February 1998

# THE RISE IN UNION WAGE PREMIA IN SOUTH AFRICA

**By**
**Julian F. Hofmeyr**
**University of Natal, Durban**

**and**
**Robert E.B. Lucas**
**Boston University**

The authors would like to acknowledge helpful comments from several people, especially Eli Berman, Stephen Donald, Jonathan Eaton, John Knight, Kevin Lang, Peter Moll, Steven Nickell and Paul Schultz. Julian Hofmeyr thanks the Institute for Development Economics, Boston University, for its hospitality while working on this paper during 1996-97 and hereby acknowledges the financial assistance of the South African Centre for Science Development towards this research. Opinions expressed in this publication and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the Centre for Science Development.

**Abstract**

The change in the premia earned by African, urban, male, union members, relative to earnings of other, comparable workers in South Africa, between 1985 and 1993, is examined using two independent, national, sample surveys. The interval from 1985 to 1993 was part of a period of rapidly rising union membership for African workers, following the Industrial Conciliation Amendment Act of 1979 which extended the definition of an employee to encompass African workers for collective bargaining. Membership now reported by the unions suggests one of the highest rates of unionization of wage employees among the developing countries Meanwhile, the simple average earnings of African employees has risen substantially, in real terms, despite extremely high rates of unemployment. In this context, our sample data show that the premia in mean earnings of African, urban, male, union members relative to other African, urban, male regular employees rose 18.5 percent from 1985 to 1993. Clearly this rise could reflect changes in composition of regular employees and of union membership in particular during this episode of rapidly rising membership. Four approaches to estimation of the union premia are applied to examine this possibility: earnings equations with a union dummy; separate earnings regimes for members and non-members; a switching model with separate earnings regimes and endogenous membership; and a multinomial logit model which extends the standard switching model to encompass sample selection into employment as well as into regular employment among the employed. A series of nesting hypotheses are conducted to explore the significance of differences between these approaches. The levels of estimated premia differ widely across the various approaches. However, there is uniform agreement that the union earnings premia, based on either the mean observed characteristics of union or non-union members, had risen by 1993 and that the change in measured characteristics of union and non-union members between sample years actually tended to diminish the premium. This suggests that the observed rise in union premium must either reflect an increase in rents to union members as compared to identical non-member workers or some change in the composition of union members which we are unable to detect in our data.

The objective of the present paper is to examine the change, from 1985 to 1993, in the earnings premia received by African, urban, male, union members in South Africa. The historical setting of this change is of particular interest, in the light of the transformation from the apartheid regime, and is therefore described briefly, before turning to outline prior related work and the approaches adopted here.

**The Setting**

The South African Industrial Conciliation Act of 1956 prohibited the registration of multiracial trade unions. By the late 1970s, a few unregistered trade unions for Africans existed but these did not effectively engage in collective bargaining. (Horner, 1974; Nel and van Rooyen, 1993). Following the recommendations of the Wiehahn Commission, the Industrial Conciliation Amendment Act of 1979 extended the definition of an employee to encompass African workers for the first time and in January 1980 the African Transport Workers' Union was the first African trade union to acquire full registration under the amended act. Union membership has grown rapidly among African workers since 1980. Indeed, membership now reported by the unions suggests one of the highest rates of unionization of wage employees among the developing countries.

Meanwhile, during the 1980s, the fraction of the African labor force without wage employment grew quickly, to exceed 50 percent by the early 1990s, even as the simple average of real wages continued to rise. (Fallon and Pereira de Silva, 1994). Many factors may have contributed to this wage drift, including rising skills of African workers and diminished job barriers as the restrictions of apartheid were eliminated (Hofmeyr, 1994). Nonetheless the enhanced role for collective bargaining would seem a natural candidate for examination, though to date it has been largely neglected in empirical studies. (See, however, Fallon and Lucas, 1997).

Certainly the high rates of unemployment and segmentation of the labor force have been issues of central concern to the new government, elected in 1994. The African National Congress has been closely aligned with the Confederation of South African Trade Unions, both before and since the election, and there has been active discussion of the need for some form of social contract to contain

the pressures of segmentation.

In this context, the present paper looks at the change in the unions earnings premia using the only two existing national sample surveys, completed since the 1979 Industrial Conciliation Amendment Act, with information on employment status, earnings and union membership for the African population. These data show that the mean earnings premium of African, urban, male, union members -- relative to other African, urban, male regular employees -- rose by 18.5 percentage points from 1985 to 1993. Clearly this rise could reflect changes in the composition of regular employees and of union membership in particular, during this episode of rapidly rising membership. The aim of this paper is therefore to examine the roles of changes in measured characteristics of employees and of pay structures in contributing to this sharp rise in the mean union earnings premium.

**Prior work**

An extensive set of estimates of union wage premia exists, for the US in particular. (For an early survey see Lewis, 1986). Most earlier estimates simply incorporated an exogenous union membership dummy into individuals' earnings equations. Later approaches attempted to take into account the endogeneity of the union-membership decision (see Maddala, 1983, 347-364). One such approach is that of Lee (1978), who examines a model with potentially distinct earnings equations for members versus non-members and endogenous switching between regimes.

Both of the surveys analyzed here have been used, in previous work, to examine the union-nonunion wage gap for South Africa. Moll (1993) uses the earlier survey to explore union membership as a dummy variable in a pooled earnings equation, plus separate earnings regimes for union and non-union members with and without endogenous switching. Schultz and Mwabu (1996) use the 1993 survey to examine the union earnings premia at different quantiles of the earnings distribution as part of a broader study of income distribution, treating union membership as exogenous.[1]

---

[1] For a generalization of this approach to Canadian and US data see Donald, Green and Paarsch (1995).

**Approach adopted here**

In this paper we seek to impose a common structure in estimating earnings equations from both surveys in order to examine how the union premium has evolved over time. Such a comparison between Moll for 1985 and Mwabu and Schultz for 1993 is not possible since quite different specifications and estimation strategies are adopted.

Secondly, we explore some extensions to the prevailing methodology. Somewhat surprisingly, although corrections for sample selection bias from self-selection into wage employment are common in estimating earnings equations, these corrections do not seem to have been combined with models of endogenous union membership and switching between earnings regimes, either in the South African context or elsewhere. Given the high open rates of unemployment in South Africa, such sample selection effects could well prove important. In addition, however, the development literature has long emphasized segmentation between formal and informal labor markets, with union membership being only one of the contributing dimensions. (Calvo, 1978).

The present paper therefore considers three endogenous outcomes and their implications for estimation of the union differential — the outcomes with respect to possessing employment, to entry into more formal employment, and to union membership. The alternative approaches adopted to these issues are described in **Section 1** and the data are described in greater detail in **Section 2**. After laying out the specifications in **Section 3,** the resultant estimates are presented in **Section 4.**

Finally, **Section 5** turns to the central question of the paper --how large was the rise in union earnings premium for workers with given measured characteristics? This section also decomposes the estimated changes in premia into the role of changing measured characteristics of employees versus changing pay structures. The results, summarized in the concluding section, suggest that the observed rise in union premium must either reflect an increase in pure rents to union members as compared to identical non-member workers or some change in the composition of union members which we are unable to detect in our data.

## 1. Alternative Approaches to Estimation.

As mentioned in the introduction, several approaches to estimation of union wage premia have been applied to US data and it is probably fair to say that there is closer agreement about the magnitude of the premium than the preferred method.[2] Four of these approaches are adopted in the present context, in order to explore the robustness of our findings.

The four approaches are permutations on four components.

> First, different earnings regimes may be discernible between union and non-union members within the formal sector.
>
> Second, switching between the union and non-union earnings regimes may be treated endogenously.
>
> Third, it seems plausible *a priori* that informal sector earnings structures may be quite different from those in more formal labor markets (whether unionized or not), and division along the formal-informal boundary is also an endogenous outcome.
>
> Fourth, there is a sample-selection problem in that earnings are observed only for those who are employed, rather than for the whole population.

The issue of separate wage regimes for union and non-union members, together with endogenous switching between the two, has received substantial attention in the literature. Similarly, correction for sample selection bias is also common, though not typically applied to switching between union and non-union regimes. Finally, the separate structure of informal sector earnings has been the subject of some investigations within the development literature[3] though rarely, if ever, combined with the other two components.

For comparison purposes, the first of our four approaches is a simple, single earnings equation including an exogenous union membership dummy variable. The one compromise with the four issues raised above is exclusion of persons with earnings derived from casual and self employment, though this first approach makes no adjustment for sample selection. Perhaps precisely because this method

---

[2] See Fuchs, Krueger and Poterba (1997).

[3] See, for instance, Blau (1986).

is particularly simple, inheriting none of the potential drawbacks of the remaining methods, it is the approach favored by Lewis (1986) in his review of the literature on union premia.

The second method extends the first to incorporate the potential for distinct earnings regimes among union and non-union members, though without endogenous switching. In other words, earnings equations are estimated, by ordinary least squares, independently for non-union and for union members among "regular" employees (i.e. excluding casual and self employed people as before).[4]

The third approach adds in endogenous switching between the union and non-union earnings regimes. In particular, a probit model of union membership is estimated and the inverse Mills ratio, computed from the resultant coefficients, is added as an explanatory variable to the equations from the previous method.[5]

The final approach addresses the problems of switching between the formal and informal sectors and sample selection with respect to employment and non-employment, while retaining endogenous union membership. Specifically, this is tackled by specifying a multinomial logit model (MNL) of employment status for individuals, where employment status refers to four[6] categories: not employed, casual or informal self employment, regular employment of non-union members, and regular employment of union members. Separate wage equations are then estimated for the last two categories, incorporating the inverse Mills ratio calculated from the MNL to correct for sample-selection effects.

More formally, the last of these models is given by:[7]

---

[4] More precise discussion of these alternative employment categories is postponed until **Section 2**.

[5] In the South African context, this is the method deployed by Moll (1993) for the 1985 sample.

[6] In the 1993 sample five categories are distinguished, as discussed in the following section.

[7] See Greene (1992: pp.618-620).

1. $w_{ie} = \beta_e x_{ie} + \varepsilon_{ie}$ e=1,..,n,u,..M
2. $\pi_{ik} = \gamma_k z_{ik} + \eta_{ik}$ k=0,1,..,n,u,..M

where

$w_{ie}$ is the natural logarithm of earnings for person i in employment category e
n and u refer to non-union and union regular employees respectively,
x and z are vectors of exogenous variables,
$\beta$ and $\gamma$ are row vectors of coefficients and
$\varepsilon$ and $\eta$ are error terms which are assumed to be jointly normally distributed with zero means and covariance matrix

$$\begin{bmatrix} \sigma_\varepsilon^2 & \sigma_{\varepsilon\eta} \\ \sigma_{\varepsilon\eta} & \sigma_\eta^2 \end{bmatrix}$$

Earnings $w_{ie}$ are observed only if the individual falls into category e. $\pi$ is a latent variable which we do not observe; we observe only p which is a polychotomous variable with values 0 through M such that p is equal to k if the $k^{th}$ category applies, which is assumed to happen when $\pi_{ik} > \pi_{ic}$ (c=0 through M for $c \neq k$).

A two-stage procedure is adopted, in which $\gamma$ is estimated first by fitting a multinomial logit form to the polychotomous employment status variable. The inverse Mills ratio, $\mu_{ic}$, for each observation is then computed from these estimates and in the second stage the earnings equations

3. $w_{ie} = \beta_e x_{ie} + \phi_e \mu_{ie} + \varepsilon_{ie}$ e=n,u

are estimated by ordinary least squares. In this latter stage, the parameters on the Mills ratios provide estimates of the covariance between the errors in the particular earnings equation and the multinomial decision equation ($\sigma_{\varepsilon\eta}$).

The third approach outlined above simply replaces the multinomial logit model of employment status with a binomial probit model of union membership.[8] Denoting the ensuing Mills ratio by $\lambda_{ie}$, the estimated earnings equation may then be written for this third model as

4. $w_{ie} = \beta_e x_{ie} + \psi_e \lambda_{ie} + \varepsilon_{ie}$ e=n,u

We should like to be able to compare the various models statistically, though equations (3) and (4) are not nested. For a simple comparison we therefore consider estimating the union of the two as

---

[8] See Lee (1973) and Moll (1993).

5. $w_{ie} = \beta_e x_{ie} + \phi_e \mu_{ie} + \psi_e \lambda_{ie} + \varepsilon_{ie}$ e=n,u

and testing (3) and (4) as alternative nested forms within (5). The second approach sketched above omits the issues of sample selection and endogenous switching of union membership entirely and estimates the earnings equations

6. $w_{ie} = \beta_e x_{ie} + \varepsilon_{ie}$ e=n,u

by ordinary least squares. Obviously, (6) can be tested as nested within (3) imposing the condition that $\phi_e = 0$ or nested within (4) with $\psi_e = 0$. Finally, the most restrictive case is that with a single earnings equation incorporating a union dummy variable $u_{ie}$, where $u_{ie}=1$ if e = u, 0 otherwise

7. $w_{ie} = \beta x_{ie} + \delta u_{ie} + \varepsilon_{ie}$ e=n,u

Clearly (7) may be tested as nested within (6) constraining $\beta_n = \beta_u$ except for the intercept.

Estimates of (3) through (7), together with the associated employment status MNL and union membership probit equations, as well as the various nesting tests, are discussed in **Section 4**, after a brief description of the data and specification of the x and z vectors.

## 2. The Data.

The data are taken from two sources:

(i) the income and expenditure survey conducted by the Bureau of Market Research (BMR) at the University of South Africa in 1985. The survey covers some 4,600 African households comprising roughly 12,300 individuals, ages 16-65, from the major urban areas of South Africa.

(ii) the 1993 household survey conducted by the South African Labour and Development Research Unit (SALDRU) in conjunction with the World Bank. The survey covers some 9,000 households comprising roughly 24,000 individuals of all races, in the age range 16-65, from the whole of South Africa, including the then nominally independent states -- Transkei, Venda, Bophuthatswana and Ciskei (TVBC).

The intersection of the two sample frames is used to define a common sample here. In particular, this

confines analysis to the common major urban areas (outside of TVBC).[9] Whites are excluded from the BMR survey and in fact we chose to focus on African males alone for the present purposes.

The SALDRU, 1993, survey provides the following categories of employment: regular, casual, self-employed farmer and other self-employed.[10] "Regular" employment refers to anyone who had a regular ('permanent') wage- or salary-paying job, and includes self-employed professionals. "Casual" covers those having casual or temporary wage employment. Unfortunately there is no clear distinction between the formal and informal sectors; however, the vast majority of those in regular jobs are likely to have a more formal employment relationship, while the majority of those in the "other self-employed" category are likely to be in what is typically deemed to be the informal sector. Casual employees could be in either sector. A number of individuals engaged in more than one occupation: where one of these is a regular job, the individual is classified in this group. Where he engages in both casual wage work and has some sort of self-employment, he is classified as being in the type of work which yielded the greater income in the previous month.[11]

The BMR, 1985, survey classifies individuals as either being employed (with a regular income), self-employed (or engaged in home production) or not employed. Those doing casual or temporary work are probably classified as having no job or some sort of self employment; it is unlikely that they would appear as having a regular income or occupation. It would seem reasonable to assume then that those classified as employed are formal-sector workers, while those engaged in self employment are in the

---

[9] In principle the choice of urban or rural location ought also to be treated as endogenous. However, this is not possible, given the nature of the BMR data for 1985 which are confined to urban areas alone.

[10] Self-employed farmers were excluded from the analysis since the numbers were very small given the urban character of the sample.

[11] A further problem arises in the case of the "other self-employed": where more than one individual in a household was involved in the activity, only the total household income from this activity was recorded, together with the three household members who devoted most time to it. It was therefore assumed that the individual recorded first, who spent most time of the three, was self-employed, while the others were unpaid family helpers. Unpaid family helpers were initially included as a separate category, but the ensuing model of employment status proved incapable of distinguishing between these helpers and those without any form of employment. Unpaid family helpers are therefore pooled with the not-working group.

informal sector.[12]

Combining these distinctions with information on union membership[13], four categories of employment status are delineated for the 1985 sample and five for 1993. To place in context the four estimation methods, outlined in the previous section, it will help to note the resultant employment status of the urban, African males in our two samples:

**Table 2.1. Employment Status in the Two Samples**
(Percent)

| | **1985** | **1993** |
|---|---|---|
| Not employed | 42.7 | 43.7 |
| Informal self-employment | 2.8 | 4.3 |
| Casual wage employment | | 6.4 |
| Regular employment: not a union member | 39.8 | 24.3 |
| Regular employment: union member | 14.7 | 21.3 |

Sample-selection effects are usually considered to be important for women, since a relatively large proportion of working-age women do not participate in the labor market in many societies. However, in this South African context the fraction of adult males not employed is extraordinarily large[14] and sample selection effects may thus, potentially, be quite important for men too. Moreover, although casual and self employment is small by the standards of some developing countries, the portion of adult urban males in this category is far from negligible. It would seem that if one is to account for endogenous determination of union status, then there is at least as good a *prima facie* case for treating the labor-market participation decision and the formal-informal division in the same way.

---

[12] This sample is restricted to Africans, and is therefore unlikely to contain many self-employed professionals, particularly in the earlier, 1985, survey.

[13] The 1993 survey asked union membership directly. In the expenditure portion of the 1985 survey, information was collected on payment of union dues, thus permitting definition of a union membership dummy.

[14] Of those not employed (ages 16-70) in 1993:
12 % report being unemployed and looking for work,
29% are discouraged workers,
42% are enrolled in education,
17% are ill, disabled or retired.

In passing, it may be noted that the fraction of men reported in regular employment declined from 54.5 percent in the 1985 sample to 45.6 percent in 1993. Obviously, many factors may have contributed to this.[15] More importantly, for our purposes, the fraction of regular employees who are union members also rose substantially, from 27 percent in 1985 to 46.7 percent by 1993. Some data on the composition of the new membership will be presented in **Section 5.**

## 3. Specification.

Two sets of equations require specification: those on employment status (the probit for the switching model and the more general multinomial logit); plus the earnings equations.

### Employment Status.

The categories of the polychotomous dependent variable for the MNL equations follow the employment status measures in **Table 2.1**, using those not employed as the reference group. For the union membership probit the dependent variable is the last two of these categories -- regular employees who are union members and those who are not. The explanatory variables for the MNL are a subset of those adopted for the union membership probit, omitting from the MNL those variables which exhibit single values for one of the non-regular employee categories. In particular, the variables omitted from the MNL model thus exclude measures which are not applicable to any one category, such as sector of employment for the unemployed.

The choice of explanatory variables for inclusion must recognize that we are essentially dealing with a reduced-form model with complex criteria for selection into the various components. This reduced form reflects not only the decision of the individual to opt for a particular type of employment, but also the decision of the employer to hire him. It seems probable that the South African unions do less screening, yet even the outcome with respect to union membership among regular employees may

---

[15] It is unlikely that all of this decline can be attributed to the ambiguity in treatment of casual employees in the earlier, BMR survey. See Fallon and Pereiras de Silva (1994) and Fallon and Lucas (1996).

be influenced both by employees' decisions and those of employers in organized and uncovered sectors. This suggests that the appropriate explanatory variables are ones relevant to decisions both by employers and employees.

Employer's choices, among African males, are presumably dominated by correlates of productivity, including the standard human capital measures such as education and experience. Educational qualifications are represented here by three dummy variables for completion of primary schooling, secondary schooling, and any form of tertiary education, respectively.[16] The omitted or reference category is non-completion of primary schooling. Experience is proxied by age minus an allowance for duration of formal education, and this variable is entered in quadratic form, to allow for some degree of non-linearity in keeping with other typical findings in human capital studies.

On the individual's side, employment choices are presumably affected by both tastes and availability of unearned income. Tastes may be conditioned by many personal attributes, including age and education. Thus the educational qualification and experience measures, already mentioned in respect of productivity, may also be correlated with differing attitudes toward work. On the other hand, unearned income per capita in a workers' household is more likely to influence the individual's choices than those of the employer, and the income effect is hypothesized to reduce the choice of more onerous employment.[17] In addition, it is frequently hypothesized that the responsibilities associated with marriage alter attitudes toward work and a dummy for being single (or married, the omitted category) is accordingly included, though again this measure could trigger differing choices among employers too, if they believe that married workers are more reliable on average, for instance. For similar reasons, a dummy variable is included for whether the man is not head of the sampled

---

[16] The BMR, 1985 survey gives completion of standard 9 or 10 the same code. In this case we are forced to regard standard 9 or 10 as completion of secondary school, though strictly speaking this is correct only for standard 10. Tertiary education refers to completion of teacher training, nursing, diploma at a technical institution or university degree.

[17] 'Unearned' income refers here to all household income not earned by the man in question. This includes earnings of other household members as well as property and transfer incomes of the family, measured in thousands of Rands per week per person. Unearned income is not instrumented here, though this may not be entirely satisfactory.

household. In anticipation that there may be some correlation among unobserved factors leading to union membership of various household members, or that having a member in a unionized job may make it easier for other members to obtain such jobs, a further dummy variable appearing in both the MNL and probit is whether any other household member reports being a union member.

In addition to these explanatory variables, ten dummy variables are included, in both the MNL and probit, for the separate urban areas in the sample (Johannesburg, the eleventh urban area is the omitted category). These measures allow for variations in the structure of production, and hence differences in patterns of labor demand, while labor supply is imperfectly mobile across locations.

In the context of the union membership probit equation, the data permit three other sets of explanatory measures to be included. The first is a dummy variable for whether the man contributes to a pension scheme (1985 survey) or has pension rights as part of his package (1993 survey). Although these variables are not quite the same, they both serve as proxies for larger firm size, which can be expected to be positively correlated with union membership. Second, the reported occupations of regular employees are aggregated into four groups here — professionals, clerical and sales, mid-occupations (skilled and semi-skilled) and unskilled.[18] The first three are represented by dummy variables, 'unskilled' being the omitted, reference group. Finally the industry of employment is coded into seven major sectors, represented by dummy variables, manufacturing being the omitted group in this case.[19] Obviously, occupation, sector of employment and the pension variable are not relevant for men who are not employed and these measures are consequently excluded from the MNL specification of employment status.

---

[18] The first two categories are distinguished in both surveys. However the remainder are categorized as skilled, semi-skilled or unskilled in the BMR survey whereas the SALDRU survey classifies most of the remainder by function (e.g. mining occupations), rather than by skill. Aggregating the SALDRU survey categories into mid-occupations versus unskilled, for comparability with the BMR survey, thus required exercise of some judgement.

[19] In the 1993 survey industrial sector is coded at the one-digit level into 16 categories which have been collapsed into 7 here to correspond with the 1985 survey. From both samples, agricultural workers and domestic servants are excluded in this study, as wage determination is quite different in these sectors. The numbers in these latter groups are in any event tiny — for agricultural workers because this is an urban sample, and because domestic service is dominated by women (and has virtually no union members).

**The Earnings Equations.**

In all cases, the earnings equations are reported only for regular employees, omitting casual and self-employment. This is largely because of our ultimate interest in the pay differential for union versus non-union regular employees and especially the way this has changed over time, but also reflects the questionable nature of the data on earnings in the informal sectors in the 1985 survey.[20] The dependent variable of these estimated earnings equations is the natural logarithm of gross pay from the worker's main activity plus bonus, profit share and the value of in-kind payments, all converted to a weekly basis. In studies for the industrialized economies, it is usual to adopt hourly earnings, (rather than earnings per week used here), so as to abstract from the effect of variations in hours worked. Unfortunately, however, the BMR (1985) survey reports neither hourly earnings nor usual working hours and the SALDRU (1993) survey data on hours worked in an average day are subject to considerable reporting error.

The explanatory variables included in each of the earnings equations are a subset of those from the employment status and union membership equations, plus a set of interactions. The first groups of variables appearing in the earnings equations are the standard human capital measures: education (in dummy form, as before), experience and experience squared. The dummy variables for a single (unmarried man) and for one who is not head of household are included, to allow for the possibility that perceived differences in attitude toward work translate into lower pay. The sectoral dummies are included, since it seems well established that different industries offer quite distinct pay scales in many countries.[21] The vector of dummies for the major occupations is also included, on the grounds that these groupings may well reflect broad skill differentials. Finally, the dummy variables for the major conurbations are incorporated, though as usual it is not apparent whether any observed differences would reflect cost of living differentials, the consequences of more specific industrial structure, or some other components. Three vectors of interactions were explored: experience and its square by education; experience and its square by sector; and education by sector. The only set which proved

---

[20] For estimates of earnings equations in the informal sectors from the 1993 survey see Hofmeyr (1998).

[21] See Lang and Dickens (1992).

to be significant at the 5 per cent level (using an F-test) was education by sector and these interactions are therefore reported in each of the estimated earnings equations in the following section.

## 4. Results of Estimation.

This section turns first to the issue of nesting of the alternative forms, prior to outlining some of the more important features emerging from the coefficient estimates.

### Nesting Tests

The results of the various nesting tests are summarized in **Table 4.1**.

**Table 4.1 Nesting Hypotheses Tests for Earnings Equations**

| Estimation Method | Nested Within | Test Statistic | 1985 Non-Union | 1985 Union | 1993 Non-Union | 1993 Union |
|---|---|---|---|---|---|---|
| (7) Union Dummy | (6) Separate OLS | F | 1.258 | | 2.112 | |
| (6) Separate OLS | (3) MNL Mills | Chi-Sq | 6.44 | 9.66 | 0.37 | 2.31 |
| (6) Separate OLS | (4) Switching | Chi-Sq | 85.08 | 43.64 | 69.90 | 54.74 |
| (3) MNL Mills | (5) Both Mills | Chi-Sq | 33.36 | 3.47 | 17.25 | 7.30 |
| (4) Switching | (5) Both Mills | Chi-Sq | 0.20 | 11.57 | 0.05 | 0.04 |

An F-test is used to test the simplest OLS model with union dummy, (7), as nested within the OLS model (6) which permits a full interaction between all coefficients and the union dummy. In the 1993 sample, the nested form is rejected at the 99 percent confidence level (with degrees of freedom 38, 2888), whereas in 1985 (40, 772 degrees of freedom) confidence in rejecting the constrained form is much lower. In other words, this suggests that the union and non-unionized sectors did not offer significantly different pay profiles in 1985 (other than a pure union differential) whereas by 1993 pay structures differed significantly.

A likelihood ratio test is adopted in testing the separate OLS equations (6) as nested within the specifications including the inverse Mills ratio derived from either the multinomial employment status model or the union probit switching model. In the 1985 sample, the OLS excluding the Mills ratios

is clearly rejected as a constraint on either of these alternative specifications. In the 1993 case, the OLS form is strongly rejected in favor of the switching model but not in favor of that with sample selection correction derived from the multinomial logit model.

Finally, on a likelihood ratio test, the estimates from the MNL model are rejected as nested within the encompassing model with both Mills ratios (5), whereas the switching model is not, with the notable exception of union members in 1985 where both patterns are reversed.

On balance, the 1993 tests on nesting the OLS estimates within the MNL and switching models, plus the majority of results on nesting within the encompassing model favor the switching model. There are two major possible reasons that the multinomial logit model, which combines sample selection correction with endogenous union membership, is outperformed by the switching model which incorporates the latter — endogenous union membership — only. Either sample selection correction effects are indeed unimportant, or our model and data are insufficient to detect these effects. In view of the extraordinary non-employment rates noted in **Table 2.1** it seems relatively implausible that sample selection effects are entirely negligible.

A possible reason that the MNL model may not detect any sample selection effects is if the model does not fit well. The frequencies for actual and predicted categories from the probit and MNL estimators are shown in **Table 4.2**, with the probit results in bold.[22] The predicted categories are those having highest probability. The MNL does well in predicting the not-employed; about 90 percent of those not employed in 1985 are predicted to fall in this category, and in 1993 the comparable figure is over 85 percent. On the other hand, the MNL does less well in discerning the much smaller group in informal employment, tending to assign the self-employed as regular non-union employees and the casual workers in the 1993 sample as not employed (presumably reflecting an ambiguous commitment to the labor market on the part of many of the latter group). However, the MNL model does do quite well in discerning those in regular employment (nearly 79 percent are

[22] The total observations available for the probit estimates are somewhat fewer than for the regular employees in the MNL case because of missing values on the additional explanatory variables in the probit.

correctly allocated to this category in 1985 and over 76 percent in 1993), which suggests that the sample selection into regular employment is substantially correct.

**Table 4.2 Actual and Predicted Outcomes of MNL and Probit Estimates**

| | 1985 | | | | | 1993 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | PREDICTED | | | | | PREDICTED | | | | | |
| ACTUAL | Not Employed | Self Employed | Non-Union | Union | TOTAL | Not Employed | Self Employed | Casual Worker | Non-Union | Union | TOTAL |
| Not Employed | 2112 | 0 | 212 | 30 | 2354 | 827 | 2 | 0 | 94 | 47 | 970 |
| Self Employed | 21 | 0 | 132 | 1 | 154 | 22 | 2 | 0 | 54 | 17 | 95 |
| Casual Worker | | | | | | 60 | 0 | 0 | 49 | 34 | 143 |
| Non-Union | 453 | 0 | 1697<br>**2049** | 41<br>**115** | 2191<br>**2164** | 153 | 1 | 0 | 297<br>**316** | 89<br>**107** | 540<br>**423** |
| Union | 186 | 0 | 547<br>**610** | 78<br>**198** | 811<br>**808** | 88 | 0 | 0 | 147<br>**119** | 238<br>**316** | 473<br>**435** |
| TOTAL | 2772 | 0 | 2588 | 150 | 5510 | 1150 | 5 | 0 | 641 | 425 | 2221 |

**Probit results in bold**

**Percent Correct Predictions on Union Membership**

| | 1985 | | 1993 | |
|---|---|---|---|---|
| | Non-Union | Union | Non-Union | Union |
| MNL | 77.5 | 9.6 | 55.0 | 50.3 |
| MNL conditional on regular employment | 97.6 | 12.5 | 76.9 | 61.8 |
| Probit | 94.7 | 24.5 | 74.7 | 72.6 |

On the other hand, the MNL estimates are outperformed by the probit in terms of predicting union membership among the regular employees, as may be seen from the lower panel of **Table 4.2**. The MNL actually does well in terms of assigning the non-union members, and indeed the probit does not improve on these predictions conditional upon being predicted to be a regular employee in the MNL. Rather, it is in discerning the union members themselves where the MNL is dominated. This is presumably largely because of the additional explanatory variables available to the probit (the pension contribution dummy, occupation and sector dummies which must be omitted from the MNL given the absence of values on these measures for men who are not employed). However, it is important

to note that it is in the 1985 case that the MNL does particularly poorly in predicting union membership, whereas it is in the 1993 sample that the MNL inverse Mills correction to the earnings equations proves statistically insignificant. (See **Table 4.1**).

A second possible reason that the MNL model may fail to detect sample selection effects is if we have too few identifying variables included in the selection equation but excluded from the earnings equation. (See Lewis, 1986). Indeed, only two such variables are available in the data for the MNL model, namely household income per capita and the dummy for another union member present in the household. For the probit model an additional measure is available — the pension contribution dummy.

Thus, on balance, the union membership probit is not dramatically superior in terms of fit. Rather, the distinction between the statistical confidence in the Mills correction from the probit, as distinct from the MNL case, simply points to the sensitivity of these corrections to precise specification.

**Earnings Equations Estimates.**

Some features of the estimated earnings equations are worth noting at this juncture, especially for reference when we turn to the union differentials in **Section 5.** In describing the earnings equations results, this section focuses upon the estimates from the switching model which are presented in **Table 4.3**.[23] The multinomial logit model of employment status and the union membership probit equation are relegated to the appendix as **Tables A.1** and **A.2**[24], while the remaining three approaches

---

[23] In all tables of regression results, t-statistics for a zero null hypothesis immediately follow coefficient estimates.

[24]Those without employment of any sort form the omitted, or reference, category for the dependent variables in the MNL equations and the coefficients for the non-employed category are normalized to zero, implying that the coefficients presented give log-odds ratios relative to this group. In the probit estimates the omitted, reference group is the set of non-union members.

to the earnings equations appear as **Tables A.3** through **A.5**.[25] It may, however, be noted at the outset that the general properties of the estimates derived from the MNL model and the OLS on separate regimes differ little from those in the switching model. Indeed most of the exceptions occur in cells with only small numbers of observations, such as in the interactions between the tertiary education dummy and the sectoral dummies.

[25] Standard errors are computed using a robust estimator, to correct for potential heteroskedasticity in the data and in the two-stage MNL procedure. In contrast, the switching model reported in Table 4.3 adopts a maximum likelihood estimator of the two equation switching model, with built in adjustment for heteroskedasticity. However, for the test on nesting the switching model within the encompassing model, summarized in Table 4.1, both are estimated on the same basis as the earnings equations in the appendix. See Maddala (1983: 252-255) and Greene (1992: 618-620).

## Table 4.3 Earnings Equations from Switching Model

| | 1985 | | | | 1993 | | | |
|---|---|---|---|---|---|---|---|---|
| | Non-Union | | Union Members | | Non-Union | | Union Members | |
| Constant | 3.940 | 124.79 | 4.179 | 86.50 | 4.786 | 28.30 | 5.617 | 40.46 |
| Primary Education | 0.088 | 2.32 | 0.095 | 2.15 | 0.120 | 1.07 | -0.048 | -0.68 |
| Secondary Education | 0.240 | 4.26 | 0.297 | 4.10 | 0.533 | 3.65 | 0.093 | 0.95 |
| Tertiary Education | 0.862 | 4.16 | 0.158 | 0.62 | 1.575 | 5.16 | -0.265 | -0.68 |
| Experience (yrs) | 0.018 | 5.28 | 0.018 | 3.86 | 0.037 | 3.89 | 0.025 | 3.02 |
| Experience $^2$/1000 | -0.323 | -4.81 | -0.311 | -3.24 | -0.805 | -4.13 | -0.384 | -2.37 |
| Single | -0.113 | -4.05 | -0.073 | -1.90 | -0.035 | -0.46 | -0.029 | -0.44 |
| Not Household Head | -0.107 | -3.88 | -0.086 | -2.14 | -0.161 | -2.24 | 0.080 | 1.16 |
| Professional | 0.558 | 8.26 | 0.467 | 3.60 | 0.403 | 2.39 | 0.269 | 2.10 |
| Clerical and Sales | 0.416 | 10.91 | 0.242 | 4.40 | 0.166 | 1.72 | 0.203 | 2.44 |
| Mid-Occupations | 0.385 | 16.95 | 0.328 | 10.18 | 0.074 | 1.15 | 0.055 | 1.00 |
| Mining | 0.296 | 4.69 | -0.234 | -1.21 | -0.130 | -0.71 | -0.182 | -1.97 |
| Electricity, Gas, Water | 0.045 | 0.33 | 0.103 | 0.64 | -0.075 | -0.27 | -0.018 | -0.07 |
| Construction | 0.063 | 1.15 | -0.084 | -1.18 | 0.205 | 1.57 | -0.150 | -1.12 |
| Trade | -0.073 | -1.54 | -0.083 | -1.30 | 0.067 | 0.47 | -0.011 | -0.07 |
| Transport | 0.021 | 0.33 | 0.048 | 0.55 | 0.090 | 0.65 | -0.324 | -2.86 |
| Services | -0.090 | -1.94 | -0.137 | -1.60 | 0.013 | 0.08 | -0.027 | -0.22 |
| Primary * Mining | -0.072 | -0.88 | 0.309 | 1.40 | 0.385 | 1.73 | 0.105 | 1.03 |
| Secondary * Mining | -0.047 | -0.41 | 0.568 | 2.29 | -0.110 | -0.24 | 0.137 | 0.91 |
| Tertiary * Mining | -0.494 | -1.34 | | | -0.097 | -0.18 | 0.634 | 1.16 |
| Primary * EGW | 0.028 | 0.15 | -0.260 | -1.29 | 0.439 | 1.35 | -0.016 | -0.05 |
| Secondary * EGW | 0.454 | 1.17 | | | 0.157 | 0.29 | 0.297 | 0.89 |
| Primary * Construction | -0.088 | -1.27 | -0.065 | -0.67 | -0.255 | -1.60 | 0.355 | 2.08 |
| Secondary * Construction | 0.106 | 0.95 | 0.246 | 1.36 | -0.379 | -1.66 | -0.361 | -1.44 |
| Tertiary * Construction | -0.177 | -0.40 | | | | | | |
| Primary * Trade | -0.005 | -0.09 | -0.013 | -0.18 | 0.003 | 0.02 | 0.059 | 0.36 |
| Secondary * Trade | 0.241 | 2.82 | 0.116 | 0.99 | -0.366 | -1.78 | 0.123 | 0.56 |
| Tertiary * Trade | 0.144 | 0.55 | 0.849 | 1.89 | 0.775 | 1.40 | 0.922 | 1.89 |
| Primary * Transport | 0.091 | 1.20 | -0.072 | -0.67 | 0.060 | 0.34 | 0.394 | 2.85 |
| Secondary * Transport | -0.049 | -0.44 | 0.093 | 0.61 | -0.194 | -0.88 | 0.517 | 2.65 |
| Tertiary * Transport | -0.103 | -0.34 | | | | | | |
| Primary * Services | 0.088 | 1.62 | -0.017 | -0.17 | 0.003 | 0.02 | 0.159 | 1.00 |
| Secondary * Services | 0.326 | 4.53 | 0.219 | 1.81 | 0.031 | 0.16 | 0.295 | 1.76 |
| Tertiary * Services | 0.019 | 0.09 | 0.456 | 1.39 | -0.688 | -2.11 | 0.907 | 2.17 |
| Pretoria | -0.027 | -0.72 | -0.065 | -1.15 | 0.256 | 2.20 | 0.277 | 2.77 |
| East-West Rand | -0.083 | -2.32 | -0.046 | -0.90 | 0.032 | 0.43 | -0.065 | -1.00 |
| Vaal Triangle | -0.077 | -1.94 | -0.110 | -2.03 | 0.079 | 0.86 | 0.000 | -0.01 |
| Cape Peninsular | -0.147 | -3.38 | -0.149 | -2.66 | -0.032 | -0.31 | -0.144 | -1.44 |
| Durban | -0.148 | -3.62 | -0.247 | -4.73 | 0.062 | 0.69 | 0.008 | 0.09 |
| Pietermaritzburg | -0.322 | -6.16 | -0.285 | -4.76 | -0.219 | -1.48 | -0.016 | -0.13 |
| Port Elizabeth | -0.238 | -5.23 | -0.186 | -2.86 | -0.071 | -0.46 | -0.338 | -2.50 |
| Bloemfontein | -0.201 | -4.72 | -0.396 | -4.82 | -0.035 | -0.34 | -0.089 | -0.87 |
| Klerksdorp | -0.135 | -3.22 | -0.143 | -1.36 | 0.579 | 2.98 | 0.337 | 2.55 |
| Free State Gold | -0.131 | -3.03 | -0.207 | -2.99 | -0.216 | -1.31 | 0.008 | 0.09 |
| Inverse Mills Ratio | -0.316 | -6.11 | 0.052 | 1.25 | -0.368 | -4.16 | -0.206 | -3.04 |
| Number Observations | 2164 | | 808 | | 423 | | 435 | |
| R Bar Squared | .442 | | .352 | | .400 | | .222 | |
| F Statistic | 39.93 | | 11.99 | | 7.70 | | 3.94 | |
| Error Correlation with Probit | -.719 | | .150 | | -.713 | | -.533 | |

T-statistics follow coefficient estimates

In each of the four estimates in **Table 4.3** the usual pattern is discerned of earnings rising with experience then ultimately declining. However these profiles differ between the two samples. In 1985, the implied turning points occur at roughly similar levels of experience for the union and non-union members (around 28 years), and there is little difference in curvature between the two groups. In contrast, by 1993, the turning point in earnings for union members is after some 33 years of experience which is about ten years later than for non-members. Moreover, the earnings-experience profile is much flatter among union members in 1993, indicating wage compression within the unions.

Within each of the four estimates in **Table 4.3**, earnings generally rise with level of skill associated with the occupational group, as might be anticipated. Within both samples the range from unskilled (the reference group) to professional is somewhat smaller among union members, again reflecting wage compression by the unions. In addition, this range seems to have narrowed, within the union and non-union categories, from 1985 to 1993 (though this could also reflect the aforementioned lack of comparability of the unskilled groups as between the two samples).

The educational and sectoral dummies (with manufacturing omitted as the reference group) are included separately and in interacted form. The coefficients on the education dummies themselves thus refer to the effects of education within manufacturing, with 'primary education incomplete' omitted as the reference group. Among non-union members, earnings rise steadily with education completed as one might expect, and indeed the estimated returns to education rose in 1993 as compared to 1985. In 1985, the returns to primary and secondary completion, among union members in manufacturing, are fairly similar to those for non-members.[26] On the other hand, the data indicate no significant gain in earnings among union members in manufacturing upon completion of primary or secondary education, by 1993, which would be consistent with the aim of South Africa's unions to achieve a more egalitarian wage distribution.

On the whole, these patterns persist in the other sectors also, though the returns to completion of

---

[26] The number of African union members in manufacturing with a tertiary education, in 1985, was quite tiny.

both primary and secondary school do prove somewhat greater among union members in the mining sector in 1985 and in transport in 1993. Indeed, the sectoral dummies suggest that earnings, *ceteris paribus*, differ only to a small extent across sectors in both samples. In part, this may reflect lack of precision in our estimates, and it is notable that the sectoral differences appear somewhat more significant after controlling for sample selection in the multinomial case (see **Table A.3**), though most of the absolute differences appear small by international standards, irrespective of estimation method.

Finally, **Table 4.3** includes a set of dummies for the main metropolitan areas with Johannesburg as the omitted category. In 1985, almost all of the urban areas are estimated to have had significantly lower pay than did Johannesburg and indeed the data indicate substantial variations among the included metropolitan areas themselves (even for union workers). On the other hand, at least the differences from Johannesburg are estimated to be somewhat smaller in the 1993 sample, among both union and non-union workers, though it remains true that point estimates for some areas (such as Klerksdorp) still indicate significant, remaining, geographical earnings dispersion.

An important theme which runs through much of the foregoing account is the wage compression which appears in the 1993 data, especially among union members. However, it remains to consider, in the following section, what role this compression played in the earnings premium received by union members.

## 5. The Change in Union Earnings Premium.

This section first raises some methodological issues before turning to the results on union earnings premia implied by the foregoing results. A final subsection then addresses decomposing the change in the premium between the two samples.

### Three Methodological Issues.

First, it should be emphasized that our objective is to evaluate the change in premia which union members earn as compared to other regular employees. The estimation of union premia has been the subject of a considerable body of literature. Nonetheless, it should be borne in mind that this is not

the same thing as estimating the effect of expanded union membership on the cost of regular employees, if only because the pay of non-members may also have been changed by the growth in unions. Moreover, the current focus does not address the premia which regular union members earn relative to those in more informal employment, which might be considered more closely akin to the shadow premium wage. In the present context this latter comparison has not been undertaken in large part because of the problems with the data on earnings in the less formal sectors, mentioned earlier.

Second, brief note may be made of an ambiguity in defining a premium in those contexts where separate pay structures are permitted in estimates for union members and non-members respectively. One may either ask how much less union members would earn if they were not members, based on the mean characteristics of members, or how much more non-members would earn if they joined a union, based on the mean characteristics of non-members. Both sets of results are reported in the following subsection.

Third, two alternative treatments of the Mills ratio terms are considered here in predicting earnings. If the Mills ratio term is set to zero in making these projections, then the interpretation of the estimated union premium is simply the gap in earnings predicted, using either the mean observable characteristics of union or non-union members, ($x_{ic}$ in the earnings equations). The purpose of including the Mills ratio in the initial estimates is then only to obtain unbiased estimates of the coefficients on the remaining included variables. However, a correlation between the errors in the earnings equations and the decision equation (the MNL on employment status or the union membership probit) may well arise, and hence a non-zero coefficient on the Mills ratio, from the role of unobserved attributes of the individual which are common to both the decision and earnings equations. In this case, including the relevant mean value for the Mills ratio in predicting earnings provides a different interpretation of the resultant premium in which both the observed vector of attributes ($x_{ic}$) and any unobserved attributes reflected in the covariance of errors ($\sigma_{\varepsilon\eta}$) are held constant in comparing earnings. In other words, including the mean Mills ratio offers a potentially broader comparison, though one where precisely what is held constant in comparing is of necessity unknown. Both forms are reported here.

**Results on the Estimated Premium**

For the urban African males in our samples, the percentage by which the simple average earnings of union members exceed those of other regular employees, who are not union members, rose from 8 percent in 1985 to 26.5 percent in 1993. The premia estimated from the various permutations on our four estimation approaches — with characteristics of union members or non-members, and with the Mills ratio set to the mean value or to zero-- are given in **Table 5.1**. The range of these estimates proves to be wide with some values more plausible than others. However, there is no clear pattern as to whether the characteristics of union members or non-members, or to the mean Mills ratio being included, generates the larger estimates.

**Table 5.1 Estimated Union Premia**
**(Percent)**

| | 1985 | | | | | 1993 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Mills Ratio Set to** | | **Zero** | | **Mean** | | | **Zero** | | **Mean** | |
| **Characteristics of Union Member or Non-Member** | | **Non-Mem** | **Member** | **Non-Mem** | **Member** | | **Non-Mem** | **Member** | **Non-Mem** | **Member** |
| **Difference in Sample Means** | 8.0 | | | | | 26.5 | | | | |
| **OLS with Union Dummy** | 6.4 | | | | | 19.8 | | | | |
| **OLS on Separate Regimes** | | 6.2 | 5.9 | | | | 21.9 | 15.2 | | |
| **MNL Model** | | -6.1 | -3.9 | -5.0 | -1.9 | | 56.1 | 44.7 | 23.7 | 17.1 |
| **MNL excluding 'Other Union Member in HH'** | | 32.8 | 34.5 | 8.3 | -8.7 | | 99.6 | 89.8 | 23.6 | 22.2 |
| **Switching Model** | | 12.6 | 17.5 | -2.3 | 71.9 | | 80.5 | 82.5 | 63.7 | 100.6 |

The estimates from the multinomial logit sample selection model and the probit switching model of union membership differ substantially from the differences in the sample means, which may either reflect the importance of sample selection and endogeneity of union membership or imprecision in these models. Certainly the estimated negative premia from the MNL model in 1985 are implausible and the extreme sensitivity of these estimates to precise specification should be emphasized. For

instance, a second set of estimates from the MNL model are reported in **Table 5.1** in which a single variable -- 'other union member in household' is omitted from the employment status equations and the implied premia change very substantially for several of the possible permutations.

One systematic pattern does emerge clearly from the results in **Table 5.1**. In pair-wise comparisons, the premium is larger in 1993 than in 1985, in every single case. The switching model and the MNL model with the Mills ratio excluded generate particularly large estimated premia for 1993, (in which year, it may be recalled, both models performed relatively well in predicting union membership). As a result, the increase in the estimated premium from 1985 to 1993 is even larger than the increase in the difference in sample means for these approaches, though not in the case of the two OLS models.

**Decomposing the Change in the Premium over Time**.

The results presented in **Table 5.1** are based on the characteristics of workers within the relevant sample. Thus, some or all of the increment to the estimated premium between the two sample years may have reflected a change in composition of the labor force in regular employment and of the rapidly rising union membership in particular.

In our two samples, the rate of union membership among regular employees rose from 27 percent in 1985 to 46.7 percent by 1993. (See **Table 5.2**). The rise in membership rates was particularly steep for those with no education and for those at the opposite end of the spectrum with a tertiary education (though the sample size is small among the latter). Corresponding to this, sharp rises in membership occurred among those with unskilled jobs and within the professions, though union rates among clerical and sales workers also rose rapidly. In addition, the mean level of experience of union members rose slightly between the two samples while that of other regular employees fell. Perhaps most importantly though, the changes in union membership varied substantially across sectors, rising very rapidly in mining and manufacturing but actually falling slightly in construction according to our data.

**Table 5.2 Union Membership Rates Among Regular Employees**
(African urban males. Percent)

| | 1985 | 1993 |
|---|---|---|
| **OVERALL** | 27.0 | 46.7 |
| **Primary Educ Incomplete** | 25.0 | 53.1 |
| **Primary Education** | 29.4 | 52.0 |
| **Secondary Education** | 26.4 | 45.8 |
| **Tertiary Education** | 11.4 | 39.3 |
| **Unskilled Occupations** | 22.6 | 41.5 |
| **Mid-Occupations** | 30.9 | 54.0 |
| **Clerical and Sales** | 25.0 | 46.3 |
| **Professional** | 13.8 | 52.6 |
| **Manufacturing** | 36.7 | 61.5 |
| **Mining** | 9.2 | 85.4 |
| **Electricity, Gas and Water** | 43.8 | 47.6 |
| **Construction** | 26.5 | 24.7 |
| **Trade** | 30.8 | 34.2 |
| **Transport** | 27.5 | 44.6 |
| **Services** | 15.2 | 37.2 |

For the present purposes of evaluating the change in the union premium, it is these simple correlations which matter. It should be noted though, in passing, that not all of these patterns are discerned in the multivariate results from either the union membership probit or the MNL employment status estimates. (See **Table A.1** and **A.2**). For instance, the probit estimates reveal no statistically significant effect of either education or experience on union membership, *ceteris paribus*. In contrast, the MNL results do suggest a positive effect of both education and experience (through about 30 years of experience) on union membership among regular employees in 1993 though not in 1985. In addition, the MNL results clearly indicate the higher odds of being in regular employment as education rises in both sample years, while the chances of being in employment at first rise then taper off with greater experience, with implied turning points in the range of some 20 to 25 years of experience. Other factors which prove significant in affecting union membership in the probit model include the pension contribution dummy, being head of household, having another union member in

the household, sector of employment, and to some extent urban location.

Changes in the mean characteristics of union members and of other regular employees presumably contributed to the estimated changes in the union premium noted in **Table 5.1**, though it is not obvious whether the increment would be less, or even greater, in the absence of these changes in composition. To examine this, the change over time in the log ratio of earnings of members relative to other regular employees may be decomposed as follows:

8. $(w_{u1} - w_{n1}) - (w_{u0} - w_{n0}) = (b_{u1} - b_{n1})(x_1 - x_0) + [(b_{u1} - b_{n1}) - (b_{u0} - b_{n0})]\, x_0$

or

$$(w_{u1} - w_{n1}) - (w_{u0} - w_{n0}) = (b_{u0} - b_{n0})(x_1 - x_0) + [(b_{u1} - b_{n1}) - (b_{u0} - b_{n0})]\, x_1$$

where subscripts u and n represent union and non-union categories respectively, as before
subscripts 0 and 1 represent the 1985 and 1993 samples respectively

$w_{et}$ = predicted mean log earnings for e = n,u in time period t=0,1
$b_{et}$ = the row vector of estimated coefficients for the earnings equation
for e = n,u and t =0,1
$x_t$ = the mean characteristics of either union members or other regular employees
in time period t.

The results of this decomposition for the OLS on separate regimes, the multinomial logit model (with and without 'other union member in household') and for the switching model are presented in **Table 5.3**.[27]

[27] In these particular results, the Mills ratio terms are set to zero, though the general pattern of results is not much affected by including these terms decomposed in the same fashion.

**Table 5.3 Decomposition of Change in Estimated Log Earnings Differential**

| | Characteristics of Mean Non-Member | | | | |
|---|---|---|---|---|---|
| | $(b_{u1}-b_{n1})(x_1-x_0)$ | $[(b_{u1}-b_{n1})-(b_{u0}-b_{n0})]x_0$ | $(b_{u0}-b_{n0})(x_1-x_0)$ | $[(b_{u1}-b_{n1})-(b_{u0}-b_{n0})]x_1$ | Sum |
| **OLS Separate Regimes** | -0.048 | 0.186 | -0.028 | 0.165 | 0.138 |
| **MNL Model** | -0.051 | 0.559 | -0.017 | 0.525 | 0.508 |
| **MNL excl. 'other union mem'** | -0.069 | 0.476 | -0.024 | 0.431 | 0.407 |
| **Switching Model** | -0.052 | 0.524 | -0.014 | 0.486 | 0.472 |
| | | | | | |
| | **Characteristics of Mean Union Member** | | | | |
| | $(b_{u1}-b_{n1})(x_1-x_0)$ | $[(b_{u1}-b_{n1})-(b_{u0}-b_{n0})]x_0$ | $(b_{u0}-b_{n0})(x_1-x_0)$ | $[(b_{u1}-b_{n1})-(b_{u0}-b_{n0})]x_1$ | Sum |
| **OLS Separate Regimes** | -0.113 | 0.198 | -0.044 | 0.129 | 0.084 |
| **MNL Model** | -0.137 | 0.547 | -0.073 | 0.483 | 0.410 |
| **MNL excl. 'other union mem'** | -0.129 | 0.473 | -0.082 | 0.427 | 0.344 |
| **Switching Model** | -0.049 | 0.489 | -0.081 | 0.521 | 0.440 |

There is uniform agreement with respect to the sign of $(b_{ut}-b_{nt})(x_1-x_0)$, no matter whether t refers to the estimated coefficients for 1985 or 1993, and irrespective of whether x refers to the characteristics of union members or other regular employees. In each instance this term proves negative in **Table 5.3.** In other words, the previously noted changes in composition of the labor force — no matter whether it is union members or other regular employees in question — served to reduce the level of the average union premium according to these estimates.

There is also uniform agreement that the sign of $[(b_{u1}-b_{n1})-(b_{u0}-b_{n0})]\,x_t$ is positive. It was the change in pay structures among union members relative to non-members that resulted in the overall rise in estimated premia, not the change in composition of employees at least in-so-far as indicated by our measures of workers' characteristics. It may be noted that contribution of the change in coefficients on experience and its square, as well as on the education dummies (separately or combined with the sectoral interaction), served to diminish the rise in union premia for each permutation in **Table 5.3.** Apparently the tendency toward wage compression, across age and education levels, among union members by 1993, noted in **Section 4**, had the net effect of lowering the union earnings differential.

But these effects were obviously more than offset by the changes in other coefficients.

## 6. Concluding Reflections.

This paper explores the change in earnings premium received by African, urban, male union members in South Africa, relative to other regular employees with comparable measured characteristics, at two points in time during a period of very rapidly expanding union coverage. The four approaches adopted to estimation allow for a single earnings regime with union dummy variable, for the possibility of different earnings regimes among union members as compared to non-members, for endogenous switching between such regimes, and for sample selection arising from lack of employment and from division between regular employees and informal workers.

The results suggest that by 1993 the pay structure of union members exhibited compression of earnings relative to non-members, specifically with respect to experience and education, whereas in 1985 the covered and uncovered sectors differed less in earnings patterns as reflected in the specifications adopted here. Allowing for endogenous switching between earnings regimes nonetheless significantly affected estimates for both sample years. On the other hand, the attempt to allow for sample selection into regular employment, using a multinomial logit model of employment status, hardly altered statistical confidence in the earnings estimates derived form the switching model. Given the high level of non-employment within these samples, as well as significant participation in informal employment, we doubt that sample selection issues are really irrelevant. However, the estimates with sample selection prove very sensitive indeed to precise specification of the model, as does the estimated union premium.

Several permutations on the premium are explored, derived from: the four approaches to estimating the earnings equations; whether characteristics of non-members are used to predict how much more such workers would earn as union members, or if union members characteristics are used, instead, to evaluate how much more they enjoy in the covered sectors; and whether the Mills ratio terms from the union membership probit or multinomial logit employment status equations are set to zero or to

the sample mean (potentially allowing for the role of unobserved effects resulting in correlated error structures). Considerable disparity is found in the estimates among these permutations, yet no obvious, systematic pattern emerges as to which permutations offer larger estimates. However, there is uniform agreement that the premium had risen by 1993 and both models incorporating the Mills ratio correction suggest a rise in premium which was even larger than the increment to the union differential in sample mean earnings.

A decomposition of the change in estimated premium between samples shows that the change in measured characteristics of the work force tended to diminish the premium. Instead, the rise in estimated premia is attributable entirely to the change in the structure of the estimated earnings regimes. This latter effect is despite the role of wage compression, across education and experience levels among union members by 1993, in diminishing the wage premium. In other words, the observed rise in union premium must either reflect an increase in rents to union members as compared to identical non-member workers or some change in the composition of union members which we are unable to detect in our data.

## References

Blau, David .M. "Self-employment, earnings, and mobility in Peninsular Malaysia", World Development, v.14, 839-852, July 1986.

Calvo, Guillermo A. "Urban unemployment and wage determination in LDCs: trade unions in the Harris-Todaro model," International Economic Review, v.19, 65-81, February, 1978.

Donald, Stephen G., David A. Green and Harry J. Paarsch. "A semi-parametric estimator of distribution functions with covariates: and application to earnings and wage distributions between Canada and the United States", mimeo, July 1995.

Fallon, Peter and Robert E.B. Lucas. "South African labor markets: adjustment and inequalities". Mimeo. World Bank: Washington DC. 1997.

Fallon, Peter and Luiz A. Pereiras de Silva. South Africa: Economic Performance and Policies. World Bank Southern Africa Department Discussion Paper No. 7. 1994.

Fuchs, Victor R., Alan B. Krueger and James M. Poterba "Why do economists disagree about policy? The roles of beliefs about parameters and values", Working Paper No. 389, Industrial Relations Section, Princeton University, August, 1997.

Greene, William H. LIMDEP: User's Manual and Reference Guide Version 6.0. Econometric Software: New York. 1992

Hofmeyr, Julian F. An analysis of African wage movements in South Africa. Research Monograph No. 9, Economic Research Unit, University of Natal Durban. 1994.

Earnings differentials between the unionised, formal and informal sectors in South Africa. Occasional Paper No. 33, Economic Research Unit, University of Natal Durban. 1998.

Horner, Dudley. "Registered unions in South Africa", (with unpublished addendum on African Trade Unions), South African Institute of Race Relations, Johannesburg, 1974.

Lang, Kevin and William T. Dickens. "Labor market segmentation, wage dispersion and unemployment", Working Paper No. 4073, National Bureau of Economic Research, 1992.

Lee, Lung-Fei. "Unionism and relative wages: a simultaneous model with qualitative and limited dependent variables". International Economic Review, v.19, 415-433, 1978.

Lewis, H. Gregg. "Union relative wage effects", in Orley Ashenfelter and Richard Layard, eds., Handbook of Labor Economics, Vol. II. North-Holland: Amsterdam. 1986.

Maddala, G.S. Limited Dependent and Qualitative Variables in Econometrics. Cambridge University Press: Cambridge. 1993.

Moll, Peter G. "Black South African unions: relative wage effects in international perspective", Industrial and Labor Relations Review, v.46, 245-261, January, 1993.

Nel, P.S. and P.H. van Rooyen. South African Industrial Relations: Theory and Practice. J.L. van Schaik Academic Press: Pretoria. 1993.

Schultz, T. Paul and Germano Mwabu, "The effects of labor unions on income distribution in South Africa", mimeo, prepared for presentation at the meetings of the European Society for Population Economics, Uppsala, Sweden, June 13-15, 1996.

# APPENDIX

## Table A.1 Multinomial Logit Model of Employment Status

1985

| | Self-Employed | | Regular Employee Non-Union | | Regular Employee Union Member | |
|---|---|---|---|---|---|---|
| **Constant** | -3.036 | -5.35 | 1.230 | 5.42 | -0.368 | -1.29 |
| **Primary Education** | 0.524 | 2.50 | 0.403 | 4.06 | 0.620 | 5.13 |
| **Secondary Education** | 1.565 | 5.14 | 1.026 | 7.21 | 0.959 | 5.44 |
| **Tertiary Education** | -9.529 | -0.05 | 2.260 | 5.78 | 1.178 | 2.38 |
| **Experience (yrs)** | 0.294 | 8.43 | 0.196 | 17.00 | 0.205 | 13.29 |
| **Experience $^2$ /1000** | -6.092 | -9.35 | -4.811 | -20.15 | -5.125 | -16.04 |
| **Household Income** | -0.280 | -2.54 | -0.380 | -7.13 | -0.395 | -5.68 |
| **Single** | -0.669 | -2.20 | -0.578 | -4.17 | -0.546 | -3.30 |
| **Not Household Head** | -2.944 | -8.97 | -2.867 | -21.09 | -3.155 | -18.69 |
| **Other Union Member in HH** | -0.900 | -1.89 | -0.480 | -3.55 | 1.257 | 9.08 |
| **Pretoria** | 0.498 | 1.56 | -0.096 | -0.62 | 0.079 | 0.38 |
| **East-West Rand** | 0.002 | 0.01 | -0.258 | -1.61 | 0.232 | 1.16 |
| **Vaal Triangle** | -0.152 | -0.40 | -0.299 | -1.63 | 0.395 | 1.82 |
| **Cape Peninsular** | -0.707 | -1.52 | -0.334 | -1.90 | 0.496 | 2.34 |
| **Durban** | -0.746 | -1.86 | -0.543 | -3.35 | 0.240 | 1.22 |
| **Pietermaritzburg** | -0.293 | -0.66 | -0.640 | -3.06 | 0.488 | 2.06 |
| **Port Elizabeth** | 0.165 | 0.46 | -0.584 | -3.04 | -0.131 | -0.54 |
| **Bloemfontein** | -0.531 | -1.06 | 0.033 | 0.17 | -0.318 | -1.10 |
| **Klerksdorp** | -0.015 | -0.03 | 0.639 | 3.53 | -0.395 | -1.32 |
| **Free State Gold Fields** | -0.784 | -1.56 | -0.145 | -0.74 | -0.157 | -0.61 |

**Number of Observations** 5510: **Chi-Squared** 3756 (DF=57)

1993

| | Self-Employed | | Casual | | Regular Non-Union | | Regular Union Member | |
|---|---|---|---|---|---|---|---|---|
| **Constant** | -2.231 | -3.55 | -0.559 | -1.21 | -0.126 | -0.37 | -1.289 | -3.26 |
| **Primary Education** | 0.273 | 1.03 | -0.197 | -0.89 | 0.067 | 0.44 | 0.272 | 1.59 |
| **Secondary Education** | 0.634 | 1.61 | 0.317 | 1.09 | 0.797 | 3.80 | 1.012 | 4.34 |
| **Tertiary Education** | -10.789 | -0.04 | 1.325 | 1.80 | 2.206 | 4.12 | 2.304 | 3.81 |
| **Experience (yrs)** | 0.218 | 4.96 | 0.116 | 3.75 | 0.158 | 7.58 | 0.204 | 8.22 |
| **Experience $^2$ /1000** | -4.658 | -5.27 | -3.602 | -5.07 | -3.977 | -9.08 | -4.734 | -9.31 |
| **Household Income** | 0.405 | 0.70 | -2.859 | -2.42 | -0.778 | -1.66 | -5.049 | -5.42 |
| **Single** | 0.175 | 0.51 | -0.591 | -2.12 | -0.902 | -4.64 | -1.005 | -4.54 |
| **Not Household Head** | -2.687 | -7.00 | -1.708 | -5.97 | -1.961 | -9.88 | -2.472 | -10.56 |
| **Other Union Mem. in HH** | -0.489 | -1.26 | -0.301 | -1.06 | -0.946 | -4.36 | 1.025 | 5.30 |
| **Pretoria** | -0.327 | -0.68 | -0.144 | -0.31 | 0.230 | 0.75 | 0.352 | 0.99 |
| **East-West Rand** | -0.981 | -2.88 | 0.420 | 1.55 | 0.065 | 0.31 | 0.827 | 3.59 |
| **Vaal Triangle** | -0.955 | -2.20 | -0.258 | -0.67 | -0.021 | -0.08 | 0.006 | 0.02 |
| **Cape Peninsular** | -2.108 | -2.03 | 0.294 | 0.71 | 0.929 | 3.36 | 0.933 | 2.81 |
| **Durban** | -1.029 | -2.27 | 0.017 | 0.05 | 0.184 | 0.77 | 0.379 | 1.38 |
| **Pietermaritzburg** | -1.124 | -1.44 | -12.742 | -0.05 | -0.031 | -0.08 | 0.402 | 0.96 |
| **Port Elizabeth** | -1.138 | -2.00 | -0.283 | -0.68 | -0.620 | -1.87 | 0.147 | 0.42 |
| **Bloemfontein** | -0.719 | -1.59 | -1.383 | -2.43 | 0.205 | 0.78 | -0.359 | -1.05 |
| **Klerksdorp** | 0.413 | 0.47 | -11.892 | -0.04 | 1.548 | 2.92 | 1.470 | 2.55 |
| **Free State Gold** | -1.978 | -1.84 | -0.938 | -1.17 | -0.233 | -0.53 | 2.297 | 6.00 |

**Number of Observations** 2221: **Chi-Squared** 1476 (DF=76). **T-statistics follow coefficient estimates**

## Table A.2 Probit Model of Union Membership

| | 1985 | | 1993 | |
|---|---|---|---|---|
| **Constant** | -1.053 | -6.28 | -0.545 | -1.72 |
| **Primary Education** | 0.040 | 0.62 | 0.000 | 0.00 |
| **Secondary Education** | -0.041 | -0.40 | -0.101 | -0.62 |
| **Tertiary Education** | -0.531 | -1.90 | -0.391 | -1.07 |
| **Experience (yrs)** | -0.005 | -0.49 | -0.001 | -0.06 |
| **Experience $^2$/1000** | -0.015 | -0.08 | 0.176 | 0.44 |
| **Household Income** | -0.044 | -1.09 | -1.125 | -1.78 |
| **Pension Contribution Dummy** | 0.607 | 9.07 | 0.880 | 7.12 |
| **Single** | 0.005 | 0.06 | -0.014 | -0.08 |
| **Not Household Head** | -0.227 | -2.62 | -0.449 | -2.69 |
| **Other Union Member in HH** | 1.106 | 12.04 | 1.130 | 6.94 |
| **Professional** | -0.235 | -1.04 | 0.429 | 1.33 |
| **Clerical and Sales** | -0.050 | -0.46 | 0.190 | 0.93 |
| **Mid-Occupations** | 0.006 | 0.09 | 0.073 | 0.54 |
| **Mining** | -0.261 | -1.51 | 0.342 | 1.59 |
| **Electricity, Gas, Water** | 0.400 | 1.67 | -0.337 | -1.02 |
| **Construction** | -0.180 | -1.77 | -0.747 | -4.16 |
| **Trade** | -0.042 | -0.56 | -0.680 | -4.11 |
| **Transport** | -0.179 | -1.75 | -0.430 | -2.54 |
| **Services** | -0.539 | -6.62 | -0.799 | -4.94 |
| **Pretoria** | 0.096 | 0.84 | 0.092 | 0.38 |
| **East-West Rand** | 0.237 | 2.27 | 0.070 | 0.46 |
| **Vaal Triangle** | 0.276 | 2.48 | -0.206 | -1.06 |
| **Cape Peninsular** | 0.589 | 5.02 | 0.180 | 0.79 |
| **Durban** | 0.440 | 4.10 | 0.067 | 0.34 |
| **Pietermaritzburg** | 0.710 | 5.62 | 0.097 | 0.31 |
| **Port Elizabeth** | 0.255 | 1.95 | 0.489 | 1.52 |
| **Bloemfontein** | -0.131 | -0.90 | 0.012 | 0.05 |
| **Klerksdorp** | -0.429 | -2.70 | 0.179 | 0.52 |
| **Free State Gold** | -0.014 | -0.10 | 0.656 | 2.46 |
| **Number Observations**<br>**Chi Squared**<br>**Degrees Freedom** | 2972<br>515.1<br>29.0 | | 855<br>285.5<br>29.0 | |

**T-statistics follow coefficient estimates**

## Table A.3 Earnings Equations With Sample-Selection Correction from Multinomial Logit Model

| | 1985 | | | | 1993 | | | |
|---|---|---|---|---|---|---|---|---|
| | Non-Union | | Union Members | | Non-Union | | Union Members | |
| **Constant** | 3.947 | 114.28 | 3.984 | 71.01 | 4.907 | 20.02 | 5.668 | 30.82 |
| **Primary Education** | 0.106 | 3.25 | 0.114 | 2.55 | 0.169 | 1.32 | -0.055 | -0.71 |
| **Secondary Education** | 0.260 | 4.21 | 0.304 | 4.01 | 0.562 | 3.69 | 0.068 | 0.69 |
| **Tertiary Education** | 0.829 | 5.05 | 0.113 | 1.65 | 1.898 | 8.78 | -0.286 | -1.76 |
| **Experience (yrs)** | 0.023 | 6.26 | 0.022 | 4.68 | 0.042 | 4.36 | 0.021 | 2.29 |
| **Experience $^2$/1000** | -0.450 | -5.85 | -0.409 | -3.95 | -0.886 | -4.24 | -0.281 | -1.46 |
| **Single** | -0.115 | -4.30 | -0.081 | -2.19 | -0.085 | -1.18 | 0.002 | 0.02 |
| **Not Household Head** | -0.217 | -4.43 | -0.160 | -3.42 | -0.221 | -2.52 | 0.127 | 1.46 |
| **Professional** | 0.542 | 8.23 | 0.497 | 4.45 | 0.486 | 3.26 | 0.315 | 2.33 |
| **Clerical and Sales** | 0.425 | 12.44 | 0.237 | 4.42 | 0.215 | 2.43 | 0.223 | 3.13 |
| **Mid-Occupations** | 0.402 | 18.81 | 0.329 | 10.43 | 0.096 | 1.70 | 0.061 | 1.18 |
| **Mining** | 0.232 | 4.42 | -0.227 | -1.40 | 0.061 | 0.34 | -0.152 | -2.08 |
| **Electricity, Gas, Water** | 0.090 | 0.87 | 0.092 | 0.40 | -0.127 | -0.83 | -0.054 | -0.16 |
| **Construction** | 0.034 | 0.84 | -0.072 | -1.23 | 0.067 | 0.46 | -0.247 | -1.89 |
| **Trade** | -0.088 | -2.09 | -0.077 | -1.14 | -0.076 | -0.48 | -0.104 | -0.42 |
| **Transport** | -0.001 | -0.01 | 0.050 | 0.69 | 0.063 | 0.42 | -0.364 | -3.57 |
| **Services** | -0.155 | -3.22 | -0.122 | -1.51 | -0.133 | -0.91 | -0.085 | -0.64 |
| **Primary * Mining** | -0.064 | -0.98 | 0.286 | 1.60 | 0.305 | 1.41 | 0.105 | 1.18 |
| **Secondary * Mining** | -0.018 | -0.18 | 0.560 | 2.15 | -0.149 | -0.59 | 0.143 | 0.98 |
| **Tertiary * Mining** | -0.411 | -2.44 | | | -0.454 | -1.76 | 0.646 | 3.71 |
| **Primary * EGW** | 0.027 | 0.18 | -0.266 | -1.07 | 0.383 | 1.80 | -0.035 | -0.09 |
| **Secondary * EGW** | 0.468 | 3.91 | | | 0.126 | 0.61 | 0.310 | 0.80 |
| **Tertiary * EGW** | | | | | -0.386 | -0.81 | | |
| **Primary * Construction** | -0.105 | -1.86 | -0.068 | -0.69 | -0.312 | -1.73 | 0.317 | 1.91 |
| **Secondary * Construction** | 0.115 | 1.26 | 0.253 | 1.09 | -0.357 | -1.34 | -0.348 | -1.65 |
| **Tertiary * Construction** | -0.118 | -0.69 | | | | | | |
| **Primary * Trade** | -0.010 | -0.20 | -0.015 | -0.19 | -0.015 | -0.08 | 0.092 | 0.36 |
| **Secondary * Trade** | 0.234 | 2.72 | 0.110 | 0.73 | -0.392 | -1.68 | 0.124 | 0.45 |
| **Tertiary * Trade** | 0.172 | 0.86 | 0.819 | 6.22 | 0.426 | 1.69 | 0.911 | 2.54 |
| **Primary * Transport** | 0.075 | 0.82 | -0.068 | -0.74 | -0.015 | -0.08 | 0.392 | 2.70 |
| **Secondary * Transport** | -0.057 | -0.46 | 0.103 | 0.80 | 0.045 | 0.24 | 0.526 | 1.94 |
| **Tertiary * Transport** | -0.036 | -0.17 | | | -0.983 | -4.67 | | |
| **Primary * Services** | 0.075 | 1.29 | -0.010 | -0.10 | -0.067 | -0.37 | 0.139 | 0.83 |
| **Secondary * Services** | 0.320 | 4.06 | 0.230 | 2.01 | -0.254 | -1.12 | 0.278 | 1.52 |
| **Tertiary * Services** | 0.073 | 0.40 | 0.506 | 1.49 | -0.812 | -3.96 | 0.832 | 4.73 |
| **Pretoria** | -0.022 | -0.61 | -0.062 | -1.07 | 0.295 | 2.25 | 0.271 | 2.42 |
| **East-West Rand** | -0.070 | -1.95 | -0.032 | -0.69 | 0.071 | 0.94 | -0.083 | -1.10 |
| **Vaal Triangle** | -0.054 | -1.50 | -0.088 | -1.64 | 0.064 | 0.72 | -0.020 | -0.21 |
| **Cape Peninsular** | -0.098 | -2.67 | -0.120 | -2.04 | 0.019 | 0.17 | -0.163 | -1.81 |
| **Durban** | -0.101 | -2.62 | -0.225 | -4.74 | 0.051 | 0.60 | -0.027 | -0.33 |
| **Pietermaritzburg** | -0.253 | -4.21 | -0.251 | -5.02 | -0.224 | -1.57 | -0.025 | -0.19 |
| **Port Elizabeth** | -0.226 | -5.37 | -0.186 | -3.46 | -0.047 | -0.39 | -0.336 | -1.90 |
| **Bloemfontein** | -0.203 | -5.16 | -0.418 | -5.31 | -0.014 | -0.14 | -0.086 | -0.88 |
| **Klerksdorp** | -0.137 | -3.40 | -0.169 | -1.82 | 0.701 | 7.50 | 0.365 | 3.43 |
| **Free State Gold** | -0.132 | -3.22 | -0.205 | -2.76 | -0.155 | -0.92 | -0.043 | -0.39 |
| **Inverse Mills Ratio** | 0.141 | 2.26 | 0.157 | 2.99 | 0.091 | 0.64 | -0.119 | -1.45 |
| **Number Observations** | 2164 | | 808 | | 423 | | 435 | |
| **R Bar Squared** | .433 | | .359 | | .372 | | .209 | |
| **F Statistic** | 38.59 | | 12.31 | | 6.68 | | 3.72 | |
| **Error correlation with MNL** | .119 | | .180 | | .039 | | .102 | |

**T-statistics follow coefficient estimates**

# Table A.4 Earnings Equations Excluding Sample-Selection Correction

| | 1985 | | | | 1993 | | | |
|---|---|---|---|---|---|---|---|---|
| | Non-Union | | Union Members | | Non-Union | | Union Members | |
| **Constant** | 4.049 | 143.20 | 4.237 | 109.01 | 5.015 | 29.32 | 5.438 | 45.22 |
| **Primary Education** | 0.110 | 3.37 | 0.094 | 2.12 | 0.172 | 1.34 | -0.036 | -0.47 |
| **Secondary Education** | 0.255 | 4.12 | 0.303 | 3.94 | 0.556 | 3.63 | 0.111 | 1.18 |
| **Tertiary Education** | 0.800 | 4.86 | 0.184 | 2.76 | 1.842 | 9.21 | -0.251 | -1.49 |
| **Experience (yrs)** | 0.018 | 5.61 | 0.019 | 4.00 | 0.039 | 4.52 | 0.027 | 3.22 |
| **Experience $^2$/1000** | -0.343 | -5.31 | -0.316 | -3.16 | -0.819 | -4.52 | -0.407 | -2.27 |
| **Single** | -0.110 | -4.10 | -0.072 | -1.95 | -0.068 | -1.00 | -0.018 | -0.22 |
| **Not Household Head** | -0.123 | -4.82 | -0.092 | -2.27 | -0.180 | -2.97 | 0.077 | 0.94 |
| **Professional** | 0.542 | 8.25 | 0.449 | 3.98 | 0.489 | 3.25 | 0.322 | 2.31 |
| **Clerical and Sales** | 0.426 | 12.50 | 0.242 | 4.48 | 0.216 | 2.43 | 0.215 | 3.04 |
| **Mid-Occupations** | 0.403 | 18.86 | 0.324 | 10.28 | 0.100 | 1.77 | 0.058 | 1.12 |
| **Mining** | 0.235 | 4.47 | -0.215 | -1.31 | 0.059 | 0.32 | -0.154 | -2.10 |
| **Electricity, Gas, Water** | 0.090 | 0.85 | 0.088 | 0.36 | -0.129 | -0.86 | -0.025 | -0.07 |
| **Construction** | 0.036 | 0.87 | -0.070 | -1.20 | 0.071 | 0.48 | -0.230 | -1.74 |
| **Trade** | -0.082 | -1.95 | -0.081 | -1.20 | -0.075 | -0.48 | -0.100 | -0.40 |
| **Transport** | 0.002 | 0.03 | 0.056 | 0.78 | 0.066 | 0.45 | -0.368 | -3.66 |
| **Services** | -0.152 | -3.15 | -0.113 | -1.38 | -0.137 | -0.94 | -0.066 | -0.50 |
| **Primary * Mining** | -0.072 | -1.10 | 0.306 | 1.66 | 0.299 | 1.39 | 0.098 | 1.10 |
| **Secondary * Mining** | -0.036 | -0.35 | 0.563 | 2.05 | -0.136 | -0.54 | 0.122 | 0.85 |
| **Tertiary * Mining** | -0.439 | -2.63 | | | -0.444 | -1.71 | 0.668 | 3.76 |
| **Primary * EGW** | 0.032 | 0.21 | -0.255 | -0.98 | 0.382 | 1.80 | -0.053 | -0.13 |
| **Secondary * EGW** | 0.479 | 3.95 | | | 0.106 | 0.52 | 0.256 | 0.63 |
| **Tertiary * EGW** | | | | | -0.394 | -0.83 | | |
| **Primary * Construction** | -0.107 | -1.89 | -0.071 | -0.72 | -0.319 | -1.78 | 0.290 | 1.72 |
| **Secondary * Construction** | 0.095 | 1.05 | 0.241 | 1.09 | -0.360 | -1.34 | -0.379 | -1.69 |
| **Tertiary * Construction** | -0.160 | -0.93 | | | | | | |
| **Primary * Trade** | -0.016 | -0.31 | -0.013 | -0.16 | -0.011 | -0.06 | 0.095 | 0.36 |
| **Secondary * Trade** | 0.227 | 2.64 | 0.113 | 0.74 | -0.397 | -1.71 | 0.120 | 0.44 |
| **Tertiary * Trade** | 0.135 | 0.68 | 0.867 | 6.52 | 0.449 | 1.79 | 0.934 | 2.47 |
| **Primary * Transport** | 0.071 | 0.78 | -0.075 | -0.80 | -0.071 | -0.39 | 0.397 | 2.76 |
| **Secondary * Transport** | -0.066 | -0.53 | 0.084 | 0.63 | -0.252 | -1.11 | 0.537 | 2.00 |
| **Tertiary * Transport** | -0.066 | -0.31 | | | -0.791 | -3.92 | | |
| **Primary * Services** | 0.072 | 1.23 | -0.017 | -0.16 | -0.011 | -0.07 | 0.117 | 0.71 |
| **Secondary * Services** | 0.317 | 4.01 | 0.208 | 1.79 | 0.048 | 0.25 | 0.259 | 1.42 |
| **Tertiary * Services** | 0.047 | 0.26 | 0.445 | 1.29 | -0.946 | -4.67 | 0.822 | 4.70 |
| **Pretoria** | -0.013 | -0.37 | -0.069 | -1.17 | 0.283 | 2.19 | 0.294 | 2.68 |
| **East-West Rand** | -0.054 | -1.50 | -0.054 | -1.16 | 0.078 | 1.03 | -0.036 | -0.50 |
| **Vaal Triangle** | -0.030 | -0.86 | -0.119 | -2.23 | 0.056 | 0.64 | -0.002 | -0.02 |
| **Cape Peninsular** | -0.076 | -2.18 | -0.164 | -2.82 | -0.012 | -0.13 | -0.136 | -1.50 |
| **Durban** | -0.069 | -1.93 | -0.260 | -5.54 | 0.042 | 0.50 | -0.012 | -0.14 |
| **Pietermaritzburg** | -0.210 | -3.76 | -0.307 | -6.46 | -0.232 | -1.62 | 0.014 | 0.10 |
| **Port Elizabeth** | -0.201 | -4.96 | -0.189 | -3.50 | -0.026 | -0.23 | -0.304 | -1.72 |
| **Bloemfontein** | -0.217 | -5.57 | -0.387 | -4.80 | -0.041 | -0.43 | -0.101 | -1.04 |
| **Klerksdorp** | -0.172 | -4.69 | -0.125 | -1.38 | 0.675 | 8.38 | 0.404 | 3.78 |
| **Free State Gold** | -0.135 | -3.32 | -0.210 | -2.85 | -0.094 | -0.66 | 0.078 | 1.00 |
| **Number Observations** | 2164 | | 808 | | 423 | | 435 | |
| **R Bar Squared** | 0.432 | | 0.352 | | 0.373 | | 0.206 | |
| **F Statistic** | 39.24 | | 12.26 | | 6.84 | | 3.75 | |

**T-statistics follow coefficient estimates**

# Table A.5 Earnings Equations with Union Dummy

| | 1985 | | 1993 | |
|---|---|---|---|---|
| **Constant** | 4.080 | 175.90 | 5.182 | 48.44 |
| **Union Member** | 0.062 | 3.86 | 0.181 | 5.43 |
| **Primary Education** | 0.108 | 4.05 | 0.039 | 0.58 |
| **Secondary Education** | 0.278 | 5.65 | 0.293 | 3.38 |
| **Tertiary Education** | 0.596 | 3.71 | 0.767 | 1.02 |
| **Experience (yrs)** | 0.018 | 6.78 | 0.032 | 5.11 |
| **Experience $^2$/1000** | -0.335 | -6.19 | -0.598 | -4.54 |
| **Single** | -0.103 | -4.74 | -0.028 | -0.53 |
| **Not Household Head** | -0.114 | -5.28 | -0.093 | -1.86 |
| **Professional** | 0.539 | 9.08 | 0.409 | 3.75 |
| **Clerical and Sales** | 0.380 | 13.00 | 0.237 | 4.01 |
| **Mid-Occupations** | 0.384 | 21.46 | 0.085 | 2.08 |
| **Mining** | 0.211 | 4.23 | -0.055 | -0.73 |
| **Electricity, Gas, Water** | 0.102 | 0.94 | -0.126 | -0.65 |
| **Construction** | 0.005 | 0.16 | -0.079 | -0.82 |
| **Trade** | -0.076 | -2.09 | -0.161 | -1.36 |
| **Transport** | 0.019 | 0.31 | -0.171 | -1.96 |
| **Services** | -0.156 | -3.65 | -0.134 | -1.35 |
| **Primary * Mining** | -0.057 | -0.93 | 0.057 | 0.66 |
| **Secondary * Mining** | 0.002 | 0.02 | -0.093 | -0.67 |
| **Tertiary * Mining** | -0.213 | -1.26 | 0.131 | 0.16 |
| **Primary * EGW** | -0.135 | -1.01 | 0.292 | 1.26 |
| **Secondary * EGW** | 0.456 | 3.80 | 0.196 | 0.87 |
| **Tertiary * EGW** | | | 0.540 | 0.58 |
| **Primary * Construction** | -0.098 | -1.97 | -0.084 | -0.67 |
| **Secondary * Construction** | 0.104 | 1.22 | -0.216 | -1.05 |
| **Tertiary * Construction** | 0.054 | 0.32 | | |
| **Primary * Trade** | -0.025 | -0.56 | 0.126 | 0.93 |
| **Secondary * Trade** | 0.180 | 2.38 | -0.099 | -0.58 |
| **Tertiary * Trade** | 0.352 | 1.85 | 0.493 | 0.56 |
| **Primary * Transport** | 0.035 | 0.49 | 0.199 | 1.80 |
| **Secondary * Transport** | -0.033 | -0.34 | 0.153 | 0.88 |
| **Tertiary * Transport** | 0.158 | 0.73 | 0.488 | 0.65 |
| **Primary * Services** | 0.069 | 1.33 | 0.054 | 0.44 |
| **Secondary * Services** | 0.290 | 4.39 | 0.169 | 1.29 |
| **Tertiary * Services** | 0.222 | 1.20 | -0.067 | -0.09 |
| **Pretoria** | -0.024 | -0.78 | 0.282 | 3.25 |
| **East-West Rand** | -0.052 | -1.78 | -0.001 | -0.02 |
| **Vaal Triangle** | -0.051 | -1.72 | 0.009 | 0.14 |
| **Cape Peninsular** | -0.097 | -3.14 | -0.099 | -1.46 |
| **Durban** | -0.128 | -4.46 | 0.033 | 0.53 |
| **Pietermaritzburg** | -0.237 | -6.26 | -0.080 | -0.72 |
| **Port Elizabeth** | -0.192 | -5.75 | -0.174 | -1.61 |
| **Bloemfontein** | -0.250 | -7.12 | -0.061 | -0.88 |
| **Klerksdorp** | -0.182 | -5.39 | 0.529 | 6.50 |
| **Free State Gold** | -0.149 | -4.20 | 0.026 | 0.38 |
| **Number Observations**<br>**R Bar Squared**<br>**F Statistic** | 2972<br>.415<br>48.98 | | 858<br>.315<br>9.94 | |

**T-statistics follow coefficient estimates**