Applied Economics, 2005, **37**, 1693–1704

Routledge
Taylor & Francis Group

Trade union membership and earnings in Kenyan manufacturing firms

Damiano Kulundu Manda[a], Arne Bigsten[b,*] and Germano Mwabu[a]

[a]*Department of Economics, University of Nairobi, Kenya*
[b]*Department of Economics, Göteborg University, Göteborg 405 30, Sweden*

This study analyses the effect of trade unions on male earnings in the Kenyan manufacturing sector using a regression method, which takes into account endogeneity of the union status of workers. In contrast to earlier studies of the Kenyan labour market that report a negative effect of unions on wages, a positive effect is found. The study further shows that elite workers tend to abstain from union membership.

I. Introduction

Does it pay to be a union member in Africa? Some studies have found a negative membership effect, and have tried to come up with plausible explanations for this counter-intuitive result. This study resolves this paradox by analysing the effect of union membership on wages in Kenyan formal sector manufacturing. Kenya is fairly typical for Africa, with 32% of formal sector employees being union members.[1] In Kenyan manufacturing, about 40% of formal sector workers are union members.[2]

The trade unions' main aim is to improve the welfare of its members by negotiating higher earnings for its members, better conditions of service, and increasing job tenure. To achieve these objectives, unions need a strong financial base and bargaining power. However, in Kenya, unions are generally financially weak and are occasionally faced with internal leadership struggles. Their main source of finance is membership fees, which are not high enough to sustain drawn-out strikes. Also, their ability to strike is limited by control of union activities by the government through legislation, and by threat of de-registration. Still, workers can carry out a legal strike after exhausting all avenues of arbitration and issuing a strike notice 21 days in advance.

Restrictions on unions' activities thus limit their ability to use strikes to achieve gains for the employees. However, unions can still influence the level of earnings and working conditions of their members through collective bargaining, although the unions sometimes lack resources to hire negotiators with skills and abilities to match the full-time negotiators of the employers. Nevertheless, union leaders are dependent on union members for support, and they press hard for workers' demands and sometimes manage to achieve substantial results (Sandbrook, 1975; Bigsten, 1984).

The empirical knowledge of whether and how trade unions actually affect earnings in Africa is still sparse. Worldwide, most empirical studies concerning the influence of unions on wages are based on the experiences of industrialized countries, and few attempts have been made to estimate the relationship for less developed countries, especially in Sub-Saharan Africa.

*Corresponding author. E-mail: Arne.Bigsten@economics.gu.se

[1] The percentage of formal sector workers unionised is 20% in Cameroon, 21% in Ivory Coast, 27% in Ghana, 35% in Tanzania, 33% in Zambia and 42% in Zimbabwe (Mazumdar and Mazaheri, 2002, pp. 36–7).

[2] For example, in Tanzania the union density in manufacturing is 43%.

Applied Economics ISSN 0003–6846 print/ISSN 1466–4283 online
http://www.tandf.co.uk/journals
DOI: 10.1080/00036840500118903

The studies done in industrialized countries find unions to have a significant positive effect on wages, but with considerable variation in the estimated size of the effect (Lewis, 1986; Booth, 1995).

In Africa, Schultz and Mwabu (1998) found a negative union effect on wages for certain categories of white workers in South Africa, while the effect for black workers was positive. In Kenya, the study by Johnson (1971) found that union wages were 30% higher than non-union wages in private industries, 11% higher in the public sector. However, Johnson failed to consider factors, such as firm size, which might be correlated with union membership. In another Kenyan study, House and Rempel (1976), using OLS estimation techniques, found the union-effect to be negative, but statistically insignificant. However, they did not include other wage determinants, such as education and experience, in their analysis. Manda (1997) found that trade-union wages in Kenyan manufacturing were about 22% lower than for non-union workers. Thus, there is need for a fresh empirical analysis of the effect of trade unions on earnings in Kenya.

The three Kenyan studies differ in a number of ways. For instance, some of the studies based their analysis on wages and others on earnings (wages plus allowances and other benefits). If union workers receive more allowances than non-union workers, then the analyses of union-effects based on wages or on earnings are bound to be different. In some African countries, such as Zambia (Anderson, 1993), unions have pushed for increased non-taxable allowances and other benefits, since such benefits would entirely benefit the worker. Such action would be expected in a country like Kenya, which has one of the highest tax rates in the region. In addition, negotiating for higher allowances and benefits for workers would not be quite as involved as negotiating for higher wages, since employers also stand to gain in terms of tax exemptions, if they can prove to the government that they actually pay workers non-wage benefits. In Kenyan manufacturing, non-wage benefits constituted about 18% of total remuneration in manufacturing in 1995.

The time periods for which the earlier Kenyan studies were carried out also differ substantially, and this could of course partly explain the differences in the results obtained, but the main shortcoming of the three studies is that they assume that trade unions exert a unilateral and exogenous effect on wages. They therefore estimated a single-equation model without taking account of the endogeneity of union membership and earnings, which may bias the estimates (see for instance, Schmidt and Strauss, 1996). It is shown that when an appropriate estimation method is applied, the results change dramatically.

This paper analyses the effect of unions on earnings using merged data on firms and workers from the manufacturing sector in Kenya. It further investigates whether there are any major differences in the characteristics of union and non-union workers. Section II presents the analytical models and describes the data, while Section III presents the empirical results of our analysis, and Section IV summarizes the paper and draws conclusions.

II. Model and Data

Model properties and estimation strategy

An appropriate model for estimating union-status and union-membership premium is the switching-regression model with endogenous switching (Lee, 1978), which provides a unified framework for testing selectivity and pooling. The model has recently been used by Lanfranchi *et al.* (2002) to study compensating differentials due to shift-work. The same approach is used, since it also deals effectively with the problems related to the empirical estimation of the union-earnings effect. For a more detailed discussion of the econometric issues involved, see Lee (1978) and Lanfranchi *et al.* (2002).

The study starts from two earnings functions, one for union workers (Equation 1a) and the other non-union workers (Equation 1b)

$$W_{ui} = \beta_1' X_{1i} + u_{1i} \tag{1a}$$

$$W_{ni} = \beta_2' X_{2i} + u_{2i} \tag{1b}$$

where W_{ui} are log hourly earnings for union workers and W_{ni} are log hourly earnings for non-union workers. Vector X includes standard explanatory variables, while u_1 and u_2 are error terms. The Hausman test is used to test for the pooling of data on union and non-union members.

Equation 2 represents union membership choice

$$\begin{aligned} I_i^* &= \alpha_0 + \delta Z_i + \gamma(W_{ui} - W_{ni}) + v_i \\ I_i &= 1 \quad \text{if } I_i^* \geq 0 \text{ and } 0 \text{ otherwise} \end{aligned} \tag{2}$$

where I_i^* is a latent variable determining whether union membership, I_i, is observed, and $(W_{ui} - W_{ni})$ is our estimate of the union premium. The earnings Equations 1a and 1b and the choice Equation 2 constitute a switching regression model with endogenous switching and can be used to estimate the union wage premium as well as its effect on union-status.

It is quite likely that the error terms in Equations 1 and 2 are correlated, giving rise to selectivity bias. This needs to be corrected for this in the wage

equations and to do this the estimation strategy of Lanfranchi *et al*. (2002) was used. The reduced-form union-status choice equation is

$$I_i^* = \delta Z_i + \gamma(\beta_u - \beta_n)X_i + \gamma(u_{1i} - u_{2i}) + v_i. \quad (3)$$

Assuming common values for the β's and the u's, expression 3 can be re-written in compact form as $I^* = \gamma\check{Z} + v$. Following Maddala (1983) conditional expected earnings are computed using Equations 4a and 4b.

$$E(W_u|U = 1) = \beta_u X + \sigma_{u_1,v}[(\phi(\gamma\check{Z})/\Phi(\gamma\check{Z})] \quad (4a)$$

$$E(W_n|U = 0) = \beta_n X - \sigma_{u_2,v}[(\phi(\gamma\check{Z})/(1 - \Phi(\gamma\check{Z}))] \quad (4b)$$

where $\phi(\gamma\check{Z})$ and $\Phi(\gamma\check{Z})$ are, respectively, the density functions and the distribution functions of the standard normal, evaluated at $\gamma\check{Z}$. Including $+\sigma_{u1,v}\phi/\Phi$ and $-\sigma_{u2,v}\phi/(1-\Phi)$ when estimating the respective earnings-equations control for selectivity.

To estimate the model one may either use full-information maximum likelihood (ML) or a two-step procedure. Davidson and MacKinnon (1993) show that the two-step procedure is appropriate if one wants to test for selectivity-bias. In the estimations they therefore proceed with the two-stage procedure as follows. First, the reduced-form union-choice Equation 3 is estimated using ML methods to get γ and then $\phi(\gamma\check{Z})$ and $\Phi(\gamma\check{Z})$ are computed. Second, the conditional earnings Equations 4a and 4b are estimated. Using the estimated β_u's and β_n's one can compute the union premium $(W_{ui*} - W_{ni*})$ for each worker. Third, the structural-form choice-equation (Equation 2) is estimated using ML.

If one finds that $\sigma_{u_{1,v}}$ and $\sigma_{u_{2,v}}$ are statistically significant in the earnings-equations, then correction for sample-selection is indeed needed. Pooling can be tested using an empirical specification suggested by Maddala (1983), where the estimated earnings function is:

$$E(W) = \beta_n X + (\beta_u - \beta_n)X\Phi(\gamma\check{Z}) + (\sigma_{u_1,v} - \sigma_{u_2,v})\phi(\gamma\check{Z}). \quad (5)$$

If $\beta_u = \beta_n$ for all X except the constant term, then Equation 5 collapses to the treatment-effects model

$$E(W) = \beta X + \alpha\Phi(\gamma\check{Z}) + (\sigma_{u_1,v} - \sigma_{u_2,v})\phi(\gamma\check{Z}) \quad (6)$$

where α measures the earnings difference between union and non-union workers.

We can compare the treatment-effects model versus the separate earnings-equations using an F-test. In this case, the restricted model (Equation 6) is the treatment-effects model. If the treatment effects model is rejected, it means it is believed that returns to specific characteristics differ between union and non-union workers.

Data and related econometric issues

The data used in the study come from the Regional Programme on Enterprise Development (RPED) survey of the manufacturing sector in Kenya (see Bigsten and Kimuyu, 2001, for details). The data-set is a three-year panel survey of more than 200 manufacturing sector firms carried out by a team of researchers from the Departments of Economics of the Göteborg University and the University of Nairobi in collaboration with the Kenya Association of Manufacturers in the years 1993–1995. Whereas the data on firms is a panel (with a few replacements in 1994 and 1995), data on individuals is not a panel, although some of the workers were interviewed more than once during the first, second and third waves. This study uses information on both employers and employees. The data set is thus a matched employer-employee data set. The survey covered firms in four manufacturing sub-sectors (food, textiles, wood and metal) in four major urban centres in Kenya (Nairobi, Mombasa, Nakuru, and Eldoret).

The labour market sections of the survey provide information on both firm and individual characteristics. The information on firms was elicited from representatives of each firm and includes total number of firm employees, total wage bills, profits, firm ownership, and the proportion of firm employees in trade unions among other things. This provides more accurate information on, firm size, for instance, than would be the case had individual workers provided such information. The information on individuals was obtained through interviews, with at most ten employees randomly chosen from a list of workers of each firm. The employees interviewed provided information on, for example, earnings, education level, previous experience, tenure, union status, age, sex and hours of work.

As mentioned above some of the workers were interviewed more than once during the three waves. Theoretically, this could mean that there is correlation among the error terms of repeated observations of individuals. This potential problem was mitigated by including a dummy variable to capture the effect of individuals who were interviewed more than once. Also, since there is no (observable) panel-nature on individual workers in the data, the estimates of any possible union premium are likely to be an upper bound due to the fact that one may not be able to control for other worker-level individual factors that may be correlated with union status. Studies for the USA such as Robinson (1989), Jakubson (1991)

and Neumark and Wachter (1995) and , have shown that when fixed effects are not taken into account, the estimated union premium is likely to be an upper bound.

The variables included in the analysis are defined as follows. The dependent variables are union status for the probit equation, and the natural logarithms of hourly wages and of hourly earnings (wages plus allowances and other benefits)[3] for the earnings functions. Hourly wages (earnings) are computed by dividing monthly wages (earnings) by the total number of hours of work in a month (i.e., hours of work in a week multiplied by 4.3 weeks in a month). Allowances and benefits include things such as house and transport allowances, food and clothing allowances, and bonuses. The independent variables include dummies for union status (predicted), location, sector, education, and training (on-the-job training and general-training). General training can be considered to be equivalent to vocational training and is not firm-specific, while on-the-job training is firm-specific training carried out at the workplace. Dummies for occupation, and survey periods (waves 1 through 3) were also included. The wave dummies are included to capture the time-effects on wage-rates that might arise due to changes in macroeconomic factors and in policies. Specific time variant factors were also controlled for using sector-specific time dummies.

Other independent variables are a measure of firm-size (number of employees in a firm), age of the employee and its square, tenure, and dummies for firm-ownership status, and for relative of the owner of the firm. Other independent variables specifically in the union-membership equation are weekly hours of work (capturing working conditions in a firm) and the value of recent investments in equipment per employee in a firm. These two variables are used as identifiers for the probit equation. In using these variables as instruments for union membership, it is assumed that the overall objective of a union is to defend and improve working conditions and the living standards of its members (House and Rempel, 1976). In Kenya the *statutory* number of hours of work in one week are 45 (including a lunch-break of one hour). It is expected that in firms where unions are effective the weekly hours of work will not exceed 45 hours. Firms in which weekly hours of work exceed 45 can be seen as imposing longer hours of work on their employees. One of the union's aims is therefore assumed to be to ensure that its members are not engaged for longer hours than officially permitted. Individuals observed to be working excessively for long hours are therefore thought likely to be in non-unionized firms, to be non-union members, or to be in firms where only weak unions operate. We use 'weekly hours of work' and its square as the first set of instruments for the union status of a worker.

The second variable that is used to identify the union-status of a worker is the value per worker (in Kenya Shillings) of the most recent investment in equipment undertaken by a firm. The use of this variable is based on the hold-up theory (McDonald and Solow, 1981). The possibility of re-negotiation for both the wage and employment levels forms the basis of the hold-up phenomenon in wage-employment negotiations. According to hold up theory, the labour market equilibrium, which is attained through the bargaining process, may be either efficient or inefficient (see also Nickell and Wadhwani, 1990). If the bargaining equilibrium is efficient, the contract between the firm and representatives of the workers is assumed to be complete and not open to re-negotiation. However, if the equilibrium is inefficient, the possibility of re-negotiation is open. A new investment by the firm opens the possibility for workers to renegotiate the wage and employment agreement so as to get a share of the investment-induced profits. Firm-investments thus here increase union activity and membership. The possibility of profitable re-negotiation is particularly high if the new investment is reversible only at high cost (Azam and Ris, 2001). It is assumed that new equipment investment in Kenya is only reversible at high cost given that such equipment is normally imported. One thus expects union activities to intensify following new investment by a firm. However, whether the activity translates into a high pay or better employment depends on the bargaining strength of the unions.

The validity of these two instruments is formally tested for and it is checked whether the associated over-identifying restrictions are satisfied (Wooldridge, 2002). The key variables in the analysis are fully defined in Appendix Table A1.

The analysis is restricted to male workers for two reasons. First, male workers dominate the manufacturing sector workforce, with 86% of the respondents being male. Second, sometimes there are contrasting results for union-premium for males and females. Workers in informal sector firms are also included. About 90% of the respondents in the data set are formal manufacturing-sector workers. Formal firms

[3] Overtime pay is not included.

Table 1. Means of selected variables by union status, manufacturing sector

Variables	Workers		
	All workers	Union	Non-union
Hourly earnings	22.34	19.97	23.71
Hourly wages	18.81	16.01	20.43
Allowances and benefits	3.43	3.86	3.28
Weekly hours of work	46.9	45.6	46.7
Weekly hours of work squared	2229	2100	2300
Age	33.98	35.94	32.85
Age squared	1198	1346	1118
Primary dummy	0.45	0.48	0.45
Secondary dummy	0.39	0.32	0.39
University dummy	0.02	0.01	0.02
On-the-job training	0.26	0.28	0.24
General training	0.18	0.14	0.21
Production	0.65	0.74	0.60
Firm size (workers)	144.00	194.01	113.13
Tenure (years)	7.91	9.79	6.82
Located in Nairobi	0.65	0.74	0.61
Located in Mombasa	0.15	0.12	0.17
Located in Nakuru	0.11	0.08	0.12
Located in Eldoret	0.09	0.08	0.10
Food sector	0.21	0.25	0.20
Wood sector	0.21	0.21	0.22
Textile sector	0.28	0.21	0.28
Metal sector	0.3	0.28	0.31
Union membership	0.39		
Union density	43.48	67.22	29.75
Employee interviewed more than once	0.2	0.21	0.2
Relative of the owner of the firm	0.04	0.01	0.06
Recent investment per worker in a firm	45885	77577	45886
Employee in foreign-owned firm	0.26	0.28	0.25
Total number of observations	2764	1023	1741

are firms that are included in the official registry of firms. Based on this, the sample consists of 2764 observations, of which 1023 observations are union workers and 1741 non-union workers.

III. Results

Descriptive statistics

The means of the variables used in the analysis are shown in Table 1. The difference between hourly earnings and hourly wages for all male workers in the formal manufacturing sector is about Kshs 3 (US$ 1 = Kshs 57 in 1995). Thus on average, allowances and other benefits form about 18% of total male earnings. The variation in hourly earnings (standard deviation = 23) is higher than in hourly wages (standard deviation = 16). A large proportion of the respondents have primary and secondary education, but only a few have university education; 86% had at least primary education. This reflects the impact of the rapid expansion of the education system (Manda, 1997). Average tenure for male workers is about eight years. About 65% of the employees work in firms located in Nairobi, and about 39% are union members. This is higher than the national estimate of about a third of all workers, and shows that manufacturing is relatively more unionized than other sectors of the economy. About 26% of the employees work in foreign-owned firms.

Trade union membership is not compulsory in Kenya, so a question is whether there is any distinct group of workers who are in fact union members. An attempt is made to answer this question using information in columns 2 and 3 of Table 1. As shown, hourly earnings and wages are higher for non-union than for union workers, but union workers receive slightly higher allowances and other benefits than do non-union workers. Thus the difference between hourly earnings and hourly wages are higher for union than for non-union workers.

Weekly hours of work are also lower for union than for non-union workers. The average age of union workers is about 36 years, three years higher than for non-union workers. While the proportion of

Table 2. Probit estimates for union membership, marginal effects

	Reduced form probit		Structural-form probit#		Structural-form probitΨ	
	Marginal effects	*t*-stat.	Marginal effects	*t*-stat.	Marginal effects	*t*-stat.
Primary dummy	−0.0210	−0.71	0.0475	1.36	0.0435	1.26
Secondary dummy	−0.1102**	−3.44	0.0274	0.55	0.0331	0.65
University dummy	−0.2061**	−2.77	−0.0720	−0.73	−0.0041	−0.04
Age	0.0376**	4.51	0.0565**	5.76	0.0583**	5.79
Age squared	−0.0005**	−4.67	−0.0007**	−5.65	−0.0007**	−5.79
Tenure	0.0078**	3.98	0.0113**	5.17	0.0100**	4.84
On-the-job training	0.0558*	2.04	0.1357**	3.87	0.1655**	4.10
General training	−0.0698**	−2.5	0.0381	0.92	0.0504	1.46
Firm size	0.0001**	2.66	0.0001**	4.15	0.0001**	4.11
Located in Nairobi	0.072	1.85	0.1736**	3.70	0.2152**	3.98
Located in Mombasa	0.0405	0.87	0.0546	1.16	0.0661	1.38
Located in Nakuru	0.0725	1.48	0.0337	0.68	0.0459	0.94
Metal sector	−0.0890*	−1.94	−0.2983**	−4.16	−0.2597**	−4.06
Wood sector	0.0265	0.52	−0.0553	−1.01	0.0074	0.14
Textile sector	0.0898	1.89	−0.0102	−0.19	0.0886	1.85
Union density	0.0047**	17.93	0.0061**	13.52	0.0073**	10.09
Relative of firm owner	−0.2424**	−4.77	−0.1655**	−2.69	0.0177	0.17
Interviewed more than once	0.0469	1.69	0.0904**	2.95	0.1234**	3.53
Employee in foreign-owned firm	0.0183	0.79	0.1637**	3.6	0.2019**	3.71
Production-work	0.1608**	7.31	0.1426**	6.27	0.1975**	8.16
Weekly hours of worker	−0.0288**	−2.07	−0.0019	−0.12	−0.0019	−0.12
Weekly hours of work squared	0.0003	1.53	2.80e-06	0.02	2.83e-06	0.02
Recent investment per worker in a firm	8.26e-05**	2.65	9.80e-06	0.27	9.80e-06	0.27
Union premium	–		0.8288**	3.75	0.9453**	3.75
Sector-specific time dummies	Yes		Yes		Yes	
Waves	Yes		Yes		Yes	
Log likelihood	−1419		−1412		−1412	
Chi-square	804		817		817.49	
Significance level	0.0		0.0		0.0	
Pseudo R^2	0.22		0.224		0.223	
Number of observations	2764		2764		2764	

Notes: *Significance level is 0.05; **Significance level is 0.01.
#Union-premium based on earnings.
ΨUnion-premium based on wages.

union workers with only primary education is higher than for non-union workers, the proportion of non-union workers with secondary and university education is higher. Tenure is considerably higher for union than for non-union workers. About 14% of union workers and 21% of non-union workers have received general training. Among production workers 74% are union members, compared to 65% of all workers, so that production workers constitute a dominant group of union workers.

Estimation results

This section starts by reporting results for the reduced-form probit (Equation 3) for union membership choice. The estimated marginal effects and *t*-values are reported in columns 1 and 2 of Table 2. The results show that the probability of being a union member declines as the level of education of the worker increases. The coefficients for age and age-squared have the expected signs and are statistically significant. Thus age raises the probability that a male worker would be in a union, but at a decreasing rate. The probability of union membership increases by 0.008 per year of tenure. A worker who has received on-the-job training is 6% more likely to be a member of a union, whereas general training reduces the probability of being a union member by about the same amount. Employees in larger firms, are more likely to be union members.

Location of the firm does not have a significant effect on the probability of being a union member. Relative to the food sector, male employees in the metal sector has a significantly lower probability of being union members. Employees in more-unionized firms are more likely to be union members than employees in less-unionized firms. The foreign ownership of firms does not have a significant

effect on the probability of union membership. Production work increases the probability of union membership.

The coefficients on the number of weekly hours of work and its square have negative and positive signs, respectively, reflecting a U-shaped relationship between the number of hours worked and union membership. The probability of being a union member thus decreases as the number of weekly hours increases, but at a declining rate. The probability of being in a union decreases up to 48 weekly hours of work, after which it increases. The U-shaped relationship is interpreted to mean that workers join the union for protection against long hours of work. However, this result should be taken with caution because the coefficient on the quadratic term is significant only at the 10% level.

There is need to stress that in the sample the number of hours worked per week by salaried workers is exogenous because it is determined by the firm. As a consequence, workers generally get the same salary per month irrespective of the number of hours worked. Generally, firms have an incentive not to compensate full-time workers for overtime work unless under union pressure.

A large recent investment in equipment per employee in a firm raises the probability that employees in that firm would be union members. These two variables (weekly hours of week and recent equipment investment) are used as identifiers of the union-status of workers. A likelihood-ratio test rejects the null hypothesis that the effects of the two identifiers are not jointly significant in the probit equation (see row 1, Table A3), indicating that the two variables jointly affect union-membership status.

The picture that emerges from the results is fairly clear. Elite workers tend to abstain from union membership, while those in 'non-elite' occupations such as 'production workers' tend to be union members. Also, employees in larger and more organized firms, and these with longer tenure, are more likely to be union members.

The last four columns of Table 2 report the marginal effects and the associated *t*-statistics for the structural probit model (Equation 2). The union-premium variable is obtained by taking the difference between predicted wages (or earnings) for union and non-union wages (or earnings). In column 3, the union-premium variable is obtained by taking the difference between predicted earnings from the union and non-union earnings equations, while in column 5 the union premium is obtained by taking the difference between predicted wages from the union and non-union wage equations.

The union-premium variable has a positive and statistically significant coefficient at the 1% level in both columns 3 and 5, which confirms that union membership choice is positively influenced by the earnings-differential between union and non-union members. The probability of being in a union increases by 0.82% to 0.95% when the union premium increases by 1 percentage point.

The signs of the estimated variable coefficients used in the structural-form probit are the same as those in the reduced form probit except for the primary and secondary education dummies and the general training dummy, all of which are now positive but not statistically significant even at the 5% level. The magnitudes of the marginal effects, however, generally differ slightly between the structural and reduced form probit models, with some of the estimated coefficients more statistically significant, and others less so. For instance, the secondary and university education dummies are not significant in the structural-form probit, and the size of their coefficients declines. On the other hand, the magnitude of some of the coefficients for location is larger, and, for Nairobi, is statistically significant in the structural form probit equation. Also, the significance level of the marginal effect of foreign ownership firms on union status is higher in the structural-probit equations.

Table 3 shows the results of the estimated wage (and earnings) equations. The results from the reduced-form probit are used to control for sample selection when estimating the earnings functions for union and non-union workers. A test for heteroscedasticity shows that it is not a serious problem for the union and non-union wage-equations or for the union earnings-equation, but it is a serious problem for the non-union earnings equation (see Table 3), so the standard errors for heteroscedasticity are corrected using White's 1980 method. The validity of pooling data is also tested for on union and non-union workers using Hausman's specification test. The test rejects pooling, and therefore separate wage and earnings equations for union and for non-union members are estimated. The earnings results for union and for non-union workers based on wages and earnings are shown in columns 1–4 of Table 3. The Lagrange Multiplier test is also used to test for over-identification of the two instruments (identifiers) in the model. The test rejects the null that the instruments are not valid.

All the coefficients on primary, secondary, and university education dummy variables are positive and statistically significant except for the primary education dummy, which is not statistically significant even at the 5% level in the union equations. The estimated coefficients for the education dummies

Table 3. Earnings equations: sample-selection and treatment-effects models: dependent variable is natural logarithm of hourly earnings and wages (*t*-statistics in parentheses)

	Sample selection model#		Sample selection modelΨ		Treatment effects model	
	Union	Non-union	Union	Non-union	All workers #	All workers Ψ
Constant	2.2138**	1.0104**	2.1623**	1.2598**	1.3339**	1.5305**
	(7.72)	(4.64)	(7.88)	(5.89)	(7.82)	(8.95)
Primary dummy	0.0583	0.1208**	0.0712	0.1100**	0.0952**	0.0903*
	(1.50)	(2.82)	(1.91)	(2.61)	(2.86)	(2.69)
Secondary dummy	0.2437**	0.2851**	0.2306**	0.2632**	0.2910**	0.2610*
	(4.74)	(5.68)	(4.68)	(5.34)	(7.84)	(7.01)
University dummy	1.1878**	1.1349**	0.9065**	0.9302**	1.2019**	0.9828**
	(6.61)	(10.35)	(5.28)	(8.64)	(12.85)	(10.46)
Age	−0.0145	0.0464**	−0.0083	0.0343**	0.0300**	0.0235**
	(−0.96)	(3.81)	(−0.57)	(2.86)	(3.34)	(2.61)
Age-squared	0.0003	−0.0004**	0.0002	−0.0003*	−0.0002**	−0.0002**
	(1.40)	(−2.63)	(0.98)	(−2.03)	(−2.00)	(−1.65)
	(1.40)	(−2.63)	(0.98)	(−2.03)	(−2.00)	(−1.65)
Tenure	−0.0034	0.0098**	−0.0038	0.0041	0.0037	0.0008
	(−1.04)	(3.04)	(−1.20)	(1.28)	(1.65)	(0.35)
On-the-job training	−0.0480	0.1097**	−0.0061	0.1234**	0.0580*	0.0772**
	(−1.20)	(2.94)	(−0.16)	(3.37)	(1.97)	(2.61)
General training	0.1581**	0.2152**	0.1754**	0.2404**	0.2092**	0.2271**
	(3.48)	(5.64)	(4.05)	(6.42)	(6.73)	(7.27)
Firm size	0.0001	0.0002**	0.00002	0.0001**	0.0001**	0.0001*
	(1.36)	(4.40)	(0.54)	(2.63)	(3.73)	(2.19)
Located in Nairobi	0.2836**	0.4859**	0.2094**	0.3838**	0.4357**	0.3497**
	(4.37)	(10.26)	(3.37)	(8.26)	(10.64)	(8.49)
Located in Mombasa	0.1602*	0.2111**	0.1395	0.1829**	0.2191**	0.2011**
	(2.10)	(3.91)	(1.91)	(3.45)	(4.59)	(4.18)
Located in Nakuru	−0.0276	−0.0115	−0.1067	−0.0862	−0.0038	−0.0720
	(−0.36)	(−0.20)	(−1.44)	(−1.54)	(−0.08)	(−1.43)
Metal sector	0.4291**	0.0408	0.3226**	0.1006	0.2002**	−0.1943**
	(6.03)	(0.61)	(4.73)	(1.52)	(3.72)	(−3.59)
Wood sector	0.1110	0.0510	0.0754	0.0868	0.0753	0.0870
	(1.54)	(0.70)	(1.09)	(1.22)	(1.28)	(1.47)
Textile sector	0.2162**	0.2069**	0.1145	0.1872**	0.1775**	0.1415**
	(3.19)	(2.89)	(1.76)	(2.67)	(3.28)	(2.60)
Union density	−0.0059**	0.0013	−0.0054**	0.0004	−0.0016**	−0.0019**
	(−4.78)	(1.25)	(−4.59)	(0.36)	(−4.08)	(−5.33)
Related to firm owner	0.1207	0.0024	−0.0435	0.0449	0.0813	0.0840
	(0.82)	(0.04)	(−0.31)	(0.68)	(1.47)	(1.52)
Interviewed more than once	−0.0796*	0.0230	−0.0666	0.0282	−0.0102	−0.0067
	(−2.05)	(0.61)	(−1.80)	(0.76)	(−0.34)	(−0.23)
Employee in foreign-owned firm	−0.0002	0.1864**	−0.0108	0.1448**	0.1231**	0.0937**
	(−0.01)	(5.69)	(−0.35)	(4.50)	(4.80)	(3.62)
Production worker dummy	−0.3558**	−0.1955**	−0.3520**	−0.1950**	−0.2825**	−0.2724**
	(−6.83)	(−4.66)	(−7.06)	(−4.73)	(−10.76)	(−10.51)
Union premium					0.4328**	0.4772**
					(7.27)	(10.18)
Selection term	1.2523**	−0.0242	1.1703**	−0.313	−0.3487**	−0.4101**
	(4.84)	(−0.10)	(4.74)	(−1.33)	(−11.60)	(13.67)
Sector-specific time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Waves	Yes	Yes	Yes	Yes	Yes	Yes
σ_{ui}	0.430	0.552	0.412	0.543	0.560	0.576
Number of observations	1023	1741	1023	1741	2764	2764

Notes: *Significant at 0.05 level; **Significant at 0.01 level.

are higher as the level of education increases. This structure of earnings means that private average returns to education increase as the level of education rises. Age and age-squared have the expected positive and negative signs but are only statistically significant in the non-union wage and earnings equations, suggesting that general work-experience is more beneficial to non-union than to union workers.

Tenure has a negative but statistically insignificant effect on earnings. However, for the non-union earnings-equation, tenure has a positive and statistically significant effect.

The coefficient on the *general* training dummy is positive and statistically significant. This coefficient is also consistently higher than for on-the-job-training probably due to the fact that firms try to recover training costs from those who benefit from specific training, and at the same time pay more to those with general training to reduce worker turnover. It also seems that non-union workers benefit more from on-the-job training than do union workers.

Another notable variable in the regression is firm size, which has a positive and statistically significant effect on non-union earnings. Thus those working in larger firms receive higher earnings than those working in smaller firms. Relative to working in Eldoret, working in Nairobi and Mombasa generally seems to have a statistically significant positive effect on earnings.

For union workers, those in metal, wood, and textile sectors receive higher wages and earnings than did those in the food sector, but this is not statistically significant for the wood sector, or for textile sector wages. Non-union workers in the metal, wood and textile sectors also receives higher earnings than do those in the food sector, but this is statistically significant only for the textile sector. Non-union workers may free-ride, in the sense that their incomes are significantly increased when union-density goes up (see Neumark and Wachter, 1995, for US evidence on this union effect). Non-union workers employed in foreign-owned firms receive higher earnings than do union workers employed in the same firms. Consistently across the estimated equations, production workers receive significantly lower earnings than do those in positions such as management, administration and supervision. The selection-term shows that there is positive selection into unions and negative selection into non-union, though only statistically significant in the union equations.

Columns 5 and 6 of Table 3 report on estimation of the treatment-effects model (Equation 6). As indicated earlier, the Hausman test rejects the null hypothesis that the wage (and earnings) coefficients are the same for union and non-union workers. This means that data for the two sub-samples cannot be pooled when estimating the union effect.

The wage and earnings equations for union and non-union sub-samples (Table 3) are used to compute the union premium, obtained by taking the difference between predicted wages or earnings using the union and non-union wages (and earnings) equations reported in columns 1–4 of Table 3. The average union premium in the sample ranges from 4 to 5%, statistically significant at the 1% level. This finding shows that it is important to take into account selectivity and endogeneity of union membership in estimating the wage (and earnings) effects of union status. It is also necessary to avoid pooling of data on union and non-union workers when measuring the union premium. If, for example, one were to correct for sample selection and the endogeneity of union membership, but erroneously pool the data, the union premium would be overestimated, in this case from 4–5% to about 43–48% (see columns 5 and 6 of Table 3).

These results are quite different from those obtained from OLS (see Table A2). The OLS results show that the union effect on earnings is negative. Thus it is important to take into account the endogeneity of union membership when analysing the effect of union membership on earnings. These findings strongly contradict earnings effects of union membership previously reported for Kenya (House and Rempel, 1976; Manda, 1997). Both of these OLS studies found the effect of unionization on earnings to be negative. The negative effect reported in these studies is attributed to the failure to correct for sample selection and the endogeneity of the union status. When appropriate estimation methods are used, the conclusion reached by House and Rempel (1976) that the negative effect of unions on earnings is due to substitution between the real wage and other union-provided goods such as job security, employment, and fringe benefits, is not supported by the data.

IV. Conclusions

The study has examined the effect of trade unions and other factors such as education, firm size, and tenure on wages and earnings. Using individual wages or earnings as the dependent variables, similar results were obtained on impact of union membership. The results show that earlier results indicating a negative effect of union membership on wages and earnings are due to mis-specification, that is, for lack of attention to the endogeneity of the workers' union-status and to the sample selectivity problem. When these problems are taken into account in the regression for male manufacturing workers, there is a significantly positive effect of union membership on wages and earnings. The union-status equation showed that elite-workers tend to abstain from union membership, while production workers tend to be union members. Unions thus primarily benefit the less-skilled labour, a finding also reported by

Schultz and Mwabu (1998) for South Africa. The study has also shown that one of the reasons why workers become union members is to get a positive wage premium, in addition to getting protection from excessively long hours of work and from arbitrary job loss.

Acknowledgement

We would like to thank Henry Ohlsson and Rick Wicks for very useful comments.

References

Anderson P. (1993) Labour market structure in a controlled economy: the case of Zambia, PhD Thesis, Göteborg University.

Azam J.-P. and Ris, C. (2001) Rent-sharing, hold up and manufacturing sector wages in Cote d'Ivoire, World Bank, Working Paper No. 2600.

Bigsten, A. (1984) *Education and Income Determination in Kenya*, Gower, Aldershot.

Bigsten, A. and Kimuyu, P. (Eds) (2001) *Structure and Performance of Manufacturing in Kenya*, Palgrave, London.

Booth, A. (1995) *The Economics of the Trade Union*, Cambridge University Press, Cambridge, New York and Melbourne.

Davidson, R. and MacKinnon, J. G. (1993) *Estimation and Inference in Econometrics*, Oxford University Press, Oxford.

House, W. J. and Rempel, H. (1976) The impact of unionisation on negotiated wages in the manufacturing sector in Kenya, *Oxford Bulletin of Economics and Statistics*, **38**, 111–23.

Jakubson, G. (1991) Estimation and testing of the union wage effect using panel data, *The Review of Economic Studies*, **58**, 971–91.

Johnson, G. E. (1971) The determination of individual hourly earnings in urban Kenya, Discussion Paper No. 115, IDS, Nairobi.

Lanfranchi, J., Ohlsson, H. and Skalli, A. (2002) Compensating wage differentials and shift work preferences, *Economics Letters*, **74**, 393–98.

Lee, L. F. (1978) Unionism and wage rates: a simultaneous equation model with qualitative and limited dependent variables, *International Economic Review*, **19**, 415–33.

Lewis, H. G. (1986) Union relative wage effects: a survey in *Handbook of Labour Economics*, Vol. 2, (Eds) O. Ashenfelter and R. Layard, North-Holland, Amsterdam, pp. 1139–81.

Maddala, G. S. (1983) *Limited Dependent and Qualitative Variable in Econometrics*, Cambridge University Press, Cambridge.

Manda, K. D. (1997) Labour supply, returns to education, and the effect of firm size on wages: the case of Kenya, PhD Thesis, Göteborg University.

Mazumdar, D. and Mazaheri, A. (2002) *Wages and Employment in Africa*, Ashgate, Aldershot.

McDonald, J. M. and Solow, R. M. (1981) Wage bargaining and employment, *American Economic Review*, **71**, 896–908.

Neumark, D. and Wachter, M. L. (1995) Union effects on non-union wages: evidence from panel data on industries and cities, *Industrial and Labor Relations Review*, **49**, 20–38.

Nickell, S. J. and Wadhwani, S. (1990) Insider forces and wage determination, *Economic Journal*, **100,** 496–509

Robinson, C. (1989) The joint determination of union status and union wage effects: some tests of alternative models, *Journal of Political Economy*, **97**, 639–67.

Sandbrook, R. (1975) *Proletarians and African Capitalism: The Kenyan Case*, Cambridge University Press, Cambridge.

Schmidt, P. and Strauss, R. (1976) The effect of union on earnings and earnings on unions: a mixed logit approach, *Internal Economic Review*, **17**, 204–12

Schultz, T. P. and Mwabu, G. (1998) Labour unions and the distribution of wages and employment in South Africa, *Industrial and Labour Relations Review*, **51**, 680–703.

Wooldridge, J. M. (2002) *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge MA.

Appendix

Table A1. Definition of variables

Variable	Definition
Hourly wages	Natural logarithm of hourly wages (salary and wages)
Hourly earnings	Natural logarithm of hourly earnings (wages plus allowances and other benefits such as bonus)
Primary dummy	1 if individual has primary education, 0 otherwise
Secondary dummy	1 if individual has secondary education, 0 otherwise
University dummy	1 if individual has university education, 0 otherwise
Age	Age of the worker in years
Age-squared	The square of age of the worker in years
Tenure	Total number of years the individual has worked in the current firm
On-the-job training	1 if individual has had on-the-job training, 0 otherwise
General training	1 if individual has had general training (e.g., vocational training), 0 otherwise

(*continued*)

Table A1. Continued

Variable	Definition
Firm size	The total number of workers in the firm
Located in Nairobi	1 if individual works for firm located in Nairobi, 0 otherwise
Located in Mombasa	1 if individual works for firm located in Mombasa, 0 otherwise
Located in Nakuru	1 if individual works for firm located in Nakuru, 0 otherwise
Metal sector	1 if individual works for firm in metal sector, 0 otherwise
Wood sector	1 if individual works for firm in wood sector, 0 otherwise
Textile sector	1 if individual works for firm in textile sector, 0 otherwise
Union density	The proportion of union workers in a firm
Interviewed more than once	1 if individual was interviewed more than once, 0 otherwise
Employee in foreign-owned firm	1 if individual works in a foreign owned firm, 0 otherwise
Production worker	1 if individual works as a production worker, 0 otherwise
Recent firm investment in equipment per worker	The value in Kenya shillings of recent investment in equipment per worker in the firm divided by 1000
Weekly hours of work	Number of hours worked in a week
Weekly hours of work squared	The square of number of hours of worked in a week
Union membership	1 if individual is a union member, 0 otherwise
Relative of firm owner	1 if individual is a relative of the owner of the firm, 0 otherwise
Sector specific time dummies	(Six sector specific time dummies (two for wood, two for food and two for textile)
Waves	(Two dummies)

Table A2. OLS earnings equations: dependent variables are natural logarithm of hourly wages and hourly earnings

	Wage equation		Earnings equation	
	Coefficient	*t*-value	Coefficient	*t*-value
Constant	1.2524**	8.51	1.1014**	7.25
Union membership	−0.2023**	−8.85	−0.1449**	−6.13
Primary dummy	0.0805**	2.75	0.0858**	2.84
Secondary dummy	0.1930**	5.99	0.2324**	6.99
University dummy	0.8389**	10.44	1.0801**	13.01
Age	0.0418**	5.43	0.0455**	5.72
Age-squared	−0.0005**	−4.40	−0.0005**	−4.28
Tenure	0.0055**	2.85	0.0078**	3.88
On-the-job training	0.1225**	4.4	0.0895**	3.38
General training	0.1738**	6.44	0.1611**	5.78
Firm size	0.0001**	3.84	0.0002**	5.2
Located in Nairobi	0.3802**	10.57	0.4610**	12.39
Located in Mombasa	0.2096**	5.01	0.2257**	5.23
Located in Nakuru	−0.0533	−1.21	0.0123	0.27
Metal sector	0.1354**	2.86	0.1501**	3.07
Wood sector	0.0923	1.79	0.0792	1.49
Textile sector	0.1979**	4.17	0.2254**	4.6
Union density	0.0011**	3.99	0.0010**	3.48
Interviewed more than once	0.0258	1	0.0155	0.58
Employee in foreign owned firms	0.1152**	5.14	0.1416**	6.11
Production worker dummy	−0.1832**	−8.34	−0.2078**	−9.17
Relative of owner of the firm	−0.0437	−0.91	0.0261	−0.57
Waves	Yes		Yes	
Sector specific time dummies	Yes		Yes	
Adjusted R^2	0.3242		0.382	
Number of observation	2764		2764	

Notes: *Significant at the 0.05 level.
**Significant at the 0.01 level.

Table A3. Tests

Item tested	Type of test	Test value	Prob. value
1. Value of excluded instruments	Likelihood ratio test	χ^2 (3) = 51.94	0.0000
2. Test for pooling union and non-union data	Hausman test	χ^2 (28) = 1208.47	0.0000
3. Over-identification	Lagrange multiplier test	χ^2 (DF) = value	χ^2-critical
i. Union earnings equation	Lagrange multiplier test	χ^2 (31) = 21.07	43.77
ii. Union wage equation	Lagrange multiplier test	χ^2 (31) = 22.19	43.77
iii. Non-union earnings equation	Lagrange multiplier test	χ^2 (31) = 29.94	43.77
iv. Non-union wage equation	Lagrange multiplier test	χ^2 (31) = 13.93	43.77