fact is reflected in the values for the discrimination coefficient (δ_g) for the two years. The difference of the absolute value of the discrimination coefficient from 1995 to 1993 is negative only for Whites.

5 Conclusions

In most academic work the results obtained hinge to an important degree on underlying assumptions. The present paper is no exception to this rule. Important assumptions will thus be summarised before major findings are pointed out.

The focus has been on discrimination among full time wage earning employees. Other forms of discrimination, labelled pre-labour market discrimination, have at best been controlled for (e.g. dummy variables for occupational categories) and at worst not controlled for owing to lack of data (e.g. quality of education). Further, the applied decomposition method estimates discrimination as a residual between observed wage differentials and the wage component explained by productivity characteristics. Misspecification of the earnings functions will thus induce bias into the discrimination estimates. A more fundamental assumption is that the presence of discrimination has only distributional effects. In other words, the wage bill is regarded as constant whether discrimination is present or not. In addition, to make discrimination estimates operational, one needs a notion of what the wage structure would be like in the absence of discrimination. In the present paper, average characteristics based on a pooled sample[20] containing all racial groups is used as a proxy for the non-discriminatory wage structure. Lastly, certain ad hoc assumptions were unavoidable when attempting to make data from 1993 congruent with data from 1995.

African males were subject to discrimination in the form of underpayment in both 1993 and 1995. However, the largest share of the observed wage differential for this group was due to below average productivity characteristics. For Coloured and Asian males, substantial wage differentials were mostly explained by above or below average productivity characteristics (relative to the average male worker). An exception was found for Asian males in 1995, in whose case above average productivity characteristics accounted for some 78 per cent of the observed wage differential. Based on discrimination coefficients, it was found that discrimination was most remarkable towards white males. Whites are overpaid, illustrated by the fact that in 1993 and 1995, productivity characteristics could explain no more than approximately 65 per cent and 54 per cent of the observed wage differential, respectively.

Among females, discrimination exhibits similar patterns to those found for males. Differing productivity compared to the average female worker explains by far the greatest part of the observed wage differentials for Africans, Coloureds and Asians. But productivity as the explanatory force is not as predominant among Africans as among Coloureds and Asians. The results for Asian females in 1995 are an exception. In 1995 Asian productivity accounted for only 62 per cent of the wage differential. Again, the data suggest that the most striking discriminatory effect is defined by White overpayment. For White females, above average productivity characteristics account for 71 per cent and 66 per cent of the wage differential in 1993 and 1995, respectively.

In sum, using data from 1993 and 1995, variation in discrimination trends among racial groups seems to be the rule. At the least, I cannot conclude that discrimination is declining as an overall trend for all racial groups. But what do the differences when comparing the 1993 results with the 1995 results really reflect? Two possible explanations seem plausible: the differences reflect either changes over time or data inconsistencies. If the time span of two years is too short to pick

[20] Discriminatory wages are present in the pooled sample.

up considerable changes in the underlying wage structure, then variation in discriminatory trends probably reflects data inconsistencies. A more likely scenario is that changing discrimination coefficients incorporate both changes over time and problems with the data. Unfortunately, it is not clear how to separate these effects. Despite these problems, one finding stands out, namely that Whites are overpaid relative to their measured productivity. However, some of this effect could be due to superior quality of education, which is not controlled for.

6 References

Altonji, J. G. and Blank, R. M. (1999), 'Race and Gender in the Labour Market', in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Vol 3C, Elsevier, Amsterdam, pp 3143-3260.

Arrow, K. (1973A), 'The Theory of Discrimination', in O. Ashenfelter and A. Rees (eds), *Discrimination in Labor Markets*, Princeton University Press, New Jersey, pp 3-33.

Arrow, K. (1973B), 'Some Mathematical Models of Race in the Labour Market', in A. Pascal (ed), *Racial Discrimination in Economic Life*, Lexington Books, D. C. Heath and Company, Lexington, Massachusetts pp 187-205.

Becker, G. (1971), '*The Economics of Discrimination*', Second edition, Chicago: University of Chicago Press.

Bergmann, B. (1973), comments to Oaxaca's article: 'Sex Discrimination in Wages', in O. Ashenfelter and A. Rees (eds), *Discrimination in Labor Markets*, Princeton University Press, New Jersey, pp 152-154.

Cotton, J. (1988), 'On the Decomposition of Wage Differentials', *Review of Economics and Statistics*, Vol 70, pp 236-243.

Deaton, A. (1997), *'The Analysis of Household Surveys: A Microeconometric Approach to Development Policy'*, Baltimore: John Hopkins University Press.

Gujarati, D. (1995), '*Basic Econometrics'*, Singapore: McGraw-Hill.

Halvorsen, R. and Palmquist, R. (1980), 'The Interpretation of Dummy Variables in Semilog-Arithmic Equations', *American Economic Review*, Vol 70, No. 3, pp 474-475.

Hinks, T. (1999), 'Racial Wage Discrimination and the End of Apartheid in South Africa: A Multilateral Approach', *Discussion Paper Series,* No. 77, Middlesex University Business School.

Hinks, T. and Allanson, P. (2000), 'Will the End of Apartheid Spell the End of Racial Discrimination in South Africa? A Multilateral Analysis of Racial Differentials in the Early Post-Apartheid Period', *Working Draft*.

Knight, J. B. and McGrath M. D. (1987), 'The Erosion of Apartheid in the South African Labour Market: Measures and Mechanisms', *Applied Economics Discussion Paper*, No. 35, Institute of Economics and Statistics, Oxford University.

Marshall, R. (1974), 'The Economics of Racial Discrimination-A Survey', *Journal of Economic Literature*, Vol 12, pp 849-872.

Moll, P. (1998), 'Discrimination is declining in South Africa but inequality is not', *Working Paper 5*, South African Network for Economic Research.

Neumark, D. (1988), 'Employers' Discriminatory Behaviour and the Estimation of Wage Discrimination', *Journal of Human Resources,* Vol 23, pp 279-295.

Oaxaca, R. (1973A), 'Male Female Wage Differentials in Urban Labor Markets', *International Economic Review*, Vol 14, pp 693-709.

Oaxaca, R. (1973B), 'Sex Discrimination in Wages', in O. Ashenfelter and A. Rees (eds), *Discrimination in Labor Markets*, Princeton University Press, New Jersey, pp 124-152.

Oaxaca, R. and Ransom, M. (1988), 'Searching for the Effect of Unionism on the Wages of Union and Non-Union Workers', *Journal of Labor Research,* Vol 9, pp 139-148.

Oaxaca, R. and Ransom, M. (1994), 'On Discrimination and the Decomposition of Wage Differentials', *Journal of Econometrics*, Vol 61, pp 5-21.

Sapsford, D. and Tzannatos, Z. (1993), '*The Economics of the Labour Market*', Mackays of Chatham plc Chatham, Kent.

Sydsæter, K., Strøm, A. and Berck, P. (1999), '*Economists' Mathematical Manual'*, Berlin: Springer-Verlag.

7 Appendix

The appendix contains regression results for individual racial groups. Tables A1-A4 state the estimated regression coefficients for the four male racial groups. Figures A1 and A2 illustrate the male wage-education profiles for the years 1993 and 1995. Figures A3 and A4 show the wage-experience patterns in 1993 and 1995. The remaining tables and figures contain the same kind of information as described above, but for females. Based on the estimates in this appendix bilateral discrimination coefficients, as described in Section 2, can be estimated.

TABLE A1 Estimated Coefficients from the African Male Earning Functions

	1993 African Male Earnings Regression					1995 African Male Earnings Regression				
Variable	Coef.	Std. Err.	t	P>|t|	Dum Int	Coef.	Std. Err.	t	P>|t|	Dum Int
yrsed	0,022	0,017	1,295	0,197		-0,005	0,009	-0,535	0,596	
yrsedsq	0,004	0,001	3,278	0,001		0,005	0,001	7,249	0,000	
exp	0,048	0,006	8,152	0,000		0,031	0,003	10,082	0,000	
expsq	-0,001	0,000	-6,694	0,000		0,000	0,000	-7,029	0,000	
locdum1	-0,173	0,094	-1,843	0,067	-0,159	-0,139	0,032	-4,356	0,000	-0,130
locdum2	0,000	Base Case			0,000	0,000	Base case			0,000
uniond1	0,092	0,071	1,293	0,198	0,096	0,196	0,018	11,116	0,000	0,217
uniond2	0,000	Base Case			0,000	0,000	Base case			0,000
inddum1	-0,616	0,197	-3,131	0,002	-0,460	-0,821	0,053	-15,468	0,000	-0,560
inddum2	-0,124	0,148	-0,836	0,404	-0,116	-0,052	0,051	-1,026	0,313	-0,051
inddum3	0,000	Base Case			0,000	0,000	Base case			0,000
inddum4	0,105	0,122	0,863	0,389	0,111	0,204	0,060	3,395	0,002	0,227
inddum5	-0,195	0,101	-1,938	0,054	-0,177	-0,149	0,041	-3,665	0,001	-0,138
inddum6	-0,173	0,064	-2,689	0,008	-0,159	-0,189	0,028	-6,832	0,000	-0,172
inddum7	-0,031	0,070	-0,447	0,655	-0,031	0,066	0,031	2,150	0,039	0,068
inddum8	-0,006	0,173	-0,036	0,971	-0,006	-0,054	0,039	-1,371	0,180	-0,052
inddum9	-0,004	0,081	-0,051	0,959	-0,004	0,000	0,030	0,004	0,997	0,000
occdum1	0,580	0,106	5,479	0,000	0,786	0,562	0,047	11,957	0,000	0,754
occdum2	0,966	0,165	5,871	0,000	1,627	0,740	0,101	7,335	0,000	1,096
occdum3	0,518	0,077	6,740	0,000	0,678	0,292	0,022	13,226	0,000	0,340
occdum4	0,307	0,086	3,555	0,000	0,359	0,346	0,031	10,993	0,000	0,414
occdum5	-0,449	0,231	-1,945	0,053	-0,362	0,405	0,160	2,527	0,017	0,500
occdum6	0,228	0,081	2,812	0,005	0,256	0,302	0,030	10,137	0,000	0,352
occdum7	0,258	0,060	4,284	0,000	0,294	0,251	0,018	13,631	0,000	0,285
occdum8	0,000	Base Case			0,000	0,000	Base case			0,000
provdum1	0,073	0,129	0,562	0,575	0,075	0,023	0,044	0,528	0,601	0,024
provdum2	0,287	0,211	1,360	0,176	0,333	-0,154	0,115	-1,341	0,189	-0,143
provdum3	-0,118	0,128	-0,928	0,355	-0,112	-0,183	0,040	-4,516	0,000	-0,167
provdum4	0,000	Base Case			0,000	0,000	Base case			0,000
provdum5	-0,194	0,084	-2,308	0,022	-0,177	-0,392	0,039	-10,104	0,000	-0,325
provdum6	0,172	0,117	1,466	0,144	0,188	-0,014	0,046	-0,300	0,766	-0,014
provdum7	0,018	0,083	0,217	0,829	0,018	0,072	0,062	1,155	0,257	0,074
provdum8	-0,102	0,126	-0,806	0,421	-0,097	-0,151	0,070	-2,164	0,038	-0,140
provdum9	0,118	0,114	1,029	0,305	0,125	0,017	0,041	0,420	0,677	0,017
_cons	5,675	0,150	37,738	0,000		6,099	0,059	103,016	0,000	

F=32.79
R-Squared= 0.6080
N=1446

F=1482.75
R-Squared= 0.66718
N=7442

TABLE A2 Estimated Coefficients from the Coloured Male Earning Functions

Variable	1993 Coloured Male Earnings Regression					1995 Coloured Male Earnings Regression				
	Coef.	Std. Err.	t	P>|t|	Dum Int	Coef.	Std. Err.	t	P>|t|	Dum Int
yrsed	-0,007	0,037	-0,177	0,860		-0,001	0,019	-0,029	0,977	
yrsedsq	0,005	0,003	2,042	0,048		0,005	0,001	4,350	0,000	
exp	0,038	0,012	3,112	0,004		0,042	0,004	10,263	0,000	
expsq	-0,001	0,000	-1,690	0,100		-0,001	0,000	-6,678	0,000	
locdum1	-0,707	0,198	-3,579	0,001	-0,507	-0,170	0,086	-1,972	0,060	-0,156
locdum2	0,000	Base Case			0,000	0,000	Base Case			0,000
uniond1	0,263	0,104	2,535	0,016	0,301	0,185	0,031	5,955	0,000	0,203
uniond2	0,000	Base Case			0,000	0,000	Base Case			0,000
inddum1	0,147	0,160	0,918	0,364	0,158	-0,420	0,097	-4,316	0,000	-0,343
inddum2	0,673	0,355	1,897	0,066	0,960	0,078	0,186	0,416	0,681	0,081
inddum3	0,000	Base Case			0,000	0,000	Base Case			0,000
inddum4	-0,248	0,152	-1,634	0,111	-0,219	0,236	0,128	1,851	0,077	0,266
inddum5	-0,132	0,127	-1,037	0,307	-0,123	-0,183	0,051	-3,588	0,001	-0,167
inddum6	-0,030	0,095	-0,320	0,751	-0,030	-0,121	0,037	-3,309	0,003	-0,114
inddum7	0,191	0,140	1,362	0,182	0,211	0,061	0,086	0,715	0,482	0,063
inddum8	0,025	0,289	0,085	0,933	0,025	-0,100	0,060	-1,655	0,111	-0,095
inddum9	0,130	0,125	1,038	0,306	0,138	-0,022	0,045	-0,479	0,637	-0,021
occdum1	0,554	0,129	4,281	0,000	0,740	0,774	0,077	10,050	0,000	1,169
occdum2	1,295	0,176	7,341	0,000	2,652	0,713	0,113	6,291	0,000	1,039
occdum3	0,249	0,150	1,656	0,106	0,282	0,385	0,080	4,787	0,000	0,469
occdum4	0,485	0,151	3,209	0,003	0,624	0,435	0,044	9,882	0,000	0,545
occdum5	-0,170	0,276	-0,616	0,542	-0,156	0,493	0,284	1,735	0,096	0,638
occdum6	0,486	0,101	4,824	0,000	0,626	0,364	0,059	6,225	0,000	0,439
occdum7	0,034	0,126	0,273	0,787	0,035	0,296	0,043	6,899	0,000	0,345
occdum8	0,000	Base Case			0,000	0,000	Base Case			0,000
provdum1	-0,049	0,108	-0,451	0,655	-0,048	-0,271	0,088	-3,092	0,005	-0,238
provdum2	-0,667	0,324	-2,056	0,047	-0,487	-0,511	0,125	-4,092	0,000	-0,400
provdum3	0,092	0,137	0,672	0,506	0,096	-0,426	0,084	-5,077	0,000	-0,347
provdum4	0,000	Base Case			0,000	0,000	Base Case			0,000
provdum5	0,000	No Obs			0,000	-0,572	0,195	-2,928	0,007	-0,436
provdum6	0,000	No Obs			0,000	0,077	0,222	0,344	0,734	0,080
provdum7	0,000	No Obs			0,000	0,159	0,171	0,931	0,361	0,173
provdum8	-0,597	0,426	-1,402	0,169	-0,450	-0,326	0,155	-2,105	0,046	-0,278
provdum9	0,590	0,264	2,238	0,032	0,804	0,129	0,096	1,350	0,189	0,138
_cons	5,933	0,236	25,162	0,000		6,187	0,154	40,112	0,000	

F=21.62
R-Squared=0.5522
N=271

F=9711.82
R-Squared=0.6626
N=1926

TABLE A3 Estimated Coefficients from the Asian Male Earning Functions

Variable	1993 Asian Male Earnings Regression Coef.	Std. Err.	t	P>|t|	Dum Int	1995 Asian Male Earnings Regression Coef.	Std. Err.	t	P>|t|	Dum Int
yrsed	-0,052	0,036	-1,443	0,173		-0,001	0,074	-0,009	0,993	
yrsedsq	0,008	0,002	3,200	0,007		0,007	0,003	2,233	0,036	
exp	0,068	0,011	6,331	0,000		0,055	0,007	7,985	0,000	
expsq	-0,001	0,000	-4,401	0,001		-0,001	0,000	-4,245	0,000	
locdum1	0,728	0,231	3,149	0,008	1,071	-0,385	0,092	-4,178	0,000	-0,319
locdum2	0,000	Base Case			0,000	0,000	Base case			0,000
uniond1	-0,023	0,051	-0,457	0,655	-0,023	0,062	0,049	1,271	0,216	0,064
uniond2	0,000	Base Case			0,000	0,000	Base case			0,000
inddum1	0,000	No Obs			0,000	-0,184	0,255	-0,721	0,478	-0,168
inddum2	-0,538	0,173	-3,104	0,008	-0,416	-0,405	0,147	-2,767	0,011	-0,333
inddum3	0,000	Base Case			0,000	0,000	Base case			0,000
inddum4	0,379	0,140	2,719	0,018	0,461	0,267	0,126	2,119	0,045	0,307
inddum5	-0,203	0,117	-1,735	0,106	-0,184	-0,113	0,124	-0,912	0,371	-0,107
inddum6	-0,266	0,143	-1,857	0,086	-0,234	-0,134	0,063	-2,118	0,045	-0,125
inddum7	0,165	0,093	1,773	0,100	0,179	0,131	0,069	1,889	0,072	0,140
inddum8	0,179	0,113	1,576	0,139	0,195	-0,032	0,084	-0,386	0,703	-0,032
inddum9	0,046	0,177	0,259	0,800	0,047	-0,055	0,078	-0,704	0,488	-0,053
occdum1	-0,159	0,182	-0,877	0,397	-0,147	0,382	0,113	3,377	0,003	0,465
occdum2	-0,142	0,137	-1,035	0,320	-0,133	0,485	0,123	3,953	0,001	0,624
occdum3	-0,230	0,139	-1,655	0,122	-0,205	0,167	0,082	2,042	0,053	0,181
occdum4	-0,416	0,177	-2,348	0,035	-0,341	0,198	0,132	1,493	0,149	0,219
occdum5	0,000	No Obs			0,000	0,891	0,286	3,120	0,005	1,437
occdum6	-0,088	0,133	-0,660	0,521	-0,084	0,141	0,072	1,945	0,064	0,151
occdum7	-0,564	0,053	-10,561	0,000	-0,431	0,096	0,085	1,135	0,268	0,101
occdum8	0,000	Base Case			0,000	0,000	Base case			0,000
provdum1	-0,363	0,115	-3,148	0,008	-0,304	-0,310	0,136	-2,269	0,033	-0,266
provdum2	0,000	No Obs			0,000	-0,486	0,049	-9,976	0,000	-0,385
provdum3	0,000	No Obs			0,000	-0,122	0,169	-0,721	0,478	-0,115
provdum4	0,000	Base Case			0,000	0,000	Base case			0,000
provdum5	0,000	No Obs			0,000	0,000	No Obs			0,000
provdum6	-0,341	0,216	-1,573	0,140	-0,289	0,451	0,174	2,585	0,017	0,570
provdum7	0,000	No Obs			0,000	0,000	No Obs			0,000
provdum8	-0,275	0,197	-1,393	0,187	-0,240	0,243	0,120	2,018	0,055	0,275
provdum9	0,000	No Obs			0,000	0,207	0,088	2,338	0,028	0,229
_cons	6,803	0,181	37,648	0,000		6,039	0,454	13,312	0,000	

1993: F=21.34, R-Squared=0.5708, N=141

1995: F=35.92, R-Squared=0.4900, N=589

TABLE A4 Estimated Coefficients from the White Male Earning Functions

	1993 White Male Earnings Regression					1995 White Male Earnings Regression				
Variable	Coef.	Std. Err.	t	P>|t|	Dum Int	Coef.	Std. Err.	t	P>|t|	Dum Int
yrsed	-0,090	0,044	-2,030	0,050		0,143	0,096	1,494	0,148	
yrsedsq	0,007	0,002	3,784	0,001		-0,001	0,004	-0,365	0,718	
exp	0,063	0,012	5,259	0,000		0,071	0,005	13,920	0,000	
expsq	-0,001	0,000	-3,552	0,001		-0,001	0,000	-12,508	0,000	
locdum1	0,125	0,149	0,838	0,407	0,133	0,004	0,078	0,058	0,954	0,004
locdum2	0,000	Base Case			0,000	0,000	Base case			0,000
uniond1	0,028	0,065	0,428	0,671	0,028	0,112	0,041	2,751	0,011	0,119
uniond2	0,000	Base Case			0,000	0,000	Base case			0,000
inddum1	-0,625	0,250	-2,505	0,017	-0,465	-0,283	0,120	-2,350	0,027	-0,246
inddum2	0,117	0,117	0,998	0,325	0,124	0,111	0,080	1,385	0,178	0,118
inddum3	0,000	Base Case			0,000	0,000	Base case			0,000
inddum4	-0,083	0,100	-0,829	0,413	-0,079	0,034	0,065	0,523	0,605	0,035
inddum5	-0,067	0,114	-0,585	0,562	-0,065	-0,062	0,067	-0,918	0,367	-0,060
inddum6	-0,146	0,184	-0,792	0,434	-0,136	-0,197	0,040	-4,922	0,000	-0,179
inddum7	-0,135	0,083	-1,636	0,111	-0,126	0,045	0,080	0,566	0,576	0,046
inddum8	-0,011	0,129	-0,083	0,934	-0,011	0,048	0,042	1,136	0,267	0,049
inddum9	-0,113	0,126	-0,903	0,373	-0,107	-0,072	0,024	-2,977	0,006	-0,069
occdum1	2,468	0,202	12,222	0,000	10,802	0,601	0,143	4,194	0,000	0,824
occdum2	2,522	0,187	13,509	0,000	11,456	0,796	0,152	5,248	0,000	1,217
occdum3	2,125	0,198	10,718	0,000	7,374	0,305	0,137	2,217	0,036	0,356
occdum4	2,293	0,197	11,624	0,000	8,909	0,286	0,123	2,326	0,028	0,330
occdum5	2,966	0,299	9,924	0,000	18,415	0,513	0,236	2,171	0,040	0,670
occdum6	2,168	0,189	11,453	0,000	7,739	0,427	0,111	3,855	0,001	0,532
occdum7	2,295	0,213	10,754	0,000	8,921	0,254	0,117	2,172	0,040	0,289
occdum8	0,000	Base Case			0,000	0,000	Base case			0,000
provdum1	-0,028	0,120	-0,236	0,814	-0,028	-0,097	0,056	-1,723	0,097	-0,093
provdum2	0,000	No Obs			0,000	-0,175	0,072	-2,443	0,022	-0,161
provdum3	-0,327	0,316	-1,034	0,308	-0,279	-0,129	0,059	-2,174	0,039	-0,121
provdum4	0,000	Base Case			0,000	0,000	Base case			0,000
provdum5	0,000	No Obs			0,000	-0,122	0,065	-1,874	0,073	-0,114
provdum6	0,046	0,091	0,503	0,618	0,047	0,074	0,052	1,421	0,168	0,077
provdum7	-0,351	0,095	-3,703	0,001	-0,296	-0,087	0,100	-0,871	0,392	-0,083
provdum8	-0,235	0,129	-1,818	0,077	-0,210	0,058	0,068	0,846	0,406	0,060
provdum9	0,059	0,114	0,519	0,607	0,061	0,125	0,044	2,808	0,010	0,133
_cons	5,503	0,302	18,238	0,000		5,595	0,620	9,027	0,000	

F=139.67
R-Squared=0.4305
N=365

F=2793.54
R-Squared=0.4977
N=1770

FIGURE A1

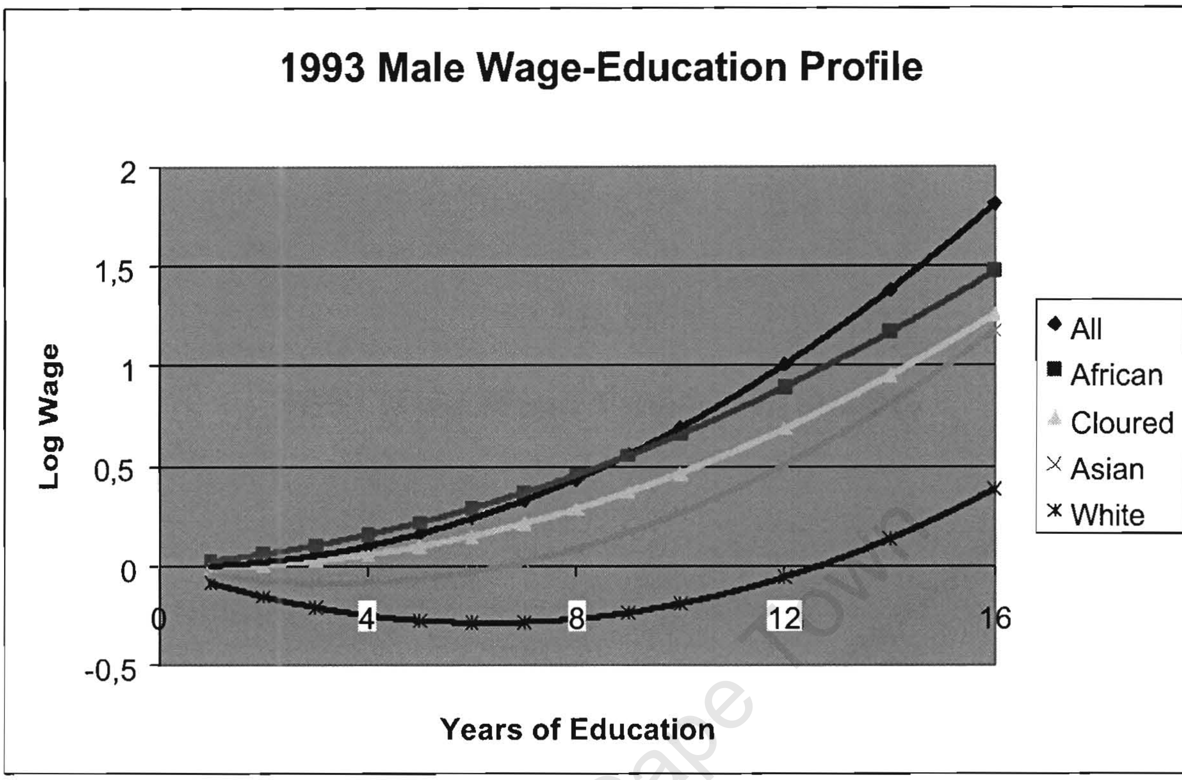


FIGURE A2

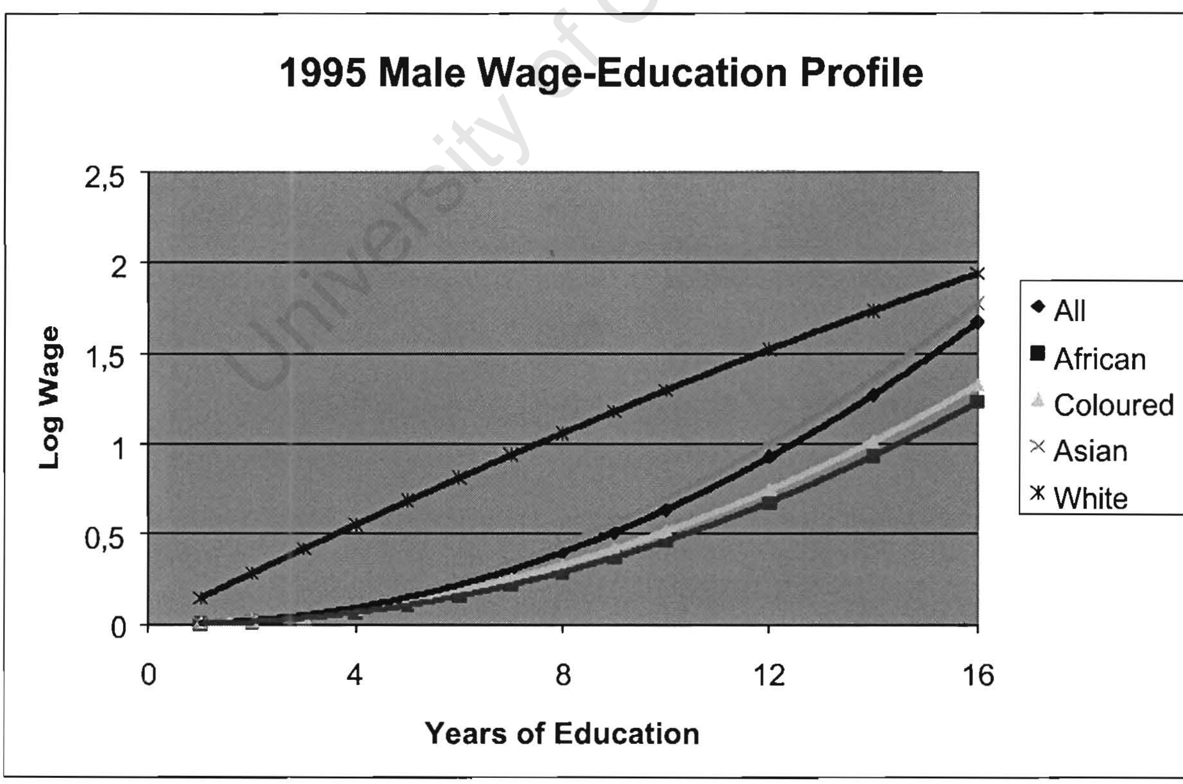


FIGURE A3

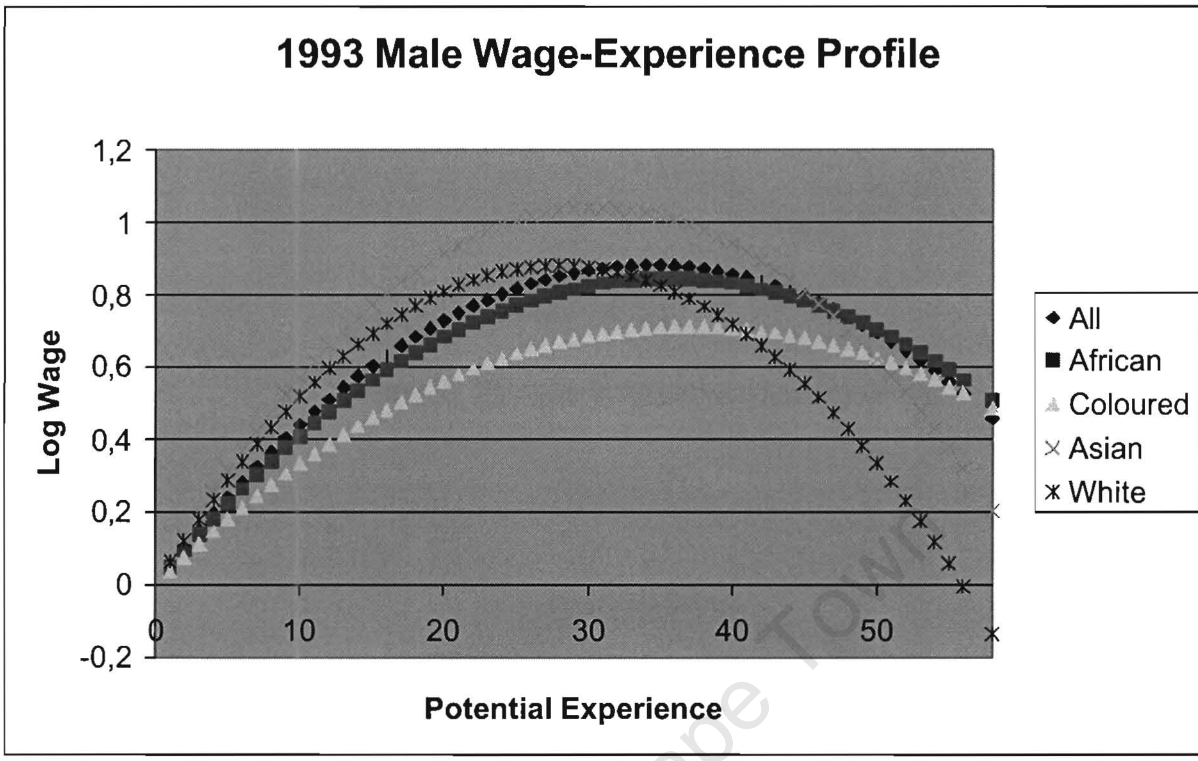


FIGURE A4

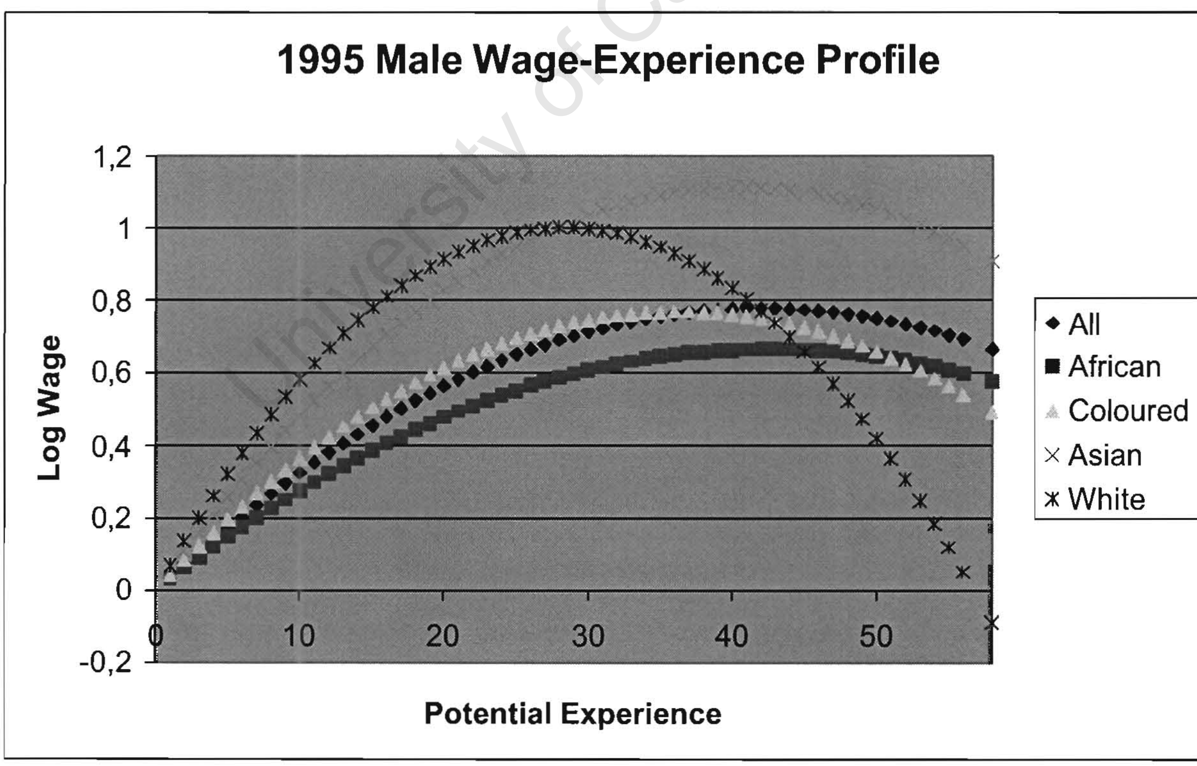


TABLE A5 Estimated Coefficients from the African Female Earning Functions

Variable	1993 African Female Earnings Regression					1995 African Female Earnings Regression				
	Coef.	Std. Err.	t	P>|t|	Dum Int	Coef.	Std. Err.	t	P>|t|	Dum Int
yrsed	0,035	0,022	1,597	0,112		0,020	0,009	2,290	0,029	
yrsedsq	0,003	0,002	1,765	0,079		0,004	0,001	7,034	0,000	
exp	0,033	0,007	5,082	0,000		0,023	0,004	5,641	0,000	
expsq	0,000	0,000	-3,483	0,001		0,000	0,000	-2,809	0,008	
locdum1	-0,155	0,087	-1,789	0,075	-0,143	-0,122	0,032	-3,801	0,001	-0,115
locdum2	0,000	Base Case			0,000	0,000	Base case			0,000
uniond1	0,397	0,047	8,350	0,000	0,487	0,203	0,023	8,921	0,000	0,225
uniond2	0,000	Base Case			0,000	0,000	Base case			0,000
inddum1	-0,394	0,109	-3,627	0,000	-0,326	-0,615	0,080	-7,694	0,000	-0,459
inddum2	-0,375	0,186	-2,018	0,045	-0,313	-0,064	0,205	-0,312	0,757	-0,062
inddum3	0,000	Base Case			0,000	0,000	Base case			0,000
inddum4	1,536	0,124	12,346	0,000	3,645	0,212	0,161	1,315	0,198	0,236
inddum5	0,064	0,206	0,310	0,757	0,066	0,097	0,191	0,506	0,616	0,101
inddum6	-0,074	0,093	-0,793	0,429	-0,071	-0,143	0,046	-3,145	0,004	-0,133
inddum7	0,225	0,105	2,138	0,034	0,253	0,113	0,102	1,107	0,277	0,119
inddum8	0,264	0,145	1,824	0,070	0,302	0,179	0,054	3,323	0,002	0,197
inddum9	-0,252	0,078	-3,218	0,002	-0,222	0,094	0,045	2,081	0,046	0,098
occdum1	1,452	0,122	11,896	0,000	3,273	0,644	0,032	20,434	0,000	0,905
occdum2	1,010	0,333	3,035	0,003	1,745	0,583	0,153	3,798	0,001	0,791
occdum3	0,617	0,108	5,729	0,000	0,853	0,330	0,036	9,295	0,000	0,391
occdum4	0,005	0,083	0,061	0,951	0,005	0,083	0,050	1,663	0,106	0,087
occdum5	-0,135	0,139	-0,969	0,334	-0,126	-0,179	0,246	-0,728	0,472	-0,164
occdum6	-0,111	0,466	-0,238	0,812	-0,105	-0,002	0,095	-0,017	0,987	-0,002
occdum7	0,281	0,083	3,397	0,001	0,325	0,166	0,048	3,475	0,001	0,180
occdum8	0,000	Base Case			0,000	0,000	Base case			0,000
provdum1	0,355	0,087	4,084	0,000	0,426	0,009	0,069	0,130	0,898	0,009
provdum2	0,546	0,086	6,372	0,000	0,726	-0,277	0,093	-2,991	0,005	-0,242
provdum3	-0,077	0,114	-0,671	0,503	-0,074	-0,119	0,042	-2,827	0,008	-0,112
provdum4	0,000	Base Case			0,000	0,000	Base case			0,000
provdum5	-0,163	0,259	-0,629	0,530	-0,151	-0,488	0,062	-7,840	0,000	-0,386
provdum6	-0,115	0,108	-1,066	0,288	-0,109	-0,024	0,057	-0,423	0,675	-0,024
provdum7	-0,053	0,073	-0,727	0,468	-0,052	0,002	0,047	0,044	0,965	0,002
provdum8	0,161	0,104	1,553	0,122	0,175	-0,054	0,054	-1,004	0,323	-0,052
provdum9	0,293	0,097	3,010	0,003	0,340	0,151	0,042	3,591	0,001	0,163
_cons	5,116	0,159	32,144	0,000		5,758	0,094	61,543	0,000	

1993: F=110.91, R-Squared=0.6799, N=957

1995: F=2535.08, R-Squared=0.6592, N=3272

TABLE A6 Estimated Coefficients from the Coloured Female Earning Functions

	1993 Coloured Female Earnings Regression					1995 Coloured Female Earnings Regression				
Variable	Coef.	Std. Err.	t	P>|t|	Dum Int	Coef.	Std. Err.	t	P>|t|	Dum Int
yrsed	-0,029	0,050	-0,571	0,572		0,009	0,022	0,414	0,683	
yrsedsq	0,006	0,004	1,699	0,099		0,004	0,002	2,390	0,026	
exp	0,062	0,013	4,867	0,000		0,030	0,006	5,281	0,000	
expsq	-0,001	0,000	-4,151	0,000		0,000	0,000	-3,358	0,003	
locdum1	-0,752	0,258	-2,910	0,006	-0,528	-0,256	0,071	-3,638	0,002	-0,226
locdum2	0,000	Base case			0,000	0,000	Base case			0,000
uniond1	0,317	0,071	4,464	0,000	0,372	0,214	0,043	4,972	0,000	0,238
uniond2	0,000	Base case			0,000	0,000	Base case			0,000
inddum1	-0,008	0,175	-0,047	0,963	-0,008	-0,498	0,072	-6,870	0,000	-0,392
inddum2	0,000	No Obs			0,000	0,554	0,093	5,971	0,000	0,739
inddum3	0,000	Base case			0,000	0,000	Base case			0,000
inddum4	0,000	No Obs			0,000	-0,023	0,205	-0,112	0,912	-0,023
inddum5	-1,244	1,165	-1,068	0,293	-0,712	0,112	0,085	1,328	0,199	0,119
inddum6	-0,040	0,085	-0,475	0,638	-0,039	-0,281	0,043	-6,544	0,000	-0,245
inddum7	-0,075	0,117	-0,644	0,524	-0,073	0,191	0,113	1,694	0,105	0,211
inddum8	0,671	0,115	5,855	0,000	0,956	0,155	0,072	2,137	0,045	0,167
inddum9	-0,262	0,099	-2,644	0,012	-0,230	-0,133	0,054	-2,433	0,024	-0,124
occdum1	1,238	0,227	5,455	0,000	2,448	0,781	0,118	6,623	0,000	1,184
occdum2	0,667	0,391	1,704	0,098	0,948	0,813	0,119	6,855	0,000	1,254
occdum3	0,615	0,143	4,295	0,000	0,850	0,328	0,041	7,975	0,000	0,388
occdum4	0,788	0,266	2,962	0,006	1,200	0,232	0,079	2,942	0,008	0,261
occdum5	0,999	0,239	4,183	0,000	1,717	0,696	0,117	5,955	0,000	1,005
occdum6	0,679	0,212	3,202	0,003	0,972	0,108	0,081	1,331	0,198	0,114
occdum7	0,406	0,125	3,244	0,003	0,500	0,133	0,048	2,784	0,011	0,143
occdum8	0,000	Base case			0,000	0,000	Base case			0,000
provdum1	0,105	0,092	1,138	0,263	0,110	-0,265	0,070	-3,811	0,001	-0,233
provdum2	-0,282	0,123	-2,285	0,029	-0,246	-0,586	0,083	-7,045	0,000	-0,443
provdum3	-0,019	0,149	-0,125	0,901	-0,018	-0,311	0,078	-3,991	0,001	-0,267
provdum4	0,000	Base case			0,000	0,000	Base case			0,000
provdum5	0,000	No Obs			0,000	-0,343	0,161	-2,127	0,045	-0,290
provdum6	-1,280	0,290	-4,409	0,000	-0,722	-0,210	0,104	-2,017	0,057	-0,189
provdum7	0,000	No Obs			0,000	-0,303	0,099	-3,069	0,006	-0,261
provdum8	0,000	No Obs			0,000	-0,245	0,084	-2,926	0,008	-0,217
provdum9	0,762	0,133	5,730	0,000	1,142	0,098	0,080	1,218	0,237	0,103
_cons	5,406	0,280	19,286	0,000		6,264	0,146	43,005	0,000	

1993	1995
F=792.67	F=165.15
R-Squared=0.6320	R-Squared=0.6476
N=259	N=1090

TABLE A7 Estimated Coefficients from the Asian Female Earning Functions

Variable	1993 Asian Female Earnings Regression Coef.	Std. Err.	t	P>|t|	Dum Int	1995 Asian Female Earnings Regression Coef.	Std. Err.	t	P>|t|	Dum Int
yrsed	-0,263	0,073	-3,625	0,006		0,014	0,037	0,387	0,702	
yrsedsq	0,018	0,004	4,345	0,002		0,005	0,002	2,798	0,010	
exp	0,061	0,009	6,606	0,000		0,044	0,009	4,916	0,000	
expsq	-0,001	0,000	-4,995	0,001		-0,001	0,000	-3,367	0,003	
locdum1	0,100	0,178	0,562	0,588	0,105	-0,207	0,040	-5,122	0,000	-0,187
locdum2	0,000 Base Case				0,000	0,000 Base case				0,000
uniond1	0,174	0,085	2,052	0,070	0,191	0,111	0,037	3,034	0,006	0,117
uniond2	0,000 Base Case				0,000	0,000 Base case				0,000
inddum1	0,000 No Obs				0,000	0,216	0,088	2,451	0,022	0,241
inddum2	0,000 No Obs				0,000	-0,385	0,134	-2,874	0,009	-0,320
inddum3	0,000 Base Case				0,000	0,000 Base case				0,000
inddum4	0,515	0,087	5,890	0,000	0,674	0,000 No Obs				0,000
inddum5	0,356	0,087	4,091	0,003	0,427	-0,365	0,201	-1,822	0,082	-0,306
inddum6	-0,067	0,121	-0,555	0,592	-0,065	-0,098	0,096	-1,018	0,319	-0,093
inddum7	-0,017	0,093	-0,178	0,862	-0,016	0,399	0,287	1,389	0,178	0,491
inddum8	0,401	0,183	2,189	0,056	0,493	0,078	0,157	0,493	0,627	0,081
inddum9	0,119	0,137	0,866	0,409	0,126	0,021	0,117	0,178	0,861	0,021
occdum1	1,648	0,355	4,646	0,001	4,198	0,357	0,144	2,489	0,020	0,429
occdum2	1,833	0,257	7,126	0,000	5,254	0,251	0,238	1,053	0,303	0,285
occdum3	1,260	0,296	4,260	0,002	2,525	0,089	0,091	0,987	0,334	0,094
occdum4	1,194	0,320	3,733	0,005	2,300	-0,011	0,232	-0,047	0,963	-0,011
occdum5	0,000 No Obs				0,000	0,000 No Obs				0,000
occdum6	1,101	0,351	3,136	0,012	2,006	-0,055	0,085	-0,646	0,525	-0,053
occdum7	1,037	0,326	3,178	0,011	1,821	-0,109	0,099	-1,092	0,286	-0,103
occdum8	0,000 Base Case				0,000	0,000 Base case				0,000
provdum1	0,000 No Obs				0,000	-0,119	0,125	-0,952	0,351	-0,112
provdum2	0,000 No Obs				0,000	-0,513	0,144	-3,569	0,002	-0,401
provdum3	0,000 No Obs				0,000	0,282	0,133	2,124	0,045	0,325
provdum4	0,000 Base Case				0,000	0,000 Base case				0,000
provdum5	0,000 No Obs				0,000	0,000 No Obs				0,000
provdum6	0,000 No Obs				0,000	0,176	0,313	0,564	0,579	0,193
provdum7	0,000 No Obs				0,000	0,000 No Obs				0,000
provdum8	0,000 No Obs				0,000	-0,086	0,122	-0,701	0,491	-0,082
provdum9	0,000 No Obs				0,000	0,237	0,059	4,046	0,001	0,267
_cons	5,830	0,153	38,210	0,000		6,056	0,192	31,594	0,000	

1993: F=4.13, R-Squared=0.6952, N=89

1995: F=441.45, R-Squared=0.6034, N=306

TABLE A8 Estimated Coefficients from the White Female Earning Functions

Variable	1993 White Female Earnings Regression Coef.	Std. Err.	t	P>|t|	Dum Int	1995 White Female Earnings Regression Coef.	Std. Err.	t	P>|t|	Dum Int
yrsed	-0,082	0,029	-2,882	0,007		0,079	0,081	0,980	0,336	
yrsedsq	0,006	0,002	2,924	0,006		0,001	0,003	0,441	0,663	
exp	0,034	0,010	3,425	0,002		0,043	0,003	13,493	0,000	
expsq	-0,001	0,000	-2,464	0,019		-0,001	0,000	-10,478	0,000	
locdum1	0,198	0,117	1,699	0,099	0,219	0,084	0,080	1,042	0,307	0,087
locdum2	0,000	Base Case			0,000	0,000	Base case			0,000
uniond1	-0,086	0,095	-0,911	0,369	-0,083	0,129	0,043	2,983	0,006	0,137
uniond2	0,000	Base Case			0,000	0,000	Base case			0,000
inddum1	-0,452	0,198	-2,282	0,029	-0,364	-0,118	0,118	-1,006	0,324	-0,112
inddum2	-0,196	0,183	-1,073	0,291	-0,178	0,067	0,153	0,438	0,665	0,069
inddum3	0,000	Base Case			0,000	0,000	Base case			0,000
inddum4	0,019	0,114	0,169	0,867	0,019	0,037	0,118	0,317	0,754	0,038
inddum5	0,103	0,201	0,512	0,612	0,109	-0,115	0,178	-0,649	0,522	-0,109
inddum6	-0,063	0,122	-0,517	0,608	-0,061	-0,255	0,055	-4,593	0,000	-0,225
inddum7	0,298	0,154	1,941	0,061	0,348	-0,232	0,125	-1,859	0,075	-0,207
inddum8	-0,021	0,122	-0,171	0,865	-0,021	-0,061	0,061	-0,999	0,327	-0,059
inddum9	-0,036	0,116	-0,308	0,760	-0,035	-0,150	0,046	-3,303	0,003	-0,140
occdum1	2,354	0,184	12,819	0,000	9,526	0,612	0,205	2,985	0,006	0,844
occdum2	2,221	0,183	12,104	0,000	8,214	0,711	0,187	3,811	0,001	1,036
occdum3	1,934	0,178	10,874	0,000	5,917	0,431	0,226	1,909	0,068	0,539
occdum4	1,949	0,210	9,292	0,000	6,022	0,340	0,229	1,483	0,151	0,405
occdum5	0,000	No Obs			0,000	0,000	No Obs			0,000
occdum6	1,989	0,390	5,106	0,000	6,308	0,225	0,152	1,479	0,152	0,253
occdum7	1,482	0,343	4,321	0,000	3,400	0,228	0,348	0,655	0,519	0,256
occdum8	0,000	Base Case			0,000	0,000	Base case			0,000
provdum1	0,025	0,063	0,399	0,693	0,026	0,020	0,057	0,348	0,731	0,020
provdum2	0,000	No Obs			0,000	-0,116	0,057	-2,019	0,054	-0,109
provdum3	-0,031	0,120	-0,262	0,795	-0,031	-0,115	0,055	-2,073	0,049	-0,108
provdum4	0,000	Base Case			0,000	0,000	Base case			0,000
provdum5	0,000	No Obs			0,000	-0,024	0,046	-0,518	0,609	-0,023
provdum6	0,027	0,044	0,608	0,548	0,027	-0,046	0,105	-0,441	0,663	-0,045
provdum7	-0,217	0,072	-2,998	0,005	-0,195	-0,025	0,063	-0,400	0,693	-0,025
provdum8	-0,403	0,104	-3,855	0,000	-0,332	-0,058	0,091	-0,639	0,529	-0,057
provdum9	0,269	0,034	7,816	0,000	0,309	0,195	0,064	3,057	0,005	0,215
_cons	5,368	0,230	23,375	0,000		5,760	0,561	10,273	0,000	

1993: F=24.80, R-Squared=0.5184, N=265

1995: F=117.06, R-Squared=0.3719, N=1154

FIGURE A5

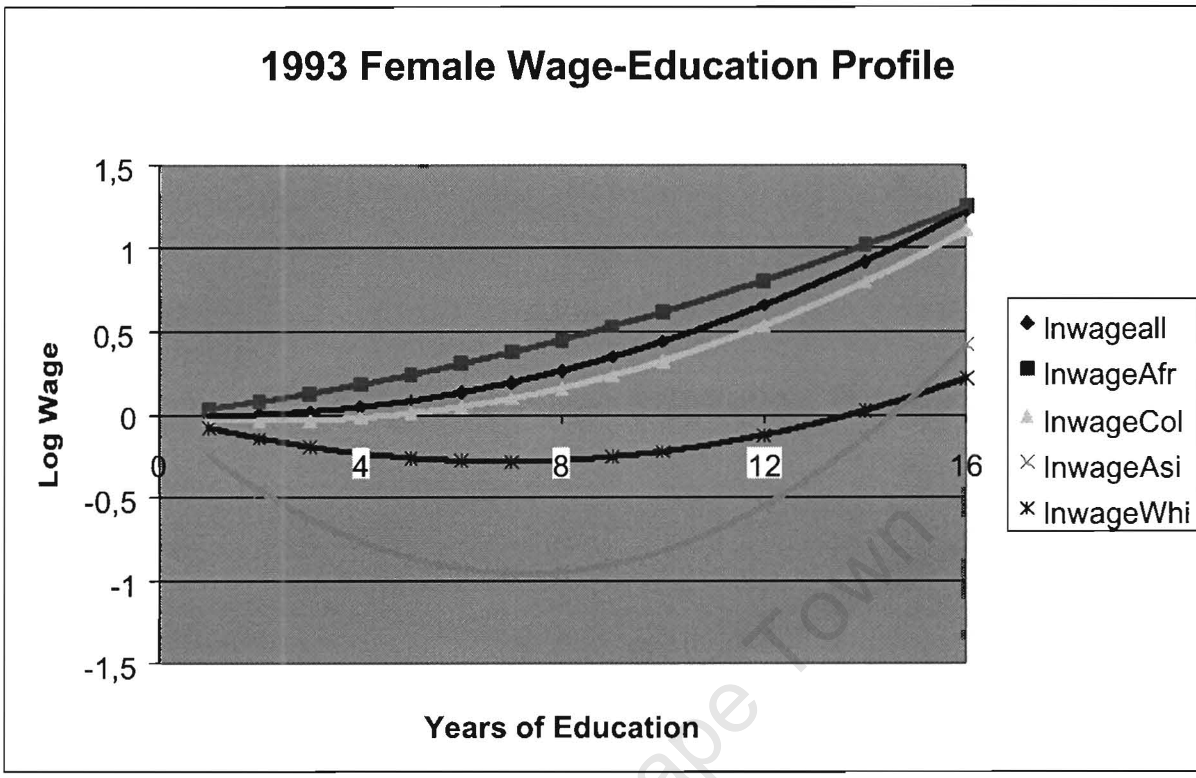


FIGURE A6

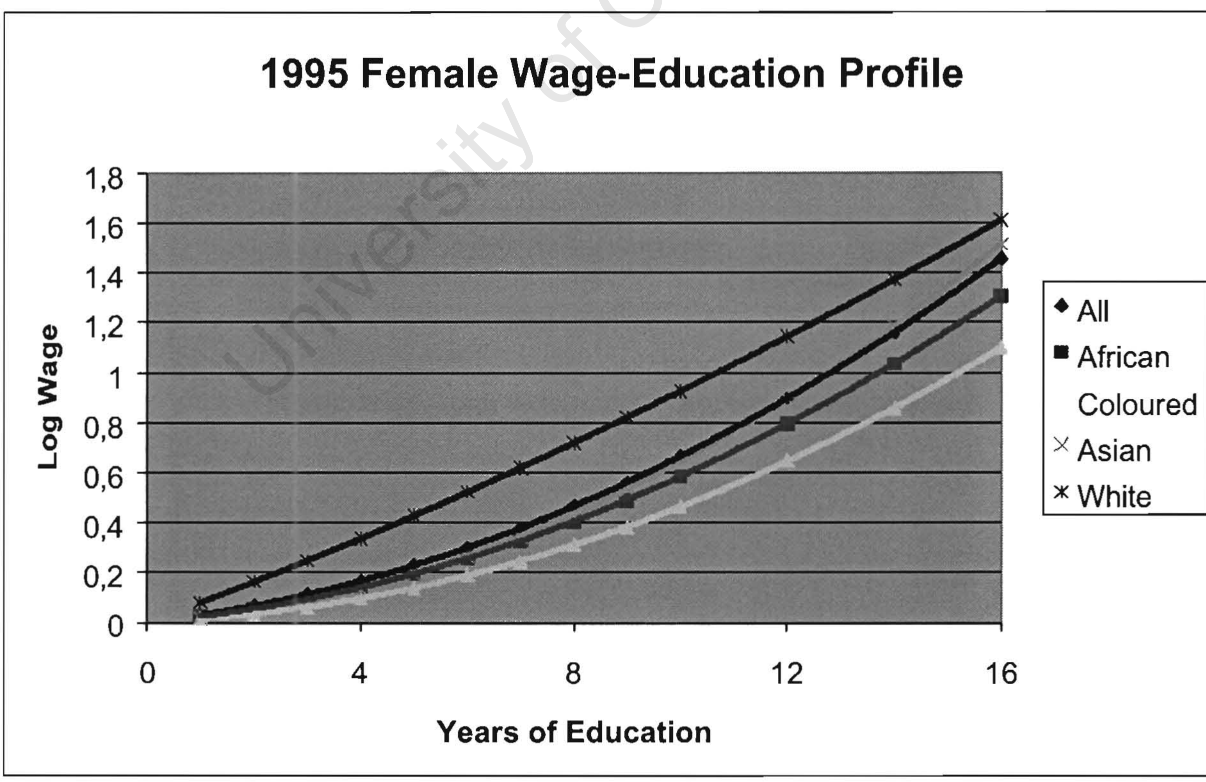


FIGURE A7

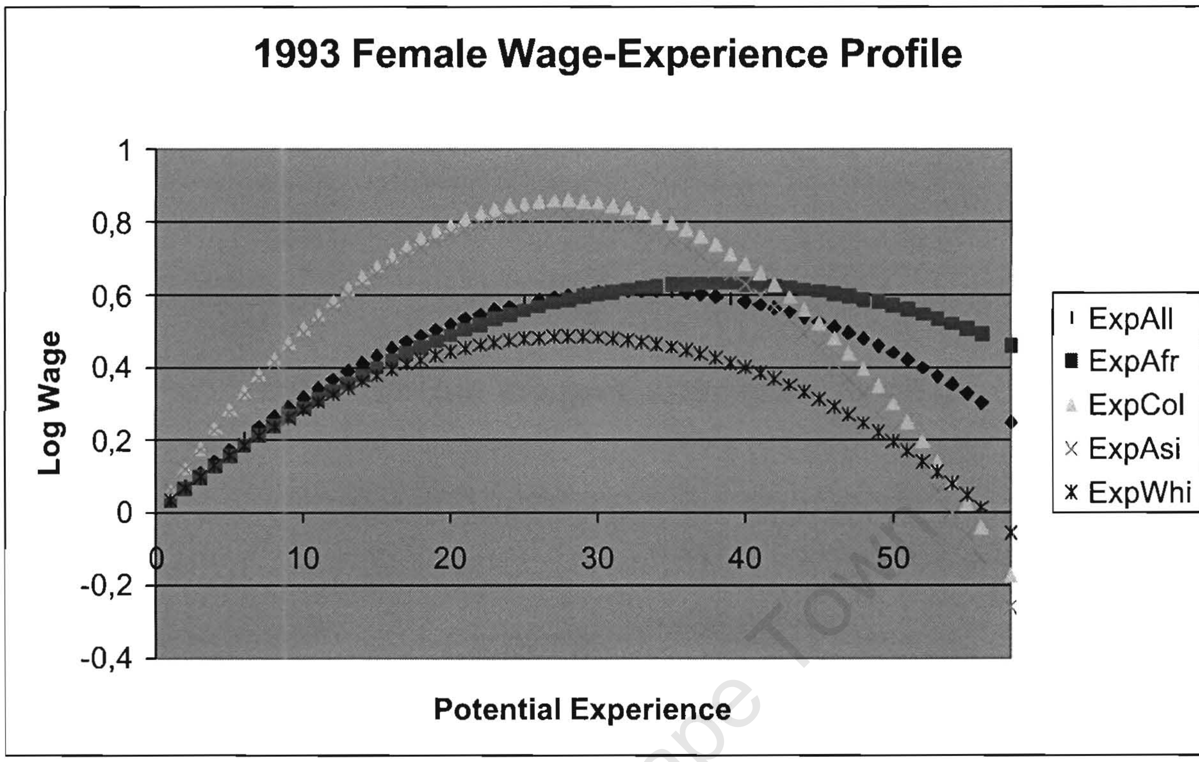


FIGURE A8