

An Analysis of the First Wage of Youth in Cape Town

Murray Leibbrandt & Cecil Mlatsheni

1. Introduction

In the context of high youth unemployment in South Africa, youth who have jobs have surmounted a prevalent obstacle. In addition, long duration of unemployment is a serious concern for youth. In a preliminary analysis of the youth of Cape Town we found that there is significant racial inequality in securing the first job and in the duration of unemployment prior to obtaining the job. Of interest in this paper is to ascertain the implications of the apparent Cape Town job acquisition dynamics on the first wage. This paper thus considers determinants of the level of wages of the first job and explains the racial and gender differences in wages of the first job that youth hold. The paper begins with a descriptive analysis followed by a multivariate analysis.

In the multivariate analysis, duration of unemployment is included as an explanatory variable. The aim of this is to determine whether or not the time taken to find the first job has an effect on the wages of that job. A positive relationship between time taken to secure the first job and wages of the first job would support a reservation wage hypothesis; those individuals who spend a longer period searching, do so by choice in order to find a higher paying job. A negative relationship between duration of unemployment and wages would on the other hand lend credence to the argument that unemployment is involuntary and that those individuals who take long to find employment are either less productive or unlucky, such that after a long search spell they accept whatever wages they are offered.

The multivariate analysis investigates the factors that lead to differences in the first wage and then proceeds to determining how much of the observed disparities in the first wages of youth are explained by these productive factors and how much remain unexplained. The latter analysis is carried out via decompositions by race. For the sake of completeness, a gender decomposition analysis is also carried out and the results thereof are included in Table A1.1 and Table A1.2 in Appendix 1. The investigation into the dynamics around the wages of the first job could also be considered a proxy for job worth. It is therefore an investigation of whether the patterns of inequality in job access between African and

Coloured youth carry over to remuneration once a job is secured. It is clear that White youth achieve the best results by far in terms of all aspects of the transition from schooling to work and they have been excluded from this analysis in order to focus the discussion on the more interesting comparison of two races that have large segments that are not obviously socioeconomically that different to each other; namely Africans and Coloureds.

Empirical research on the first wages and wage discrimination

There has not been an analysis of the wages of the first job in South Africa. However, useful insights can be gathered from such an analysis. It provides information on the situation at the point of transition from schooling to work, that is, before post-work-entry factors influence the situation of youth. It can shed light on whether or not there is labour market inequality at the point of work acquisition, as it is known that there is definitely inequality further down the employment track of individuals. Researchers in the United States have conducted research in this area over a number of decades and as such, this research can be informative for the purposes of the analysis carried out below.

A study that is most related to the analysis in this paper, in as far as focusing on the first wage, is that of Eckstein and Wolpin (1999). They conduct a study of the effect of racial discrimination on the first job wage offers, with a focus on White and Black youth. They use the 1979 youth cohort of the National Longitudinal Surveys of Labor Market Experience in the United States and control for schooling, gender, age and race. The study finds significant racial differences in wages of both high-school dropouts and high-school graduates. In terms of race decompositions carried out, Eckstein and Wolpin (1999) find that discrimination could account for the entire racial wage differential for both high-school dropouts and high-school graduates, concluding that the bound on the extent of discrimination is therefore not informative. One useful caution raised in this paper is that an individual's wage is determined by two unobservables, namely the skill bundle and its per-unit valuation. Therefore, measuring wage discrimination as the wage differential at a point in time, controlling for a small set of observable characteristics that are a priori related to skill bundles (e.g. schooling, work experience) will misstate the extent of discrimination if measured characteristics explain only part of the racial skills differentials. However, the paper acknowledges that although there have been strides in expanding the presumed

correlates of skills, it is unlikely that we will ever be able to directly measure skill bundles or collect significantly more-convincing proxies. Another note of caution arising from Eckstein and Wolpin (1999) is that observed wages often do not correspond with offered wages, a factor which makes it difficult to measure wage discrimination accurately. They argue that if Blacks face higher job search costs they will accept lower-wage jobs even if wage offers are not discriminatory and Blacks are equally productive.

Neal and Johnson (1996) conduct a study of the role of premarket factors in wage differences between Blacks and Whites using the National Longitudinal Survey of youth in the United States. However, their study is not of wages of youth at the point of transition from schooling to work but that of wages of individuals in their late twenties. They use premarket factors to explain wages at a latter part of the employment path of youth. The study controls for race, age, Armed Forces Qualifying Test (AFQT) scores, and schooling (in some specifications). The main findings of this study are that (1) there are large returns to measured skills, (2) there is a strong relationship between wages and test score measures of achievement or aptitude, (3) there is evidence of a Black-White wage gap that the authors attribute more to skill gaps between Blacks and Whites even though they do observe some evidence of labour market discrimination. A central point of discussion in Neal and Johnson (1996) is the use of test scores to capture individual skill.

Like Eckstein and Wolpin (1999), Neal and Johnson (1996) also highlight the difficulty in measuring worker skill and uses test scores as a measure to overcome this problem. They argue that some of the controls used in wage studies, such as occupation, postsecondary schooling, part-time work, marital status, geographic location, and actual labour market experience are subject to worker choice and could be contaminated by current labour market discrimination. Therefore controlling for them could misstate the wage effects of current discrimination. Blau and Kahn (2005) echo this view in as far as controlling for occupation is concerned. They argue that explanatory variables such as test scores and education may affect wages both directly, holding occupation and industry constant, and indirectly, through their effect on representation in higher-paying industries and occupations. They conclude then that coefficients from regressions which exclude industry and occupation variables thus shed light on the total effect of these variables. Returning to

Neal and Johnson (1999), the study criticises the use of years of schooling as a sole measure of worker skill as it is a measure of an input rather than an outcome. Years of schooling as a measure may systematically overstate the relative skills of Blacks. Neal and Johnson (1996) cite evidence from standardised tests that Black children exhibit lower levels of achievement than White children in the same grade do. Very importantly then the implications they draw from this are that studies that rely on schooling only as a measure of skill will likely overstate the effect of current labour market discrimination on wages and confuse barriers in acquiring human capital with barriers when entering the labour market.

Mason (1998) reports on studies which raise controversial views on the causes of the racial wage gap in the United States. There are proponents of the labour quality hypothesis, that racial discrimination in the labour market is not a major issue if one properly controls for interracial differences in labour quality. The argument is that discriminatory behaviour may affect premarket accumulation of skills, but once individuals enter the labour market they are paid based on their productive attributes (e.g. schooling, ability, experience, tenure on the job). Preferences for discrimination in primary institutions that produce skills (schools, families, and neighbourhoods), however, can lead to interracial differences in skill accumulation. Mason (1998) argues that premarket discrimination is not viewed by proponents of the labour quality hypothesis as a major source of interracial differences in skill accumulation. He goes on to cite studies that point to the alleged inferiority of African Americans in cognitive ability, family structure, and the market functionality of community culture as sources of premarket discrimination. The postulation goes on to say that the inferior labour quality that results is not captured by traditional explanatory variables and that there are skills which are not observed by statistical analysts but are easily observed by employers and which are in larger supply among Whites than African Americans. Therefore, premarket discrimination in the non-competitive areas mentioned above may lower African-American test scores. Regardless of the ultimate source of interracial differences in premarket skills, whether biology, discrimination, or culture, labour quality theorists claim that after adjusting for differences in test scores, interracial differences in wages become inconsequential. However, Mason (1998) cites other studies that find race to be important even after controlling for test scores which would suggest the presence of some discrimination.

The review above highlights some of the debates around first wages and interracial wage discrimination from the United States, a country which has grappled with these issues for decades. These debates influence to some extent the choice of variables in the analysis below.

Data and methodology

The Cape Area Panel Survey (CAPS) is used to estimate a wage equation for a sample that consists of African and Coloured youth. The wage variable is adjusted for inflation to set all wages at a 2002 real equivalent. A unique aspect of this analysis is that the wages used are not wages at the time of the survey but at the time of obtaining the first job.

When analysing wages, sample selection bias is generally an important consideration. On the supply side, however, it is under conditions where the characteristics of the unemployed are materially different to those of the employed that sample selection bias is most likely to arise. In the case of voluntary unemployment, it is foreseeable that the unemployed could possess some characteristics that are materially different to the employed and they would thus choose not to work at the level of wages on offer. In such a case, the effect of individual characteristics on wages would be distorted if this selection effect was not accounted for. In the case of predominantly involuntary unemployment, it follows that if jobs were available then unemployed individuals would take them up. In such an environment there would therefore be no strong justification for assuming that analysing only the sample of individuals for which wages are observed necessarily biases the results. In the South African context of Black youth in an environment of mass unemployment and socio-economic deprivation, it is fair to assume that the bulk of unemployment in this sample is involuntary. The relationship between duration of unemployment and wages of the first job observed in the descriptive analysis and the modelled relationship later sheds further light on the accuracy of the above postulation. The negative relationship between duration of unemployment and wages observed in the analysis below is more indicative of involuntary unemployment rather than voluntary unemployment.

Caution should be added though, that sample selection bias could emanate from the demand side as well if employers select those individuals who possess positive unobservable characteristics. Given these views, the presence of selection bias is investigated via a Heckman sample selection model at the beginning of the multivariate analysis. It should be noted, however, that the variable capturing age at the time of the first job (which has a statistically significant coefficient in the regression that follows) is a credible control for ability in the wage equation, since the first wage is unaffected by experience. This specification neatly handles the endogeneity problem associated with unobserved ability in the ordinary least squares regressions.

The Heckman correction for sample selection bias model:

$$y_i = x_i\beta + u \rightarrow \text{outcome equation}$$

$$z_i^* = w_i' \alpha + e_i;$$

$$z_i = 1 \text{ if } z_i^* > 0; 0 \text{ otherwise}$$

$$\Pr(z_i = 1) = \Phi(\alpha_i w_i) \rightarrow \text{selection equation}$$

In the selection equation, z_i^* is a dependent variable signifying whether an individual has ever worked or not, w_i' is a vector of covariates for unit i , α is a vector of coefficients, and e_i is a random disturbance term. Individuals who have ever had a job are selected. The fact that individuals may have entered the labour market under different economic conditions is controlled for by a variable capturing the unemployment rate in Cape Town at the time of labour market entry.

Similarly in the outcome equation, y_i is a dependent variable signifying wages, while x_i , β and u represent vectors of covariates, coefficient and the error term respectively.

Given that Heckman model is relatively sensitive to specification and the added complexity that would arise when trying to include duration of unemployment in this modelling, further

analysis is carried out using ordinary least squares regression. This modelling allows for the direct inclusion of duration of unemployment in the wage equation. It was also argued above that unobserved ability, which could be the main source of sample selection bias in this case, could be captured by the variable for age at the time of the first job, given that the first wage is unaffected by experience. Controlling for education and literacy and numeracy test scores, a positive relationship between age at which the first job is acquired and wages of the first job would possibly signify acquisition of unobserved productive. It is also reassuring to observe that the coefficients of most of the explanatory variables in the ordinary least squares model do not change much both in terms of magnitude and significance when compared with the outcome equation of the Heckman model. The wage equation is specified as follows:

$$\log w_i = \beta_0 + \beta_i x_i + \eta_i$$

X_i is a vector of the following individual characteristics used in both the Heckman model and the ordinary least squares model:

- race,
- gender,
- age at the time of first job,
- combined literacy and numeracy scores,
- education,
- time to first job

As is common in many wage studies, race, gender and age are expected to have a significant influence on wages. Education is an obvious characteristic to control for. However, education does not necessarily capture ability accurately (Neal & Johnson, 1996) especially for those individuals who do not have tertiary qualifications. Therefore, literacy and numeracy test scores are also included in the model in order to capture another dimension of individual ability. This is not to say that education is not useful as a predictor of wage differences but rather that it does not capture the full range of skills sets. Students leaving school with grade 12 qualifications for example, would have markedly different labour market related skills sets, depending on which schools they attended. The wages these

individuals would ordinarily be expected to command would also differ according to their productivity. Literacy and numeracy test scores contribute toward capturing these differences in individual ability and their inclusion results in a better specified model, especially given that the individuals in this survey were given the same tests. It is also true that type of industry and occupation are important influences on wages, however, these variables have been omitted from this analysis in order to obtain the full effect of education and test scores. Besides affecting wages directly, when controlling for industry and occupation, educational attainment and test scores also have an indirect affect through their effect on representation in industry and occupation (Blau & Kahn, 2005). Omitting industry and occupation variables therefore results in coefficients that shed light on the total effect of education and test scores (Blau & Kahn, 2005).

A descriptive analysis of wages of the first job

Level of education is expected to have an impact on the wages that first-time jobholders can obtain. When looking at the educational attainment of Africans and Coloureds at the time of securing the first job (Table 1) it is notable that the differences by race are not very large. The figures in Table 1 indicate column-wise percentages of educational attainment of youth. Amongst the combined sample of Coloured and African youth, 52% had achieved matric at the time of their first job. Amongst Africans, 54% had matric at the time of their first job while the corresponding figure for Coloureds is 51%. That Coloured youth gain exposure to the labour market relatively earlier than African youth is evidenced by the finding that 20% of Coloured youth have grade 1-9 by the time of their first job compared to only 15% of African youth.

Table 1. Average education level at the onset of the first job

	African	Coloured	Total
Grade 1-9	15.07	20.25	18.06
Grade 10-11	21.14	19.01	19.91
Grade 12	54.43	50.71	52.28
Diploma	6.44	7.82	7.23
Degree	2.92	2.22	2.51
Total	100	100	100

Source: CAPS Wave 1-4

Notes: Own calculations using survey weights

An analysis of age reinforces the notion that more Coloured youth obtain their first jobs at younger ages than African youth. Table 2 reflects column-wise percentages of age at the time of obtaining the first job. In Table 2, of those youth who have secured a first job, only 14% of African youth obtain their first job by age 19 while in contrast, 39% of Coloured youth secure their first job by this age. Stated differently, only 4% of those Coloured youth who have found a first job obtain their first job beyond age 24 while the corresponding figure for Africans is 14%. Overall, in the sample of 16 to 28 year olds, most youth (70%) who have found a first job do so between the ages of 19 and 23.

Table 2. Age at the time of obtaining the first job

Age	African	Coloured	Total
16	0.64	2.04	1.45
17	2.8	5.74	4.5
18	3.68	10.92	7.87
19	6.61	20.54	14.67
20	13.09	16.28	14.94
21	15.76	16.1	15.95
22	16.14	10.45	12.85
23	16.14	7.68	11.24
24	10.93	6.29	8.24
25	10.55	3.33	6.37
26	3.05	0.65	1.66
27	0.51	0	0.21
28	0.13	0	0.05
Total	100	100	100

Source: CAPS Wave 1-4

Notes: Own calculations using survey weights

An analysis of wages of the first job provides useful information in that these wages relate to reward prior to any labour market experience. It is reasonable to assume that any variation in wages if not caused by work experience should then be the result of differences in productive characteristics acquired prior to labour market entry, whether through schooling and/or community. There are indeed marked differences in wages of the first job by race, gender and education as can be seen in Table 3. The mean wage of the first job of

Coloured youth is 50% higher than that of African youth. Similarly, the median wage of Coloured youth is 43% higher than that of African youth.

Table 3. Mean and median monthly wages for the first job

	Mean wage - job 1	Median wage - job 1
African	1568	1400
Coloured	2344	2000
Female	1976	1500
Male	2178	2000
Grade 1-9	1634	1500
Grade 10-11	1713	1500
Grade 12	2271	2000
Diploma	3679	2800
Degree	4339	3800

Source: CAPS Wave 1-4

Notes: Own calculations using survey weights

An analysis of the gender wage gap reveals that the mean first job wage for males is 10% higher than that of females while the median wage is 33% higher. With respect to the relationship between educational achievement at the time of the first job and wages of the first job, mean starting wages increase with educational attainment. Achieving matric as opposed to just grade 10 or 11 is associated with a 33% higher mean wage and median wage.

Table 4. Distribution of monthly wages of the first job by race

Race	Quintiles of the wage of the first job					Total
	1	2	3	4	5	
African	31.56	29.23	20.22	11.48	7.51	100
Coloured	10.58	18.08	25.58	20.73	25.03	100
Total	19.95	23.06	23.18	16.6	17.21	100

Source: CAPS Wave 1-4

Notes: Own calculations using survey weights

As mentioned above, there are differences in first wage by race. In fact, the distribution of the first wage by race reflects a skewed pattern. Table 4 shows that amongst Coloured youth, 46% of them have first wages that are in the top two quintiles whereas only 19% of African youth locate in the top two quintiles. Furthermore, for Coloured youth the distribution of wages is fairly even in the across the top three quintiles whereas it is heavily skewed towards the lowest two quintiles for Africans. Around 60% of African youth locate in the lowest two quintiles of first wages.

The analysis now turns to the relationship between duration of unemployment and the wage of the first job. To begin, Table 5 indicates that most Coloured youth who find jobs do so within a year, 56% within 6 months. In contrast, only 39% of African youth who secure a first job do so within 6 months. The gender difference is less stark with 56% of males finding the first job within 6 months compared to 50% of females. In total, 73% of males who find a first job do so within a year compared to 72% of females.

Table 5: Months unemployed leading up to the first job by race and gender

Months to first job	African	Coloured	Female	Male	Total Gender
6	38.86	56.16	46.99	55.17	51.07
12	19.33	19.13	21.6	16.77	19.19
18	14.96	13.07	12.26	15	13.63
24	10.49	7.03	9.95	6.14	8.05
30	5.55	2.85	4.71	2.58	3.64
36	6.29	1.4	2.81	2.87	2.84
42	4.52	0.35	1.68	1.47	1.58
Total	100	100	100	100	100

Table 6 reflects a comparison of first wage, split into quintiles, with duration of unemployment prior to the first job. The last column indicates that most individuals (47%) who get a job do so within 6 months. Another 20% get a job within 7 months to a year. Looking at the distribution of unemployment duration within wage quintiles, a similar proportion of individuals (around 46%) get a job within 6 months and around 65% get a job within a year across the first four quintiles. For individuals in the highest quintiles, a slightly higher proportion of them (58%) get a job within 6 months. Looking at these statistics, one may be tempted into thinking that the unemployment situation is not that bad given the proportion of individuals who secure jobs quickly. However, this analysis is of those individuals who secure jobs, which are 43% of the total for non-studying Coloureds and Africans.

Table 6: Distribution of wages according to months unemployed leading up to the first job

Months to first job	Quintiles of wages of the first job					Total
	1	2	3	4	5	
6	45.83	46.67	44.57	46.38	58.33	47.38
12	20.83	20.51	22.28	23.19	11.46	20.36
18	10.12	13.85	16.85	14.49	11.46	13.57
24	8.93	7.18	9.24	7.25	10.42	8.45
30	5.36	3.59	3.8	2.9	2.08	3.71
36	5.36	5.13	2.17	2.9	5.21	4.1
42	3.57	3.08	1.09	2.9	1.04	2.43
Total	100	100	100	100	100	100

Source: CAPS Wave 1-4

Notes: Own calculations using survey weights

Table 7 presents a different angle on these statistics. It indicates that of those individuals who secure jobs within 6 months or 7 to 12 months, the wages they command are more or less evenly distributed for the lowest three quintiles. There does not seem to be a clear pattern of wage advantage for individuals who secure jobs quickly. For individuals who get jobs within 6 months, 32% of them are in the top two quintiles, compared to 27% for 7 to 12 months, 29% for 13 to 18 months, and 30% for 19 to 24 months. Taken together with the results of the table above, the picture that emerges is that even though most high paying jobs are found relatively quickly, the jobs that are found quickly are distributed across the wage spectrum. Duration of unemployment may not necessarily be an important determinant of the first wage, instead productive qualities as reflected by literacy and numeracy test scores, for example, may be more important. The multivariate analysis that follows sheds light on these relationships.

Table 7: Quintiles of wages by months unemployed leading up to the first job

Months to first job	Quintiles of wages of the first job					Total
	1	2	3	4	5	
6	20.81	24.59	22.16	17.3	15.14	100
12	22.01	25.16	25.79	20.13	6.92	100
18	16.04	25.47	29.25	18.87	10.38	100
24	22.73	21.21	25.76	15.15	15.15	100
30	31.03	24.14	24.14	13.79	6.9	100
36	28.13	31.25	12.5	12.5	15.63	100
42	31.58	31.58	10.53	21.05	5.26	100
Total	21.51	24.97	23.56	17.67	12.29	100

Source: CAPS Wave 1-4

Notes: Own calculations using survey weights

Looking at the distribution of test scores achieved by race, not reflected here, reveals that Coloured youth achieve higher results than African youth. Sixty percent of Coloured youth score positive standardised combined literacy and numeracy scores compared to only 30% of African youth. An interesting question that follows from this finding is whether better able individuals, as reflected by higher test scores, have an advantage in terms of finding employment quicker. Table 8 indicates that 62% of the highest scoring individuals find their first job within 6 months, compared to 50% for individuals in the 4th quintile, and around 48% in each of the lowest three quintiles. These statistics reflect a clear advantage for the top test score quintile but not much of a difference across the other quintiles. Overall, there seems to be a weak relationship between test scores and duration of unemployment.

Table 8: Distribution of literacy and numeracy test scores across months of the first job

Time to first job	Quintiles of combined literacy and numeracy scores					Total
	1	2	3	4	5	
6	48.74	46.88	47.81	49.64	62.3	49.86
12	16.58	21.48	22.37	22.1	15.57	20.26
18	13.07	14.84	12.28	12.68	10.66	12.95
24	9.05	8.2	6.58	7.25	7.38	7.68
30	5.03	4.3	4.39	4.35	1.64	4.16
36	4.02	3.13	4.39	2.54	1.64	3.24
42	3.52	1.17	2.19	1.45	0.82	1.85
Total	100	100	100	100	100	100

Source: CAPS Wave 1-4

Notes: Own calculations using survey weights

Table 9: Distribution of literacy and numeracy test scores across months of the first job

Time to first job	Quintiles of combined literacy and numeracy scores					Total
	1	2	3	4	5	
6	18	22.26	20.22	25.42	14.1	100
12	15.07	25.11	23.29	27.85	8.68	100
18	18.57	27.14	20	25	9.29	100
24	21.69	25.3	18.07	24.1	10.84	100
30	22.22	24.44	22.22	26.67	4.44	100
36	22.86	22.86	28.57	20	5.71	100
42	35	15	25	20	5	100
Total	18.41	23.68	21.09	25.53	11.29	100

Source: CAPS Wave 1-4

Notes: Own calculations using survey weights

However, there does seem to be clear relationship between test score achievement and the first wage as reflected in Table 10. The expectation is that individuals who score higher obtain better paying jobs. Indeed, looking at the 5th quintile of test score achievement this notion is confirmed. Of the individuals in the highest quintile of test score achievement, 44% have the highest paying first jobs. In contrast, of the individuals who achieve the lowest test scores, only 7% work in the highest paying first jobs.

Table 10: Relationship between literacy and numeracy test scores and wages of the first job

Quintiles of test scores	Quintiles of wages of the first job					Total
	1	2	3	4	5	
1	28.97	29.91	23.36	10.9	6.85	100
2	24.37	28.39	25.13	12.81	9.3	100
3	18.15	23.51	23.21	20.24	14.88	100
4	14.25	17.5	23	21	24.25	100
5	9.3	9.3	18.02	19.19	44.19	100
Total	19.91	22.99	23.11	16.66	17.33	100

Source: CAPS Wave 1-4

Notes: Own calculations using survey weights

The descriptive analysis above has revealed a number of useful points. Around 50% of youth have matric at the time of the first job. A relatively higher proportion of Coloured youth start work at a young age. In addition, mean and median wages are higher for Coloured youth and for males. Amongst African youth, the distribution of the first wage is skewed

towards lower wage quintiles whereas for Coloured youth it is more balanced across wage quintiles. Furthermore, across wage quintiles most jobs are obtained within 6 months although a slightly higher proportion of the highest paying jobs are obtained within 6 months. That said, there is no clear wage advantage to securing jobs quickly. Therefore, duration of unemployment may not necessarily be an important determinant of the first wage, instead productive qualities as reflected by literacy and numeracy test scores, for example, may be more important. With regard to test scores, a higher proportion of the highest scoring individuals find jobs within 6 months. There is also a clear positive relationship between test scores and wages of the first job. Notwithstanding the value of the above descriptive analysis, the analysis proceeds to test the relationships between these variables more formally in a multivariate context.

Empirical modelling

It was discussed above that the problem of sample selection bias should be considered when modelling wages. However, it was also argued that this bias may very well not rise in the labour market in question, more so when dealing with the first job. South Africa is plagued by very high youth unemployment as well as high economic deprivation and as such, involuntary unemployment is likely to be the dominant feature amongst non-White youth especially. South African data sets also reflect significant numbers of discouraged workers. The results of the selection and outcome regressions are presented in Table 11.

Table 11. Heckman regression of first wage

Variables	Outcome		Selection	
	Coefficients	Marginal effects	Coefficients	Marginal effects
		0.376**		0.122**
Coloured	0.309*** (0.053)	*	0.343*** (0.101)	*
Male	0.117** (0.046)	0.202**	0.443*** (0.0987)	0.155** *
Age at first job	0.031** (0.012)	0.031**		
Lit-num test scores	0.006 (0.030)	0.037	0.162** (0.0639)	0.058**
Incomplete secondary	-0.016 (0.050)	-0.016	-0.628*** (0.106)	- 0.217** *
Tertiary	0.263** (0.120)	0.263**	0.423 (0.305)	0.135
Unemployment rate			0.296*** (0.0350)	0.106** *
Lambda	-0.375*** (-0.057)			
Constant	6.817*** (0.271)		-6.004*** (0.771)	
Observations	539		836	
/athrho	-0.876*** (-0.162)			
/Insigma	-0.631*** (-0.045)			
Rho	-0.705 (-0.081)			
Sigma	0.532 (-0.024)			

LR test of independence of equations (rho = 0): $\chi^2(1) = 13.71$ Prob > $\chi^2 = 0.0002$

Standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

The choice of explanatory variables was discussed earlier. The unemployment rate in Cape Town at the time of labour market entry is the exclusion restriction in the selection equation. This variable controls for the fact that the youth under consideration will have entered the labour market in various years with the economy and the labour market at varying stages of activity.

The result of the likelihood ratio test suggests that, the assumption that the errors of the wage equation are uncorrelated with those of the selection equation, can be rejected. It seems then that it is appropriate to correct for sample selection bias.

The reported marginal effects suggest that Coloured youth have wages that are 38% higher than African youth, holding all else constant. In addition, males have a 20% wage advantage over females. Each additional year in age results in a 3% wage advantage. Having a tertiary qualification as opposed to matric only, raises wages by 26%. Furthermore, the literacy and numeracy test scores affect entry into wage employment but not wages. Similarly, obtaining matric gives one an advantage over non-matriculants in as far as employment is concerned but not in wages.

Attention now turns to assessing the relationship between duration of unemployment and wages as well as decomposition of the racial wage gap and gender wage gap. For this more involved analysis, use of ordinary least squares regression will yield results that are more intelligible than under the Heckman model. Use of ordinary least squares for this analysis has been argued above. Specifically, it was argued that unobserved ability, which could be the main source of sample selection bias in this case, could be captured by the variable for age at the time of the first job, given that the first wage is unaffected by work experience. On the supply side, it was argued that voluntary unemployment would probably not be a significant feature for youth entering the labour market under conditions of mass unemployment and relative economic deprivation.

The role that duration of unemployment plays in the first wage of youth is of special interest. The analysis here begins with an investigation of whether youth who are at a disadvantage in terms of unemployment duration, experience a further disadvantage in the

form of lower wages. In regressing duration of unemployment on first wage, one also gains a sense of the likelihood of voluntary versus involuntary unemployment as discussed earlier, the findings of which have a bearing on the motivation given earlier of the possible absence of a significant sample selection issue for this labour market, on the labour supply side at least.

The analysis then moves on to report on the determinants of wages amongst Coloureds, Africans, males and females, using ordinary least squares regression. Certain variables are successively introduced in three specifications in order to highlight some interesting features. The full specification is then run by race in preparation for the decomposition analysis that follows.

A key interest in this analysis is the effect of duration of unemployment on the wages of the first job. The results reported in Table 12 indicate a negative relationship between wage of the first job and time taken to secure that job. The omitted comparison variable is time to first job of zero to 6 months. In comparison to duration of unemployment of zero to 6 months only 7 to 12 months is significant. The results suggest that individuals who obtain a job at some stage between 7 to 12 months of unemployment have wages that are 19% lower than those who obtain their first job between zero to 6 months of unemployment, significant at the 5% level. The negative relationship between first wages and duration of unemployment is interesting, however this will be discussed further when interpreting the full specification of the model.

Table 12: Regression of first wage and duration of unemployment

Time to job 7-12 months	-0.185** [0.075]
Time to job 13-18 months	-0.021 [0.069]
Time to job 19-24 months	-0.12 [0.104]
Time to job 25 months plus	-0.147* [0.088]
Constant	7.644*** [0.038]
Observations	857
R-squared	0.016

Robust standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: Own calculations using survey weights

Looking at the first set of results in Table 13, regression one excludes education in order to compare the change in the effect of test scores that results when education is introduced into regression two. Similarly, the variables representing duration of unemployment are once more introduced in the last regression. Starting with race however, it is clear from all three specifications that race plays an important part in determining the first wage. Coloured youth earn first wages that are 33% higher than African youth.

Gender is also significant, being male is associated with wages that are 13% higher. As discussed previously, controlling for age at the time of the first job rather than age at the time of the survey yields more accurate results of the effect of age on the first wage. Each additional year in age at the time of the first job results in first wages that are 4% higher. Having controlled for school-level skills, the significant effect of age is rather interesting. It may be capturing undocumented or unobservable abilities that youth acquire post school.

The combined literacy and numeracy scores also have a significant effect on the first wage of youth, however, the magnitude of the coefficient of this variable is reduced when controlling for education and duration as well. In the first specification, each standard deviation increase in the test scores results in a 15% increase in wages. This effect drops to 9% when controlling for education and drops further in magnitude to 6% and significance

level to 10% when introducing unemployment duration as well. It is interesting that the effect of the literacy and numeracy test scores is significant even after controlling for education. This result implies that the test scores reflect productive characteristics that are not captured by level of education attainment. Besides innate ability, quality of education is also a possible contributing factor.

Table 13: Determinants of the first wage using the full specification

	I	II	III
Coloured	0.382*** [0.030]	0.379*** [0.033]	0.333*** [0.055]
Male	0.145*** [0.029]	0.169*** [0.033]	0.134*** [0.049]
Age at first job	0.043*** [0.007]	0.031*** [0.007]	0.035** [0.014]
Lit-Num test scores	0.152*** [0.019]	0.088*** [0.023]	0.055* [0.033]
Incomplete Secondary		-0.175*** [0.035]	-0.091* [0.051]
Tertiary		0.244*** [0.086]	0.408*** [0.129]
Time to job 7-12 months			-0.132** [0.060]
Time to job 13-18 months			-0.154* [0.080]
Time to job 19-24 months			-0.136 [0.096]
Time to job 25 months plus			-0.008 [0.085]
Constant	6.376*** [0.149]	6.689*** [0.169]	6.644*** [0.311]
Observations	1608	1271	556
R-squared	0.206	0.233	0.184

Robust standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: Own calculations using survey weights

With respect to education, specification II reflects that individuals with incomplete secondary education earn wages that are 18% lower than individuals with matric only. On the other hand, individuals who have above matric educational qualifications earn wages that are 24% higher than individuals with matric only. When unemployment duration is

introduced in specification III, the effect of education changes somewhat. The wages that individuals who do not attain matric can expect to obtain are 8% lower than those with matric and this result is significant at the 10% level of significance only. On the other hand, individuals with qualifications that exceed matric have predicted earnings that are 41% higher than matriculants.

The duration of unemployment leading up to the first job is significant at duration of up to 18 months. Individuals who were unemployed between 7 months and 12 months before obtaining their first jobs are predicted to receive wages that are 13% lower than individuals who secure a first job within 6 months. This result is significant at the 5% level of significance. Individuals who are unemployed for a period of between 13 and 18 months before obtaining the first job have wages that are 15% lower than individuals who are unemployed for 6 months or less, significant at the 10% level of significance only. The duration variable is not significant at unemployment duration exceeding 18 months. The results of the duration variable suggest that reservation wages do not play a significant role with respect to the first job. If individuals chose to wait for higher wage offers then longer unemployment duration would be associated with higher wages once the job is obtained. This result lends support to the notion that unemployment is involuntary with regard to the first job at least. Of course, behaviour related to the first job may be very different to that of subsequent jobs. Young school leavers wanting to get a foothold in the labour market are likely to be less choosy than those individuals who are already seeking their 2nd or 3rd jobs.

The full specification regression analysed above is now run by race for the purposes of the decomposition analysis that is to follow. Analysing race separately, reporting is on the results of both the model which excludes duration and the one that includes it. Looking at the specification which excludes the unemployment duration variable first, Table 14 reflects that gender is a significant determinant of the first wage. Amongst Coloured youth, males can expect to have first wages which are 14% than females, whereas amongst African youth the male advantage is larger at 25%. The effect of each additional year in age at which the first job is obtained is 3% amongst Coloured youth and 4% amongst African youth, consistent with the full specification reported on earlier.

The effect of education within the two races is also similar although individuals with more than matric qualifications have a greater wage advantage amongst Coloured youth than African youth. Specifically Coloured youth who have greater than matric qualifications have wages that are estimated to be 26% higher than Coloured youth with matric, the corresponding figure for Africans is 20%.

Furthermore, the coefficients of the test scores are similar across the specifications. Within race, each standard deviation increase in test scores results in a 9% increase in wages amongst Coloured youth compared to a 7% increase amongst African youth, these are similar to the 10% effect found in the full specification.

Table 14: Regression of the first wage by race excluding duration of unemployment

	Coloured	African	Pooled
Male	0.140*** [0.042]	0.247*** [0.046]	0.158*** [0.040]
Age at first job	0.028*** [0.009]	0.039*** [0.012]	0.028*** [0.008]
Incomplete Secondary	-0.168*** [0.045]	-0.198*** [0.050]	-0.101** [0.044]
Tertiary	0.257** [0.107]	0.200** [0.095]	0.372*** [0.080]
Lit-Num test scores	0.098*** [0.031]	0.070** [0.029]	0.143*** [0.025]
Constant	7.144*** [0.199]	6.496*** [0.274]	7.008*** [0.183]
Observations	750	521	1378
R-squared	0.15	0.156	0.168

Robust standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: Own calculations using survey weights

Turning to the model specification that controls for duration of unemployment, Table 15 shows that amongst Coloured youth the effect of gender is statistically insignificant. Many of the other variables also become insignificant under this specification. However, Coloured youth with greater than matric level of schooling are estimated to enjoy wages that are 48% higher than individuals with matric only. Furthermore, amongst Coloured youth, longer duration of unemployment prior to the first job is estimated to decrease the wage of the first job. Securing a job within 12 months or 18 months compared to within 6 months is estimated to decrease wages, significant at the 10 %level. This finding implies that individuals who spend a longer time searching for the first job probably do not do so by choice as they end up with lower wages.

Table 15: Regression of the first wage by race including duration of unemployment

	Coloured	African	Pooled
Male	0.084 [0.064]	0.244*** [0.066]	0.123** [0.060]
Age at first job	0.029 [0.019]	0.042* [0.022]	0.031* [0.016]
Incomplete Secondary	-0.085 [0.068]	-0.11 [0.070]	0.070 [0.075]
Tertiary	0.476*** [0.153]	0.155 [0.182]	0.462*** [0.122]
Lit-Num test scores	0.051 [0.046]	0.075* [0.044]	0.122*** [0.038]
Time to job1-12mth	-0.148* [0.078]	-0.085 [0.087]	-0.191*** [0.070]
Time to job1-18mth	-0.194* [0.109]	-0.07 [0.104]	-0.157** [0.078]
Time to job1-24mth	-0.19 [0.133]	0.009 [0.112]	-0.153 [0.111]
Time to job1-25mth plus	0.208** [0.096]	-0.086 [0.116]	-0.145 [0.114]
Constant	7.120*** [0.390]	6.433*** [0.484]	6.908*** [0.331]
Observations	323	233	609
R-squared	0.129	0.13	0.128

Robust standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: Own calculations using survey weights

The discussion above highlights the importance of the selected explanatory variables in explaining differences in wages of youth in their first jobs. The results have been reported for specifications of the race regressions that exclude the duration variable and specifications that include it. The coefficients of the duration variable suggest that the longer the duration of unemployment the lower the wage of the first job secured. This effect is significant up to 18 months duration. This result suggests that reservation wages is not the driving force behind the negative relationship between duration of unemployment and wages of the first job. Most likely, it is the less productive individuals that take longer to find a job. Productivity in the model above is captured by the schooling variable and the literacy and numeracy scores. If the duration variable is capturing the same thing, to an extent this would explain the loss of significance of some of the variables when it is included. This analysis will now be taken further to determine how much of the racial gap in

wages is due to the observable characteristics discussed above and how much is unexplained by these characteristics.

Decomposition analysis

The analysis is carried out along the lines of Oaxaca's method of analysing group wage differences (Oaxaca, 1973) based on linear regressions. It consists in decomposing the average wage gap into some "characteristics" and "returns" components in the following way:

$$\bar{W}_1 - \bar{W}_2 = \beta(\bar{X}_1 - \bar{X}_2) + (\beta_1 - \beta)\bar{X}_1 + (\beta - \beta_2)\bar{X}_2$$

Where W is the average wage, X is the vector of individual productive characteristics introduced in the respective equations and β_i is the associated vector of coefficients. β is the non-discriminatory set of coefficients. The main difficulty here is to determine β the structure that would prevail in the absence of discrimination.

If $\beta = \Omega\beta_1 + (I - \Omega)\beta_2$ where Ω is a weighting matrix and, I is the identity matrix, any assumption regarding β reduces to an assumption about Ω . The results are presented for the choice of three assumptions regarding Ω : $\Omega = 1$, $\Omega = 0$ (Oaxaca's index) and $\Omega = (X_1'X_1)^{-1}X_1'X_2$ (Oaxaca and Ransom's index).

In addition to reporting on the results of three assumptions of omega (the weighting matrix) this section also reports on decompositions with the duration variable excluded from the regression and with the duration variable included. In looking at the race decomposition first (Table 16) the estimated racial wage gap in the model excluding duration is 0.41. With the assumption of a weighted omega, the results of the decomposition indicate that only 16% of the racial wage gap is due to differences in productive characteristics. Furthermore, Coloured advantage makes up 19% of the gap while 65% of the wage gap is due to African disadvantage. Therefore, the majority of the wage gap is due to labour market disadvantage faced by Africans.

Under the assumption that omega equals 1, that is, assuming that the non-discriminatory wage structure is that faced by Coloureds, the results indicate that 9% of the wage gap is due to differences in characteristics while 91% is due to some other unexplained factors. Similarly, when taking African wages as the non-discriminatory wage structure, omega equal to zero, 2% of the difference in wages is explained by characteristics while 98% remains unexplained. The clear message given by these statistics is that very little of the gap in first wages between these two races is explained by the observed characteristics of these individuals.

Table 16: Race decomposition excluding duration variable

Results	Coefficient	Percentage
Omega = 1		
Characteristic	0.037811	9.31%
Coefficient	0.368232	90.69%
Omega = 0		
Characteristic	0.007231	1.78%
Coefficient	0.398811	98.22%
Omega = weighted		
Productive	0.064782	15.95%
Advantage	0.079079	19.48%
Disadvantage	0.262181	64.57%
Raw	0.406042	100%

Source: CAPS Wave 1-4

Notes: Own calculations using survey weights

In moving from the specification that excludes unemployment duration to the one that includes the duration variable, the racial wage gap drops to 0.33 as can be seen in Table 17. Under the assumption of weighted omega, 19% of the racial wage gap is as a result of productive characteristics, 24% is a result of Coloured advantage while 57% is due to African disadvantage. The decomposition results of the two specifications are similar in that both report African youth disadvantage making up more than half of the estimated racial wage gap. Furthermore, with omega assumed to be one and or omega assumed to be zero the results indicate racial wage gap is accounted for entirely by unobservable factors.

Table 17: Race decomposition with duration of unemployment included

Results	Coefficient	Percentage
Omega = 1		
Characteristic	-0.00402	-1.21%
Coefficient	0.33734	101.21%
Omega = 0		
Characteristic	-0.00983	-2.95%
Coefficient	0.343149	102.95%
Omega = weighted		
Productive	0.061972	18.59%
Advantage	0.080944	24.28%
Disadvantage	0.1904	57.12%
Raw	0.333316	100%

Source: CAPS Wave 1-4

Notes: Own calculations using survey weights

A gender decomposition was also conducted and the results, reported in Table A1.1 and Table A1.2 in Appendix 1, indicate that the total gender wage gap in the specification that excludes unemployment duration is .144 while in the specification that includes duration the gender gap is .116. The results indicate that with all three assumptions of omega, worker characteristics do not explain the wage differences between males and females. Under the weighted omega assumption, 54% of the gender wage gap is explained by male advantage in both the specification with duration and the one without. Female disadvantage accounts for 55% and 61% of the gender wage gap under the model with duration and without duration respectively, again assuming weighted omega.

The analysis above reflects significant racial and gender differences in the level of the first wage of Coloured and African youth in Cape Town. Observing that race and gender inequality is present right at the outset of the employment history of youth is instructive of itself. These differences are evident even after controlling for educational attainment, literacy and numeracy test scores, gender and age. The race decomposition gives an estimate of this racial inequality in the wage of the first job. This strong result probably

stems from historical influence, such as past job reservation policies, in the Cape Town labour market. The stronger gender effect is harder to explain even though finding male advantage is common across many studies.

The results of the outcome equation in the Heckman regression in as far as race and gender are concerned are very similar to those of the ordinary least squares regression above in that they have a statistically significant effect of a large magnitude. Considering the other explanatory variables as well, the conclusion reached is that the results of the regression that is corrected for sample selection bias are not very different from those of the OLS regression above. There is therefore confidence in the assumptions made above and the OLS analysis carried out above as being reasonably accurate.

Conclusion

This study has looked at the factors that affect the level of the first wage of African and Coloured youth in Cape Town. This is a unique study in that it provides a sense of the relative positions of the youth of these two races at the outset of their labour market participation. The extent of initial disadvantage experienced by African youth may be a precursor to persistent labour market inequity down the line. A Heckman model controlling for sample selection bias was presented as well as a number of specifications of ordinary least squares regressions. The two models yielded very similar results.

One particular object of interest in this study was the effect the duration of unemployment prior to the first job would have on the wage of the first job. The analysis in this paper suggests that youth who take a longer time to find employment also tend to work for a lower wage. This effect is statistically significant for employment duration of up to 18 months.

Furthermore, a decomposition analysis was carried out in order to ascertain how much of the racial gap in first wages is explained by individual characteristics and how much is unexplained. A gender analysis was also included in Appendix 1. The period of the survey (2002 – 2006) is fortuitous as it coincides with the most admirable economic growth rates in South Africa in recent times. The analysis is therefore carried out for a period where the

economy and labour market could be thought of as functioning relatively well as opposed to the crisis affected periods just prior to and after the time of the panel survey.

The study has revealed statistically significant findings in racial differences in the wages youth earn in their first jobs. Even after controlling for educational attainment and literacy and numeracy test scores, Coloured youth start out with higher wages than African youth. Similarly, young males start out with higher wages than young females. These results are robust as they hold up under OLS estimation and under correction for possible selection bias. When decomposing the estimated race and gender differences it was found that very little of these were explained by individual characteristics and that most of the race and gender differences were unexplained.

A number of factors could possibly explain these findings. Around half of the population of Cape Town consists of Coloured individuals while the African population makes up less than a quarter. Network effects are likely then to be more synergistic among Coloured youth and this could explain their relative success in finding better paying jobs at the outset. In addition to this, the legacy of past job reservation policies may feed into these network effects. If the model above was near perfectly specified, one could look to discrimination as a possible cause of the observed findings. What is most striking is that even after controlling for individual aptitude and ability via the literacy and numeracy tests, the race advantage remains. Quality of schooling is a factor that should be considered. It is not necessarily an obvious consideration because there are very poorly run and resourced schools in both African and Coloured communities. However, anecdotal evidence of African learners seeking places at Coloured schools indicates that generally quality of schooling may be better than in African schools on average.

References

- Blau, F. & Kahn, L. (2005). Do cognitive test scores explain higher US wage inequality? *The Review of Economics and Statistics*. 87(1): 184-193.
- Eckstein, Z. & Wolpin, I. (1999). Estimating the effect of racial discrimination on first Job wage offers. *The Review of Economics and Statistics*. 81(3): 384-392.
- Heckman, J. (1979). "Sample selection bias as a specification error". *Econometrica*, 47 (1): 153-61.
- Mason, P. (1998). Race, cognitive ability, and wage inequality. *Challenge*. 41(3): 63-84.
- Neal, D. & Johnson, R. (1996). The role of premarket factors in black-white wage differences. *Journal of Political Economy*. 104(5): 869-895.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor market. *International Economic Review*. 14(3): 693-709.
- Oaxaca, R. & Ransom, M. (1994). On discrimination and the decomposition of wage differentials. *Journal of Econometrics*. 61(2): 5-21.

Appendix 1:

Table A1.1. Gender decomposition without duration of unemployment

Variables	Coefficient Estimates	Percentage
Omega = 1		
Characteristics	-0.01869	-12.99%
Coefficient	0.162638	112.99%
Omega = 0		
Characteristics	-0.03294	-22.88%
Coefficient	0.176882	122.88%
Omega = weighted		
Productive	-0.02257	-15.68%
Advantage	0.078338	54.42%
Disadvantage	0.08818	61.26%
Raw	0.143944	100%

Source: CAPS 2002-2006

Notes: Own calculations using survey weights

Omega = 1 for Male

Table A1.2. Gender decomposition with duration of unemployment

Variables	Coefficient Estimates	Percentage
Omega = 1		
Characteristics	-0.00776	-6.67%
Coefficient	0.124047	106.67%
Omega = 0		
Characteristics	-0.02778	-23.89%
Coefficient	0.144068	123.89%
Omega = weighted		
Productive	-0.01063	-9.14%
Advantage	0.062805	54.01%
Disadvantage	0.064113	55.13%
Raw	0.11629	100%

Source: CAPS 2002-2006

Notes: Own calculations using survey weights

Omega =1 for Male

Table A1.3. Determinants of the first wage using the smallest subset of observations

	I	II	III
Coloured	0.368*** (0.058)	0.347*** (0.055)	0.333*** (0.055)
Male	0.132*** (0.048)	0.140*** (0.047)	0.134*** (0.049)
Age at 1st job	0.052*** (0.015)	0.033** (0.014)	0.035** (0.014)
Lit Num test scores	0.089*** (0.034)	0.054 (0.034)	0.055* (0.033)
Incomplete Secondary		-0.086* (0.051)	-0.091* (0.051)
Tertiary		0.414*** (0.131)	0.408*** (0.129)
Time to job 7-12 months			-0.132** (0.060)
Time to job 13-18 months			-0.154* (0.080)
Time to job 19-24 months			-0.136 (0.096)
Time to job 25 months plus			-0.008 (0.085)
Constant	6.196*** (0.321)	6.626*** (0.319)	6.644*** (0.311)
Observations	556	556	556
R-squared	0.137	0.168	0.184

Robust standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Source: CAPS 2002-2006

Notes: Own calculations using survey weights

