Imperfect information and the racial wage-gap for South African men

Abstract

The job-seeker-vacancy matching relationship is characterised by imperfect information and uncertainty. Employers are imperfectly informed about the productivity of job applicants, who in turn face imperfect information about the availability and nature of job vacancies. The resulting uncertainty leads to inefficiencies in job search and matching which may contribute to unemployment. When faced with asymmetric information about job-applicants, employers may resort to statistical discrimination and base hiring and promotion decisions partly on easily observable characteristics. This would disadvantage high productivity workers who are unable to send credible signals to employers. There are indications that this information asymmetry is particularly acute in the South African labour market (Levinsohn, 2007, and Abel et al. 2017) and that this may be one of the reasons for very high unemployment.

This paper investigates the impact of this kind of statistical discrimination as a determinant of wage gaps between races, age groups and education levels. A theoretical model of worker productivity uncertainty and employer learning is constructed for the South African labour market. The model combines insights from statistical discrimination and learning models to produce testable predictions regarding the impact of imperfect information on wage differentials. The predictions of this model are tested using reduced-form and structural approaches and South African data.

Both the reduced-form and structural estimates provide evidence in support of the hypothesis that South African employers engage in statistical discrimination based on race, age and educational attainment when making employment and wage decisions. Black, young and men with completed secondary school were found to have greater *ex ante* uncertainty around their expected productivity and thus benefitted more from employer learning. The greater uncertainty, it is argued, is driven by low and variable quality of pre-tertiary education that has reduced the potency of the matric certificate as a signal of worker's expected productivity.

1. Introduction

Two decades have passed since South Africa's political transition in 1994, yet large racial differences in earnings persist. Earnings differences are the single largest contributors to the high income inequality that has become synonymous with South Africa. The inequality of labour market earnings is deeply rooted in apartheid era policies that generated persistent differences in human capital. Underlying the differences in human capital is low quantity and quality of schooling received by blacks together with low labour market returns compared to whites (see for example Burger & Jafta, 2006; Burger & van der Berg, 2011; Branson & Leibbrandt, 2013; and Branson, Garlick, Lam and Leibbrandt, 2012). Labour market discrimination has also played role in the earnings inequality between blacks and whites (see for example Allanson, Atkins and Hinks, 2002).

This paper aims to extend the literature on earning differences by investigating an often cited but scarcely researched market imperfection that is a defining feature of the South African labour market. Employers face a great deal of uncertainty when predicting and assessing the expected productivity of job applicants. There are indications that this information asymmetry is particularly acute in the South African labour market (Levinsohn, 2007, Schoer, Rankin and Roberts, 2014; Schoer & Rankin, 2011; Duff & Fryer, 2005; and Abel, Burger and Piraino, 2017) and that this may be one of the reasons for the country's very high unemployment rate. We argue and provide empirical evidence that uncertainty in worker productivity systematically differs by race, age and level of school attainment. Additionally, we show that systematic differences in worker uncertainty have an impact on initial wages and on subsequent wage growth.

When faced with asymmetric information about job applicants, employers may resort to statistical discrimination by making hiring, wage offers and promotion based partly on easily observable characteristics. We investigate the impact of this kind of labour market discrimination as a determinant of wage inequality. This allows us to take the literature on labour market discrimination a step further than existing studies that only indirectly measure labour market discrimination as a residual component of an earnings regression. Addressing differences in human capital is an important long-term policy objective, however in the short-term there may be measurable success achieved if policy was targeted at other causes of earnings differences. The findings of this paper provide the necessary empirical evidence to guide such short-term policymaking. The short-term policies could, *inta alia*, include general skills assessment and certification by the Department of Labour's labour centres.

In this paper, we construct a theoretical model of worker uncertainty that allows the uncertainty to be resolved over the employment spell with the current employer. The model combines insights from traditional models of statistical discrimination (e.g. Aigner & Cain, 1977) and employer learning to produce testable predictions regarding the impact of information asymmetry on earnings differences. The model relies on differences in the variance of productivity to generate labour market discrimination by allowing the variance of productivity to be a function of observable characteristics. This represents a departure from traditional models of statistical discrimination that rely on differences in the accuracy of the signal of productivity to generate labour market discrimination. Differences in the variance of productivity are motivated by the large variation in learner performance and school quality that are well documented in the South African schooling literature (see for example van der Berg, 2007; van der Berg, 2008; and Branson & Leibbrandt, 2013).

To test the model's predictions, we perform reduced form and structural estimation using South African labour market data. Both the reduced form and structural estimates provide evidence in support of the hypothesis that South African employers engage in statistical discrimination based on race, age and schooling attainment when making employment and wage decisions. Black, young and men with completed secondary schooling are found to have greater *ex ante* uncertainty around their expected productivity and thus benefit more from employer learning. We further show that this uncertainty is driven by variation in worker productivity that is in turn driven by low and variable quality of pre-tertiary schooling received by many South Africans.

2. Background and context: South African literature

Earnings differences between black and white men in South Africa follows a racial hierarchy that is evident in other labour market outcomes. The literature has pointed to two main explanations for the persistence of high earnings inequality in favour of white men. Firstly, there is a large and persistent human capital differential between the two groups that favours white men. Under the apartheid government, "the schooling system was divided along racial lines" with "unequal educational funding, support and management" (Branson & Leibbrandt, 2013:7). As a result, blacks acquired fewer years of educational attainment compared to whites. The quality of the education received by blacks was also inferior (Moll, 1998; Lam, Ardington & Leibbrandt, 2011; van der Berg, 2007; Burger & van der Burg, 2011; and Branson & Leibbrandt, 2013). Access to schooling for blacks however began to increase in the last years of the apartheid regime and a period of convergence in educational attainment emerged between the two groups. The quality of schooling in black schools, on the other hand, remained low and possibly deteriorated further as the expansion in access to schooling gained momentum after the political transition (Moll, 1998; van der Berg, 2007; and Branson & Leibbrandt, 2013).

The second major explanation advanced in the literature for the persistence of earnings differences between the two racial groups relates to the evidence of labour market discrimination suffered by suffered by black workers. This evidence is based on decomposition techniques that decompose the earnings difference between black and white men into a component reflecting differences in human capital and a residual component. The residual component is then used as a measure of the extent of labour market discrimination. Allanson, Atkins and Hinks (2000) find that differences in human capital account for about two thirds of the difference in earnings between black and white men with the reaming third accounted for by labour market discrimination suffered by black men. The labour market discrimination component has been found to be persistent even after the implementation of affirmative action policies (Burger & Jafta, 2006).

The above literature on earnings differences has several important shortcomings. The bulk of this literature has been focused on improving our understanding of human capital differentials and how these differentials can be overcame with the use of government policy. The key policy recommendation has been improving the education system. Fixing the education system is an important policy objective. However, a policy objective that can only be significantly realised in the long-term. In the short-term, we may have to look at other policies that perhaps target other causes of earnings differences. This paper aims to provide the necessary empirical evidence to guide such short-term policymaking.

The labour market discrimination literature, on the other hand, offers no concrete policy recommendations. Burger and Jafta (2006) provide evidence, which indicates that affirmative action policies have been largely ineffective in combating labour market discrimination. Part of the problem is that this literature almost exclusively only deals with measuring the potential magnitude of the effect of labour market discrimination on earnings differences. No attempts are made investigate the nature, cause and transmission mechanism of discrimination in the labour market on the one hand and labour market outcomes on the hand. Furthermore, this literature only measures labour market discrimination indirectly as the residual of an earnings regression after controlling for human capital differences. This paper aims to take the literature a step further than existing studies by modelling labour market discrimination as a process driven by information asymmetries deeply rooted in the South African schooling system.

There are indications that information asymmetry is particularly acute in the South African labour market (Levinsohn, 2007, Schoer et al., 2014; Schoer & Rankin, 2011; Duff & Fryer, 2005; and Abel et al. 2017). The assessment of the potential productivity of job-seekers is plagued by uncertainty on the employers' side. There is a large body of evidence which indicates that employers' ability to accurately assess the potential productivity of job-seekers may systematically differ by race, age and level of educational attainment. The large variation in both learner performance and school quality between and within population groups can be plausibly linked to the uncertainty faced by employers when assessing the productivity of job-seekers. There is a research gap in our literature regarding the impact of this uncertainty on labour market outcomes.

The literature on the economics of education in South Africa points to two channels through which uncertainty regarding the potential productivity of job-seekers may operate. Firstly, the evidence points to a lack of credible signals of worker productivity for job-seekers who never completed secondary schooling. In addition, there is a weakening of the secondary school certificate as a signal of a job-seeker's potential productivity for those who do not go on to obtain tertiary school qualification (Duff & Fryer, 2005; Schoer & Rankin, 2011; Schoer et al., 2014; and Levinsohn, 2007). The low quality of schooling received by the majority of leaners coupled with the massive increase in the supply of job-seekers with a secondary school certificate are the main culprits for the problems associated with the signals of worker productivity.

Variation in the accuracy of the signals of worker productivity is an incredibly important source of uncertainty for employers when they evaluate the suitability of job-seekers for employment and the wage level to set. Variation in the accuracy of the signals of worker productivity is the key motivation for many models of statistical discrimination. In this paper, we however focus on another equally important source of uncertainty in the assessment of worker productivity that is neglected in the international literature and that is of immense importance for the South African case. The evidence on the South African education system points to greater variation in schooling outcomes for blacks compared to whites. The greater variation in schooling outcomes for blacks feeds into greater variation in labour market productivity. Greater variation in productivity, in turn, leads to greater uncertainty on the part of the employer.

In addition to lower learner performance, black schools have higher variation in learner performance (van der Berg, 2002, 2007 & 2008). Within the black population, inequality in terms of educational attainment and in the quality of education has been growing. This inequality appears to strongly correlate with socio-economic status (van der Berg, 2002 & 2008; and Branson & Leibbrandt, 2013). Those with lower socio-economic statuses and residing in rural areas tend to have lower schooling outcomes. The effect of the low and variable quality of black schools meant that black learners acquired fewer numerical and comprehension skills (Moll, 1998).

Chamberlain and van der Berg (2002); Burger and van der Berg (2011); Branson, Ardington, Lam and Leibbrandt (2013); and Lam et al. (2011) have expressed concerns regarding the effective level of learning and cognitive gains achieved due to the high variation in learner performance and the low and variable quality of schooling received by mainly black learners. The performance of black learners is poor and highly variable even when compared to other African and developing countries that dedicate fewer resources to education (van der Berg, 2007). Furthermore, black learners also have higher grade repetition rates compared to their white counterparts (Lam, Leibbrandt and Mlatsheni, 2007). This further undermines educational attainment for black learners and further widen the variation in productivity.

Consequently, our task in this paper is to construct a theoretical model that models uncertainty regarding worker productivity as being driven by variation in worker productivity across groups. Before fleshing out the details of the theoretical model, in the next section we review the theoretical and empirical literature on statistical discrimination.

3. Statistical discrimination: theory and empirical evidence

Information asymmetry and its impact on the assessment of worker productivity is a subject of many theoretical models in the labour economics literature (e.g. job-matching models – Jovanovic, 1979; implicit contract models – Harris and Holstrom, 1982; and adverse selection models – Salop and Salop, 1976). This paper, however, finds its theoretical roots in the statistical discrimination models first pioneered by Phelps (1972) and Arrow (1973), and further developed by Aigner and Cain (1977) and Lundberg and Startz (1983). According to the statistical discrimination model, when employers are faced with information asymmetry regarding the skills set and expected productivity of job-seekers, then employers have an incentive to use easily observable characteristics to distinguish among workers if these characteristics are correlated to productivity. This leads to an outcome where "the average wage of a group is not proportional to its average productivity" and this in turn constitutes economic discrimination (Aigner and Cain, 1977:178).

Early models of statistical discrimination (e.g. Phelps, 1972), were premised on employers' perceived differences in average productivity between black and white workers. The modern models on the other hand, rely on differences in employers' ability to assess the productivity of job-seekers from different racial or gender groups. The key features of these models are set out in Aigner and Cain's (1977) influential paper. Black and white workers are assumed to have the same average productivity. The employer however is not able to observe a worker's actual productivity. The employer must rely on a noisy signal of productivity when assessing a worker's suitability for employment. The signal of productivity can range from information contained in a resume and job application forms, to a test score from a placement evaluation. The key assumption for this model and many subsequent models of statistical discrimination is that the signal of productivity is less informative for black workers. Lang (1986) motivates this assumption with research originating from sociolinguistics and sociology that highlights cultural and communication (verbal and non-verbal) differences between blacks and whites. Lang argues that these differences make it harder for managers and supervisors to assess workers belonging to groups other their own.

Aigner and Cain (1977) showed that a model with only the above features fails to generate an equilibrium outcome that constitutes economic discrimination. To demonstrate labour market discrimination, the authors make one further key assumption. They assume that employers are risk averse. With this additional assumption, the employer's employment decision and wage offer is also a function of the conditional variance of productivity. Subsequently, the model shows that blacks receive lower wages on average even though their average productivity is similar to that of whites. This constitutes labour market discrimination since the average wages of blacks are not proportional to their average productivity (Aigner and Cain, 1977).

Lundberg and Startz (1983) extended the statistical discrimination model to incorporate a human capital investment option. This represented an improvement on the Aigner and Cain (1977) model that took pre-labour market human capital investment as given. The human capital investment is assumed to be costly, unobservable and undertaken prior to entering the labour market. The assumption that blacks have a less informative productivity signal is maintained. With black workers' productivity measured with greater error and human capital investments unobservable, black workers, receive lower returns to their human capital. As a result, blacks have a lower incentive to invest on human capital. The lower investment on human capital leads to lower average wages even though blacks start with average productivity that is equal to that of whites.

The models by Aigner and Cain (1977) and Lundberg and Startz (1983) have been influential but have also attracted some criticism. Assuming that employers are risk averse is hugely unpopular (Lang, 1986). In the next section, we provide possible justification for this assumption and motivate why it is a reasonable assumption for the South African labour market. Additionally, we conduct a simple test that lends credence for the assumption of employer risk aversion. Lundberg and Startz's (1983) assumption of unobservable human capital investment and the result that blacks acquire less human capital than equally comparable whites is at odds with available US empirical evidence (Lang, 1983; and Oettinger, 1996).

The statistical discrimination literature has traditionally been theoretical in nature with very few empirical contributions. The earlier models offered very few predictions that could be tested empirically. The recent models, in contrast, have a dynamic structure and incorporate the employer learning hypothesis. These extensions have made the recent models of statistical discrimination more amenable to empirical testing.

Oettinger (1996) extended the statistical discrimination 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. One of the key predictions from the model relies on racial differences in the estimated returns to tenure, labour market experience and job mobility between black and white men. The model showed that groups that are statistically discriminated against should have lower estimated wage returns to labour market experience and job mobility, and higher estimated wage returns to tenure compared to groups that suffer no statistical discrimination. Oettinger (1996), Lewis and Terrel (2001) and Goldsmith, Hamilton and Darity (2006) provide evidence consistent with this prediction.

Altonji and Pierret (2001) devised an alternative empirical test for statistical discrimination that relies on employers' ability to learn about the true productivity of their workers by observing their output. As employers learn about the true productivity of their workers, the coefficients on the easy to observe correlates of productivity in wage regression should fall while the coefficients on the hard to observe correlates should rise. Using U.S. data on young people, Altonji and Pierret (2001) find evidence of young workers being statistically discriminated against on the basis of education. Interestingly, the authors find no evidence of statistical discrimination on the basis of race even though race is a good candidate for an easy to observe correlate of productivity.

Pinkston (2006) uses the framework developed by Altonji and Pierret (2001) and demonstrates that black men in the US have less credible labour market signals compared their white counterparts when these workers enter the labour market. Strobl (2003) apply this framework to a developing country. The author uses matched employer-employee data from Ghana and provides evidence in support of the statistical discrimination model.

4. Theoretical model

In this section, we develop a statistical discrimination model that incorporates learning by employers for the South African labour market. The model developed here is in the spirit of and follows the formulation of Aigner and Cain (1977). Our model, however defers from Aigner and Cain's (1977) model as it relies on differences in the variance of productivity, as opposed to differences in the accuracy of signal of productivity, to generate labour market discrimination. We allow the variance of productivity to be a function of observable characteristics. Group variation in productivity is motivated by low and variable learner performance and school quality within the black population and between the two population groups.

4.1 Model setup

Suppose individual worker productivity, y, is determined as:

$$y = \alpha + \boldsymbol{\theta s} + u\sigma_u \qquad (1)$$

where α is a constant, $\boldsymbol{s}$ is a vector of observable determinants of worker productivity, $\boldsymbol{\theta}$ is a vector of parameters capturing the effect of $\boldsymbol{s}$ on worker productivity. Worker productivity is also a function of other factors that are assumed to be unobservable and thus captured by the model error term, u. As in Aigner and Cain (1977), y is assumed to be normally distributed with zero mean and a standard deviation of σ_u.

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

$$\hat{y}_t = y + e_t\sigma_e \qquad (2)$$

e_t captures the noise or error in employers' assessment of worker productivity and is assumed to be normally distributed with zero mean and a standard deviation of one. Employers can observe $\boldsymbol{s}$, which includes all individual attributes that can be obtained from résumés and job application forms, but not u. At period t, the firm can observe all the previous signals $(\hat{y}_0, \dots, \hat{y}_t)$, which can be used to form an expectation of the worker's productivity: $E(y|\boldsymbol{s}, \hat{y}_0, \dots, \hat{y}_t)$. We further assume that firms are risk-averse as they may dislike the uncertainty that results from variation in worker productivity. To incorporate risk-aversion by employers, we follow Aigner and Cain (1977) by allowing the firm's hiring decision and wage setting to be a function of not only the worker's conditional expected productivity but also the conditional variance y, written as $Var(y|\boldsymbol{s}, \hat{y}_0, \dots, \hat{y}_t)$.

Allowing firms to be risk-averse and to want to minimise uncertainty resulting from variation in worker productivity is a reasonable characterisation of the South African labour marker. Evidence from the behavioural and experimental economics literature, using laboratory games, has shown a systematic pattern of distrust and bias against black participants by white participants (Burns, 2006; and van der Merwe & Burns, 2008). Willingness by white participants to enter into strategic interactions with black participants in these games is impeded by the distrust and bias that appears to be motivated by racial stereotypes (Burns, 2006; and van der Merwe & Burns, 2008). Consequently, strategic interactions between black and white men in the labour market, like engaging in an employment relationship, may be subject to the same systematic pattern of distrust and bias. There is also a growing perception by firms that once entered into, employment contracts are costly and burdensome to terminate in South Africa (Levinsohn, 2007, and Schoer & Rankin, 2011). This further heightens the level of risk since the majority of South African job-seekers receive low and variable quality of schooling. Other labour market rigidities like adherence to minimum wages, affirmative action legislation, and bargaining council agreements that are in operation in the South African labour market further compound the risk factor involved in the employment process. Our empirical analysis provides an explicit test for the assumption of risk aversion by employers.

The firm's wage offer at period t to a worker with observables $(\boldsymbol{s}, \hat{y}_0, \dots, \hat{y}_t)$ is given by

$$w_t = E(y|\boldsymbol{s}, \hat{y}_0, \dots, \hat{y}_t) - \delta Var(y|\boldsymbol{s}, \hat{y}_0, \dots, \hat{y}_t) \qquad (3).$$

δ is a parameter that captures the importance of worker uncertainty on the firms wage offer. According to equation (3), wages depend positively on expected worker productivity and negatively on its conditional variance. Thus, high variance workers incur a wage penalty and the importance of that penalty depends on δ. Risk neutral hiring decisions is a special case of equation (3) in which $\delta = 0$.

It follows from equation (3) that in period 0 the worker will earn

$$w_0 = E(y|\boldsymbol{s}, \hat{y}_0) - \delta Var(y|\boldsymbol{s}, \hat{y}_0) \qquad (4),$$

Solving for the conditional expectation and variance of productivity yields the following expression

$$w_0 = \alpha + \boldsymbol{\theta s} + \left(\frac{\sigma_u^2}{\sigma_u^2+\sigma_e^2}\right)(\hat{y}_0 - \alpha - \boldsymbol{\theta s}) - \delta\left\{\sigma_u^2\left(\frac{\sigma_e^2}{\sigma_u^2+\sigma_e^2}\right)^2 + \sigma_e^2\left(\frac{\sigma_u^2}{\sigma_u^2+\sigma_e^2}\right)^2\right\} \qquad (5),$$

More generally, equation (5) is expressed as follows for the period t wage[1]

$$w_t = \alpha + \boldsymbol{\theta s} + \left(\frac{\sigma_u{}^2}{\sigma_u{}^2+\left(\frac{1}{t+1}\right)\sigma_e{}^2}\right)\left[\left(\frac{1}{t+1}\right)(\hat{y}_0 + \cdots + \hat{y}_t) - \alpha - \boldsymbol{\theta s}\right] - \delta\left\{\left(\frac{\left(\frac{1}{t+1}\right)\sigma_e{}^2}{\sigma_u{}^2+\left(\frac{1}{t+1}\right)\sigma_e{}^2}\right)^2 \sigma_u{}^2 + \left(\frac{\sigma_u{}^2}{\sigma_u{}^2+\left(\frac{1}{t+1}\right)\sigma_e{}^2}\right)^2 \left(\frac{1}{t+1}\right)\sigma_e{}^2\right\} \quad (6)$$

Wages are therefore a function of the observable determinants of productivity, $\boldsymbol{\theta s}$, and the final two terms on the right-hand side of equation (6) represent the conditional expectation of productivity and the (negative) conditional variance of productivity. Expected productivity is updated based the difference between productivity signals that employers receive via interviews and evaluations of the worker's productivity and the unconditional expected productivity: $\left(\frac{1}{t+1}\right)(\hat{y}_0 + \cdots + \hat{y}_t) - \alpha - \boldsymbol{\theta s}$, This difference is weighted by the variances of worker productivity, σ_u^2, and noise in employers assessment of worker productivity, σ_e^2. The additional information gleaned from such signals becomes less important as the worker's tenure increases, which is consistent with Lange's (2007) prediction that employer learning about worker productivity is front-loaded and revealed early in an employment spell.

The conditional variance of productivity term in equation (6) tends to zero, since each successive period of employment provides more information and resolves uncertainty about the productivity of the employee. Equation (6), therefore, locates the sources of individual wage growth in two places. Firstly, wages may grow as the firm acquires positive new information about individual productivity relative to the group average,$(\hat{y}_0 + \cdots + \hat{y}_t) - \alpha - \boldsymbol{\theta s}$. Secondly, the additional information removes the uncertainty regarding the worker's productivity and thus the penalty attached to such uncertainty diminishes and allows the wage to converge on the worker's expected productivity.

4.2 Model predictions

We follow Aigner and Cain (1977) who define discrimination as differences in earnings across groups that is not related to differences in average ability between the groups. As such, consider two groups of workers (group 1 and group 2) with equal average productivity, expressed as $E(y_1) = E(y_2) = \mu$. Suppose that the former has a larger conditional variance of productivity, $\sigma_{u1}^2 > \sigma_{u2}^2$, given the signal of productivity, $\hat{y}_t$. We denote group 1 as the disadvantaged group, the group that suffers labour market discrimination in the form of statistical discrimination.

[1] The reader is referred to the appendix for the derivation of equations (5) and (6)

Assuming that group 1 workers have a larger conditional variance of productivity implies that employers face greater uncertainty in predicting the expected productivity of workers from this group. Risk averse employers will attach a larger penalty against group 1 workers because of the greater dispersion around their productivity. Given that the two groups were assumed to have equal average productivity, equation (5) predicts lower initial average wages for group 1 workers compared to group 2 workers. The lower initial average wages for group 1 workers is not related to average productivity differences between the two groups and therefore constitutes statistical discrimination.

As the employer views the worker's output on the job, the employer will acquire new information about individual worker productivity relative to the group average. The arrival of new information will help resolve the *ex ante* uncertainty and allow the employer to update his initial assessment of the worker's productivity. As the uncertainty gets resolved, the wage penalty associated with the uncertainty falls. Our model therefore predicts greater subsequent relative wage growth for group 1 workers. This is a standard result of the employer learning hypothesis and suggests that changes in wages result from the arrival of new (positive) information regarding worker productivity (Kahn and Lange, 2014, and Sicilian, 1995). Assuming greater dispersion in productivity, $\sigma_{u1}^2 > \sigma_{u2}^2$, for group 1 workers, allows the employer to bridge the informational gap in expected productivity for the individual worker relative to the group average. This information gap is smaller for group 2 workers because this group is more homogenous and thus the arrival of information for the individual does not constitute new information.

Employer learning will lead to greater subsequent wage growth if the arrival of new information permits positive updating of the employer's initial assessment of worker productivity. Not all group 1 workers will turn out to be good workers and earn a positive subsequent assessment of their productivity. This is implied by the greater variation in productivity for this group. Consequently, if employers continue to set wages equal to the uncertainty adjusted expected productivity in each period, the model predicts that the variance of wages will increase more rapidly over the employment spell for group 1 workers compared to the group 2 workers.

In section two, we reviewed a large body evidence, which indicated that there is greater variation in labour market productivity for black men driven by greater variation in educational attainment, low and variable learner performance, and low and variable quality of pre-tertiary schooling received by this group. In light of the preceding discussion, black men are considered to be group 1 (disadvantaged group) and white men to be group 2. Accordingly, our theory model makes the following predictions:

a) Black men will have lower initial wages, conditional on expected productivity;
b) Black men should experience greater within-firm wage growth; and
c) The variance of wages for black men should increase more rapidly over the employment spell.

In other words, black men is the group that suffers statistical discrimination. We will however also investigate statistical discrimination based on age and level of educational attainment. The next section discusses the data used for the empirical analysis and descriptive analysis.

5. Data and descriptive analysis

The empirical analysis in the next section makes use of the individual cross-sectional waves of the Labour Force Surveys (LFS) together with the panel version – Labour Force Survey Panel (LFSP) collected 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 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 LFS 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 approximately 30 000 households was maintained in each wave.

Stats SA's LFSP is the first nationally representative panel dataset of the South African labour market. It was constructed from the LFS cross-sectional surveys running from September 2001 to March 2004 (Stats SA, 2006). Individuals were only linked after the collection and release of the surveys, since the surveys were designed as a rotating panel of dwelling units rather than individuals (Stats SA, 2006).

Although our identification of statistical discrimination is mainly achieved from the cross-sectional correlation between tenure, wages and other the easily observable individual attributes, we use the panel structure to address inconsistencies in the tenure variable. Light and Brown (1992) provide detailed motivation for the use of panel data for this purpose. The algorithm used to construct an internally consistent tenure variable is described in Burger (2016).

The estimation sample is restricted to black and white men between the ages of 18 to 60 working in formal, private sector firms. Workers in subsistence agriculture and those reporting to be self-employed were also excluded from the analysis. Workers with a reported tenure value that is above 20 years are also dropped from the analysis.

In the previous section, our theory model predicted (prediction C) that the variance of wages for black men should increase with tenure at a more rapid pace compared to their white counterparts. Figure 1 below, plots the standard deviation of the log of hourly wage against tenure. The blue curve for black men is positively slopped indicating that the variance for this group increases with tenure. On the hand, the red curve for white men is negatively slopped indicating that the variance for this group decreases with tenure. For both groups we conducted statistical hypothesis testing of the equality of the variances at one year and ten years of tenure. Levene's robust test statistic[2] (Levene, 1960) was used and for the both groups the null hypothesis that the variance at one and at ten years of tenure was rejected at the 99% level of significance. In other words, the positive relationship in the blue curve and negative relationship in the red curve between the variance of wages and tenure are statistically significant. This evidence is consistent with our model's prediction. In the next section we conduct, formal testing of the model's other predictions.

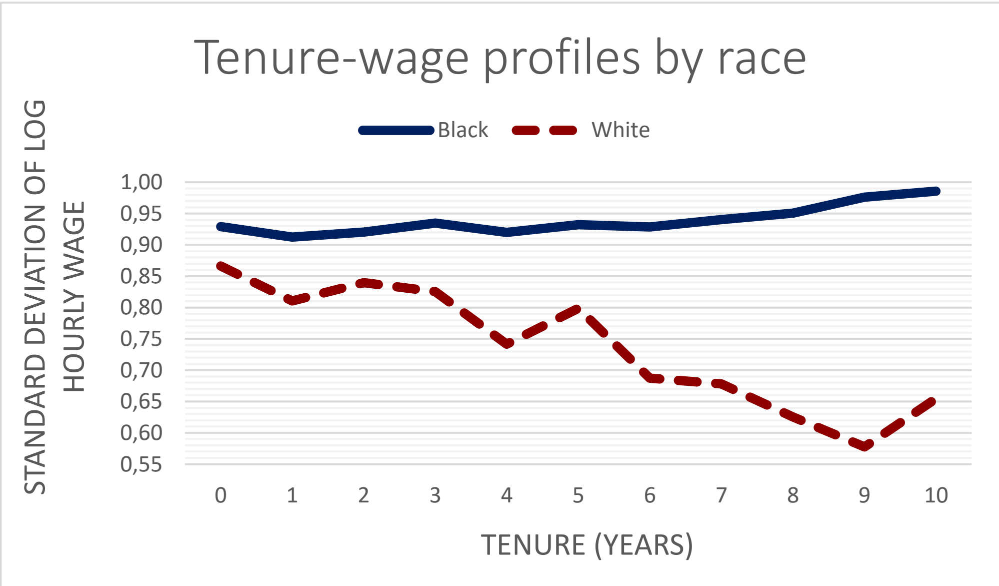

Figure 1

6. Empirical analysis

In this section, we report and discuss the results from reduced form and structural estimation of the theory model.

6.1 Reduced form estimates

The reduced form estimation is based on ordinary least squares (OLS) and exploits group-specific variation in the average wage gain due to having accumulated at least one year of tenure as an indication of employer learning. The estimated employer learning is then used to draw

[2] The hypothesis testing was implemented with Stata's "robvar" command.

inferences about uncertainty regarding worker productivity for groups defined by race, age and educational attainment. Assuming that employers set wages equal to the uncertainty-adjusted expected productivity, conditional on a worker's productivity signal, initial average wages for disadvantaged workers (black, young and men with low levels of educational attainment) will be lower relative to their 'true' productivity. This is due to the greater difficulty of predicting these workers expected productivity. The lower initial wage relative to 'true' productivity represents the penalty that a risk-averse employer will attach to the wages of workers whose productivity is more uncertain.

If employers continue to equate wages to the uncertainty-adjusted expected productivity in each period, then as the uncertainty around the worker's productivity is resolved, actual wages should converge to true productivity. Therefore, there should be more rapid wage growth for disadvantaged workers since employers were more uncertain about their productivity. Since employer learning occurs early in an employment spell (Lange, 2007), we should observe a steeper wage tenure profile for these workers. This is indeed the prediction made and empirically tested by Oettinger (1996), Lewis and Terrell (2001), and Goldsmith et al. (2006) for the U.S. labour market.

We proceed to estimate a pooled OLS wage (hourly wages in logs) regression controlling for human capital, demographic and job characteristics. We then add a dummy variable equal to one if tenure is larger or equal to one, and zero otherwise. Adding this dummy variable in our hourly log wage regression ensures that the average 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). This tenure dummy variable is then interacted with race, age (dummies), and education in order to allow for heterogeneous short-term returns across groups that potentially differ in productivity variance.

Table 1 presents the results for the pooled OLS wage regressions, but only coefficients for our key variables of interest are reported[3]. Other variables that are controlled for but not shown in Table 1 include province, firm size, wave fixed effects, household head status, and geographical classification of residence (i.e. rural vs urban), occupation and industry dummies.

In Table 1, we specify schooling as a spline with knots at 7 years of schooling (*primary* – completed primary schooling), 11 years of schooling (*secondary* – incomplete secondary schooling), 12 years of schooling (*matric* – completed secondary schooling), and then the last knot represents those with more than 12 years of schooling (*tertiary*). We specify a separate

[3] The reader is referred to the appendix section for the full list of coefficients – Table A1.

dummy variable for individuals with 12 years of schooling plus diploma or certificate not obtained from a university (*diploma + certificate*).

Table 1: Log hourly wage regression, pooled OLS

Variables	Model 1a	Model 1b	Model 1c	Model 1d
Primary	0.026 (0.003)***	0.026 (0.003)***	0.026 (0.004)***	0.026 (0.003)***
Secondary	0.075 (0.005)***	0.076 (0.005)***	0.077 (0.005)***	0.075 (0.005)***
Matric	0.151 (0.016)***	0.152 (0.016)***	0.153 (0.016)***	0.107 (0.028)***
Diploma&Certificate	0.231 (0.029)***	0.232 (0.029)***	0.231 (0.029)***	0.284 (0.044)***
Tertiary	0.146 (0.016)***	0.146 (0.016)***	0.148 (0.017)***	0.174 (0.024)***
Black	-0.749 (0.018)***	-0.816 (0.039)***	-0.749 (0.018)***	-0.749 (0.018)***
Tenure	0.036 (0.004)***	0.036 (0.004)***	0.037 (0.004)***	0.036 (0.004)***
Tenure^^{2}	-0.001 (0.0002)***	-0.001 (0.0002)***	-0.001 (0.0002)***	-0.001 (0.0002)***
Potential Experience	0.023 (0.002)***	0.024 (0.002)***	0.026 (0.002)***	0.023 (0.002)***
Potential Experience^^{2}	-0.0003 (0.00004)***	-0.0003 (0.00004)***	-0.0003 (0.00004)***	-0.0003 (0.00004)***
Tenure dummy (one year)	0.025 (0.018)	-0.046 (0.041)	-0.005 (0.024)	-0.032 (0.056)
Black*Tenure dummy		0.085 (0.040)**		
Age (18-24)*Tenure dummy			0.045 (0.031)	
Age (25-30)*Tenure dummy			0.050 (0.021)**	
Age (31-35)*Tenure dummy			0.011 (0.017)	
No Matric*Tenure dummy				0.044 (0.057)
Matric*Tenure dummy				0.100 (0.060)*
Intercept	1.435 (0.057)***	1.488 (0.065)***	1.392 (0.062)***	1.446 (0.057)***
R^2	0.64	0.64	0.64	0.64
N	27,118	27,118	27,118	27,118

Other control variables: occupation, industry, rural/urban classification, province of residence, household head status, firm size, and wave dummies.
Robust standard errors are contained in parentheses.
** p<0.1; ** p<0.05; *** p<0.01*

We observe from model 1a that the wage return to incomplete secondary schooling (7.7%) is more than double that of primary schooling (2.6%). Completing secondary schooling (matric) brings about a further doubling of the wage returns to schooling (16.3%). A year of tertiary schooling increases the wage returns by a further 15.7%. While a post-secondary diploma or certificate, increases wages 26% regardless of how long an individual took to complete it.

Holding the level of schooling and other variables controlled for in our wage regression constant, black men earn significantly lower wages compared to white men. This is captured by the negative and very large coefficient (-0.749) on the black dummy variable.

An additional year of tenure appears to be more valuable than an additional year of potential experience. In model 1b to model 1d, we try to determine whether the wage return due to the accumulation of the first year of tenure differs by race, age and level of schooling completed.

In model 1b, the coefficient on the tenure dummy variable interacted with black is large and statistically significant. There is an 8.9% additional wage return for black men that does not accrue to white men. This suggests that, with everything else held constant, black men enjoy much rapid average wage growth within the first year of an employment spell relative to white men.

The tenure dummy variable is interacted with age dummies in model 1c. Men that are older than 35 years are the omitted category and form the comparison group for the age-tenure dummy interaction variables. The age-tenure dummy interaction is large and statistically significant only for men aged between 25 and 30. For these men, there is a 5% additional wage return that accrue to them in the first year of an employment spell.

In model 1d, we specify schooling dummy variables for incomplete secondary school and less, and for completed secondary schooling. These school dummy variables are then interacted with the tenure dummy variable. Men with more than 12 years of schooling are omitted category and serve as the comparison group for the school-tenure dummy interaction variables. The school-tenure dummy interaction is large and statistically significant only for men with 12 years of schooling. Relative to other men, these men enjoy an additional 10.5% wage return from accumulating the first year of an employment spell.

The results in Table 1 reveal that workers that are black, aged between 25 to 30, and have 12 years of school (completed secondary) enjoy greater wage growth in the first year of an employment spell, relative to their respective counterparts. We interpret this as evidence consistent with the presence of greater *ex ante* uncertainty about the productivity of these workers. The uncertainty is driven by greater variation in productivity for these workers and possibly by these workers having less informative productivity signals. Consequently, they benefit the most from ‘employer learning’ and uncertainty being resolved.

In Table 1, we implicitly assumed that employer learning and the resolving of worker uncertainty takes place within the first year of an employment spell. To test the robustness of our results, we re-ran all the regressions[4] in Table 1 but instead focused on the wage gain due to the accumulation of the first two years and first six months of an employment spell. These alternative specifications did not alter the pattern of the results presented in Table 1.

A further robustness check performed is correcting for sample selection bias that usually arises in wage regressions based on South African labour data. With high unemployment and the likelihood of obtaining employment varying by race, level of schooling and age, wage earners are very unlikely to be a random sample of the working-age population. We address[5] this issue by running a Heckman sample selection model[6]. This too did not alter the pattern of the results presented in Table 1.

In Table 2 and 3 below, we explore alternative channels through which employer learning and the resolving of worker uncertainty can operate. Faced with uncertainty about the potential productivity of a worker, a firm may choose to hire that worker on a contract or on a non-permanent basis. If the worker proves to be a good hire with the arrival of positive new information regarding his productivity, the firm could then award that worker with a written contract or permanent employment and this could include or not include a raise in the worker's wage.

In Table 2, we estimate the same regression as in Table 1 but with contract (takes on a value of one if a worker has written contract of employment, zero otherwise) as the dependent variable. In Table 3, the dependent variable is permanent (takes on a value of one if employment is on a permanent basis, zero if it is casual, fixed-term, or seasonal). The results reported in these tables are consistent with and provide credence to the evidence reported in Table 1. Black men, those aged 25 to 35, and those with 12 or less years of schooling (i.e. incomplete and complete secondary schooling) are more likely to be offered a written contact and to transition into permanent employment in the first year of an employment spell relative to their respective counterparts. This is conditional on being employed without a written contract or on a non-permanent basis in the previous period. It is however, worth pointing out that men with incomplete secondary schooling have a stronger likelihood of obtaining a written contact and to transition into permanent employment compared to men with complete secondary schooling. A possible explanation for this could be that a lack of credible signals of productivity might be

[4] These results are contained in Table A2 (two years) and Table A3 (six months) in the appendix.
[5] The results are contained in Table A4 in the appendix.
[6] We specified three exclusion restrictions – presence of a social grant recipient in the household, number of elderly people in the household, and number of children in the household.

a bigger concern for employers than the variation in the productivity for men with incomplete secondary schooling. This would then lead to a stronger preference by employers for utilising non-wage mechanisms (i.e. hiring on a non-employment basis or without a written contract) to insulate them against the uncertainty regarding worker productivity for workers with incomplete secondary schooling.

Table 2: Written contract of employment regression, pooled OLS

Contract (yes=1; no=0)	Model 2a	Model 2b	Model 2c	Model 2d
Tenure dummy (one year)	0.064 (0.012)***	-0.016 (0.023)	0.054 (0.015)***	-0.017 (0.028)
Black*Tenure dummy		0.096 (0.023)***		
Age (18-24)*Tenure dummy			0.004 (0.023)	
Age (25-30)*Tenure dummy			0.016 (0.014)	
Age (31-35)*Tenure dummy			0.022 (0.011)**	
No Matric*Tenure dummy				0.104 (0.029)***
Matric*Tenure dummy				0.055 (0.032)*
R^2	0.19	0.19	0.19	0.19
N	26,716	26,716	26,716	26,716
Other Controls	Yes	Yes	Yes	Yes

Table 3: Permanent employment regression, pooled OLS

Permanent (yes=1; other=0)	Model 3a	Model 3b	Model 3c	Model 3d
Tenure dummy (one year)	0.136 (0.012)***	-0.042 (0.020)**	0.119 (0.015)***	-0.010 (0.027)
Black*Tenure dummy		0.213 (0.021)***		
Age (18-24)*Tenure dummy			0.008 (0.020)	
Age (25-30)*Tenure dummy			0.027 (0.013)**	
Age (31-35)*Tenure dummy			0.030 (0.010)***	
No Matric*Tenure dummy				0.185 (0.028)***
Matric*Tenure dummy				0.104 (0.031)***
R^2	0.23	0.24	0.23	0.23
N	27,020	27,020	27,020	27,020
Other control	Yes	Yes	Yes	Yes

Other control variables: race dummy, schooling spline, tenure, tenure-squared, potential experience, potential experience-squared, occupation, industry, rural/urban classification, province of residence, household head status, firm size, and wave dummies.

Robust standard errors are contained in parentheses.

** p<0.1; ** p<0.05; *** p<0.01*

The empirical evidence provided thus far reveal that workers from disadvantaged groups earn lower wages and employed on non-permanent contracts. However, these workers experience more rapid subsequent wage growth and greater likelihood of transitioning to permanent employment contracts during the first year of employment. We interpret this evidence as being consistent with our hypothesis and thus supportive of our model's predictions. The remainder of this section will report and discuss results from the structural estimation of the theory model.

6.2 Structural estimation results

Up to this point, we have provided empirical evidence based on reduced form estimation of our theory model's predictions. These results provided empirical evidence in favour of greater *ex ante* uncertainty regarding the productivity of black men, men that are between the ages of 25 and 35, and men that have 12 years and less of completed schooling. The uncertainty regarding these disadvantaged groups could be driven by lack of credible signals of productivity, by greater variation in productivity, or by both. To the extent that uncertainty in worker productivity affects earnings differences, it is thus vital for policy-making that we establish which of these factors is the key driver. The reduced form estimation could not distinguish between lack of credible signals of productivity or greater variation in worker productivity. With the structural estimation of the theory model, we are able to distinguish between these sources of uncertainty. In our structural estimation, we hold constant the variation in the noise or error in employers' assessment of worker productivity and allow the variation in worker productivity to vary with the worker's observable characteristics. This allows us to offer precise and well-targeted policy recommendations.

The structural parameters of the theory model are estimated using maximum likelihood estimation. From equation (6) above, we take the expectation conditional on $\boldsymbol{s}$ and t. This gives us the expected log hourly wage function to be estimated:

$$E(w_t|\boldsymbol{s},t) = \alpha + \boldsymbol{\theta s} - \delta\left\{\sigma_u^2\left(\frac{\left(\frac{1}{t+1}\right)\sigma_e^2}{\sigma_u^2+\left(\frac{1}{t+1}\right)\sigma_e^2}\right)^2 + \left(\frac{\sigma_u^2}{\sigma_u^2+\left(\frac{1}{t+1}\right)\sigma_e^2}\right)^2\left(\frac{1}{t+1}\right)\sigma_e^2\right\} \quad (7).$$

The observable determinants of worker productivity ($\boldsymbol{s}$) are assumed to be schooling, potential experience, race and tenure. We then allow the variance of worker productivity to be determined by this function $\sigma_u^2 = \exp(\rho + \boldsymbol{\lambda}\tilde{\boldsymbol{s}})$, where $\tilde{\boldsymbol{s}}$ is a vector of the workers' observable characteristics that we allow to determine the variance of worker productivity. Included in $\tilde{\boldsymbol{s}}$ is school dummies (specified in the same manner as in the OLS regressions reported in Table 1 above), age dummies (specified in the same manner as in the OLS regressions reported in Table 1 above), and a race dummy for black men.

The likelihood function is given by equation (8) below:

$$L = log\left\{\emptyset\left[\frac{w_t - E(w_t|s,t)}{\sigma_u^2/sqrt\{\sigma_u^2+\left(\frac{1}{t+1}\right)\sigma_e^2\}}\right]\right\} - log\left(\frac{\sigma_u^2}{sqrt\{\sigma_u^2+\left(\frac{1}{t+1}\right)\sigma_e^2\}}\right) \qquad (8)$$

$\emptyset \rightarrow$ normal density operator.

A combination of the Newton-Raphson, Davidson-Fletcher-Powell, and Broyden-Fletcher-Goldfarb-Shanno methods were used for finding the numerical solution to maximization problem.

To distinguish between the two sources – variation in the signal of productivity and variation in productivity – of uncertainty regarding worker productivity, we normalize σ_e^2 to take on a value of one. For robustness, we experimented with different normalization values (0.5 and 2) and this did not alter the pattern of the results for the structural estimates[7].

In Table 4 below, we present results for the restricted and unrestricted versions of our model. The restricted model is essentially a Mincerian-like wage regression and represents the risk neutral special case of our model with the following restriction: $\delta = 0$, in equation (7). The unrestricted model is defined by equation (7) and estimates the expected log hourly wages and the variance of worker productivity simultaneously.

The top panel of Table 4 shows that the common parameters have very similar estimated coefficients across the two models. The one exception though is the coefficients on the tenure and tenure-squared terms. The unrestricted model shows a much smaller estimated wage return that is less concave. This suggests that the tenure-wage effect in the restricted model includes the effects of employer learning as uncertainty is resolved. Explicitly accounting for the variation in worker productivity as we do in the unrestricted model, makes the tenure-wage profile flatter and less concave as depicted by figure 2 below. In the context of our theoretical model, the non-linearity in the tenure-wage profile is appears to be largely a product of the non-linearity in the rate of employer learning.

Because the restricted model is nested within the unrestricted model, this allows us to perform a likelihood ratio test of the two models. This essentially provides a test for our employer risk averse assumption that is instrumental to our model and many other models of statistical discrimination that follow Aigner and Cain's (1977) formulation. The likelihood ratio test static is 3009 and leads us to strongly reject the null hypothesis and thus to prefer the unrestricted model. This is an important result because Aigner and Cain's model has attracted strong

[7] The results for the different normalization values of σ_e^2 are reported in Table A7 of the appendix.

criticism for the employer risk averse assumption. We have demonstrated that this assumption is empirically supported by the data, at least by the South African data.

Table 4: Structural estimation of the theoretical model, maximum likelihood estimation

	Unrestricted	Restricted
Log hourly wage		
Primary	0.071 (0.003)***	0.072 (0.003)**
Secondary	0.153 (0.004)***	0.156 (0.005)**
Matric	0.274 (0.013)***	0.279 (0.017)**
Diploma&Certificate	0.307 (0.017)***	0.306 (0.022)**
Tertiary	0.307 (0.008)***	0.304 (0.010)**
Potential Experience	0.048 (0.002)***	0.048 (0.002)**
Potential Experience^2	-0.001 (0.00003)***	-0.001 (0.00004)**
Black	-0.877 (0.013)***	-0.859 (0.017)**
Tenure	0.039 (0.006)***	0.066 (0.003)**
Tenure^2	-0.0005 (0.0003)*	-0.0015 (0.0002)**
Intercept	0.980 (0.048)***	0.748 (0.035)**
Productivity variance		
No Matric dummy	0.028 (0.010)**	
Matric dummy	0.022 (0.011)**	
Black	0.070 (0.009)***	
Age dummy (18-24)	0.082 (0.010)***	
Age dummy (25-30)	0.046 (0.008)***	
Age dummy (31-35)	-0.018 (0.008)**	
Intercept	-0.179 (0.011)***	
Delta	0.459 (0.094)***	
LR Chi^2(8)	3009.74	
N	38,493	38,493

* p<0.1; ** p<0.05; *** p<0.01

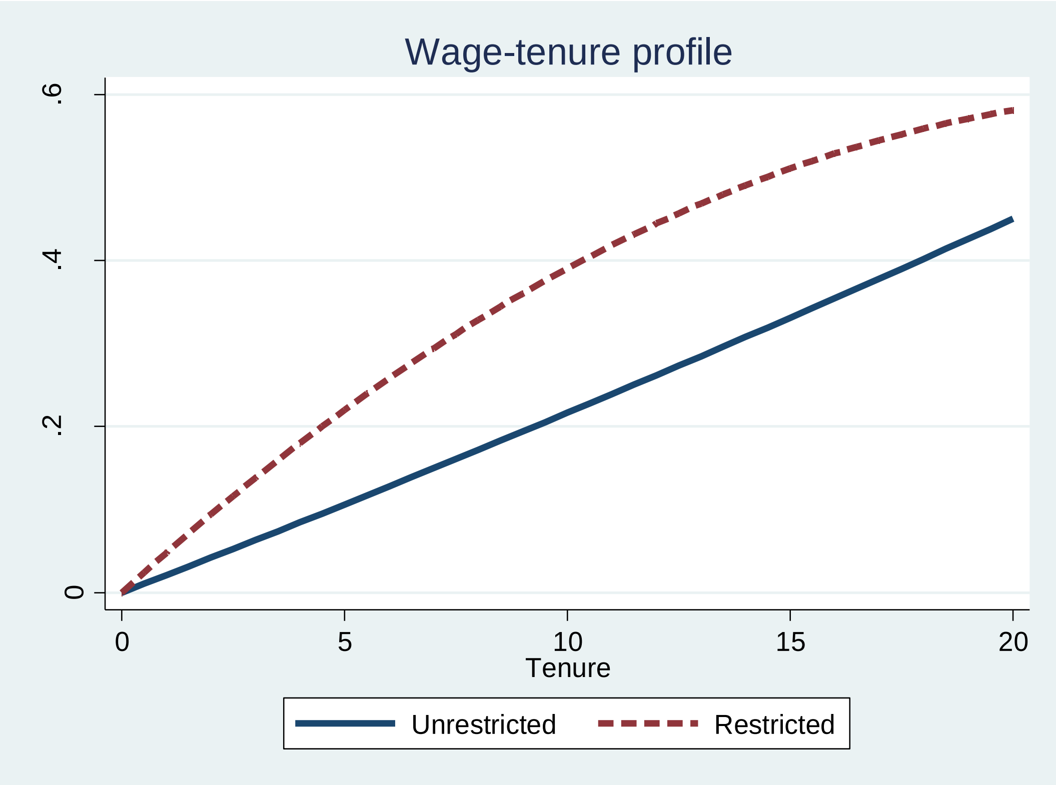

Figure 2

The bottom panel of Table 4 reports the results for the parameter estimation of the variance of worker productivity from the unrestricted model. The results suggests that, relative to men with tertiary schooling, there is greater variation in productivity amongst the men with 12 or less years of schooling. The parameter estimates for the age dummy variables suggests a negative relationship between the extent of variation in productivity and age. Men that are 24 and younger have greater variation around their productivity. This variation is almost halved for men in the 25 to 30 age category. The positive and statistically significant coefficient on black (0.07) indicates that there is greater variation in productivity amongst black men relative to their white counterparts.

These results are in accordance with our theoretical model and echo the literature discussed in section two that suggests greater variation worker productivity for disadvantaged groups. By holding the variation in the signal in productivity constant, we have demonstrated that the variation in worker productivity plays an important role in the uncertainty faced by employers when they have to assess the expected productivity of job applicants. In our model, the importance of this uncertainty is determined by delta (δ). This parameter is estimated as 0.459 and indicates the importance of the penalty incurred by workers with greater uncertainty around their expected productivity.

In Table 4, we looked at the individual effects of race, age and schooling on the variation in worker productivity. For many workers though, some of these attributes are not mutually exclusive. With this in mind, in Table 5[8] we interact our attributes of interest in order to investigate which of these c attributes reinforce or cancel each other concerning their impact on the variation of worker productivity. This would be important for policy makers tasked with drafting policies to address the uncertainty of job applicants' expected productivity.

Table 5: Structural estimation of the theoretical model, maximum likelihood estimation

	Model 5a	Model 5b	Model 5c	Model 5d
Productivity variance				
No Matric dummy	0.028 (0.010)***	0.030 (0.010)***	0.036 (0.011)***	0.028 (0.010)***
Matric dummy	0.022 (0.011)**	0.023 (0.011)**	-0.013 (0.017)	-0.010 (0.014)
Black	0.070 (0.009)***	0.040 (0.013)***	0.046 (0.013)***	0.066 (0.009)***
Age dummy (18-24)	0.082 (0.010)***	0.032 (0.024)	0.082 (0.010)***	0.059 (0.012)***
Age dummy (25-30)	0.046 (0.008)***	-0.008 (0.022)	0.045 (0.008)***	0.035 (0.009)***
Age dummy (31-35)	-0.018 (0.008)**	-0.069 (0.022)***	-0.018 (0.008)**	-0.023 (0.009)***
Black*Age (18-24)		0.059 (0.026)**		
Black*Age (25-30)		0.061 (0.024)***		
Black*Age (31-35)		0.058 (0.023)**		
Black*Matric			0.050 (0.019)***	
Matric*Age (18-24)				0.085 (0.022)***
Matric*Age (25-30)				0.049 (0.017)***
Matric*Age (31-35)				0.031 (0.018)*
Intercept	-0.179 (0.011)***	-0.154 (0.014)***	-0.163 (0.013)***	-0.170 (0.012)***
Delta	0.459 (0.094)***	0.496 (0.095)***	0.481 (0.095)***	0.452 (0.094)***
N	38,493	38,493	38,493	38,493

* p<0.1; ** p<0.05; *** p<0.01

In model 4b, we interact race with age. The parameter estimate on the black variable remains positive and statistically significant but is smaller in magnitude. The coefficients on the age dummies for ages 18 to 30 are now smaller and statistically insignificant. This suggests that black men drive the larger variation in worker productivity found in Table 4 for these young

[8] Table 5 only reports the parameter estimates for the worker productivity variance. The full set of results are to be found in Table A8 of the appendix.

workers. In model 4c, we interact race with the completed secondary schooling (i.e. matric) dummy. This too reveals that black men drive the variation in productivity for men with completed secondary schooling. Model 4d interacts the completed secondary schooling dummy with the age dummies. This reveals that the greater variation in productivity for men with completed secondary schooling is confined to younger men. These results point to young black men graduating from secondary schooling more recently as the most disadvantaged. This is consistent with this group receiving lower and more variable quality of schooling. Consequently, there is greater uncertainty about these workers and this could explain why unemployment is largely concentrated amongst this group.

7. Conclusion

This paper provides a theoretical and empirical investigation of the impact of information asymmetries between employer and job-seekers on the average wages of black and white South African men. The theoretical contribution of the paper illustrated the importance of the variation worker productivity as a source of uncertainty. The current literature on statistical discrimination almost exclusively focuses on the variation in the signal of worker productivity. Variation in the signal of productivity is important, but in the context of South Africa, variation in worker productivity is arguably a bigger concern. This is due to the immense variation learner performance and school quality that makes the task of predicting the potential productivity of job applicants insurmountable for many employers.

The theoretical model was tested with South African labour market data using reduced form and structural estimation techniques. The reduced form estimates revealed that black men, men in their youth, and men with only completed secondary schooling have more rapid within-firm wage growth in the first year of an employment spell. This was interpreted as evidence of the greater *ex ante* uncertainty regarding the expected productivity of these workers being resolved as employers view their performance on the job.

With the structural estimation, we were able to provide empirical support for the employer risk-aversion assumption. This is a key assumption for many statistical discrimination models and our results suggest that the strong criticism levelled against this assumption is not supported by the data. The parameter estimation revealed greater variation in productivity for black, young, and men with 12 and less years of schooling. This variation in productivity was however, largely driven by young black men who have recently graduated from secondary schooling. By holding the variation in the signal in productivity constant, we were able to demonstrate that the variation in worker productivity plays an important role in the uncertainty faced by employers when they have to assess the expected productivity of job applicants.

The South African literature and policy debate on earnings differences between black and white men have largely centred on addressing human capital differences. We argue that this is only achievable in the very long-term. In the short-term, addressing other causes of earnings differences like uncertainty in worker productivity may be more fruitful. The findings of this paper provide the necessary empirical evidence to guide such short-term policymaking. The short-term policies could, *inta alia*, include general skills assessment and certification by the Department of Labour’s labour centres.

There are however, some important shortcomings of our theoretical model. Our model is silent on unemployment. We essentially assume a two-state environment where unemployment as a separate state is not modelled. Unemployment is extremely high in South Africa with distribution that is skewed. We speculate that our model provides an additional reason for why we observe such high unemployment amongst young black men. Our model is also silent on the impact of other labour market rigidities like minimum wages and centralised wage bargaining. These shortcomings will serve as useful avenues for future work.

8. References

Aigner, D.J. and Cain, G.G. (1977) Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review*, 30(2): 195-187.

Altonji, J.G. and Blank, R.M. (1999) Race and gender in the labor market. In Ashenfelter, O. and Card, D. (eds.), Handbook of Labor Economics, 3C: 3143-3259. Amsterdam: Elsevier Science.

Altonji, J.G. and Pierret, C.R. (2001) Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1): 313-350.

Altonji, J.G. and Shakotko, R.A. (1987) Do wages rise with job seniority? *Review of Economic Studies*, 54: 437-459.

Arrow, K. (1973) The theory of discrimination. In Ashenfelter, O.A. and Rees, A. (eds.), *Discrimination in labor markets*, 3-33. New Jersey: Princeton University Press.

Branson, N., Ardington, C., Lam, D. and Leibbrandt, M. (2013) Changes in education, employment and earnings in South Africa – A cohort analysis. University of Cape Town, *SALDRU Working Paper*, 105.

Branson, N., Garlick, J., Lam, D. and Leibbrandt, M. (2012) Education and inequality: The South African case. University of Cape Town, *SALDRU Working Paper*, 75.

Branson, N. and Leibbrandt, M. (2013) Education quality and labour market outcomes in South Africa. *OECD Economics Department Working Papers*, 1021.

Bratsberg, B. and Terrell, D. (1998) Experience, tenure, and wage growth of young black and white men. *The Journal of Human Resources*, 33(3): 658-682.

Burger, R.P. and Jafta, R. (2006) Returns to race: Labour market discrimination in post-apartheid South Africa. *Stellenbosch University, Department of Economics Working Papers*, 04/2006.

Duff, P. and Fryer, D. (2005) Market failure, human capital, and job search dynamics in South Africa: The case of Duncan village. *DPRU Working Paper*, 05/98.

Erichsen, G. and Wakeford, J. (2001) Racial wage discrimination in SA before and after the first democratic election. University of Cape Town, University of Cape Town, *DPRU Working Papers*, 01/49.

Harris, M. and Holstrom, B. (1982) A theory of wage dynamics. *Review of Economic Studies*, 49(3): 315-333.

Jovanovic, B. (1979) Job matching and the theory of turnover. *Journal of Political Economy*, 87(5): 972-990.

Kerr, A. and Teal, F.J. (2015) The determinants of earnings inequalities: Panel data evidence from KwaZulu-Natal, South Africa. *Journal of African Economies,* 1-29.

Keswell, M. and Poswell, L. (2004) Returns to education in South Africa: A retrospective sensitivity analysis of the available evidence. *South African Journal of Economics*, 72 (4):834-860.

Lange, F. (2007) The speed of employer learning. *Journal of Labor Economics*, 25(1): 1-35.

Leibbrandt, M., Finn, A., and Woolard, I. (2012) Describing and decomposing post-apartheid income inequality in South Africa. *Development Southern Africa*, 29(1): 19-34.

Leibbrandt, M., Woolard, I., Finn, A., and Argent, J. (2010) Trends in South African income distribution and poverty since the fall of apartheid. *OECD Social, Employment and Migration Working Papers*, 101.

Levinsohn, J. (2007) *Two policies to alleviate unemployment in South Africa*. Mimeo. University of Michigan.

Lewis, D. and Terrell, D. (2001) Experience, tenure, and the perceptions of employers. *Southern Economic Journal*, 67(3): 578-597.

Lundberg, S.J. and Startz, R. (1983) Private discrimination and social intervention in competitive labor markets. *American Economic Review*, 73: 340-347.

Mwabu, G. and Schultz, T.P. (2000) Wage premiums for education and location of South African workers, by gender and race. *Economic Development and Cultural Change*, 48(2): 307-334.

Ntuli, M. and Kwenda, P. (2014) Labour unions and wage inequality among African men in South Africa. *Development Southern Africa*, 31(2): 322-346.

Oettinger, G. (1996) Statistical discrimination and the early career evolution of the black-white wage gap. *Journal of Labor Economics*, 14(1): 52-78.

Phelps, E.S. (1972) The Statistical Theory of Racism and Sexism. *American Economic Review*, 62(4): 659-661;

Rospabe, S. (2002) How did labour market racial discrimination evolve after the end of apartheid? *The South African Journal of Economics*, 70(1): 185-217.

Salop, J. and Salop, S (1976) Self-selection and turnover in the labor market. *Quarterly Journal of Economics*, 90(4): 619-627.

Schoer, V., Rankin, N., and Roberts, G. (2014) Accessing the first job in a slack labour market: Job matching in South Africa. *Journal of International Development*, 26: 1-22.

Statistics South Africa (2006). *The South African Labour Force Panel Survey methodology document*. National Statistics System Division, Pretoria

Szelewicki, M. and Tyrowicz, J. (2009) Labour market racial discrimination in South Africa revisited. *University of Warsaw Faculty of Economic Sciences, Working Paper*, 08/2009.

Van der Berg, S. (2014) Inequality, poverty and prospects for redistribution. *Development Southern Africa*, 31(2): 197-218.

9. Appendix

This section provides three sets of information that relate to the main body of this research paper. Firstly, we provide mode detailed steps for the derivation of equation (5) and (6). Secondly, we provide tables that report the coefficient estimates of all variables included in the regressions we estimate in our empirical analysis section. Lastly, we provide tables that report the results of the exercises we conduct for robustness check. These results were discussed in the main text.

9.1 Derivation of equation (5) and (6) of the theoretical model

Equation (4) above gives the period 0 wage that a worker will earn:

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

Solving first for the first term of equation (4), we have:

$$E(y|\boldsymbol{s}, \hat{y}_0) = \alpha + \boldsymbol{\theta s} + \sigma_u E(u|\boldsymbol{s}, \hat{y}_0)$$

$$= \hat{y}_0 - \sigma_e E(e_0|\boldsymbol{s}, \hat{y}_0)$$

Since employers know $\boldsymbol{s}$ and $\widehat{y_0}$ then:

$$E(e_0|\boldsymbol{s}, \hat{y}_0) = E(e_0|\hat{y}_0 - \alpha - \boldsymbol{\theta s}) = E(e_0|u\sigma_u + e_0\sigma_e)$$

In working out $E(e_0|u\sigma_u + e_0\sigma_e)$ it is useful to remember that it attempts to answer the question: suppose you observe $u\sigma_u + e_0\sigma_e$, what is your expected value of e_0. The answer needs to be expressed in terms of $u\sigma_u + e_0\sigma_e$, since this is what you observe. A good starting point is to assume that the answer is a linear function of $u\sigma_u + e_0\sigma_e$:

$$E(e_0|u\sigma_u + e_0\sigma_e) = a + b(u\sigma_u + e_0\sigma_e)$$

Where a and b are unknowns. $a = 0$, since we have assumed that u and e_0 are normally distributed with means of zero – i.e. they are centred at zero.

$$b = \frac{Cov(u\sigma_u + e_0\sigma_e, e_0)}{Var(u\sigma_u + e_0\sigma_e)}$$

$$b = \frac{\sigma_e}{\sigma_u^2 + \sigma_e^2}$$

Substituting back into $E(y|\boldsymbol{s}, \hat{y}_0) = \alpha + \boldsymbol{\theta s} + \sigma_u E(u|\boldsymbol{s}, \hat{y}_0)$, we get:

$$E(y|\boldsymbol{s}, \hat{y}_0) = \hat{y}_0 - \sigma_e\left(\frac{\sigma_e}{\sigma_u^2 + \sigma_e^2}\right)(u\sigma_u + e_0\sigma_e)$$

This simplifies to the following expression:

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

The last term of equation (4) is calculated as follows:

$$Var(y|\boldsymbol{s},\hat{y}_0) = E[y - E(y|\boldsymbol{s},\hat{y}_0)]^2$$

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

$$Var(y|\boldsymbol{s},\hat{y}_0) = \left(\frac{\sigma_e^2}{\sigma_u^2 + \sigma_e^2}\right)^2 \sigma_u^2 + \left(\frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}\right)^2 \sigma_e^2$$

Putting the simplified terms back together in equation (4) gives us equation (5):

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

Equation (6) is derived in the exact same manner as equation (5).

9.2 Full results tables and robustness check results

Table A1: Log hourly wage regression, pooled OLS

Variables	Model 1a	Model 1b	Model 1c	Model 1d
Primary	0.026 (0.003)***	0.026 (0.003)***	0.026 (0.004)***	0.026 (0.003)***
Secondary	0.075 (0.005)***	0.076 (0.005)***	0.077 (0.005)***	0.075 (0.005)***
Matric	0.151 (0.016)***	0.152 (0.016)***	0.153 (0.016)***	0.107 (0.028)***
Diploma&Certificate	0.231 (0.029)***	0.232 (0.029)***	0.231 (0.029)***	0.284 (0.044)***
Tertiary	0.146 (0.016)***	0.146 (0.016)***	0.148 (0.017)***	0.174 (0.024)***
Black	-0.749 (0.018)***	-0.816 (0.039)***	-0.749 (0.018)***	-0.749 (0.018)***
Tenure	0.036 (0.004)***	0.036 (0.004)***	0.037 (0.004)***	0.036 (0.004)***
Tenure^2	-0.001 (0.0002)***	-0.001 (0.0002)***	-0.001 (0.0002)***	-0.001 (0.0002)***
Potential Experience	0.023 (0.002)***	0.024 (0.002)***	0.026 (0.002)***	0.023 (0.002)***
Potential Experience^2	-0.0003 (0.00004)***	-0.0003 (0.00004)***	-0.0003 (0.00004)***	-0.0003 (0.00004)***
Tenure dummy (one year)	0.025 (0.018)	-0.046 (0.041)	-0.005 (0.024)	-0.032 (0.056)
Black*Tenure dummy		0.085 (0.040)**		
Age (18-24)*Tenure dummy			0.045 (0.031)	
Age (25-30)*Tenure dummy			0.050 (0.021)**	
Age (31-35)*Tenure dummy			0.011 (0.017)	
No Matric*Tenure dummy				0.044

				(0.057)
Matric*Tenure dummy				0.100
				(0.060)*
Occupation dummy 1	-0.005	-0.005	-0.006	-0.005
	(0.058)	(0.058)	(0.058)	(0.058)
Occupation dummy 2	-0.187	-0.189	-0.188	-0.187
	(0.036)***	(0.036)***	(0.036)***	(0.036)***
Occupation dummy 3	-0.456	-0.460	-0.456	-0.456
	(0.037)***	(0.037)***	(0.037)***	(0.037)***
Occupation dummy 4	-0.772	-0.774	-0.772	-0.770
	(0.036)***	(0.036)***	(0.036)***	(0.036)***
Occupation dummy 5	-0.513	-0.513	-0.513	-0.513
	(0.089)***	(0.089)***	(0.089)***	(0.089)***
Occupation dummy 6	-0.521	-0.523	-0.521	-0.521
	(0.033)***	(0.033)***	(0.033)***	(0.033)***
Occupation dummy 7	-0.592	-0.595	-0.591	-0.591
	(0.034)***	(0.034)***	(0.034)***	(0.034)***
Occupation dummy 8	-0.748	-0.749	-0.748	-0.747
	(0.035)***	(0.034)***	(0.035)***	(0.035)***
Industry dummy 1	0.834	0.833	0.837	0.835
	(0.020)***	(0.020)***	(0.020)***	(0.020)***
Industry dummy 2	0.664	0.664	0.665	0.664
	(0.021)***	(0.021)***	(0.021)***	(0.021)***
Industry dummy 3	0.841	0.841	0.840	0.839
	(0.053)***	(0.053)***	(0.053)***	(0.053)***
Industry dummy 4	0.509	0.510	0.509	0.509
	(0.024)***	(0.024)***	(0.024)***	(0.024)***
Industry dummy 5	0.512	0.512	0.514	0.513
	(0.021)***	(0.021)***	(0.021)***	(0.021)***
Industry dummy 6	0.624	0.625	0.626	0.624
	(0.029)***	(0.029)***	(0.028)***	(0.029)***
Industry dummy 7	0.579	0.578	0.580	0.579
	(0.026)***	(0.026)***	(0.026)***	(0.026)***
Industry dummy 8	0.631	0.631	0.632	0.633
	(0.032)***	(0.032)***	(0.032)***	(0.032)***
Industry dummy 9	-0.180	-0.183	-0.179	-0.177
	(0.085)**	(0.085)**	(0.085)**	(0.085)**
Industry dummy 10	0.991	0.994	0.988	0.991
	(0.202)***	(0.203)***	(0.201)***	(0.201)***
Rural dummy	-0.132	-0.133	-0.132	-0.132
	(0.014)***	(0.014)***	(0.014)***	(0.014)***
Province 1	-0.234	-0.231	-0.235	-0.234
	(0.029)***	(0.028)***	(0.028)***	(0.028)***
Province 2	-0.137	-0.136	-0.138	-0.137
	(0.033)***	(0.033)***	(0.033)***	(0.033)***
Province 3	-0.355	-0.353	-0.355	-0.355
	(0.026)***	(0.026)***	(0.026)***	(0.026)***
Province 4	-0.125	-0.123	-0.126	-0.125
	(0.025)***	(0.025)***	(0.025)***	(0.025)***
Province 5	-0.152	-0.150	-0.153	-0.152
	(0.026)***	(0.026)***	(0.026)***	(0.026)***
Province 6	-0.025	-0.023	-0.025	-0.025

	(0.024)	(0.024)	(0.024)	(0.024)
Province 7	-0.146	-0.145	-0.148	-0.147
	(0.026)***	(0.026)***	(0.026)***	(0.026)***
Province 8	-0.375	-0.373	-0.376	-0.375
	(0.030)***	(0.030)***	(0.030)***	(0.030)***
Household Head	0.127	0.127	0.127	0.127
	(0.014)***	(0.014)***	(0.014)***	(0.014)***
Firm size	0.076	0.076	0.076	0.076
	(0.004)***	(0.004)***	(0.004)***	(0.004)***
Wave dummy 1	-0.012	-0.011	-0.011	-0.012
	(0.016)	(0.016)	(0.016)	(0.016)
Wave dummy 2	-0.062	-0.062	-0.062	-0.062
	(0.017)***	(0.017)***	(0.017)***	(0.017)***
Wave dummy 3	-0.051	-0.051	-0.051	-0.051
	(0.017)***	(0.017)***	(0.017)***	(0.017)***
Wave dummy 4	0.027	0.027	0.027	0.027
	(0.018)	(0.018)	(0.018)	(0.018)
Wave dummy 5	0.027	0.027	0.027	0.027
	(0.017)	(0.018)	(0.017)	(0.017)
Intercept	1.435	1.488	1.392	1.446
	(0.057)***	(0.065)***	(0.062)***	(0.057)***
R^2	0.64	0.64	0.64	0.64
N	27,118	27,118	27,118	27,118

* p<0.1; ** p<0.05; *** p<0.01

Table A2: Robustness check – 2-year tenure dummy

Variables	Model 2a	Model 2b	Model 2c	Model 2d
Primary	0.026	0.026	0.026	0.026
	(0.003)***	(0.003)***	(0.004)***	(0.003)***
Secondary	0.075	0.076	0.076	0.075
	(0.005)***	(0.005)***	(0.005)***	(0.005)***
Matric	0.151	0.152	0.152	0.120
	(0.016)***	(0.016)***	(0.016)***	(0.022)***
Diploma&Certificate	0.231	0.233	0.231	0.263
	(0.029)***	(0.029)***	(0.029)***	(0.036)***
Tertiary	0.146	0.146	0.147	0.163
	(0.016)***	(0.016)***	(0.016)***	(0.020)***
Black	-0.748	-0.801	-0.749	-0.748
	(0.018)***	(0.029)***	(0.018)***	(0.018)***
Tenure	0.036	0.036	0.037	0.036
	(0.005)***	(0.005)***	(0.005)***	(0.005)***
Tenure^^{2}	-0.001	-0.001	-0.001	-0.001
	(0.0002)***	(0.0002)***	(0.0002)***	(0.0002)***
Potential Experience	0.023	0.024	0.025	0.023
	(0.002)***	(0.002)***	(0.002)***	(0.002)***
Potential Experience^^{2}	-0.0003	-0.0003	-0.0003	-0.0003
	(0.00004)***	(0.00004)***	(0.00004)***	(0.00004)***
Tenure dummy (2 years)	0.019	-0.047	-0.003	-0.018
	(0.020)	(0.035)	(0.024)	(0.047)
Black*Tenure dummy		0.078		
		(0.032)**		

Age (18-24)* Tenure dummy			0.010 (0.033)	
Age (25-30)* Tenure dummy			0.044 (0.021)**	
Age (31-35)* Tenure dummy			0.002 (0.017)	
No Matric* Tenure dummy				0.023 (0.046)
Matric*Tenure dummy				0.072 (0.048)
Intercept	1.443 (0.057)***	1.483 (0.060)***	1.420 (0.059)***	1.453 (0.057)***
R^2	0.64	0.64	0.64	0.64
N	27,118	27,118	27,118	27,118

Other control variables: occupation, industry, rural/urban classification, province of residence, household head status, firm size, and wave dummies.
Robust standard errors are contained in parentheses.
** p<0.1; ** p<0.05; *** p<0.01*

Table A3: Robustness check – 6-month tenure dummy

	Model 3a	Model 3b	Model 3c	Model 3d
Primary	0.026 (0.003)***	0.026 (0.003)***	0.026 (0.004)***	0.026 (0.003)***
Secondary	0.075 (0.005)***	0.076 (0.005)***	0.076 (0.005)***	0.075 (0.005)***
Matric	0.152 (0.016)***	0.151 (0.016)***	0.152 (0.016)***	0.140 (0.037)***
Diploma&Certificate	0.231 (0.029)***	0.231 (0.029)***	0.231 (0.029)***	0.295 (0.054)***
Tertiary	0.146 (0.017)***	0.146 (0.017)***	0.147 (0.017)***	0.178 (0.029)***
Black	-0.749 (0.018)***	-0.846 (0.047)***	-0.750 (0.018)***	-0.749 (0.018)***
Tenure	0.038 (0.003)***	0.038 (0.003)***	0.038 (0.003)***	0.038 (0.003)***
Tenure^2	-0.001 (0.0002)***	-0.001 (0.0002)***	-0.001 (0.0002)***	-0.001 (0.0002)***
Potential Experience	0.023 (0.002)***	0.024 (0.002)***	0.025 (0.002)***	0.023 (0.002)***
Potential Experience^2	-0.0003 (0.00004)***	-0.0003 (0.00004)***	-0.0003 (0.00004)***	-0.0003 (0.00004)***
Tenure dummy (6 months)	0.023 (0.019)	-0.068 (0.047)	0.008 (0.025)	-0.067 (0.071)
Black*Tenure dummy		0.109 (0.049)**		
Age (18-24)*Tenure dummy			0.015 (0.031)	
Age (25-30)*Tenure dummy			0.030 (0.021)	
Age (31-35)*Tenure dummy			0.006 (0.017)	

No Matric* Tenure dummy				0.093 (0.073)
Matric* Tenure dummy				0.106 (0.077)
Intercept	1.429 (0.058)***	1.508 (0.068)***	1.405 (0.064)***	1.425 (0.059)***
R^2	0.64	0.64	0.64	0.64
N	27,118	27,118	27,118	27,118

Other control variables: occupation, industry, rural/urban classification, province of residence, household head status, firm size, and wave dummies.
Robust standard errors are contained in parentheses.
** p<0.1; ** p<0.05; *** p<0.01*

Table A4: Log hourly wage regression, Heckman Sample Selection Model

	Model 4a	Model 4b	Model 4c	Model 4d
Primary	0.063 (0.003)***	0.063 (0.003)***	0.064 (0.003)***	0.064 (0.003)***
Secondary	0.120 (0.005)***	0.120 (0.005)***	0.120 (0.005)***	0.120 (0.005)***
Matric	0.187 (0.019)***	0.189 (0.019)***	0.187 (0.019)***	0.097 (0.033)***
Diploma&Certificate	0.332 (0.023)***	0.332 (0.023)***	0.332 (0.023)***	0.420 (0.039)***
Tertiary	0.270 (0.011)***	0.270 (0.011)***	0.271 (0.011)***	0.314 (0.019)***
Black	-0.712 (0.018)***	-0.819 (0.040)***	-0.711 (0.018)***	-0.711 (0.018)***
Tenure	0.049 (0.002)***	0.049 (0.002)***	0.049 (0.002)***	0.050 (0.002)***
Tenure^2	-0.001 (0.0001)***	-0.001 (0.0001)***	-0.001 (0.0001)***	-0.001 (0.0001)***
Potential Experience	0.024 (0.004)***	0.025 (0.004)***	0.024 (0.004)***	0.024 (0.004)***
Potential Experience^2	-0.0002 (0.0001)***	-0.0002 (0.0001)***	-0.0002 (0.0001)***	-0.0002 (0.0001)***
Tenure dummy (one year)	0.025 (0.017)	-0.087 (0.041)**	0.014 (0.022)	-0.043 (0.048)
Black*Tenure dummy		0.125 (0.042)***		
Age (18-24)*Tenure dummy			0.040 (0.033)	
Age (25-30)*Tenure dummy			-0.006 (0.022)	
Age (31-35)*Tenure dummy			0.031 (0.018)*	
No Matric*Tenure dummy				0.046 (0.050)
Matric*Tenure dummy				0.155 (0.054)***
Intercept	0.802 (0.092)***	0.882 (0.096)***	0.795 (0.095)***	0.821 (0.093)***
Mills (lambda)	-0.176	-0.168	-0.180	-0.175

	(0.050)***	(0.050)***	(0.050)***	(0.050)***
N	49,581	49,581	49,581	49,581

* *p*<0.1; ** *p*<0.05; *** *p*<0.01

Table A5: Written contract of employment regression, pooled OLS

	Model 5a	Model 5b	Model 5c	Model 5d
Primary	0.009 (0.003)***	0.009 (0.003)***	0.010 (0.003)***	0.009 (0.003)***
Secondary	0.014 (0.004)***	0.014 (0.004)***	0.014 (0.004)***	0.014 (0.004)***
Matric	0.036 (0.011)***	0.036 (0.011)***	0.036 (0.011)***	0.074 (0.020)***
Diploma&Certificate	0.042 (0.014)***	0.043 (0.014)***	0.042 (0.014)***	0.072 (0.024)***
Tertiary	0.009 (0.008)	0.009 (0.008)	0.009 (0.008)	0.024 (0.012)**
Black	-0.051 (0.011)***	-0.128 (0.022)***	-0.051 (0.011)***	-0.051 (0.011)***
Tenure	0.017 (0.003)***	0.017 (0.003)***	0.017 (0.003)***	0.017 (0.003)***
Tenure^2	-0.001 (0.0001)***	-0.001 (0.0001)***	-0.001 (0.0001)***	-0.001 (0.0001)***
Potential Experience	0.005 (0.001)***	0.006 (0.001)***	0.006 (0.002)***	0.006 (0.001)***
Potential Experience^2	-0.0001 (0.00003)***	-0.0001 (0.00003)***	-0.0001 (0.00003)***	-0.0001 (0.00003)***
Tenure dummy (one year)	0.064 (0.012)***	-0.016 (0.023)	0.054 (0.015)***	-0.017 (0.028)
Black*Tenure dummy		0.096 (0.023)***		
Age (18-24)*Tenure dummy			0.004 (0.023)	
Age (25-30)*Tenure dummy			0.016 (0.014)	
Age (31-35)*Tenure dummy			0.022 (0.011)**	
No Matric*Tenure dummy				0.104 (0.029)***
Matric*Tenure dummy				0.055 (0.032)*
Occupation dummy 1	0.022 (0.027)	0.022 (0.027)	0.021 (0.027)	0.022 (0.027)
Occupation dummy 2	0.009 (0.018)	0.007 (0.018)	0.009 (0.018)	0.007 (0.018)
Occupation dummy 3	0.004 (0.021)	-0.000 (0.020)	0.003 (0.020)	0.002 (0.020)
Occupation dummy 4	-0.012 (0.019)	-0.014 (0.019)	-0.013 (0.019)	-0.015 (0.019)
Occupation dummy 5	-0.056 (0.046)	-0.057 (0.046)	-0.057 (0.046)	-0.058 (0.046)
Occupation dummy 6	-0.038	-0.040	-0.039	-0.040

	(0.018)**	(0.018)**	(0.018)**	(0.018)**
Occupation dummy 7	-0.033	-0.037	-0.034	-0.037
	(0.018)*	(0.018)**	(0.018)*	(0.018)**
Occupation dummy 8	-0.082	-0.083	-0.082	-0.084
	(0.019)***	(0.019)***	(0.019)***	(0.019)***
Industry dummy 1	0.249	0.249	0.250	0.248
	(0.014)***	(0.014)***	(0.014)***	(0.014)***
Industry dummy 2	0.139	0.139	0.139	0.139
	(0.015)***	(0.015)***	(0.015)***	(0.015)***
Industry dummy 3	0.185	0.186	0.185	0.185
	(0.026)***	(0.026)***	(0.026)***	(0.026)***
Industry dummy 4	0.023	0.024	0.023	0.024
	(0.019)	(0.019)	(0.019)	(0.019)
Industry dummy 5	0.096	0.095	0.096	0.095
	(0.015)***	(0.015)***	(0.015)***	(0.015)***
Industry dummy 6	0.054	0.054	0.053	0.054
	(0.019)***	(0.019)***	(0.019)***	(0.019)***
Industry dummy 7	0.191	0.189	0.190	0.189
	(0.017)***	(0.017)***	(0.017)***	(0.017)***
Industry dummy 8	0.119	0.118	0.119	0.118
	(0.022)***	(0.022)***	(0.022)***	(0.022)***
Industry dummy 9	0.490	0.486	0.491	0.488
	(0.045)***	(0.045)***	(0.045)***	(0.045)***
Industry dummy 10	0.082	0.085	0.083	0.082
	(0.079)	(0.079)	(0.078)	(0.079)
Rural dummy	-0.022	-0.022	-0.022	-0.022
	(0.009)**	(0.009)**	(0.009)**	(0.009)**
Province 1	-0.146	-0.143	-0.147	-0.145
	(0.020)***	(0.020)***	(0.020)***	(0.019)***
Province 2	0.081	0.082	0.081	0.082
	(0.021)***	(0.021)***	(0.021)***	(0.021)***
Province 3	0.063	0.065	0.063	0.064
	(0.017)***	(0.017)***	(0.017)***	(0.017)***
Province 4	-0.037	-0.035	-0.037	-0.036
	(0.017)**	(0.017)**	(0.017)**	(0.017)**
Province 5	0.001	0.003	0.001	0.002
	(0.017)	(0.017)	(0.017)	(0.017)
Province 6	-0.000	0.002	-0.000	0.001
	(0.016)	(0.016)	(0.016)	(0.016)
Province 7	0.072	0.074	0.071	0.073
	(0.017)***	(0.017)***	(0.017)***	(0.017)***
Province 8	-0.055	-0.052	-0.055	-0.053
	(0.020)***	(0.020)***	(0.020)***	(0.020)***
Household Head	0.024	0.024	0.022	0.024
	(0.009)***	(0.009)***	(0.009)**	(0.009)***
Wave dummy 1	0.099	0.099	0.098	0.099
	(0.011)***	(0.012)***	(0.011)***	(0.012)***
Wave dummy 2	0.095	0.096	0.095	0.096
	(0.012)***	(0.012)***	(0.012)***	(0.012)***
Wave dummy 3	0.130	0.130	0.129	0.130
	(0.012)***	(0.012)***	(0.012)***	(0.012)***
Wave dummy 4	0.150	0.150	0.149	0.150

	(0.012)***	(0.012)***	(0.012)***	(0.012)***
Wave dummy 5	0.194	0.194	0.194	0.194
	(0.012)***	(0.012)***	(0.012)***	(0.012)***
Firm size dummy 1	0.035	0.034	0.035	0.035
	(0.027)	(0.027)	(0.027)	(0.027)
Firm size dummy 2	0.121	0.120	0.120	0.121
	(0.027)***	(0.027)***	(0.027)***	(0.027)***
Firm size dummy 3	0.194	0.194	0.194	0.194
	(0.026)***	(0.026)***	(0.026)***	(0.026)***
Firm size dummy 4	0.226	0.227	0.226	0.227
	(0.026)***	(0.026)***	(0.026)***	(0.026)***
Firm size dummy 5	0.273	0.273	0.273	0.274
	(0.025)***	(0.025)***	(0.025)***	(0.025)***
Intercept	0.049	0.110	0.042	0.028
	(0.041)	(0.044)**	(0.044)	(0.042)
R^2	0.19	0.19	0.19	0.19
N	26,716	26,716	26,716	26,716

* p<0.1; ** p<0.05; *** p<0.01

Table A6: Permanent employment regression, pooled OLS

	Model 6a	Model 6b	Model 6c	Model 6d
Primary	0.003	0.003	0.003	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Secondary	0.006	0.006	0.006	0.006
	(0.003)*	(0.003)*	(0.003)*	(0.003)*
Matric	0.061	0.061	0.061	0.124
	(0.010)***	(0.010)***	(0.010)***	(0.020)***
Diploma&Certificate	0.028	0.029	0.027	0.083
	(0.011)**	(0.011)***	(0.011)**	(0.021)***
Tertiary	0.002	0.002	0.002	0.031
	(0.005)	(0.005)	(0.005)	(0.011)***
Black	-0.091	-0.261	-0.092	-0.091
	(0.009)***	(0.021)***	(0.009)***	(0.009)***
Tenure	0.044	0.044	0.043	0.043
	(0.002)***	(0.002)***	(0.002)***	(0.002)***
Tenure^2	-0.002	-0.002	-0.002	-0.002
	(0.0001)***	(0.0001)***	(0.0001)***	(0.0001)***
Potential Experience	0.003	0.004	0.004	0.004
	(0.001)***	(0.001)***	(0.001)***	(0.001)***
Potential Experience^2	-0.00004	-0.0001	-0.00004	-0.0001
	(0.00002)*	(0.00002)**	(0.00003)	(0.00002)**
Tenure dummy (one year)	0.136	-0.042	0.119	-0.010
	(0.012)***	(0.020)**	(0.015)***	(0.027)
Black*Tenure dummy		0.213		
		(0.021)***		
Age (18-24)*Tenure dummy			0.008	
			(0.020)	
Age (25-30)*Tenure dummy			0.027	
			(0.013)**	
Age (31-35)*Tenure dummy			0.030	

			(0.010)***	
No Matric*Tenure dummy				0.185
				(0.028)***
Matric*Tenure dummy				0.104
				(0.031)***
Occupation dummy 1	-0.037	-0.037	-0.038	-0.037
	(0.019)**	(0.018)**	(0.019)**	(0.018)**
Occupation dummy 2	-0.033	-0.037	-0.033	-0.036
	(0.013)***	(0.012)***	(0.013)***	(0.012)***
Occupation dummy 3	-0.023	-0.032	-0.024	-0.028
	(0.014)*	(0.013)**	(0.014)*	(0.013)**
Occupation dummy 4	-0.043	-0.048	-0.044	-0.047
	(0.013)***	(0.013)***	(0.013)***	(0.013)***
Occupation dummy 5	-0.142	-0.143	-0.143	-0.145
	(0.040)***	(0.040)***	(0.040)***	(0.040)***
Occupation dummy 6	-0.062	-0.067	-0.063	-0.066
	(0.012)***	(0.012)***	(0.012)***	(0.012)***
Occupation dummy 7	-0.040	-0.048	-0.041	-0.047
	(0.012)***	(0.012)***	(0.012)***	(0.012)***
Occupation dummy 8	-0.109	-0.112	-0.110	-0.112
	(0.013)***	(0.013)***	(0.013)***	(0.013)***
Industry dummy 1	0.015	0.014	0.017	0.014
	(0.013)	(0.012)	(0.013)	(0.012)
Industry dummy 2	-0.051	-0.049	-0.050	-0.050
	(0.013)***	(0.013)***	(0.013)***	(0.013)***
Industry dummy 3	-0.050	-0.048	-0.049	-0.049
	(0.025)*	(0.025)*	(0.025)*	(0.025)**
Industry dummy 4	-0.223	-0.220	-0.223	-0.221
	(0.017)***	(0.017)***	(0.017)***	(0.017)***
Industry dummy 5	-0.047	-0.049	-0.047	-0.048
	(0.014)***	(0.013)***	(0.014)***	(0.013)***
Industry dummy 6	-0.080	-0.078	-0.080	-0.079
	(0.016)***	(0.016)***	(0.016)***	(0.016)***
Industry dummy 7	0.000	-0.002	-0.000	-0.002
	(0.015)	(0.015)	(0.015)	(0.015)
Industry dummy 8	-0.049	-0.050	-0.049	-0.049
	(0.018)***	(0.018)***	(0.018)***	(0.018)***
Industry dummy 9	0.175	0.166	0.176	0.170
	(0.040)***	(0.041)***	(0.040)***	(0.041)***
Industry dummy 10	-0.057	-0.051	-0.056	-0.057
	(0.055)	(0.057)	(0.055)	(0.056)
Rural dummy	0.008	0.007	0.008	0.008
	(0.009)	(0.009)	(0.009)	(0.009)
Province 1	-0.022	-0.014	-0.022	-0.019
	(0.017)	(0.017)	(0.017)	(0.017)
Province 2	0.000	0.002	-0.001	0.001
	(0.019)	(0.019)	(0.019)	(0.019)
Province 3	0.002	0.007	0.002	0.004
	(0.015)	(0.015)	(0.015)	(0.015)
Province 4	-0.091	-0.086	-0.092	-0.089
	(0.015)***	(0.015)***	(0.015)***	(0.015)***
Province 5	0.015	0.020	0.015	0.017

	(0.016)	(0.016)	(0.016)	(0.016)
Province 6	-0.000	0.004	-0.001	0.001
	(0.014)	(0.014)	(0.014)	(0.014)
Province 7	-0.029	-0.024	-0.030	-0.025
	(0.016)*	(0.016)	(0.016)*	(0.016)
Province 8	-0.049	-0.042	-0.049	-0.045
	(0.018)***	(0.018)**	(0.018)***	(0.018)***
Household Head	0.041	0.041	0.039	0.041
	(0.008)***	(0.008)***	(0.008)***	(0.008)***
Wave dummy 1	0.023	0.024	0.023	0.024
	(0.010)**	(0.010)**	(0.010)**	(0.010)**
Wave dummy 2	0.012	0.012	0.011	0.012
	(0.010)	(0.010)	(0.010)	(0.010)
Wave dummy 3	0.012	0.013	0.011	0.012
	(0.011)	(0.010)	(0.010)	(0.010)
Wave dummy 4	0.033	0.032	0.032	0.033
	(0.010)***	(0.010)***	(0.010)***	(0.010)***
Wave dummy 5	0.022	0.023	0.022	0.023
	(0.010)**	(0.010)**	(0.010)**	(0.010)**
Firm size dummy 1	0.039	0.038	0.039	0.040
	(0.026)	(0.026)	(0.026)	(0.026)
Firm size dummy 2	0.072	0.071	0.072	0.072
	(0.025)***	(0.025)***	(0.025)***	(0.025)***
Firm size dummy 3	0.079	0.079	0.079	0.079
	(0.024)***	(0.024)***	(0.024)***	(0.024)***
Firm size dummy 4	0.094	0.094	0.094	0.095
	(0.024)***	(0.024)***	(0.024)***	(0.024)***
Firm size dummy 5	0.086	0.087	0.086	0.088
	(0.024)***	(0.024)***	(0.024)***	(0.024)***
Intercept	0.486	0.622	0.473	0.451
	(0.037)***	(0.040)***	(0.039)***	(0.037)***
R^2	0.23	0.24	0.23	0.23
N	27,020	27,020	27,020	27,020

* p<0.1; ** p<0.05; *** p<0.01

Table A7: Structural estimation of the theoretical model, maximum likelihood estimation

	Model 7a	Model 7b	Model 7c
	$\sigma_e^2 = 1$	$\sigma_e^2 = 0.5$	$\sigma_e^2 = 2$
Log hourly wage			
Primary	0.071	0.071	0.069
	(0.003)***	(0.003)**	(0.003)**
Secondary	0.153	0.156	0.149
	(0.004)***	(0.004)**	(0.004)**
Matric	0.274	0.275	0.266
	(0.013)***	(0.013)**	(0.014)**
Diploma&Certificate	0.307	0.306	0.305
	(0.017)***	(0.017)**	(0.018)**
Tertiary	0.307	0.301	0.313
	(0.008)***	(0.008)**	(0.008)**
Potential Experience	0.048	0.048	0.044

	(0.002)***	(0.001)**	(0.002)**
Potential Experience^2	-0.001	-0.001	-0.001
	(0.00003)***	(0.00003)**	(0.00003)**
Black	-0.877	-0.867	-0.874
	(0.013)***	(0.012)**	(0.014)**
Tenure	0.039	0.051	0.006
	(0.006)***	(0.005)**	(0.008)
Tenure^2	-0.0005	-0.001	0.001
	(0.0003)*	(0.0002)**	(0.0003)*
Intercept	0.980	0.869	1.377
	(0.048)***	(0.038)**	(0.075)**
Productivity variance			
No Matric dummy	0.028	0.025	0.038
	(0.010)***	(0.012)*	(0.009)**
Matric dummy	0.022	0.024	0.028
	(0.011)**	(0.012)	(0.009)**
Black	0.070	0.078	0.067
	(0.009)***	(0.011)**	(0.008)**
Age spline (18-24)	0.082	0.043	0.119
	(0.010)***	(0.011)**	(0.008)**
Age spline (25-30)	0.046	0.031	0.062
	(0.008)***	(0.009)**	(0.006)**
Age spline (31-35)	-0.018	-0.031	-0.004
	(0.008)**	(0.009)**	(0.006)
Intercept	-0.179	-0.297	0.012
	(0.011)***	(0.013)**	(0.010)
Delta	0.459	0.660	0.555
	(0.094)***	(0.183)**	(0.069)**
N	38,493	38,493	38,493

* p<0.1; ** p<0.05; *** p<0.01

Table A8: Structural estimation of the theoretical model, maximum likelihood estimation

	Model 8a	Model 8b	Model 8c	Model 8d
Log hourly wage				
Primary	0.071	0.071	0.071	0.071
	(0.003)***	(0.003)***	(0.003)***	(0.003)***
Secondary	0.153	0.153	0.153	0.154
	(0.004)***	(0.004)***	(0.004)***	(0.004)***
Matric	0.274	0.273	0.277	0.276
	(0.013)***	(0.013)***	(0.014)***	(0.013)***
Diploma&Certificate	0.307	0.307	0.308	0.307
	(0.017)***	(0.017)***	(0.017)***	(0.017)***
Tertiary	0.307	0.308	0.310	0.308
	(0.008)***	(0.008)***	(0.008)***	(0.008)***
Potential Experience	0.048	0.048	0.048	0.048
	(0.002)***	(0.002)***	(0.002)***	(0.002)***
Potential Experience^2	-0.001	-0.001	-0.001	-0.001
	(0.00003)***	(0.00003)***	(0.00003)***	(0.00003)***
Black	-0.877	-0.881	-0.876	-0.876
	(0.013)***	(0.013)***	(0.013)***	(0.013)***
Tenure	0.039	0.037	0.038	0.039

	(0.006)***	(0.006)***	(0.006)***	(0.006)***
Tenure^2	-0.0005	-0.0004	-0.0004	-0.0005
	(0.0003)*	(0.0003)	(0.0003)	(0.0003)*
Intercept	0.980	1.000	0.989	0.974
	(0.048)***	(0.048)***	(0.048)***	(0.048)***
Productivity variance				
No Matric dummy	0.028	0.030	0.036	0.028
	(0.010)***	(0.010)***	(0.011)***	(0.010)***
Matric dummy	0.022	0.023	-0.013	-0.010
	(0.011)**	(0.011)**	(0.017)	(0.014)
Black	0.070	0.040	0.046	0.066
	(0.009)***	(0.013)***	(0.013)***	(0.009)***
Age dummy (18-24)	0.082	0.032	0.082	0.059
	(0.010)***	(0.024)	(0.010)***	(0.012)***
Age dummy (25-30)	0.046	-0.008	0.045	0.035
	(0.008)***	(0.022)	(0.008)***	(0.009)***
Age dummy (31-35)	-0.018	-0.069	-0.018	-0.023
	(0.008)**	(0.022)***	(0.008)**	(0.009)***
Black*Age (18-24)		0.059		0.085
		(0.026)**		
Black*Age (25-30)		0.061		
		(0.024)***		
Black*Age (31-35)		0.058		
		(0.023)**		
Black*Matric			0.050	
			(0.019)***	
Matric*Age (18-24)				(0.022)***
				0.049
Matric*Age (25-30)				(0.017)***
				0.031
Matric*Age (31-35)				(0.018)*
				(0.022)***
Intercept	-0.179	-0.154	-0.163	-0.170
	(0.011)***	(0.014)***	(0.013)***	(0.012)***
Delta	0.459	0.496	0.481	0.452
	(0.094)***	(0.095)***	(0.095)***	(0.094)***
N	38,493	38,493	38,493	38,493

* p<0.1; ** p<0.05; *** p<0.01