Using the tenure profile to estimate labour market signalling and employer learning for South African men

1. Introduction

Under the human capital model, the schooling phase of the life-cycle is devoted to the acquisition of skills in preparation for entry into the labour market. These skills are then adapted, enhanced, and new ones acquired during the post-school (on-the-job training) phase of the life-cycle (Becker, 1962). The increased stock and efficiency of skills enables the worker to be more productive and in turn command higher wages. The transition from school to wage employment is, however, not automatic. Labour market frictions and imperfections act as barriers to labour market entry, especially for young labour market entrants.

One important source of labour market imperfection that leads to barriers to labour market entry is the imperfect information that characterises the employer-employee relationship. Employers are often imperfectly informed about the skills set and expected productivity of workers. This uncertainty about worker productivity and its labour market implications have been well researched by economists (see for example Arrow, 1972 and 1973; Phelps, 1972; and Aigner and Cain, 1977 for early contributions). This literature has for example illustrated that if employers' ability to predict workers' productivity differs between two groups and employers are risk averse then the group with greater uncertainty around their expected productivity will earn lower average wages, *ceteris paribus* (Aigner and Cain, 1977).

This paper investigates the importance of imperfect information between employers and employees about the productivity of workers in accounting for differences in average wages between black and white, young and old, and educated and uneducated men in South Africa. The paper postulates that there is greater *ex ante* uncertainty about the productivity of black, young and matriculants men relative to their respective counterparts. And that such uncertainty is largely as a result of these groups having less informative productivity signals. The current South African literature on earnings differences focuses on differences in human capital investments and returns on those investments (*references*), and labour market discrimination (*references*). This paper hopes to contribute and complement this literature by theoretically and empirically investigating the applicability of theories of imperfect information for South African earnings differentials.

Although the South African literature on earnings differentials has largely focused on differences on human capital investments and labour market discrimination, there are a few studies have provided suggestive evidence of imperfect information having a potential role in explaining earnings (and employment) differentials (*references*). Levinsohn (2007), for example, claims that South Africa is characterised by high degrees of uncertainty around worker productivity, and that such uncertainty coupled with high firing costs may have directly contributed to the high unemployment in the country. This is echoed by the study of Schoer and Rankin (2011) that argues that uncertainty around worker productivity has led to inefficiencies in job search and matching in South Africa.

The empirical strategy for this paper exploits group variation in the average wage gain due to the accumulation of the first year of tenure as a 'measure' of employer learning. The estimated employer learning is then used to infer about the productivity signalling of black, young and less educated workers. The empirical strategy and results are presented in section four and five, respectively. A brief review of the literature on statistical discrimination and South African earnings differentials literature is summarized in section two. Section three develops a theoretical model that will help to guide the empirical analysis. Concluding remarks are contained in section six.

2. Literature review

This paper investigates the role played by imperfect information in the assessment of worker productivity as a potential explanation for group differences in average wages in South Africa. This research focus is closely aligned to and motivated by the literature on statistical discrimination and employer learning. This section of the paper will therefore provide a brief review of this literature. The section will then conclude with a review of related South African literature.

Statistical discrimination and employer learning: theory

Aigner and Cain (1977), building on earlier work by Phelps (1972) and others, developed a statistical model of labour market discrimination. This model has served as a key point of departure for later theoretical and empirical models statistical discrimination. In the Aigner and Cain model, worker productivity (ability) is unobservable. Employers observe and base their employment decision on a noisy signal of worker productivity. Employers are also able to observe the group (e.g. black or white, male or female) that a worker belongs to. A key assumption is that the productivity signal that employers receive at the hiring stage is less informative for one group (usually blacks – minority groups more generally) relative to the other group. Productivity is for all workers however assumed to be drawn from the same distribution for both groups.

Aigner and Cain (1977) showed that a model with only the above features does not generate differences in average wages between groups that it purports to explain. The authors therefore extended the model by assuming that employers are risk averse. With this additional assumption, average wages are now a function of not only the expected productivity conditional on the productivity signal but also a negative function of the conditional variance of productivity conditional on the productivity signal. Additionally, the extended model now generates group differences in average wages with the group with the nosier signal (higher productivity variance) earning lower relative wages.

The standard statistical discrimination model has over the years seen further extensions that incorporate insights from other labour market models. Lundberg and Startz (1983) introduce a human capital investment option and the investment is assumed to be unobservable and to be productivity enhancing. With the maintained assumption that minority groups (e.g. blacks) have less informative productivity signal, minority groups have less of an incentive to invest in human capital. This is because of the nosier productivity signal for these groups and the assumption that the investments are unobservable. With minorities investing less in human capital and having a less informative productivity signal, they earn lower average wages relative to whites.

Oettinger (1996) extends the model further by introducing a dynamic structure to the model that allows for uncertainty around the productivity of workers to be resolved through employers' observations of the workers' output. This extension improves on the static nature of the previous models and introduces new (and empirically testable) predictions about the wage gap between black and white workers. The empirical contributions of the model by Oettinger (1996) are considered in the next subsection together with other key empirical studies of statistical discrimination.

Statistical discrimination and employer learning: empirics

Oettinger (1996) and Lewis and Terrell (2001) develop models that suggests that statistical discrimination along based on differences in the assessment of worker productivity can empirically inferred from differing returns to tenure and experience. Specifically, these authors' models predict that if black workers' productivity is much harder to predict then black workers should have greater returns to tenure and smaller returns to experience relative to white workers. This prediction is supported by the authors' empirical analysis. Estimation of the structural parameters of the authors' models also lends support to the statistical discrimination model.

In an influential study, Altonji and Pierret (2001) show that employers statistically discriminate against young workers on the basis of easy of education which is assumed to be a correlate of productivity that is easy to observe by employers. These studies and the abovementioned theoretical models confirm the relevance and role played by statistical discrimination in understanding the earnings gap between black and whites.

South African literature

South Africa is a developing country with a labour market that is characterised by high degrees of uncertainty about worker productivity (Levinsohn, 2007). The uncertainty is driven mainly by the low and variable quality of pre-tertiary schooling that has weakened the potency of schooling qualification as a signal of productivity. In their reading of the literature and analysis of labour market data from a South African village, Duff and Fryer (2005:7) are of the view that "there is ample evidence both of considerable quality variation, and of a poor range of signals, associated with South African human capital." Schooling qualification seems to fulfil the signalling role for completed secondary and further qualifications (Duff and Fryer, 2005; Schoer and Rankin, 2011; and Van der Berg, 2014). This is a huge concern since a significant proportion of job-seekers have attained less than 12 years of schooling (i.e. completed secondary). Schooling as a productivity signal for these individuals is therefore unavailable and this is evidenced by the convex employment and wage returns to education (Van der Berg, 2014). Moreover, with the youth making up the bulk of the unemployed, these job-seekers cannot use previous work experience to signal their productivity since many of them are still searching for their first job.

3. Developing a theory model[1]

This section develops a statistical discrimination model that incorporates learning by employers. The model developed here maintains the key features of standard statistical discrimination models and explores Aigner and Cain's (1977) employer risk aversion assumption for generating between-groups earnings differentials in developing country characterised by high unemployment.

Suppose individual worker productivity, y, is determined as:

$$y = \alpha + \theta s + u\sigma_u \tag{1}$$

where $s \in \{0,1\}$, and $u \sim n(0,1)$. We for now take s to represent schooling with $s = 0 \rightarrow$ matric and $s = 1 \rightarrow$ tertiary qualification. We focus on two groups, black ('B') and white ('W') males, and further assume that black workers have a greater variance around their productivity (i.e. $\sigma_B > \sigma_W$). The latter assumption is a key feature of statistical discrimination models and captures the notion that the minority group (blacks) have a less informative productivity signal.

Equation (1) suggests that productivity is causally determined by schooling and this causal relationship is captured by θ. Productivity is also a function of a bunch of other factors that are assumed to be unobservable and thus captured by the model error term, u. The two groups are assumed to be equally productive but with greater variance of productivity for group B.

Productivity as determined by equations (1) is unobservable by employers. Employers instead observe, in every period, a noisy signal of productivity, $\hat{y}_t$

$$\hat{y}_t = y + e_t\sigma_e$$

where $e \sim n(0,1)$.

Firms can observe s but not u, and know the values of $(\alpha, \theta, \sigma_0, \sigma_1)$. At period t, the firm can observe all the signals $(\hat{y}_0, \dots, \hat{y}_t)$ and forms expectation about the productivity $E(y|s, \hat{y}_0, \dots, \hat{y}_t)$ and variance $Var(y|s, \hat{y}_0, \dots, \hat{y}_t)$ of workers. The firm is risk-averse and attaches value of

[1] Still under construction – this is a first attempt.

$$\pi(y|s, \hat{y}_0, \dots, \hat{y}_t) = E(y|s, \hat{y}_0, \dots, \hat{y}_t) - \delta Var(y|s, \hat{y}_0, \dots, \hat{y}_t)$$

to a worker with observables $(s, \hat{y}_0, \dots, \hat{y}_t)$.

This means that in period 0 the worker will earn

$$w_0 = \pi(y|s, \hat{y}_0) = E(y|s, \hat{y}_0) - \delta Var(y|s, \hat{y}_0) \qquad (2)$$

This first term on the right-hand side solves to

$$E(y|s, \hat{y}_0) = \alpha + \theta s + \left(\frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}\right)(\hat{y}_0 - \alpha - \theta s)$$

So the expected value of productivity depends on schooling $(\alpha + \theta s)$ in a causal way as determined by equation (1). The expected productivity also depends on the additional information/signal $(\hat{y}_0 - \alpha - \theta s)$ that employers receive via interviews/worker evaluations weighted by the variances of worker heterogeneity and productivity signals.

The last term of equation (2) solves to $Var(y|s, \hat{y}_0) = \left(\frac{\sigma_e^2}{\sigma_u^2+\sigma_e^2}\right)^2 \sigma_u^2 + \left(\frac{\sigma_u^2}{\sigma_u^2+\sigma_e^2}\right)^2 \sigma_e^2$, and thus the period 0 is given as follows:

$$w_0 = \alpha + \theta s + \left(\frac{\sigma_u^2}{\sigma_u^2+\sigma_e^2}\right)(\hat{y}_0 - \alpha - \theta s) - \delta\left[\left(\frac{\sigma_e^2}{\sigma_u^2+\sigma_e^2}\right)^2 \sigma_u^2 + \left(\frac{\sigma_u^2}{\sigma_u^2+\sigma_e^2}\right)^2 \sigma_e^2\right] \qquad (3)$$

In the asymptotic case equation (3) reduces to

$$plim(w_t) = \alpha + \theta s + u\sigma_u \qquad (4)$$

The variance term tends to zero, since with each successive period that an employee is interviewed/evaluated the variance around the firm's prediction of the productivity of the employee should get smaller and smaller and tend to zero when t tends to infinity.

DRAFT

4. Empirical strategy

The proposed empirical strategy for this paper relies on group variation in the wage gain due to the accumulation of the first year of tenure as a 'measure' of employer learning. The estimated employer learning is then used to infer about the productivity signalling of black, young and less educated workers. Assuming that employers set wages equal to expected productivity conditional on a worker productivity signal, initial average wages for a disadvantaged worker will be lower relative to her 'true' productivity. This is because of the greater difficulty of predicting this worker's productivity. The lower initial wage relative to 'true' productivity can also be interpreted as the cost or compensation that a risk-averse employer will have to receive to incentivise the hiring of worker whose productivity is more uncertain.

This suggests a larger gap between initial wages and 'true' productivity for disadvantaged workers. If employers continue to equate wages to expected productivity in each period then as the uncertainty around the worker's productivity is revealed by observing the worker's output, then actual wages should converge true productivity. Therefore there should be more rapid wage growth for these workers since employers were more uncertain about the productivity of these workers. Since employers learning occurs early in an employment spell (Lange, 2007), then noisy initial productivity signal should imply a steeper wage initial tenure profile.

To operationalise the above I estimate a (pooled) OLS wage regression controlling for variables that proxy for human capital, demographic and individual characteristics. I then add a dummy variable ('*oneyear'*) equal to one if tenure is larger or equal to one, and zero otherwise[2]. Adding this dummy variable in our wage regression ensures that the wage gain due to the accumulation of the first year of tenure is not restricted by the quadratic specification of tenure (Altonji and Shakotko, 1987). 'o*neyear'* is then interacted with our characteristics of interest (race, age, and education). These interactions together with *'oneyear'* are the variables of interest.

The analysis in the next section makes use of the Labour Force Surveys (LFSs) conducted by Statistics South Africa (Stats SA). The LFSs are nationally representative cross-sectional household surveys that are designed to monitor developments in the South African labour

[2] *oneyear* is set equal to 0.5 if tenure is equal to 0.5 since our data is collected bi-annually which means that tenure in our data increases in increments of 0.5.

market. The surveys were conducted twice yearly – March and September – from September 2000 to September 2007 when they were replaced by the Quarterly Labour Force Surveys. The LFSs were designed as a rotating panel of dwelling units with 20% of these units dropped in subsequent waves and replaced with new dwelling units (Stats SA, 2006). The rotations were designed in such a way that a total sample of 30 000 households was maintained in each wave. The estimation sample is restricted to black and white 15 to 64 year old males working in formal and private sector jobs. Workers in subsistence agriculture and those reporting to be self-employed were also excluded from the analysis.

DRAFT

5. Results

This section of the paper discusses the main results of our empirical analysis. The first set of results show the estimated (log) wage returns to tenure and (potential) experience. These results uncover an interesting comparative pattern of wage growth between black and white male workers. Various theories or explanations can account for these results. We however in the remainder of our empirical analysis provide evidence for a particular explanation that is based on statistical discrimination and employer learning.

Returns to tenure and experience

Table 1 presents results for the wage returns to tenure and experience for black (Panel A) and white (Panel B) male workers. The results only show the coefficients for our variables of interest and the constant. In addition to the variables of interest, all regressions control for education, province, firm size, wave fixed effects, household head, rural/urban, union status, occupation, and industry. In addition to ordinary least square, the returns to tenure and experience are also estimated using panel data methods (fixed-effects and random-effects) and Altonji and Shakotko's (1987) instrumental variables estimation procedure implemented by way of two-stage least squares and control function approach. These additional estimators address some of the endogeneity bias associated with the OLS estimation of the wage returns to tenure and experience (see Mincer and Jovanovic, 1981; Altonji and Shakotko, 1987; Abraham and Farber, 1987; and Topel, 1991).

The overwhelming picture painted by the results in Table 1 is that black men receive much greater returns to tenure and lower returns to experience compared to white men. This result is obtained regardless of the estimator we chose to focus on making it a robust result. These results are conveniently summarised by figure 1 and 2 which show the cumulative growth in log wages due to tenure and experience (calculated from the OLS wage regressions). In figure 1 the wage tenure profile for black men consistently lies above the tenure profile for white men. The opposite is true in figure 2, where the wage experience profile for white men dominates. According to figure 1, black men on average earn an estimated 30% cumulative wage growth due to 10 years of tenure. This figure is only 12% for white men. However, 10 years of labour market experience results in 25% cumulative wage growth for black men and 68% for white men.

Table 1: Returns to Tenure & Experience

			Black men		
Panel A	**OLS**	**FE**	**RE**	**2SLS**	**CF**
Tenure	0.0379	0.0312	0.0396	0.0165	0.0226
	(0.0036)***	(0.0099)***	(0.0038)***	(0.0104)	(0.0055)***
Tenure^2	-0.0011	-0.0015	-0.0012	-0.0006	-0.0012
	(0.0002)***	(0.0008)*	(0.0003)***	(0.0008)	(0.0002)***
P.Experience	0.0264	0.0163	0.0268	0.0334	0.0321
	(0.0039)***	(0.0085)*	(0.0039)***	(0.0045)***	(0.0042)***
P.Experience^2	-0.0004	-0.0002	-0.0004	-0.0005	-0.0004
	(0.0001)***	(0.0002)	(0.0001)***	(0.0001)***	(0.0001)***
Constant	1.1366	1.1598	1.0829	1.1072	1.0986
	(0.0875)***	(0.1571)***	(0.0875)***	(0.0885)***	(0.0881)***
Observations	14331	14331	14331	14331	14331
R-squared	0.59	0.08	0.06	0.59	0.59
			White men		
Panel B	**OLS**	**FE**	**RE**	**2SLS**	**CF**
Tenure	0.0187	0.0148	0.0186	-0.0049	-0.0010
	(0.0100)*	(0.0310)	(0.0112)*	(0.0369)	(0.0151)
Tenure^2	-0.0007	-0.0015	-0.0008	-0.0006	-0.0009
	(0.0007)	(0.0028)	(0.0008)	(0.0034)	(0.0007)
P.Experience	0.0730	0.0261	0.0689	0.0840	0.0803
	(0.0107)***	(0.0460)	(0.0114)***	(0.0159)***	(0.0117)***
P.Experience^2	-0.0021	-0.0002	-0.0018	-0.0023	-0.0021
	(0.0005)***	(0.0015)	(0.0005)***	(0.0008)***	(0.0005)***
Constant	1.7559	2.7097	1.6424	1.7193	1.6887
	(0.2685)***	(0.4345)***	(0.2467)***	(0.3660)***	(0.2686)***
Observations	2437	2437	2437	2437	2437
R-squared	0.46	0.11	0.07	0.45	0.46

Regressions also control for education, province, firm size, wave, household head, rural/urban, union status, occupation, and industry. Sample restricted to men between the ages of 15 and 35 who work in private and formal sector and not in subsistence agriculture or self-employment.
Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Returns to tenure for young black men being greater than those accruing to young white men and vice-versa for returns to experience are consistent with findings from other countries. Bratsberg and Terrell (1998), for example, find similar results using data from the United States.

Figure 1: Cumulative log wage growth due to tenure (Based on OLS results in column 1 of Table 1 above)

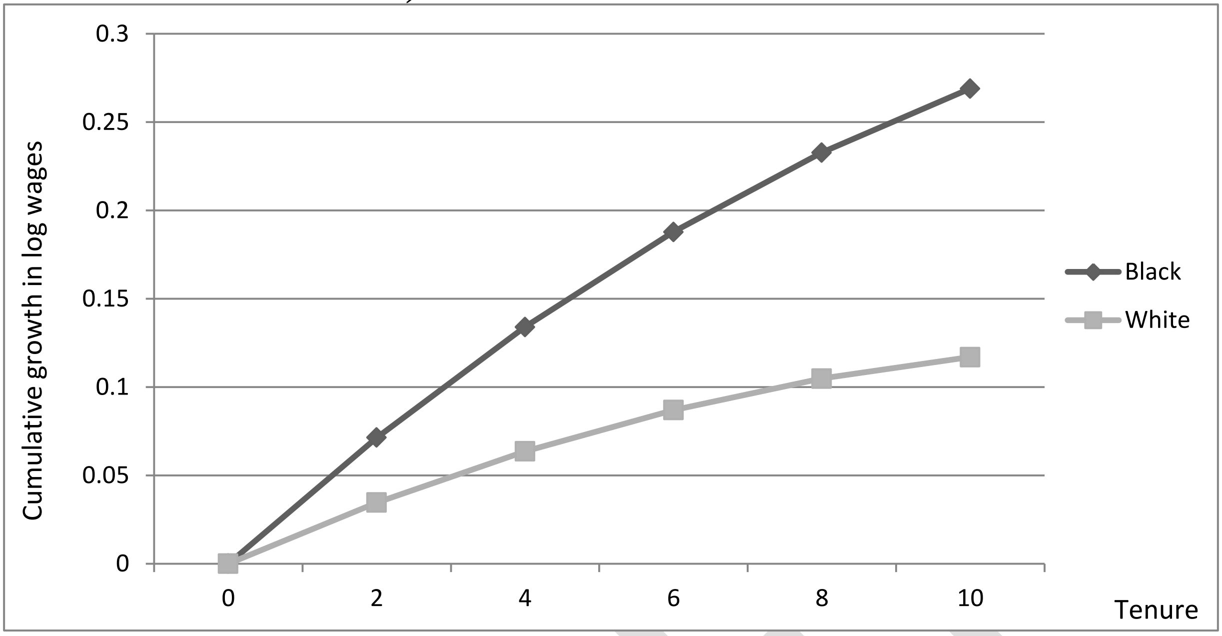
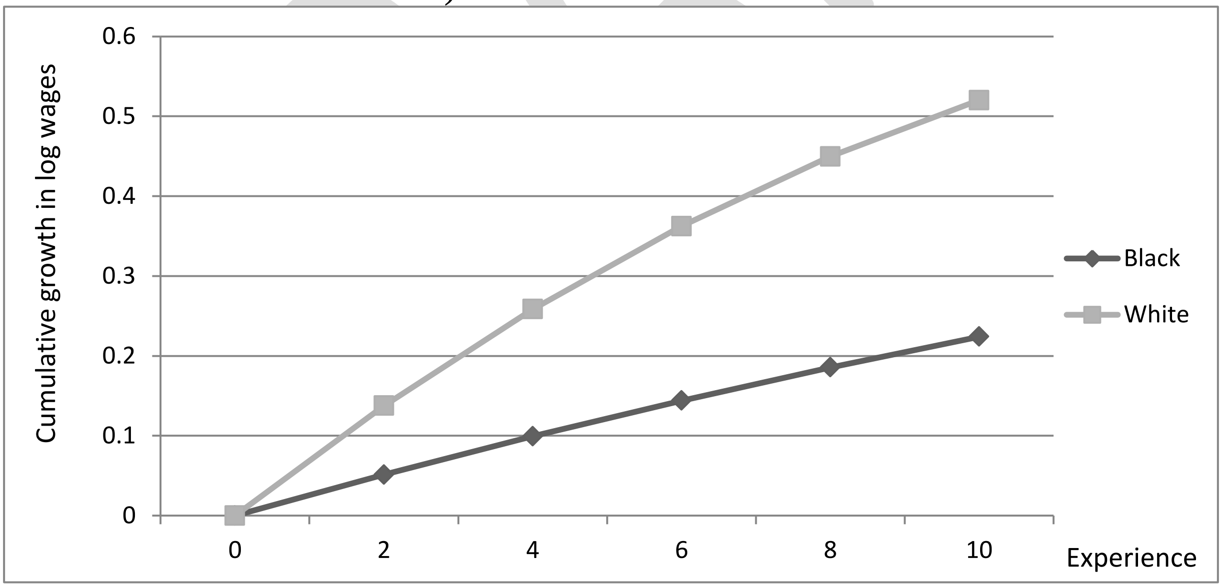


Figure 2: Cumulative log wage growth due to experience (Based on OLS results in column 1 of Table 1 above)



Because there are various explanations and theories for the empirical tenure- and experience-wage profiles, these theories and explanations may be relied upon in interpreting these results. These explanations range from the human capital theory explanation of individual productivity growth over the life-cycle that may be in embedded in the returns to tenure and experience (see for example Becker, 1962; Ben-Porath, 1967; and Mincer, 1974). Other notable explanations for the tenure- and experience-wage profiles emphasise the importance and consequences of imperfect information, implicit contracts and principal-agent considerations in the employer-employee relationship (See for example Salop and Salop, 1976; Jovanovic, 1979; Lazear, 1981; and Harris and Holmstrom, 1982). This paper focuses on imperfect information models.

Table 2: Log wage regression (Pooled OLS)

	Base	**Race**	**Age**	**Educ**	**All**
Tenure	0.045 (0.002)***	0.044 (0.002)***	0.046 (0.002)***	0.045 (0.002)***	0.046 (0.002)***
$Tenure^2$*100	-0.087 (0.005)***	-0.086 (0.005)***	-0.089 (0.005)***	-0.087 (0.005)***	-0.089 (0.005)***
Oneyear	0.041 (0.015)***	-0.120 (0.033)***	0.002 (0.021)	-0.096 (0.046)**	-0.251 (0.052)***
Oneyear*Black		0.188 (0.035)***			0.203 (0.038)***
Oneyear*Young			0.138 (0.049)***		0.125 (0.049)**
Oneyear*Below_Matric				0.132 (0.047)***	0.064 (0.050)
Oneyear*Matric				0.202 (0.051)***	0.164 (0.052)***
Black	-0.730 (0.011)***	-0.896 (0.033)***	-0.731 (0.011)***	-0.729 (0.011)***	-0.908 (0.036)***
Young (<36)	-0.400 (0.039)***	-0.393 (0.039)***	-0.502 (0.053)***	-0.398 (0.039)***	-0.486 (0.053)***
Below Matric	-2.036 (0.035)***	-2.039 (0.034)***	-2.037 (0.035)***	-2.158 (0.056)***	-2.101 (0.058)***
Matric	-0.963 (0.029)***	-0.966 (0.029)***	-0.964 (0.029)***	-1.146 (0.055)***	-1.113 (0.056)***
Constant	3.230 (0.056)***	3.364 (0.062)***	3.261 (0.058)***	3.353 (0.068)***	3.481 (0.071)***

Observations	37684	37684	37684	37684	37684
R-squared	0.58	0.58	0.58	0.58	0.58

Regressions also control for education, potential experience, province, firm size, wave, household head, and rural/urban. Sample restricted to men between the ages of 15 and 65 who work in private and formal sector and not in subsistence agriculture or self-employment.
Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3: Log wage regression (Pooled OLS) – Robustness check (two years instead of one year)

	Base	**Race**	**Age**	**Educ**	**All**
Tenure	0.044	0.044	0.045	0.045	0.045
	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***
$Tenure^2*100$	-0.086	-0.085	-0.088	-0.086	-0.088
	(0.005)***	(0.005)***	(0.005)***	(0.005)***	(0.005)***
Twoyears	0.019	-0.067	0.003	-0.066	-0.139
	(0.008)**	(0.016)***	(0.011)	(0.020)***	(0.023)***
Twoyears*Black		0.100			0.103
		(0.016)***			(0.017)***
Twoyears*Young			0.053		0.042
			(0.024)**		(0.024)*
Twoyears_Below_Matric				0.082	0.048
				(0.020)***	(0.022)**
Twoyears*Matric				0.121	0.103
				(0.022)***	(0.023)***
Black	-0.730	-0.895	-0.730	-0.729	-0.897
	(0.011)***	(0.029)***	(0.011)***	(0.011)***	(0.031)***
Young	-0.398	-0.388	-0.466	-0.393	-0.440
	(0.039)***	(0.039)***	(0.050)***	(0.039)***	(0.050)***
Below Matric	-2.036	-2.042	-2.038	-2.182	-2.131
	(0.035)***	(0.034)***	(0.035)***	(0.050)***	(0.052)***
Matric	-0.963	-0.967	-0.964	-1.170	-1.143
	(0.029)***	(0.029)***	(0.029)***	(0.048)***	(0.049)***
Constant	3.237	3.369	3.258	3.379	3.491
	(0.056)***	(0.060)***	(0.057)***	(0.064)***	(0.067)***

DRAFT

Observations	37684	37684	37684	37684	37684
R-squared	0.58	0.58	0.58	0.58	0.58

Regressions also control for education, potential experience, province, firm size, wave, household head, and rural/urban. Sample restricted to men between the ages of 15 and 65 who work in private and formal sector and not in subsistence agriculture or self-employment.
Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4: Log wage regression – Heckman selection model

	Base	**Race**	**Age**	**Educ**	**All**
Tenure	0.046	0.046	0.048	0.046	0.047
	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***
Tenure^2*100	-0.087	-0.087	-0.091	-0.088	-0.090
	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***
Oneyear	0.037	-0.133	-0.020	-0.022	-0.222
	(0.020)*	(0.050)***	(0.027)	(0.058)	(0.069)***
Oneyear*Black		0.192			0.251
		(0.051)***			(0.055)***
Oneyear*Young			0.201		0.173
			(0.065)***		(0.066)***
Oneyear*Below_Matric				0.031	-0.054
				(0.060)	(0.063)
Oneyear*Matric				0.174	0.122
				(0.067)***	(0.067)*
Black	-0.731	-0.902	-0.733	-0.730	-0.954
	(0.018)***	(0.049)***	(0.018)***	(0.018)***	(0.052)***
Young	-0.230	-0.231	-0.388	-0.231	-0.365
	(0.065)***	(0.065)***	(0.083)***	(0.065)***	(0.083)***
Below Matric	-2.033	-2.037	-2.036	-2.063	-1.992
	(0.046)***	(0.046)***	(0.046)***	(0.072)***	(0.074)***
Matric	-0.943	-0.947	-0.947	-1.098	-1.056
	(0.039)***	(0.039)***	(0.039)***	(0.072)***	(0.073)***
Lambda	-0.077	-0.070	-0.067	-0.080	-0.067
	(0.037)**	(0.037)*	(0.037)*	(0.037)**	(0.037)*
Constant	2.968	3.111	3.008	3.028	3.190

	(0.067)***	(0.077)***	(0.068)***	(0.082)***	(0.087)***
Observations	60116	60116	60116	60116	60116

Regressions also control for education, potential experience, province, firm size, wave, household head, and rural/urban. Selection equation additionally controls for number of children and elderly in the household, and presence of an old age social grant recipient in the household. Sample restricted to men between the ages of 15 and 65 who work in private and formal sector and not in subsistence agriculture or self-employment.
Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5: Employment type (permanent vs non-permanent)[3]

[3] Dependent variable → permanent (1-yes; 0-no)
gen permanent = 0 if (jlength==2 | jlength==3 | jlength==4 | jlength==5)
replace permanent = 1 if jlength==1

DRAFT

	Base	Race	Age	Educ	All
Oneyear	0.251 (0.009)***	0.021 (0.017)	0.214 (0.012)***	0.046 (0.021)**	-0.129 (0.025)***
Oneyear*Black		0.276 (0.019)***			0.228 (0.020)***
Oneyear*Young			0.133 (0.030)***		0.169 (0.030)***
Oneyear*Below_Matric				0.261 (0.023)***	0.180 (0.025)***
Oneyear*Matric				0.154 (0.026)***	0.110 (0.026)***
Black	-0.086 (0.004)***	-0.328 (0.018)***	-0.086 (0.004)***	-0.085 (0.004)***	-0.286 (0.020)***
Young	-0.217 (0.021)***	-0.206 (0.020)***	-0.315 (0.031)***	-0.205 (0.020)***	-0.324 (0.031)***
Below Matric	-0.061 (0.013)***	-0.066 (0.013)***	-0.063 (0.013)***	-0.299 (0.026)***	-0.232 (0.027)***
Matric	-0.011 (0.009)	-0.014 (0.009)	-0.012 (0.009)	-0.155 (0.026)***	-0.118 (0.026)***
Constant	0.569 (0.025)***	0.758 (0.028)***	0.599 (0.026)***	0.738 (0.030)***	0.881 (0.032)***
Observations	39772	39772	39772	39772	39772
R-squared	0.24	0.24	0.24	0.24	0.25

Regressions also control for tenure, education, potential experience, province, firm size, wave, household head, and rural/urban. Sample restricted to men between the ages of 15 and 65 who work in private and formal sector and not in subsistence agriculture or self-employment.
Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6: Written Contract (Y/N)

	Base	Race	Age	Educ	All

jlength → type employment, i.e. jlength (1-permanent; 2-fixed period contract; 3-temporary; 4-casual; 5-seasonal)

DRAFT

Oneyear	0.111 (0.009)***	-0.001 (0.017)	0.089 (0.012)***	-0.061 (0.020)***	-0.134 (0.024)***
Oneyear*Black		0.135 (0.019)***			0.079 (0.020)***
Oneyear*Young			0.080 (0.030)***		0.106 (0.030)***
Oneyear*Below_Matric				0.219 (0.021)***	0.191 (0.023)***
Oneyear*Matric				0.130 (0.024)***	0.111 (0.024)***
Black	-0.057 (0.006)***	-0.175 (0.018)***	-0.057 (0.006)***	-0.057 (0.006)***	-0.126 (0.019)***
Young	-0.239 (0.023)***	-0.233 (0.023)***	-0.298 (0.032)***	-0.229 (0.023)***	-0.304 (0.031)***
Below Matric	-0.198 (0.017)***	-0.200 (0.017)***	-0.199 (0.017)***	-0.397 (0.026)***	-0.375 (0.027)***
Matric	-0.021 (0.012)*	-0.023 (0.012)*	-0.022 (0.012)*	-0.143 (0.025)***	-0.127 (0.025)***
Constant	0.499 (0.030)***	0.591 (0.032)***	0.517 (0.030)***	0.642 (0.033)***	0.701 (0.035)***
Observations	39301	39301	39301	39301	39301
R-squared	0.16	0.16	0.16	0.17	0.17

Regressions also control for tenure, education, potential experience, province, firm size, wave, household head, and rural/urban. Sample restricted to men between the ages of 15 and 65 who work in private and formal sector and not in subsistence agriculture or self-employment.
Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

DRAFT