



2000 Annual Forum

at Glenburn Lodge, Muldersdrift

The Incidence of Unemployment in South Africa

Geeta Kingdon & John Knight

Centre for the Study of African Economies
Economics Department
University of Oxford

The Incidence of Unemployment in South Africa

by Geeta Kingdon and John Knight

September, 2000

I. Introduction

The unemployment rate in South Africa is one of the highest in the world - 39% in 1998 by the broad definition. Even according to the new official (narrow) definition, one in every four adults (26%) who wanted work and actively looked for it was unemployed. Moreover, the unemployment rates for different groups reveal great disparity in the incidence of unemployment by race, gender, age, education, and region. Given the importance of employment income in total household income in South Africa (Leibbrandt, Borat, and Woolard, 2000), the varying incidence of unemployment across different groups has important implications for the distribution of income and for the incidence of poverty.

In this paper we firstly paint a picture of the distribution of unemployment in South Africa, asking the question 'who are the unemployed?' and identifying the characteristics that make a person more likely to be unemployed. This is done by means of both descriptive statistics and the estimation of a probit equation of unemployment. The probit permits us to measure the influence of a given factor or characteristic on unemployment probability holding other factors constant. Secondly, we focus on the racial distribution of unemployment, exploring the extent to which the race gap in the probability of employment is due to the black group's inferior employment-enhancing characteristics and to employment discrimination in the labour market. While there is research in South Africa investigating racial wage discrimination (Knight and McGrath, 1977 and 1987; Rospabe, 1997; Jensen, 1999; Moll, 2000), a fuller picture of how the different races fare in the labour market needs to take account of employment discrimination as well, especially since access to employment is a strong predictor of income. Knight and McGrath (1977, 1987) showed that job discrimination was an important source of racial wage differences, but no estimates have as yet been made of the extent of employment discrimination.

Section II describes the data and Section III explores the distribution of unemployment across different groups, separating the incidence of unemployment into its two components - entry into unemployment and duration of unemployment. Section IV presents unemployment probits and Section V

decomposes the race gap in the probability of unemployment into the part explained by the differing characteristics of whites and other race groups and the unexplained residual. The final section concludes.

II. Data

We use two cross-section datasets: the October Household Survey of 1994 carried out by the Central Statistical Office, now known as Statistics South Africa (or simply StatsSA), and the other derived from an integrated household survey carried out in 1993 by the South African Labour Research Unit (SALDRU). The October Household Survey of 1994 (OHS94) is a nationally representative household survey covering 33,000 households across 1,010 clusters in 266 districts. Sampling information is available from CSS(1995). The SALDRU survey is a nationally representative household survey patterned on the World Bank's Living Standards Measurement Study surveys. It yielded a dataset covering about 9000 households across 360 clusters. A detailed account of the sampling procedure is contained in SALDRU (1994).

Some important aspects of employment and unemployment have not been captured in the two surveys. For example, no information is available on the duration of unemployment in the SALDRU survey, and only limited information exists on the employment histories of those unemployed persons who actively looked for work in the week before the survey date. Although there is a question on duration of unemployment in the OHS94 questionnaire, this question is asked only of persons currently unemployed, so that completed duration of unemployment is not known for anyone.

Both datasets are cross-section rather than panels. This limitation restricts the analysis to obtaining a snapshot of a person's unemployment situation at a point in time rather than over a period of time. This is unfortunate since a number of important questions about South African unemployment cannot be investigated, such as: Is high unemployment mainly due to a high rate of entry into unemployment or due to its long duration (*i.e.* a low rate of exit from unemployment)? Is employment probability duration-dependent? What is the completed duration of unemployment? Do probabilities of transition into employment from out-of-labour-force and unemployed states differ significantly? On the positive side, however, the strengths of the datasets lie in their nationally representative character and in their large sample size which permits reliable analysis at a high level of disaggregation.

III. The incidence of unemployment

Two definitions of unemployment are commonly utilised - the broad and the narrow. The narrowly defined unemployed are those who are currently not employed but who looked for work in the seven days (SALDRU) or four weeks (StatsSA) prior to the survey visit. The broadly defined unemployed are the narrow unemployed plus those who say they want work but did not look for work in the past week (past four weeks).

As a number of recent studies have investigated the *extent* of unemployment in South Africa (Klasen and Woolard, 1999; StatsSA, 1998), only a short discussion of the major findings suffices here as a backdrop for our further analysis. Table 1 shows that the broadly measured unemployment rate in South Africa has risen from the already high figure of 31% in 1993 to 39% in 1998. Even the narrowly measured unemployment rate in the OHSs rose from 20% in 1994 to 26% over the four-year period to 1998. These rates are extremely high by international standards (ILO 1998). The great broad-narrow discrepancy in unemployment rates suggests that a large proportion of labour force participants are not actively looking for work. Some analysts would argue that such unemployed persons are not labour force participants¹. Elsewhere (Kingdon and Knight, 1999; Kingdon and Knight, 2000a) we discuss the appropriate definition of unemployment in South Africa and argue that the broad definition is the more relevant, essentially because the non-searching are 'discouraged' workers. The broad concept of unemployment is therefore the one that we analyse in this paper.

Table 2 shows the distribution of unemployment across groups. It shows that unemployment varies dramatically by race. Africans face unemployment rates of 41% but the rate for whites is only 6%. Unemployment decreases monotonically by age, ranging from 51% for the youngest group to 17% for the eldest group. The incidence of unemployment also varies importantly by region, gender, and education. For example, people with higher education face an unemployment rate of 6% but those with primary education or less suffer an unemployment rate close to 40%. Women experience substantially higher unemployment than men. Rural unemployment rates are higher than urban rates, in contrast to the pattern in most developing countries. This is due to the segregation policies of the apartheid era which consigned

¹ For example, an ILO (1996) report argues that it is wrong to include all persons who report themselves as currently unemployed but who did not look for work in the past week/month in the belief that none was available, presumably because the lack of search is taken as an indication of lack of commitment to the labour market. It points out that in the October Household Survey of 1994, many such persons were in their 30s and 40s and had never been employed before,

millions of Africans to live in 'homelands' which were predominantly rural areas of poor land quality and little employment opportunity. These homelands effectively became labour reserves from which permanent and even temporary movement to non-homeland areas was impeded by legislative and administrative rules. Thus, high unemployment in much of rural South Africa took the form of waiting in the homelands for a formal sector job opportunity to arise outside.

Nickell (1980) suggests that unemployment incidence should be separated into two components: the chances of entering unemployment and the duration of time for which individuals remain unemployed. His argument is that these two components are determined in different ways and may be affected by different factors. We adopt this strategy for South Africa.

Firstly, consider entry into unemployment. Table 3 shows that entry into unemployment in South Africa is mainly dominated by those who have never previously held a job, *i.e.* who enter unemployment when they enter the labour force, rather than by persons who held a job and then became unemployed. Only 38% of all unemployed persons entered unemployment from the employed state. The fact that a majority of the unemployed have never held a job before is one of the most striking features of unemployment in South Africa. It is worth examining this issue more carefully.

The incidence of direct entry into unemployment (without an intervening period of work) varies by several factors. For example, it varies substantially by race. African unemployed persons are more than twice as likely as whites never to have had work. While this could be partly due to the inferior employment-enhancing characteristics of Africans *vis a vis* whites, it could also partly be due to racial discrimination in employers' hiring practices. Rural persons are more likely than urban never to have worked. This may be because there are fewer job-opportunities in rural than urban areas and because intensity of job-search is lower in the countryside owing to remoteness from centres of employment. Unemployed women are more likely than unemployed men never to have experienced a period in work. This may be due to women's lesser flexibility in terms of hours of work and the distance they are prepared to travel, or to their higher reservation wages, *ceteris paribus*, than men.

implicitly casting doubt on the notion that these were genuine labour force participants. We cannot tell from the SALDRU survey whether the non-searching unemployed had ever held a job before.

Age is an obvious important factor since the young are more likely to search rather than get 'locked-in' to an undesirable job. The young are also more able to afford unemployed job-search because they have fewer financial commitments than do older persons. Moreover, they may be more ignorant about what their skills can command in the labour market, *i.e.* may have higher reservation wages. It is more difficult to explain this phenomenon among the older age groups. For example, about 50% of all unemployed persons (61% of unemployed women and 37% of unemployed men) aged 36-64 had never worked before. This is either due to late entry into the labour force - a possible explanation for women who might delay entry until after child-bearing/rearing years - or to extraordinarily long duration of unemployment, which can be explained only by the lack of adequate jobs in the economy or by too narrow a concept of 'employment'. While the *current* labour market status of individuals - whether they are regarded as unemployed or not - is carefully determined in the October Household Surveys through a series of comprehensive questions (Bhorat, 1999), the simple single question on labour market history - *i.e.* whether the individual ever worked gainfully in the past - relied on the judgement of respondents some of whom might have been thinking of employment only as regular wage employment.

The validity of these explanations is confirmed in a multivariate context. We fitted a probit of 'ever worked before' in Table 3a. This shows that, holding other factors constant, Africans have a 35 percentage point lower probability than whites of having ever worked before for pay, profit, or family gain. Since the white probability of previous work is 70% (Table 3), the African probability is exactly half that of whites, after standardising for observed characteristics. The monotonic effect and significance of age is confirmed, as is the effect of gender. The probability of having ever worked gainfully varies importantly by whether the individual is a household head and married; this is as might be expected *a priori*. Homeland residence is associated with a 16-19 percentage point lower probability of previous work compared with non-homeland residence.

The chances of entry into unemployment from the employed state can be separated into voluntary and involuntary entry. The last column of Table 3 shows that, on average, less than a quarter of those unemployed persons who previously worked had quit work 'voluntarily' rather than because of sacking, retrenchment, illness, or end of temporary job². The fact that most people quit work involuntarily probably reflects low vacancy rates and high unemployment rates. However, there is considerable variation by

² Some caution is necessary when considering the separation of voluntary quitters from involuntary quitters. For example, a worker who knew her firm was going to fold shortly might quit 'voluntarily' before the event occurred.

group. For example, groups that are more likely to quit work voluntarily are the young, the highly educated, women, and whites.

The cost of voluntary quitting into unemployment is a function of the cost of being unemployed, which depends on the income in and out of work and on the level of one's financial commitments. Thus, for example, the young would have lower income-loss from voluntarily quitting into unemployment insofar as they are often supported by their families while unemployed. The benefits of voluntary quitting into unemployment depend on the prospect for alternative wage and job opportunities. Since the alternative wage opportunities are relatively better the lower the level of firm-specific human capital which the individual possesses, younger workers would be more likely to find that there are many firms willing to pay starting wages similar to their current earnings³. Workers liable to be sacked or made redundant include those whose productivity is on the wane but whose wages have not been commensurately reduced. Thus, we expect the old to have low incidence of voluntary quits, and this is what we observe in the final column of Table 3.

If there is scarcity of educated labour or racial discrimination by employers, more educated people, or persons belonging to the favoured racial group, will be more likely to quit voluntarily in search of better wage opportunities because their probability of re-employment is higher. This could explain why persons with higher education and whites have a higher propensity to quit voluntarily. Women are more than twice as likely as men to quit voluntarily into unemployment. Working women may quit voluntarily for child-bearing and rearing and - being usually the secondary income earners in the family - are also more likely than men to give up their work in case of family emergencies or migration of spouse.

Now consider the duration of unemployment - the second aspect of unemployment incidence. The length of time for which an individual remains unemployed depends both on the rate at which he receives job offers and on the extent to which these offers are accepted (Nickell, 1980). It is clear that most identifiable variables have an impact on both the demand and supply sides of the labour market. For example, for certain types of jobs, e.g. manual work, younger people may be more likely to receive job offers than older people - if they are seen as physically more capable. Younger people may also be more likely to accept job offers insofar as they are more flexible and have less stringent ideas about what is suitable employment.

The question from which we have obtained data on duration of unemployment was ‘How long has (name) been seeking work’. We interpret this to mean ‘how long has name been *wanting* work’ rather than ‘how long has name been *searching for* work’. This seems reasonable because the question was asked of all unemployed person and not only of those who had taken active steps to find work. The available information is from persons who are currently unemployed so that it represents *uncompleted* duration of an individual’s spell of unemployment.

The answers were recorded in categorised form (e.g. less than a month, 1-2 months, 2-6 months, 6-12 months, >12 months) rather than as a continuous variable - number of months. However, by assigning midpoints of the category, we have created a duration-of-unemployment variable ‘number of months’. It is unfortunate that the last category specified in the OHS94 duration question was ‘>12 months’ since it turns out that more than 67% of the unemployed were in this category and it seems possible that many of these suffered unemployment for much longer periods of time than a year - in other words, there is a great loss of information and of variability in the duration variable because of this truncation. However, the October Household Survey of 1997 (OHS97) includes more detailed information, in particular duration categories ‘1 to 3 years’ and ‘greater than 3 years’. In OHS97, of all unemployed persons whose duration was a year or greater, 43.7% had duration between 1 and 3 years and 56.3% of greater than 3 years. We ascribe these proportions to persons in the OHS94 category ‘greater than one year’ in the last row of Table 4. For the other rows, the corresponding proportions for the various groups are applied. The resultant measure of months of unemployment (column 2) shows very long average uncompleted duration: 27 months. Duration is seen to rise monotonically with age and to fall monotonically with education.

The duration of unemployment varies substantially by group in Table 4. It is considerably lower for the younger, well educated, and white groups than for their opposite numbers. This is consistent with the notion that, on the demand side of the labour market, groups that are deemed by employers to be more productive (younger and more educated) or more desirable (whites) will receive more job-offers. The younger group might also be more likely to accept the offers because of their greater adaptability than older people. Similarly, more educated people may have lower reservation wages (relative to their expected wages) than less well educated people because they are more knowledgeable about the worth of their

³ This explanation clearly applies more to the formal sector of the economy.

skills. Thus, some of these groups are also more likely to accept job-offers than their opposite numbers and are thus likely to quit unemployment sooner than other groups.

IV. Incidence of unemployment in a multivariate context

We wish to investigate the factors that affect the incidence of unemployment, *i.e.* to identify the characteristics of individuals that make them more likely to be unemployed. Tables 1 to 4 presented the racial, gender, regional, educational, and age distribution of unemployment in South Africa and showed certain patterns in the incidence of unemployment. However, analysis in a multivariate framework is required in order to isolate the effect of each variable holding the others constant. We utilise a standard discrete choice framework to model the states ‘unemployed’ and ‘employed’ as a binary probit.

While both supply-side (worker-related) and demand-side (employer-related) factors are responsible for an individual’s labour market state (employed or not), it is not possible in our model to distinguish between job-rationing reasons and worker preferences. For example, if being more educated increases the likelihood of being employed, this could be either because employers ration jobs by educational level of applicants in a labour surplus economy or because individuals who are more educated have more realistic reservation wages relative to their expected wages. The model is largely unable to distinguish between the constraints on and preferences for employment since their effects are not readily separable, though the inclusion of variables whose interpretation is unambiguous should help.

The only non-worker-related characteristics used in the model are (i) variables representing cost of job search, namely the condition of roads in the community; (ii) a proxy for the economic development of the community, capturing local employment opportunities and the local demand for labour; and (iii) a set of region dummies which aims to capture the effects of regional economic differences.

As unemployment is catastrophically high among Africans (41% by the broad definition in 1994) and only about half that rate among the next worst-off racial group (coloureds), we focus on the African group, though results for other minority groups are also presented, *i.e.* for the so-called coloureds, Asians, and whites. The sample contains only labour force participants and the base or reference category is the employed. Table 5 sets out the results of the pooled binary probit of unemployment for all races and separately for Africans using the OHS94 data. Table 6 presents the binary probits of unemployment

separately for coloured, Indian and white persons. A likelihood ratio test of whether it is appropriate to pool the separate races into a single equation was easily rejected⁴.

The effect of race on the probability of unemployment confirms the patterns noted earlier. The pooled model with the race dummy variables shows that, even after controlling for locational, demographic, and measured human capital characteristics such as age and education, Africans are 21, coloureds 15, and Indians 11 percentage points more likely to be unemployed than whites. The corresponding figures from a probit of unemployment fitted with SALDRU 1993 data (presented in Appendix 1) are 25, 20, and 15 percentage points respectively. It is possible that this difference reflects a reduction in racial discrimination in the employment practices of employers in the South African labour market in a period of rapid political change between 1993 and 1994.

The fact that even after the introduction of a battery of controls non-whites suffered significantly greater chances of unemployment than whites in 1993-1994 suggests either racial discrimination in employers' hiring practices or prior discrimination in the schooling system whereby blacks suffered poorer quality schooling than whites, or both. Since quality of education received in the past was governed by race (Case and Deaton, 1997), we wanted to explore whether inferior quality schooling is responsible for blacks' higher unemployment by estimating the probit models again for that subset of labour force participants for whom cognitive-skill scores were available. Despite doubts about the reliability of the test score data⁵, we nevertheless experimented with including test scores as proxies for schooling quality. The presence of test scores (literacy, numeracy, or both together) made no significant difference to the estimated coefficients of the race dummy variables. On the available, weak, evidence we cannot conclude

⁴ The unrestricted log likelihood was obtained from a pooled unemployment probit which included all the variables as well as all variables interacted with the race dummies. The restricted log likelihood was obtained from a pooled unemployment probit which included just the variables and no race interaction terms. Thus, for example, the pooling of the white and African samples was easily rejected: $\chi^2 = 1282.4$. The pooling of other races was rejected as well.

⁵ We wished to utilise cognitive skill test scores as proxies for quality of schooling received. However, there are several drawbacks associated with the test score data in the SALDRU survey. Firstly, tests were administered only to one in six of the sample households and within each of these households, it was given to only 2 members of the household, one of whom was in the age group 13-17 and one over 17. In total, 1330 individuals older than 17 took the test, but less than 500 of these are labour force participants. The test takers over the age of 17 are split 65:35 women to men. Although the descriptive material does not say so, it seems that the tests were administered at times when school children were present, but when working adults were likely not to be. As a result, the adult test takers are predominantly women and few report any wage income. As Case and Deaton (1997) point out, this selection is likely to jeopardise any general inferences from the test scores, particularly about the links between test scores and labour market outcomes such as earnings or employment.

that racial differences in unemployment probability are partly due to racial differences in educational quality, though this is clearly plausible.

The separate probits of unemployment for coloured, Indian, and white persons are presented in Table 6. It is conspicuous that in the African probit, most variables are statistically significant and the goodness-of-fit, as measured by the pseudo R-square, is better than in the probits for other race groups, particularly the whites. This is partly because there is a greater degree of variation in the dependent variable in the African sample (for instance, 40% of Africans and 7% of whites were unemployed).

The probability of unemployment decreases with age but at a diminishing rate⁶. Incumbents may be protected against competition from (young) entrants by labour market laws or institutions or by firm-specific human capital. Reservation wages may fall with age or with time spent in unemployment. Alternatively, younger people may have a greater chance of entry into unemployment because of their higher degree of job-mobility. There is support for the latter notion in Table 3 which shows that younger persons are more likely to enter unemployment voluntarily. The higher degree of job-mobility among the young is likely to result from their low levels of specific human capital, their relatively low current costs of unemployment, and their greater ease of finding another job (at least among those who ever held a job before) - which their relatively lower duration figure in Table 4 indicates.

The incidence of unemployment decreases dramatically with education for all race groups though the effect is statistically significant only for Africans and coloured persons. For example, possessing higher education reduces an African's predicted probability of unemployment to nil: the marginal effect of the higher education dummy variable is about -39 percentage points whereas the unemployment rate for Africans with no education (base category) is about 36%. Similarly, among coloureds: the marginal effect of higher education is nearly -16 percentage points and the unemployment rate for uneducated coloureds is 14%. For Indians with higher education, the predicted probability of unemployment is also nil. The comparison between Africans and all other race groups is interesting. Among Africans, education begins to

⁶ For whites there is a U-shaped relationship between education and the probability of unemployment - a relationship similar to that in OECD countries (Nickell, 1980; Steven, 1999; Blackaby and Manning, 1998, Oswald 1999). Unemployment probits fitted with AGE36-45 as the base category showed that only in the case of whites was there a significant U shape in the education-unemployment relationship. The higher incidence of unemployment among the old (>46 years old) is explained by their waning productivity which is likely to result in their greater incidence of involuntary entry into unemployment. Moreover, being less adaptable, they are more likely to have longer duration of unemployment due to their more stringent criteria for what they would consider as suitable employment.

matter to unemployment from the junior secondary level - *i.e.* Africans with 8-10 years of education have significantly lower chances of being unemployed than those with no education. However, for coloureds, education begins to matter only from the senior secondary level (11-12 years of schooling) onwards.

The role of housing tenure in predicting unemployment was first highlighted by Nickell (1980) and has been used by Hughes and McCormick (1987) and Oswald (1999), among others. This literature attributes increases in unemployment in certain OECD countries to the increased rate of home-ownership in these countries (Oswald 1999), the reason being that home-ownership (and council-housing) makes people immobile. In Oswald's unemployment probits for the UK, for example, the individual home-ownership variable has a positive coefficient. It is arguable that home-ownership can exert two opposing sorts of influence on the probability of unemployment. It may exert a *positive* effect either because homeowners are less mobile or because home-ownership may proxy household wealth and wealthier people may have higher reservation wages. Home-ownership may exert a *negative* effect on unemployment probability if it is endogenous to unemployment (*i.e.* if unemployment determines the chances of owning your own home). Table 5 shows that for Africans the former influence dominates - home ownership increases the chances of unemployment by 5.4 percentage points. However, for Indians and whites the latter effect is more relevant, their home-ownership being associated with a lower probability of unemployment (the marginal effects being about -6 and -2 percentage points respectively). The district home- ownership rate has a large positive effect on the chances of unemployment for Africans only.

The number of dependants in the household (NUMDEP) could either increase unemployment probability - because of greater child-care responsibilities, particularly for women - by making them less flexible labour force participants, or it could decrease unemployment probability because of greater economic need and the consequent lower reservation wage. Thus, one cannot predict the sign of this variable *a priori*. Among Africans, the child-care reason seems to dominate, making people significantly more likely to be unemployed, but among whites, the economic interpretation appears to apply, though the effect is considerably weaker. Gender-differentiated probits show that the effect of NUMDEP on the probability of unemployment is about twice as strong for women as for men.

We had included certain household composition variables (such as marriage and headship status) in earlier versions of the unemployment probits. Both very significantly reduced the chances of unemployment

in each race group. This is consistent with the notion of economic responsibility falling more heavily on household heads and married members. The negative effect of headship and marriage on the probability of unemployment may arise if these are taken as personal traits signalling, say, greater maturity or trustworthiness, that employers use to ration jobs. Another explanation is that married and head persons' greater economic responsibility within the household means that they have lower (more realistic) reservation wages. If so, there is an element of voluntariness in unemployment. However, our preferred specification excludes these variables on account of the strongly endogenous nature of these variables in an unemployment probit: people who are unemployed have lower chances of marrying and becoming heads of their own households.

Controlling for homeland residence, the probability of unemployment is 16 percentage points higher for urban than rural Africans⁷. Among coloureds and Indians, the unemployment chances of urban persons are 21 percentage points and 8 percentage points higher, respectively, than of their rural counterparts. This may be because urban-based job-search is considered more effective than rural-based job-search, as hypothesised in probabilistic models of labour migration. Among whites urban/rural residence has no significant impact.

Residence in a former 'homeland' still entails a substantially greater risk of unemployment than residence elsewhere. A black worker living in a homeland is about 18 percentage points more likely to be unemployed than a black worker living in a non-homeland region. This indicates that despite the considerable loosening of apartheid segregation laws by 1994, the former homeland regions still continued to act as labour reserves whose residents were at a great disadvantage in the labour market. Province dummies are included to see whether unemployment incidence varies substantially regionally. The base category is the major metropolitan area, based on Johannesburg (PWV, now known as Gauteng). Blacks in all provinces except northern and eastern Cape are significantly less likely to be unemployed than blacks in Gauteng, i.e. Gauteng acts as a magnet attracting black migrants.

In the OHS94, there is information available on distance to the nearest telephone. This is used as a proxy for the remoteness of the community. It is likely to capture aspects of the cost of job-search: we

⁷ In the SALDRU 1993 data, an unemployment probit for Africans finds the marginal effect on the urban variable to be 9.7 percentage points, i.e. substantially lower than in the OHS94 data. It is possible that between 1993 and 1994 many Africans migrated to urban areas and this swelled the urban African unemployment rate.

expect a positive sign on this variable⁸. Tables 5 and 6 show that this measure of remoteness has a highly significant positive effect on the chances of unemployment. The more remote the community, the higher the cost of job-search and, accordingly, the higher the probability of unemployment. In the African unemployment probit using SALDRU data, shown in Appendix 1, living in a cluster with impassable roads increases the chances of unemployment significantly. This too is consistent with the notion of the cost of job-search being higher in remote clusters⁹.

V Decomposition of the race gap in unemployment probability

The broad unemployment rate among Africans (41%) and whites (6%) in the OHS94 data indicates that the raw African-white race gap in unemployment rate is 35 percentage points. After standardising for observed characteristics in the pooled unemployment probit of Table 5, however, this race gap is reduced to 21 percentage points. In other words, 14 percentage points out of the 35 percentage point gap is explained by the African-white difference in observed characteristics. Thus 40% of the racial gap in the probability of unemployment is attributable to differences in measured characteristics. The unexplained residual (60%) is due to racial discrimination or to differences in the unobserved traits of blacks and whites, or to a combination of both. However, this method of inferring the extent of the unexplained gap in unemployment probability is unsatisfactory because of its restrictive assumption that the probit of unemployment is identical for blacks and whites in all respects except the intercept.

A more satisfactory method is to allow for the possibility that the coefficients of the variables differ by race and then to apply the Gomulka and Stern (1990) method of decomposing the raw race-gap in the probability of unemployment into explained and unexplained components. This method is an adaptation of

⁸ The SALDRU data have rather better community level information available which enable us to capture aspects of cost of job-search and demand for labour, both potentially important determinants of unemployment. Whether there are any roads that become impassable at certain times of the year (IMPASS) is a proxy for cost of job-search. Total number of facilities in the community (NUMFACI) - such as restaurant, post-office, bank, daily market, etc. - is a measure of the economic development of the community and, as such, at least a crude measure of the local demand for labour. It should be stated that community level information is missing on 24 clusters in the SALDRU data. As a result, we have assigned the overall mean value of NUMFACI to clusters where NUMFACI was missing. But, given the discrete (0/1) nature of the TARROAD and IMPASS variables and given the likelihood that the clusters with missing community schedules are those in remote areas, we have assigned a value of 0 for TARROAD and of 1 for IMPASS in these 24 clusters. The results are shown in Appendix 1.

⁹ The effect of NUMFACI is different for Africans and whites (results for whites not shown in Appendix 1 but available from the authors upon request): while for whites NUMFACI proxies the local demand for labour - with greater facilities

the Oaxaca method of decomposing group wage differences (Oaxaca, 1973) for the case of discrete choice models, and has been used recently (Blackaby, *et. al.*, 1998, 1999).

Let I_i^{*j} be a latent variable for the i th individual in the j th race group, where

$$I_i^{*j} = \mathbf{a}X_i + u \quad (1)$$

X is a vector of variables, \mathbf{a} is an associated vector of coefficients, and u is an error term distributed $N(0,1)$. Suppose that a binary indicator variable indicating unemployment status is given by

$$\begin{aligned} I_i^j &= 1 && \text{if } I_i^{*j} \geq 0, \text{ i.e., the individual is unemployed, and} \\ I_i^j &= 0 && \text{if } I_i^{*j} < 0, \text{ i.e., the individual is employed.} \end{aligned} \quad (2)$$

Denote the probability of observing $I_i^j=1$ by $P(I_i^j=1)$. This probability of observing unemployment is given by the cumulative normal distribution $P(\mathbf{a}, X_i)$ and it can be estimated using the familiar binary probit model. By fitting two separate unemployment probits for two different races and using the method of Gomulka and Stern (1990), it is possible to decompose the overall racial difference in unemployment probability ($\hat{I}_a - \hat{I}_w$) as:

$$\hat{I}_a - \hat{I}_w = \{\bar{P}(\mathbf{a}_a, X_a) - \bar{P}(\mathbf{a}_w, X_a)\} + \{\bar{P}(\mathbf{a}_w, X_a) - \bar{P}(\mathbf{a}_w, X_w)\} \quad (3)$$

or

$$\hat{I}_a - \hat{I}_w = \{\bar{P}(\mathbf{a}_a, X_w) - \bar{P}(\mathbf{a}_w, X_w)\} + \{\bar{P}(\mathbf{a}_a, X_a) - \bar{P}(\mathbf{a}_a, X_w)\} \quad (4)$$

where subscripts a and w refer to the African and white groups respectively, \hat{I}_a is the average of the predicted unemployment probabilities for African individuals, and \hat{I}_w the analogous average probability of unemployment for white individuals.

One feature of probit analysis is that, unlike OLS, the actual mean of the dependent variable and the predicted mean in a regression equation need not be the same. However, they are frequently close together and in our analysis the difference in the two means is relatively small. For example, in the sample of African labour force participants, the actual mean of the dependent variable (unemployed=1; employed=0) is 0.406 and the predicted mean is 0.382.

reducing the risk of unemployment - for blacks, the positive effect of NUMFACI suggests that African unemployed job-seekers migrate to clusters where there is greater demand for labour.

In what follows, we use the probit equations of unemployment for Africans and whites to obtain \mathbf{a}_a and \mathbf{a}_w . For each individual we produce the predicted probability of unemployment and then calculate the mean of the predicted probabilities summing over observations. Thus, $\bar{P}(\mathbf{a}_a, X_a)$ is the average across the sample of the predicted probabilities using African coefficients and African characteristics; $\bar{P}(\mathbf{a}_w, X_a)$ is the average across the sample of predicted probabilities using white coefficients and African characteristics, and so on. Similar computations are made for the comparison between whites and other race groups. The results of the decomposition exercise are given in Table 7.

The raw unemployment rates by race suggest most discrimination against Africans, followed by coloureds, and least discrimination against Indians. If we attribute the unexplained component to employer discrimination, then Table 7 shows that, the higher probability of unemployment that is due to discrimination is (taking the average of the two estimates in each case) 8.7, 8.3, and 5.1 percentage points for African, coloureds and Asians, respectively.

The major part of the reason why Africans have a much higher unemployment rate than whites is their lower levels of employment-enhancing characteristics such as education and their location in areas of high unemployment. However, both the lower education of Africans and their location in high unemployment areas are manifestations of pre-labour-market discrimination, for example in the schooling system - that subjected Africans to poorer access to and quality of education - and in apartheid policies that forcibly confined millions of Africans to the so-called 'homeland' regions which are very low-employment areas. Quality of education is but one unobserved factor that is likely to affect unemployment probability. Other unobserved factors omitted from the probit analysis owing to a lack of data are social networks, trust, attitudes, skills, seniority, and authority.

VI Conclusions

Unemployment is very inequitably distributed in South Africa and certain groups are much more likely to enter it and to stay in it than others. Young uneducated Africans living in homelands and remote areas are particularly vulnerable to unemployment. There are two particularly striking features of South African unemployment: firstly, the fact that rural unemployment rates are higher than urban rates is atypical among countries and is explained by historical policies restricting mobility. Secondly, it is striking that the

majority (62%) of the unemployed have never held a job before, *i.e.*, they entered unemployment from the time of entering the labour force. The very long duration of unemployment (>1 year) among a high proportion (68%) of the unemployed suggests that the demand-side of the labour market is responsible for a good part of the unemployment rate.

In the African group - the group that suffers catastrophically high unemployment rates - human capital characteristics such as education and age dramatically reduce the chances of unemployment. This is also true to a large extent for the coloured group. While the policy prescription often made from this well known fact is that education and skills should be upgraded in low education groups, such a prescription is unlikely to address the problem: unless there are more jobs in the economy, upgrading education of Africans will at best change the composition of employment in their favour.

The analysis suggests that racial differences in unemployment incidence cannot simply be dismissed as a problem of the poorer productive characteristics of the African, coloured, and Indian groups relative to the whites in South Africa. While a substantial part of the race gap in incidence of unemployment in the mid 1990s was explained by inter-group differences in characteristics, there remained a residual that could not be explained by observed characteristics. The residual may be due to employer discrimination or to inter-racial differences in unmeasured traits such as quality of education. Further research in this area incorporating data for the more recent October Household Surveys and data on quality of education would be fruitful. To explore important policy questions concerning unemployment dynamics, longitudinal datasets involving annual data will be required.

References

- Blackaby, D., K. Clark, D. Leslie, P. Murphy (1994) "Black-white male earnings and employment prospects in the 1970s and 1980s: Evidence for Britain", *Economic Letters*, Vol. 46, p273-80.
- Blackaby, D. et al (1998) "The Ethnic Wage Gap and Employment Differentials in the 1990s: Evidence for Britain", *Economics Letters*; 58(1), p97-103, January.
- Blackaby, D. et al (1999), "Unemployment among Britain's Ethnic Minorities", *Manchester School*; 67(1), p1-20, January.
- Case, Anne and Angus Deaton (1999) "School Inputs and Educational Outcomes in South Africa", *Quarterly Journal of Economics*; 114(3), p 1047-84.
- CSS (1995) "October Household Survey 1994", Statistical Release PO317, Central Statistical Service, Pretoria, March.
- Gomulka, J. and N. Stern (1990) "The Employment of Married Women in the United Kingdom 1970-1983", *Economica*, 57: p171-199, May.
- Hughes, G.A. and B. McCormick (1987) "Housing Markets, Unemployment and Labour Market Flexibility in the UK", *European Economic Review*; 31(3): 615-41, April.
- Jensen, Robert (1999) "An Early Assessment of Racial Wage Differentials in Post-Apartheid South Africa", mimeo, Harvard University.
- Kingdon, Geeta G. and John B. Knight (1999) "Unemployment and Wages in South Africa: A Spatial Approach", Working Paper WPS/99-12, Centre for the Study of African Economies, University of Oxford.
- Kingdon, Geeta G. and John B. Knight (2000a) "Are Searching and Non-Searching Unemployment Distinct States When Unemployment is High? The Case of South Africa", Working Paper WPS/2000-2, Centre for the Study of African Economies, University of Oxford.
- Kingdon, Geeta G. and John B. Knight (2000b) "Unemployment in South Africa: The Nature of the Beast", forthcoming mimeo, Centre for the Study of African Economies, University of Oxford.
- Klasen, Stephan and Ingrid Woolard (1999) "Levels, Trends, and Consistency of Employment and Unemployment Figures in South Africa", *Development Southern Africa*, 16, No. 1: 3-36.
- Knight, J.B. and M.D. McGrath (1977) "An Analysis of Racial Wage Discrimination in South Africa", *Oxford Bulletin of Economics and Statistics*, 39(4), p. 245-71.
- Knight, J.B. and M.D. McGrath (1987) "The Erosion of Apartheid in the South African Labour Market: Measures and Mechanisms", Applied Economics Discussion Paper 35, Institute of Economics and Statistics, University of Oxford.

ILO (1996) *Restructuring the Labour Market: The South African Challenge: An ILO Country Review*, International Labour Organisation, Geneva.

ILO (1998) *Yearbook of Labour Statistics*, International Labour Organisation, Geneva.

Leibbrandt, M., H. Bhorat, and I. Woolard (2000) “Understanding Contemporary Household Inequality in South Africa”, *Journal of Studies in Economics and Econometrics*, November.

Moll, Peter (2000) “Discrimination is Declining in South Africa but Inequality is Not”, *Journal of Studies in Economics and Econometrics*, November.

Nickell, Steve (1980) “A Picture of Male Unemployment in Britain”, *Economic Journal*, 90(363), p 776-94, December.

Oaxaca, R. (1973) “Male-Female Differentials in Urban Labour Markets”, *International Economic Review*, 3, pp.603-709.

Oswald (1999) “

Rospabe, Sandrine (1997) “The influence of unionisation on the racial wage gap in South Africa: A Microeconomic Study Using 1993 data”, Paper presented at a seminar at the University of Cape Town, December 1997, mimeo, Clermont Ferrand (??).

SALDRU (1994) “South Africans Rich and Poor: Baseline Household Statistics”, Project for Statistics on Living Standards and Development, South African Labour and Development Research Unit, Cape Town.

StatsSA (1998) *Unemployment and Employment in South Africa*, Statistics South Africa, Pretoria.

StatsSA (2000) Statistical Release PO317, StatsSA Homepage, <http://www.statssa.gov.za/>

Table 1
Unemployment rates in South Africa, 1993-1999

	Source	Broad definition	Narrow definition	Broad-narrow gap
1993	SALDRU	31.2	13.0*	18.2*
1994	OHS	31.5	20.0	11.5
1995	OHS	29.2	16.9	12.3
1996	OHS	35.6	21.0	14.6
1997	OHS	37.6	22.9	14.7
1998	OHS	38.6	26.1	12.5

Source: Authors' own calculations from SALDRU data; OHS figures from StatsSA (1998, p3) and StatsSA's webpage (StatsSA, 2000).

Note: * The large difference in narrow unemployment rates between SALDRU and OHS sources is due to the fact that the SALDRU survey used a reference period (for job-search) of one week whereas the OHS surveys use one of four weeks.

Table 2
Unemployment rate (%), by age, education, gender, region, and race, OHS94

	Broad definition	Narrow definition	Broad-narrow gap
Age			
16-24	51.4	37.8	13.6
25-35	35.3	23.3	12.0
36-45	25.2	14.3	10.9
46-55	21.3	11.0	10.3
55-64	16.9	8.5	8.4
Education			
none	38.7	20.1	18.6
primary	42.5	26.8	15.7
junior	35.3	23.5	11.8
secondary	28.3	19.5	8.8
higher	5.7	3.9	1.8
Gender			
male	26.2	17.3	8.9
female	40.7	25.3	15.4
Region			
rural	40.3	23.4	16.9
urban	27.9	19.1	8.8
Race			
African	41.2	26.2	15.0
Coloured	23.3	19.4	3.9
Indian	17.1	14.3	2.8
White	6.3	4.2	2.1

Source: Authors' own calculations from OHS94.

Table 3
Entry into unemployment, 1994, by age, education, gender, region, and race

	All unemployed (N)	Never worked before (%) (a)	Worked before (%) (b)	Of those who worked before, proportion who entered unemployment voluntarily (%)
Age				
16-24	4128	82.8	17.2	25.4
25-35	5245	64.6	35.4	26.4
36-45	2646	52.8	47.2	24.1
46-55	1244	47.4	52.6	22.4
55-64	338	39.8	60.2	13.3
Education				
none	1265	63.3	36.7	23.1
primary	4507	63.6	36.4	23.5
junior	4476	60.5	39.5	23.1
secondary	3056	74.9	25.1	29.1
higher	297	57.3	42.7	38.7
Gender				
male	5572	58.9	41.1	15.8
female	8029	69.9	30.9	34.2
Region				
rural	5642	72.3	27.7	24.8
urban	7959	58.4	41.6	24.4
Race				
African	10130	68.4	31.6	22.6
Coloured	2236	43.7	56.3	22.9
Indian	609	46.6	53.4	27.9
White	626	30.3	69.7	49.3
Total	13601	61.8	38.2	24.6

Source: Authors' own calculations from OHS94.

Note: a + b = 100%.

Table 3a
Binary probit of 'Ever worked before for pay, profit, or family gain', OHS94

	Coefficient	Marginal effect	Robust t-value	
age2535*	0.7087	0.267	16.14	***
age3645*	0.9158	0.352	16.61	***
age4655*	1.0202	0.389	16.29	***
age5664*	1.0870	0.408	12.19	***
Male*	0.3007	0.113	8.31	***
Household head*	0.3278	0.126	7.73	***
Married*	0.1685	0.064	5.00	***
Numdep	0.0126	0.005	1.42	
African*	-0.9260	-0.354	-10.56	***
Coloured*	-0.3958	-0.140	-3.85	***
Indian*	-0.5349	-0.177	-4.42	***
Urban homeland*	-0.0362	-0.013	-0.29	
Rural non-homeland*	0.4760	0.186	3.92	***
Urban non-homeland*	0.4297	0.160	4.45	***
Numempl	0.0237	0.009	1.15	
Primary*	0.1000	0.038	1.73	*
Junior*	0.1089	0.041	1.80	*
Secondary*	-0.1334	-0.049	-1.96	**
Higher*	0.2565	0.099	1.24	
Vocational training*	-0.3374	-0.117	-1.88	*
Lives in owned home*	-0.0342	-0.013	-0.68	
Wcape*	0.4043	0.157	3.44	***
Ncape*	0.2416	0.093	1.49	
Ecape*	0.0841	0.032	0.69	
Natal*	0.2615	0.100	2.44	**
Ofs*	0.0088	0.003	0.06	
Etv1*	0.2459	0.095	1.68	*
Ntv1*	-0.1102	-0.041	-0.62	
Nw*	0.5047	0.197	3.72	***
constant	-0.8769	----	-5.28	***
Log L		-7555.071		
Restricted Log L		-9044.858		
Pseudo R-square		0.1647		
N		13601		
Mean of dependent variable*		0.3820		

Note: the starred variables are 0/1 variables. Their mean represents the proportion of ones in the sample. Numdep is the number of dependents (aged <16 or >64); Numempl is number of employed members in the household. The base category for age is age16-24.

Table 4
Duration of unemployment, by age, education, gender, region, and race, OHS94

	Number of unemployed	Duration of unemployment (months)	% distribution of duration of unemployment					
			<1 month	1-2 months	2-6 months	6-12 months	12-36 months	>36 months
Age								
16-24	4128	21.4	7.2	5.4	9.7	20.1	35.8	21.9
25-35	5245	28.5	4.7	3.2	7.7	13.9	28.9	41.6
36-45	2646	30.4	5.2	3.1	6.4	12.7	24.7	48.0
46-55	1244	30.8	6.5	2.1	6.9	11.6	23.3	49.6
55-64	338	32.3	7.4	1.8	5.0	11.2	20.2	54.5
Education								
none	1265	30.0	5.3	3.8	6.7	11.0	27.1	46.1
primary	4507	29.8	4.8	3.3	6.5	13.1	26.8	45.5
junior	4476	27.4	5.4	3.4	8.0	14.6	30.2	38.4
secondary	3056	23.4	6.3	4.1	10.3	19.8	30.9	28.6
higher	297	16.5	22.2	5.0	8.6	24.0	22.9	17.3
Gender								
male	5572	27.3	4.7	3.6	8.3	15.1	30.1	38.2
female	8029	27.2	6.5	3.6	7.5	14.9	29.0	38.5
Region								
rural	5642	28.2	5.6	3.4	6.5	13.8	30.4	40.2
urban	7959	26.4	5.7	3.7	9.1	16.1	28.7	36.6
Race								
African	10130	28.3	4.8	3.3	7.2	14.5	29.5	40.7
Coloured	2236	21.3	6.2	5.9	12.2	19.2	33.9	22.6
Indian	609	19.7	7.6	5.5	12.5	21.8	33.6	18.9
White	626	14.6	25.9	5.8	13.8	15.4	26.2	12.9
Type of U								
Searching U	7725	26.3	4.5	4.1	9.3	16.7	29.4	36.0
Non-search U	5876	28.1	7.0	3.1	6.2	13.2	30.4	40.2
TOTAL	13601	27.2	5.7	3.6	7.9	15.0	29.6	38.2

Source: Authors' calculations from the October Household Survey, 1994. The OHS94 survey truncates the duration question at 12 months, i.e the longest duration information code provided is 'greater than 1 year'. Since 67.8% of all unemployed persons had unemployment duration of greater than 1 year, there is a great loss of information on variation of unemployment duration within this large group. For the purposes of computing column 2 'mean duration in months', the mid-points of the categories <1 month, 1-2 months, 2-6 months, 6-12 months, 1-3 years and >3 years are taken as 0.5, 1.5, 4, 9, 24, and 48 months respectively.

Table 5
Unemployment probits, OHS94, Whole sample and Africans

	Pooled			African		
	coefficient	robust t value	marginal effect	coefficient	robust t value	marginal effect
Age						
age21-25	-0.3318	-9.58 ***	-0.091	-0.2088	-4.44 ***	-0.078
age26-35	-0.7792	-23.41 ***	-0.209	-0.7196	-16.41 ***	-0.259
age36-45	-1.0890	-29.59 ***	-0.261	-1.1258	-22.86 ***	-0.368
age46-55	-1.1810	-28.24 ***	-0.243	-1.2832	-22.89 ***	-0.371
age56-64	-1.2997	-21.16 **	-0.224	-1.5460	-20.56 ***	-0.373
Education						
primary	0.0149	0.39	0.005	-0.0351	-0.82	-0.013
junior	-0.0781	-1.68 *	-0.023	-0.1798	-3.41 ***	-0.068
secondary	-0.3200	-6.03 ***	-0.091	-0.3291	-5.20 ***	-0.120
higher	-1.0376	-10.67 ***	-0.215	-1.7038	-10.91 ***	-0.393
voc diploma	-0.0098	-0.13	-0.003	0.1803	1.19	0.070
Other var						
ownship	0.0554	1.09	0.017	0.1433	2.51 ***	0.054
numdep	0.0514	8.80 ***	0.016	0.0550	7.60 ***	0.021
urban	0.4690	7.70 ***	0.134	0.4175	5.92 ***	0.158
male	-0.3578	-13.41 ***	-0.109	-0.3891	-13.96 ***	-0.148
Race						
african	0.6957	10.95 ***	0.206			
colored	0.4468	5.02 ***	0.147			
indian	0.3264	4.44 ***	0.108			
Location						
homeland	0.5458	7.01 ***	0.180	0.5675	6.56 ***	0.215
w. cape	-0.4449	-3.85 ***	-0.118	-0.3824	-3.22 ***	-0.136
n. cape	-0.0200	-0.17	-0.006	-0.1996	-1.45	-0.074
e. cape	-0.1010	-1.00	-0.030	-0.1239	-1.05	-0.047
kwazulu natal	-0.3596	-3.63 ***	-0.099	-0.5882	-5.13 ***	-0.207
free state	-0.3117	-2.24 ***	-0.084	-0.4918	-3.27 ***	-0.171
mpumalanga	-0.2168	-2.28 ***	-0.061	-0.2981	-2.68 ***	-0.108
n. province	-0.2571	-2.11 **	-0.071	-0.3473	-2.46 ***	-0.125
n.w. province	-0.3331	-3.12 ***	-0.089	-0.4737	-3.60 ***	-0.166
Community var						
disttel	0.0729	5.99 ***	0.022	0.0736	4.41 ***	0.028
downship	0.5349	4.21 ***	0.161	0.4881	3.14 ***	0.186
constant	-0.7659	-6.86 ***		-0.0888	0.67	
Ln L		-22330.51			-13875.81	
Restricted LnL		-28501.08			-16839.64	
Pseudo R^2		0.2165			0.1760	
N		47667			24929	
Mean of dependent var		0.285			0.406	

Note: The base or reference categories are as follows: age: persons aged 16-20 years old; education: persons with no education; race: whites; and province: PWV (now called Gauteng). *, **, and *** represent statistical significance at the 10%, 5% and 1% levels respectively. Disttel is distance to nearest telephone (a proxy for remoteness); downship is district home ownership rate, i.e. the proportion of households in the district that lived in owned homes.

Table 6
Unemployment probit, by race, OHS94

Variable	Coloured			Indian			White		
	coefficient	robust t-value	marginal effect	coefficient	robust t-value	marginal effect	coefficient	robust t-value	marginal effect
Age									
age21-25	-0.4764	-8.79 ***	-0.111	-0.4453	-6.06 ***	-0.076	-0.6209	-7.67 ***	-0.045
age26-35	-0.8859	-14.08 ***	-0.203	-0.8550	-11.25 ***	-0.139	-1.0733	-11.59 ***	-0.083
age36-45	-1.1094	-16.14 ***	-0.222	-0.9568	-8.23 ***	-0.151	-1.0650	-12.35 ***	-0.082
age46-55	-1.1036	-12.06 ***	-0.190	-1.1198	-11.09 ***	-0.143	-0.9169	-10.34 ***	-0.064
age56-64	-1.2739	-11.42 ***	-0.182	-1.2314	-8.64 ***	-0.124	-0.5841	-5.76 ***	-0.041
Education									
primary	0.1222	1.32	0.033	0.3215	0.79	0.076	0.0168	0.04	0.002
junior	0.0113	0.11	0.003	0.2219	0.57	0.047	0.1480	0.40	0.017
secondry	-0.2854	-2.33 ***	-0.070	-0.0268	-0.06	-0.005	-0.2440	-0.67	-0.025
higher	-0.9411	-3.19 ***	-0.159	-0.6500	-1.37	-0.097	-0.4281	-1.17	-0.039
training	0.2373	0.80	0.070	0.2146	0.94	0.048	0.1270	1.30	0.014
Other var									
ownship	0.0057	0.08	0.002	-0.2624	-3.67 ***	-0.056	-0.1504	-3.26 ***	-0.017
numdep	0.0351	2.67 ***	0.009	0.0467	2.49 ***	0.009	0.0928	4.19 ***	0.010
urban	0.9269	5.38 ***	0.210	0.4595	4.53 ***	0.077	0.0560	0.63	0.006
male	-0.2597	-3.82 ***	-0.070	-0.3991	-5.71 ***	-0.087	-0.4650	-9.83 ***	-0.052
Location									
homeland	1.0266	2.78 ***	0.367	1.2561	2.87 ***	0.414	1.3714	2.75 ***	0.344
w. cape	-0.5850	-6.66 ***	-0.153	0.0061	0.05	0.001	0.2687	2.72 ***	0.033
n. cape	0.1956	1.73 *	0.056	0.8740	2.32 **	0.264	0.1282	1.34	0.015
e. cape	-0.2997	-2.52 ***	-0.073	-0.0210	-0.16	-0.004	0.3833	4.74 ***	0.051
kwazulu	-0.3010	-1.48	-0.071	0.1398	1.47	0.028	0.2915	3.33 ***	0.036
natal									
free state	0.0560	0.25	0.015	0.1620	0.26	0.036	0.1028	0.46	0.012
mpumalanga	-0.6582	-4.42 ***	-0.126	-0.2398	-1.32	-0.043	0.1201	1.22	0.014
n. province	-0.4608	-6.29 ***	-0.098	-0.2186	-0.84	-0.039	-0.1118	-1.05	-0.011
n.w.	-0.0689	-0.37	-0.018	-0.8468	-2.76 ***	-0.103	0.1738	1.80 *	0.021
province									
Community var									
disttel	0.0438	2.02 **	0.012	0.0637	1.58	0.013	0.0992	4.07 ***	0.010
downship	0.1258	0.62	0.034	-0.0073	-0.04	-0.001	0.0446	0.19	0.005
constant	-0.3973	-2.48 ***		-0.4924	-1.26		-0.5363	-1.39	
Log L		-4344.63			-1489.16			-2010.9608	
Restricted Log L		-5239.37			-1701.73			-2276.4859	

Pseudo R^2	0.1708	0.1249	0.1166
N	9709	3972	9057
Dependent variable mean	0.230	0.153	0.069

Table 7
Decomposition of the race gap in unemployment probability

	African-white		coloured-white		Indian-white	
	Standardising by		Standardising by		Standardising by	
	African means x	White means y	Coloured means x	White means y	Indian means x	White means y
A. Total race gap in unemployment probability	0.337	0.337	0.161	0.161	0.084	0.084
B. Part of A explained by characteristics	0.280	0.220	0.094	0.062	0.032	0.035
C. Part of A not explained by characteristics	0.057	0.117	0.067	0.099	0.052	0.049
Average of x and y for race-pair	0.087		0.083		0.051	

Appendix Table 1
Unemployment probit (SALDRU 1993 data)

	Pooled				African			
	coefficient	robust t value	marginal effect	mean	coefficient	robust t value	marginal effect	mean
Age								
age21-25	-0.2713	-4.60 ***	-0.082	0.170	-0.2045	-2.87 ***	-0.075	0.173
age26-35	-0.8106	-13.96 ***	-0.233	0.331	-0.7739	-11.10 ***	-0.272	0.339
age36-45	-1.1593	-18.12 ***	-0.292	0.247	-1.1342	-14.71 ***	-0.359	0.241
age46-55	-1.2499	-18.78 ***	-0.272	0.139	-1.3106	-17.60 ***	-0.363	0.130
age56-64	-1.4605	-15.87 ***	-0.260	0.052	-1.4677	-13.95 ***	-0.355	0.054
Education								
primary	-0.0044	-0.09	-0.001	0.303	-0.0440	-0.84	-0.017	0.377
junior	-0.0641	-1.14	-0.020	0.287	-0.1458	-2.49 ***	-0.054	0.287
secondary	-0.2162	-3.70 ***	-0.066	0.202	-0.2085	-3.24 ***	-0.076	0.157
higher	-0.8391	-5.10 ***	-0.202	0.090	-1.0136	-3.90 ***	-0.287	0.035
voc diploma	-0.1803	-1.00	-0.055	0.060	-0.4502	-1.64	-0.153	0.030
Other variables								
ownship	0.1450	3.33 ***	0.046	0.659	0.2411	5.10 ***	0.089	0.662
numdep	0.0682	8.76 ***	0.022	2.204	0.0679	8.25 ***	0.026	2.509
urban	0.2705	3.80 ***	0.086	0.568	0.2554	3.39 ***	0.097	0.427
male	-0.2113	-6.33 ***	-0.068	0.542	-0.2208	-5.86 ***	-0.083	0.534
Race								
african	0.9512	6.96 ***	0.257	0.722				
colored	0.5095	3.41 ***	0.182	0.098				
indian	0.3735	2.46 ***	0.132	0.034				
Location								
homeland	0.4125	4.45 ***	0.136	0.379	0.3633	3.73 ***	0.136	0.524
w. cape	-0.1355	-1.30	-0.042	0.102	-0.1099	-0.66	-0.041	0.028
n. cape	0.5219	2.57 ***	0.190	0.013	-0.2646	-0.80	-0.094	0.003
e. cape	0.3134	3.18 ***	0.108	0.110	0.3246	2.80 ***	0.126	0.130
kwazulu natal	-0.0963	-1.06	-0.030	0.203	-0.1394	-1.29	-0.052	0.216
free state	-0.1243	-1.29	-0.038	0.077	-0.2281	-2.53 ***	-0.083	0.093
mpumalanga	-0.2121	-2.06 **	-0.064	0.090	-0.2807	-2.40 ***	-0.101	0.114
n. province	0.1139	0.92	0.038	0.083	0.0896	0.64	0.034	0.106
n.w. province	-0.1266	-1.06	-0.039	0.092	-0.1492	-1.16	-0.055	0.118
Community var								
impass	0.0007	0.21	0.000	4.451	0.0037	1.88 *	0.001	3.197
numfacil	0.1026	2.06 **	0.033	0.374	0.1139	2.13 **	0.043	0.492
cownship	0.1863	1.70 *	0.060	0.668	0.2125	1.74 *	0.080	0.680
constant	-0.9924	-5.22 ***			-0.0484	-0.36		
Ln L		-6416.24				-5306.94		
Restricted LnL		-8165.33				-6349.34		
Pseudo R^2		0.2142				0.1642		
N		13154				9496		
Mean of dependent var		0.312				0.390		

Note: The base or reference categories are as follows: Age: persons aged 16-20 years old; Education: persons with no education; Race: Whites; and Province: PWV or Gauteng.