Subsidies, selectivity and the returns to education

in urban Papua New Guinea

John Gibson and Osaiasi Koliniusi Fatai,

Department of Economics

University of Waikato

Private Bag 3105

Hamilton

NEW ZEALAND

Fax: 64-7 838-4031

Email: jkgibson@waikato.ac.nz

Running head: “Returns to education in Papua New Guinea”

Abstract

There is debate about whether the rate of return to education in developing countries declines with the level of schooling. This paper reports evidence from urban Papua New Guinea which shows that the average private rate of return to an additional year of education rises with the level of education considered. This pattern is robust to the specification of the extended earnings function and is strengthened when the measure of employee compensation includes fringe benefits. The results are not sensitive to controls for sample selectivity bias and the estimated selectivity effects are more consistent with the principle of comparative advantage than is existing evidence from developing countries. The returns to on-the-job training are also higher than indicated by previous evidence in developing countries. [*JEL:* I21, J31]

Keywords: demand for schooling, human capital, rate of return

1. Introduction

According to the influential review by Psacharopoulos (1994), rates of return to education follow the same pattern as investment in conventional capital, in declining as the investment is expanded. Hence, both private and social returns to education fall with the level of schooling. For example, Psacharopoulos reports average private rates of return of 39% for the completion of primary schooling in Asia, as opposed to returns of only 19% for the completion of secondary and higher education. This pattern of decline may support the emphasis made by institutions like the World Bank on making investments in primary education a priority for developing countries.

These findings have generated debate about whether the evidence is consistent with these patterns in sub-Saharan African countries (Bennell, 1996) and Asian countries (Bennell, 1998). The claim that private rates of return decline with the level of education is also inconsistent with the model of education in developing countries presented in a leading textbook (Todaro, 2000). In this model, the expected private returns rise with the level of education because the supply of educated labour grows faster than the supply of formal sector jobs, leading to a continuous upgrading of the educational requirements for jobs previously filled by less educated workers. Debate has even come from within the World Bank, with some supporters of greater funding for higher education arguing against the assumption that poor countries should focus predominantly on improving access to basic education (World Bank, 2000). Similarly, Curtin and Nelson (1999) have argued that concentrating public investment on primary education will perpetuate poverty.

This paper reports new evidence relevant to the debate about declining rates of return by level of education. Data from a household survey in urban Papua New Guinea are used to estimate Mincerian earnings functions, where the effect of school years on wages is allowed to vary with the level of education. A contentious debate in Papua New Guinea about public investments in education may be informed by the estimates reported here, although the education system has restructured since the survey. But the larger potential contribution of this paper is that the data are unusually detailed and allow checks on the robustness of the Mincerian earnings function. For example, information on employee remuneration includes not just wages but also employer-provided subsidies ('fringe benefits') so it is possible to see how the pattern of returns to education varies as the measure of remuneration becomes broader. Information is also available on a number of potential covariates, such as union membership and on-the-job training, which are often excluded from wage equations in developing countries. The survey also allows detailed consideration of labour force participation and sample selection issues.

Duncan (1976) first showed that the explanatory power of earnings equations rises when fringe benefits are included in a compensation measure. But there has been little subsequent work in this area, which is surprising because in the

United States, for example, non-wage compensation represents about one-quarter of total employer compensation costs (Katz and Autor, 1999). These fringe benefits are relatively more important in high wage jobs, so their exclusion may affect the estimated returns to schooling. Casual observation suggests that a wide range of fringe benefits are offered in low-income countries, including free food, housing and transport, so measuring the effect of these benefits may be particularly useful in this setting.[1]

Another feature of earnings functions for developing countries is that they often exclude measures of on-the-job training. The quadratic in potential labour market experience in the Mincerian equation is a proxy for post-school investments in human capital but earnings equations in developed countries also often include indicators of on-the-job training. Indeed, such estimates suggest that training can provide benefits to workers, firms and the overall economy (Blundell *et al.*, 1999). In contrast, in one of the few studies from developing countries, Kugler and Psacharopoulos (1989) suggest that job training has virtually no impact on earnings independent of the effects of formal education. Because public investment in human capital can be channelled either through the formal education system or by subsidising enterprise training, further evidence on the effect of training on wages in developing countries may be helpful.

The final feature of earnings equations that is important to this paper is that only a minority of the working age population in developing countries are engaged in wage employment. While ‘earnings’ can be imputed for the self-employed,[2] such estimates will include returns to business capital and risk, and are likely to be error-ridden, reducing their comparability with wages (Schultz, 1999). Whether the estimates obtained from wage-earners accurately capture the expected payoff to schooling for the population depends on the extent of sample selectivity bias. Correcting this bias requires modelling the decision to engage in wage labour, using factors that should affect reservation wages but not market wage offers (e.g., non-labour income), and assuming joint normality of the errors in the participation and wage equations (Heckman, 1979). But recent examples of this approach report counter-intuitive results, with the unobservable characteristics of high-wage workers being opposite to the characteristics that select workers into wage-earning in the first place (Horowitz and Schenzler, 1999; Siphambe, 2000). With these negative selectivity terms, the mean wage amongst the sample of workers is lower than the population mean wage that would be observed if everyone was a worker. Hence, it is worth seeing if evidence of selectivity bias in other developing countries provides more support for the principle of comparative advantage.

In addition to these potential contributions of this paper, there is one important caveat because we lack accurate estimates of the private and social costs of education at different levels in Papua New Guinea. Consequently, the “full” method of calculating the rate of return to investments in education is not emphasised in the paper. We note in passing that approximate results can be obtained by combining age-earnings profiles estimated from the survey with cost estimates originally made by Gannicott (1987) and updated by McGavin (1991). These approximations suggest that at the time of the survey in the mid-1980s, both private and social rates of return to secondary school education in urban Papua New Guinea exceeded the rates of return to primary and tertiary education (Fatai, 2001).

2. Education in Papua New Guinea

At the time of Independence in 1975, just 56% of Papua New Guineans in the primary school age groups were in schools and just 12% of those in the eligible age groups were in secondary schools (Connell, 1997). The situation was even worse at university level; the opening of the University of Papua New Guinea in 1967 allowed the number of graduates to rise from just four in 1968 to a few hundred by the mid-1970s. A second university was established in 1973 which contributed to the high cost of the tertiary sector in terms of fragmentation and lack of scale. Since then, while the number of universities has continued to grow, each has remained very small and the tertiary sector is still unable to meet the demand for skilled workers, so Papua New Guinea continues to employ a large number of expatriate professionals.

For the mid-1980s, McGavin (1991) reports participation rates of 64% for primary schooling and 15% for lower-secondary (or Provincial high) schooling. After the completion of Provincial high school (Grade 10) students could go on to a variety of other courses, including university study, although access to this higher education was limited. Participation rates were just 2% for National high school

(Grades 11 and 12) and also 2% for university and other post-secondary education (McGavin, 1991). The shortage of educated Papua New Guineas is also shown by the fact that National high school students were covered by the same scholarships and conditions (e.g., free air travel from their home villages to one of the four National high schools) as the university students.

Because of the low level of human capital in Papua New Guinea, relatively large public investments are made in education – as a percentage of national income more than double the average for developing countries (Jarrett and Anderson, 1989). However, many commentators have suggested that these public investments are not allocated to the right levels of education. On the one hand, Curtin and Nelson (1999) claim that resources are wasted by not expanding post-primary schooling rapidly enough, while Gannicott (1987) claims that resources are wasted by not concentrating on the primary sector. The claims made by Gannicott (1987) were based on 1983 estimates of the private and social rates of return that declined by education level and were especially low (1%) for social returns at the university level. However, Gannicott's analysis and the follow-up by McGavin (1991) were not based on large scale sample surveys covering both public and private sector workers and those outside the wage labour force, so there may be doubts about the conclusions.

Regardless of the evidence assembled by economists in their debates about which level of education to expand in Papua New Guinea, a major reform of the education system was begun in 1995. Under the reforms, an initial three-year elementary school is taught in vernacular language before students enter primary level education at Grade 3. For older students, Grades 7 and 8 are no longer taught at Provincial high schools, and are now taught at the primary level, which previously started at Grade 1 and terminated at Grade 6. In turn, Grade 11 and 12 that were previously taught only at the National high schools, are now taught at some Provincial high schools. The reforms have effectively expanded access to secondary schooling (especially Grades 7 and 8), perhaps because primary school graduates were failing to obtain jobs in the formal sector due to the weak growth in labour demand (Levantis, 2000).

3. The model

3.1 Conceptual framework

The basic earnings function of Mincer (1974) is used, where the logarithm of the after-tax hourly wage rate, *w* is regressed on completed years of schooling, *S* and potential years of labour market experience, *Exp* and its square. In this specification:

$$\ln w = b_0 + b_1 S + b_2 Exp + b_3 Exp^2 + u \qquad (1)$$

the regression coefficient b_1 can be interpreted as the average private rate of return to one additional year of education, under the assumption that the only (private) cost of education is foregone earnings. One drawback of the basic earnings function is that the average rate of return estimated by b_1 applies regardless of the educational level to which the additional year of schooling refers. To overcome this drawback, 'extended' earnings functions may be estimated with the continuous years of schooling variable converted into a series of intercept dummy variables referring to the completion of the main schooling cycles (Psacharopoulos, 1994). More generally, these dummy variables may be interacted to allow the estimation of separate effects for the number of years completed at each schooling level (Schultz, 1999). The most general approach is to use a string of dummy variables for each year of schooling completed, treating log earnings as a step function of years of education, with a separate step for each year (Patrinos, 1996).

3.2 Econometric issues

When only a subset of the working age population are wage earners, the model becomes more complex, with a wage equation, $w_j = \mathbf{z}_j b + u_{1j}$ where the **z**-vector contains the schooling and experience variables, and the wage for person j is observed according to the selection equation:

$$\mathbf{x}_j g + u_{2j} > 0$$
$$\text{where} \quad u_1 \sim N(0, s) \quad u_2 \sim N(0,1) \quad corr(u_1, u_2) = r\ . \tag{2}$$

Non-random selection implies that the disturbances in the two equations are not independent (i.e., $r \neq 0$). Ignoring the sample selectivity will produce biased estimates of the underlying relationship for the population. One common solution is to augment the regression equation with the Mills' ratio estimated from the selection equation (Heckman, 1979):

$$m_j = \frac{f(\mathbf{x}_j \hat{g})}{\Phi(\mathbf{x}_j \hat{g})} \tag{3}$$

with the statistical significance of the coefficient on this added variable, $l = rs$, providing a test of sample-selection bias. In this framework, the $\hat{b}$ should be unbiased estimates for the population once m_j is included in the regression equation. However, this two-step procedure is not as efficient as maximum likelihood estimation of b, g, r, and s with the selection and regression equations estimated jointly (StataCorp, 1997).

Somewhat similar to selection bias is the problem of endogenous 'treatment

which may be relevant to some of the wage equation covariates, such as union membership and public sector employment. The problem occurs if workers self-select into unions or the public sector, rather than being randomly allocated into these positions. The resulting wage equation is:

$$w_j = \mathbf{z}_j b + q d_j + u_{1j} \tag{4}$$

where q shows the effect on wages of whether or not the 'treatment' is applied (i.e., whether or not the person was a member of the union, or was placed in the public sector). The problem for estimation is that d_j may be correlated with u_{1j} potentially causing bias in *all* of the estimated coefficients of the wage equation. The correlation occurs because the treatment is applied according to the selection equation:

$$d_j = 1 \quad \text{if} \quad \mathbf{x}_j g + e_j > 0$$
$$\text{where} \qquad u_1 \sim N(0, s) \qquad e \sim N(0,1) \qquad corr(u_1, e) = r\,. \tag{5}$$

The solution is to form the Mills' ratio from a probit model for the receipt of the treatment and add this Mills' ratio to the wage equation. But unlike the usual sample-selectivity correction, the Mills' ratio is formed for both those who receive the treatment (that is, join the union or the public sector) and those who do not, because the wage is observed in each case.

4. Data

The data were collected by the Papua New Guinea Urban Household Survey, which was carried out during 1985-87, with field work in each urban area staggered over 12 months to capture any seasonal effects. Almost 1400 households containing 4000 adults were surveyed, and for each of these adults information was obtained on education levels, employment activities and incomes in the previous two weeks. Restricting attention to those aged 15-64 years reduces the sample to 3887, comprised of 3625 citizens and 262 expatriates. Because the paper aims to inform debates about the returns to different levels of education *within* Papua New Guinea, the results for overseas-educated expatriates may not be relevant, so they are excluded.[3]

Just over 40% of this sample of the urban working age population were employed for wages or salaries (n=1510). The employment to population ratio varies from only 22% for those with no schooling to 84% for those with tertiary-level education, so non-random selection into the observed sample of wage workers may affect the estimated returns to schooling.

Amongst the employees, the survey collected detailed information on both wages and 11 types of employer-provided subsidies. The most common forms of subsidies were rent and utilities, transport, food, and family education expenses. Almost one-half of the employees (47%) received some form of subsidy as part of their total remuneration, with non-wage payments more important for the most educated workers (Table 1). Almost one-third of the total remuneration of workers with tertiary education is in the form of non-wage benefits, compared with less than one-tenth for unschooled workers. This lack of proportionality means that ignoring non-wage benefits might bias the comparison of rates of return by schooling level.

(Table 1 about here)

To examine the impact of including non-wage benefits, the wage equations are estimated with two, alternate, dependent variables. The logarithm of the after tax hourly wage rate is used to provide comparability with most studies in the literature, while the broader measure includes the hourly equivalent value of the employer-provided subsidies. Descriptive statistics for the dependent and explanatory variables of the wage equations are reported in the second and fourth columns of Table 2. In addition to school years and experience, the explanatory variables include five covariates: marital status, years of tenure with the employer, sector of employment, union membership, and receipt of on-the-job training. It appears from the descriptive statistics that compared to female employees, males receive higher wages, have more potential experience and longer tenure, are more likely to be union members and to receive on-the-job training, but have fewer years of education.[4]

(Table 2 about here)

There appear to be important differences between the average characteristics of workers and those of the working age population (Table 2). Wage employees are more highly educated, with this pattern especially apparent for women, so one part of the payoff to schooling is likely to be the improved probability of obtaining a wage job. But the selection mechanisms may differ between men and women because, for example, male employees are more likely to be married than are other males, while the reverse pattern holds for females. The selection mechanism is also presumed to have two determinants that are not part of the wage equation – non-labour income and the presence of extended family members in the household. The assumption regarding extended household structure is that it enables greater participation by women due to the increased opportunities for cooperative childcare (Horowitz and Schenzler, 1999), while non-labour income should raise reservation wages.

5. Estimation results

Table 3 contains the basic earnings functions estimates. All of the coefficients are significant at the 1% level (except for the squared experience term for females) and the model explains 39% of the variation in hourly wages. This is comparable to the performance of the basic earnings function in other developing country samples (for example, Kugler and Psacharopoulos, 1989; Tannen, 1991; Siphambe, 2000). The coefficient on years of schooling implies an average private rate of return to an additional year of education of 11.1%, which is close to the average of 11.7% for lower middle income countries reported by Psacharopoulos (1994, Table 3).[5] Wages also grow with potential labour market experience until 42 years of experience have been accumulated, and thereafter experience contributes negatively to earnings.

(Table 3 about here)

The coefficient on the male intercept dummy in the first column of Table 3 suggests that hourly wages are 13.4% higher for men than for similarly educated and experienced women.[6] In the gender-disaggregated results in the final two columns, the returns to education for women appear to be higher than for men but the *t*-statistic on the difference is only 1.40 (p<0.17). Given these results and the small sample of women in wage employment, we proceed with the gender-pooled sample (although we return to testing for differences by gender below).

5.1 Including non-wage benefits

When the dependent variable of the earnings function includes the value of employer-provided subsidies as well as wages, the estimated rate of return to an additional year of education rises from 11.1% to 12.5% (Table 4).

Therefore, the exclusion of non-wage benefits from the measure of compensation used in previous studies may have caused an understatement in the estimated returns to education. Some patterns across population sub-groups might also have been distorted because gender-disaggregated models show that this rise in the returns to education occurs only for males (from 10.9% to 12.6%). This gender-specific effect occurs because in urban Papua New Guinea there is a positive correlation between fringe benefits and wages for men (r=0.42) but not for women (r=0.01).

(Table 4 about here)

5.2 Allowing non-linearities

To see whether the private rate of return to schooling varies with the level of education, several different specifications of extended earnings functions were estimated. Dummy variables were constructed for the main schooling cycles – primary, secondary, and higher education – and these were also interacted with the continuous years of schooling variable.[7] Tests of zero restrictions favoured the spline model used by Schultz (1999); the F-statistic for excluding the intercept dummies was 0.60 (p<0.62) compared with 31.1 (p<0.00) for restricting the coefficient on years of schooling to be the same across the different levels of education. More direct confrontations of these two models, using informal criteria such as maximising $\bar{R}^2$, and more formal methods such as a Davidson-MacKinnon (1981) non-nested 'J' test also favoured the model with separate slope coefficients on the years of schooling variable.[8]

The results for the preferred specification of the extended earnings function are reported in Table 4. The statistical significance of the interaction terms between school years and the dummy variables for education level indicate that the average private rate of return to an additional year of education rises with the level of education considered. For secondary schooling, the Mincerian rate of return is 9.8%, with a standard error of 0.6%, and for higher education it is 10.6% (standard error of 0.5%). Including the effect of non-wage benefits increases the degree of non-linearity in the earnings function, with the 'premium' in the Mincerian rate of return for secondary school years rising to 3.1%, and for years of higher education rising to 4.2%. The evidence from Table 4 is not consistent with the general pattern described by Psacharopoulos (1994) of declining rates of return by education level although it is consistent with the model of education presented by Todaro (2000) and with empirical evidence reported by Siphambe (2000).

The uncertainty in the literature about whether the logarithm of wages is linear, concave or convex with respect to years of schooling suggests that the results in Table 4 should be backed up by some sensitivity analyses. Therefore, Figure 1 records the results of using two other methods of introducing non-linearities into the earnings equation: adding a quadratic term in years of schooling and replacing the continuous years of schooling variable with a string of dummy variables referring to the completion of each year. The interaction of years of school with dummy variables for educational level gives the most conservative estimates of the convexity of the average private returns schedule. Thus, it is unlikely that the results presented here are biased by the particular method used to allow non-linearities in the returns to schooling.

(Figure 1 about here)

5.3 Adding more covariates

To see if the pattern of increasing returns by level of education just reflects excluded relevant variables, five additional covariates are considered, both individually and jointly: tenure with the employer, marital status, sector of employment, receipt of on-the-job training, and union status. Part of the higher rate of return to secondary and higher education may be because that extra education is correlated with other desirable characteristics (Table 5).[9] For example, when tenure is added to the model, the premium in the return to years of secondary school education falls from 3.1% to 2.7% and for higher education falls from 4.2% to 3.7% (column (i)). Although there is a further fall in the premium for secondary and higher years of education when the added covariates are considered jointly (column (vi)) the separate slope coefficients still remain statistically significant, with implied rates of return to 6.0%, 8.2%, and 9.2% for each completed school year at primary, secondary and higher education levels. Hence, the convex private returns to education appear to be a robust feature of the urban labour market in Papua New Guinea.

(Table 5 about here)

All of the added variables in Table 5 act individually to raise hourly earnings, with union membership, receipt of on-the-job-training and being married having the biggest effects, while wages also rise by 3% for each year of tenure with the current employer. The magnitude of these effects is diminished by approximately one-third when all of the variables are added at the same time, and the coefficient on the public sector variable becomes negative and weakly significant (column (vi)). While some of these effects might be endogenous, the Mills' ratios from the treatment selection equations are not significant in the wage equations, so this source of bias may not be too important (Appendix Table 1). The results in the appendix assume that selection into the public sector, into unions, and into on-the-job-training depends on the status of *other* working members of the household, due to learning and the value of personal contacts in gaining access to sectors and programs with tacit restrictions.[10] In terms of the Mincerian rate of return to schooling, the introduction of the endogenous treatment effects does not much alter the degree of convexity, slightly raising the premium for years of secondary and higher education in two out of the three equations.

The results in Table 5 for on-the-job training deserve special attention because the literature on developing countries underemphasizes training compared with the attention paid to formal education. Workers with on-the-job training have conditional mean wages that appear 21% higher than other workers, controlling for their formal education, experience and gender (column iv).

When all of the additional covariates are added, the wage effect of training is smaller (14%), but still highly significant (p<0.01). This wage effect is considerably higher than the 6% rise due to training estimated by Kugler and Psacharopoulos (1989), and unlike those previous estimates, it persists even when the training incidence variable is interacted with the formal schooling variables.[11] Such results suggest that a more comprehensive evaluation of the returns to public investment in training relative to the returns from investments in formal schooling is needed. In Papua New Guinea, the government implicitly subsidizes some training, with public sector workers almost twice as likely to receive training as their private sector counterparts,[12] so it might be economically beneficial to extend training to the whole workforce by subsidising private sector training costs.

5.4 Controlling for sample selectivity bias

The results in Tables 3-5 give the average private rates of return to years of education for *wage earners*, but whether these estimates indicate the expected private payoff to schooling for the *population* depends on the extent of sample selectivity bias. There is mixed evidence of selectivity bias; the hypothesis that the unobservable errors in the selection and wage equations are uncorrelated, i.e. r=0, is rejected in the pooled sample but not in the separate male and female samples (Table 6). The coefficient on the Mills' ratio is positive and significant at the 1% level in the pooled sample, implying that the observed wage is higher than the wage offer facing a randomly selected individual in the population. The lower wage offers received by non-participants, compared with participants, accords with the principle of comparative advantage but contrasts with the pattern found in some of developing countries where λ is negative (Tannen, 1991; Horowitz and Schenzler, 1999; Siphambe, 2000).

(Table 6 about here)

Once non-random selection into wage earning is taken account of, there is slightly more non-linearity in the private returns to education. The premium for secondary school years rises from 3.1% in Table 4 to 3.8% in the selectivity-corrected estimates in Table 6, while for higher education it rises from 4.2% to 5.3%. It also appears that separate coefficients for primary, secondary and higher education years are needed in the selection equations, suggesting an increasing effectiveness of education in raising participation probabilities. The effect of education appears to be a particularly powerful influence on female participation in wage-earning.[13] The selectivity-corrected wage equation for females does not show non-linearity in the relationship between log wages and school years, in contrast to the results for males, but tests indicated no statistically significant differences between the schooling coefficients in the wage equations for men and women.

Sample selection models can be sensitive to specification changes, especially in terms of the exclusion restrictions, which are here based on the presence of extended family members in the household and non-labour income. The results in Table 6 appear plausible, with the presence of extended family members raising female participation in wage earning and lowering it for males, while non-labour income lowers participation. But as a sensitivity analysis, the results of a different specification are reported in Table 7, with marital status added to the selection and wage equations and tenure with the employer, sector of employment, receipt of on-the-job training, and union status added to the wage equation. These additional variables lower the estimated rates of return to all levels of education by about one-quarter but otherwise do not alter any of the inferences about sample selectivity bias and the non-linearity in the earnings equations.

(Table 7 about here)

6. Conclusions

The results reported here for urban Papua New Guinea are not consistent with the claim by Psacharopoulos (1994) that rates of return to education fall with the level of schooling. While data limitations force us to concentrate on the earnings function approach to estimating the private rate of return, the data do allow many sensitivity analyses which reinforce the pattern of average private rates of return rising with the level of education. This evidence is consistent with a model where the supply of educated labour grows faster than the supply of formal sector jobs, leading to a continuous upgrading of the educational requirements for jobs (Todaro, 2000). An implication of this model is that students will demand increasingly higher levels of education, which is consistent with the reforms of the education system in Papua New Guinea that have effectively expanded access to Grades 7 and 8 for a majority of students. Because the data used here are only for urban workers, and our results are for private rates of return, we cannot answer the question of where public investments in education in Papua New Guinea should be allocated. But our results are sufficiently different from those of Gannicott (1987) and McGavin (1991) to raise some doubts about their policy prescriptions of emphasising primary education.

There are also several other contributions of our results, related to non-wage benefits, on-the-job training, and sample selectivity bias, which are potentially relevant to many other developing countries. Including non-wage benefits in the measure of employee compensation raises the estimated rate of return, especially for higher levels of education and especially for males. The receipt of on-the-job training is associated with a significant rise in wages which is independent of the effect of formal schooling, so policy makers may need to consider whether investments in human capital are best made through public schooling or by subsidising enterprise training. The estimated sample selectivity effects for urban Papua New Guinea are also more consistent with the principle of comparative advantage than is existing evidence from other developing countries.

References

Bennell, P., (1996). Rates of return to education: does the conventional pattern prevail in sub-Saharan Africa? *World Development* 24(1): 183-199.

Bennell, P., (1998). Rates of return to education in Asia: a review of the evidence. *Education Economics* 6(2): 107-120.

Blundell, R., Dearden, L., Meghir, C., Sianesi, B., (1999). Human capital investment: the returns from education and training to the individual, the firm and the economy. *Fiscal Studies* 20(1): 1-23.

Connell, J., (1997). *Papua New Guinea: The Struggle for Development* London: Routledge.

Curtin, T.R.C., Nelson, E.A.S., (1999). Economic and health efficiency of funding policy. *Social Science and Medicine* 48(11): 1599-1611.

Davidson, R., MacKinnon, J., (1981). Several tests for model specification in the presence of alternative hypotheses. *Econometrica* 49(3): 781-794.

Duncan, G., (1976). Earnings functions and non-pecuniary benefits. *Journal of Human Resources* 11(4): 462-483.

Fatai, O.K., (2001). *Returns to education and development: a case study of urban Papua New Guinea*. Unpublished thesis submitted to the University of Waikato.

Gannicott, K., Education in Papua New Guinea: a case of wasted resources. Islands/Australia Working Paper No. 87/5. Canberra: National Centre for Development Studies, Australian National University.

Heckman, J., (1979). Sample selection bias as a specification error. *Econometrica* 47(1): 153-161.

Horowitz, A., Schenzler, C., (1999). Returns to general, technical and vocational education in developing countries: recent evidence from Suriname. *Education Economics* 7(1): 5-19.

Jarrett, F., Anderson, K., (1989). *Growth, Structural Change and Economic Policy in Papua New Guinea* Canberra: National Centre for Development Studies.

Katz, L., Autor, D., (1999). Changes in the wage structure and earnings inequality. *Handbook of Labor Economics, Vol 3* (O. Ashenfelter and D. Card, eds): Amsterdam: Elsevier, pp. 1464-1553.

Kugler, B., Psacharopoulos, G., (1989). Earnings and education in Argentina: an analysis of the 1985 Buenos Aires Household Survey. *Economics of Education Review* 8(4): 353-365.

Levantis, T., (2000). *Papua New Guinea: Employment, Wages and Economic Development* Canberra: Asia Pacific Press, Australian National University.

McGavin, P., (1991). Policy evaluation of investment in education: a Papua New Guinea case study. *Economics of Education Review* 10(3): 213-226.

Patrinos, H., (1996). Non-linearities in the returns to education: sheepskin effects or threshold levels of human capital? *Applied Economics Letters* 3(3): 171-173.

Psacharopoulos, G., (1994). Returns to investment in education: a global update. *World Development* 22(9): 1325-1343.

Schultz, P., (1999). Health and schooling investments in Africa. *Journal of Economic Perspectives* 13(3): 67-88.

Siphambe, H.K., (2000). Rates of return to education in Botswana. *Economics of Education Review* 19(2): 291-300.

StataCorp (1997). *Stata Statistical Software: Release 5.0,* College Station: Stata Corporation.

Tannen, M., (1991). New estimates of the returns to schooling in Brazil. *Economics of Education Review* 10(2): 123-135.

Todaro, M.P., (2000). *Economic Development* Longman, New York, 7th Edition.

World Bank (2000). *Higher Education in Developing Countries: Peril and Promise* Washington: World Bank.

Figure 1: Returns to Years of Schooling (Compared to Nil Years)

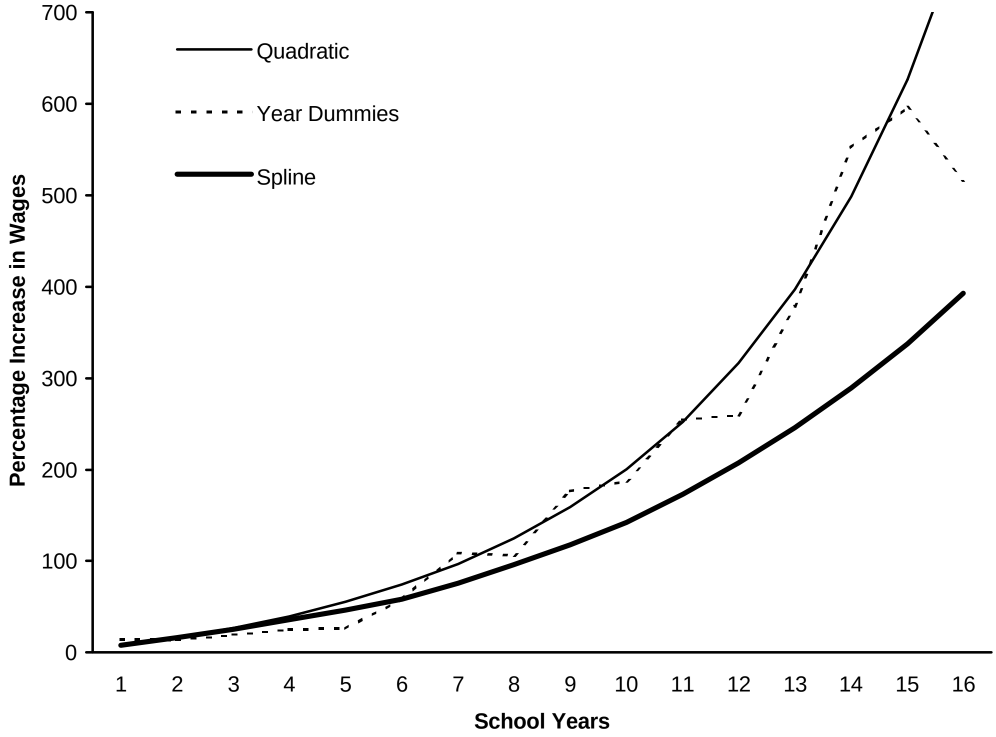

Table 1
Earnings and subsidies by schooling level in urban Papua New Guinea

Highest education level	Net hourly wage rate	Employer-provided subsidies (hourly equivalent)
No school	1.13	0.11
Primary	1.37	0.16
Secondary	1.97	0.47
Tertiary	2.52	1.21

Note: Monetary values are in 1986 Kina (K1=US$1 in 1986).

Table 2
Definition of variables, means and standard deviations

	Male		Female	
	Selection eqn	Wage eqn	Selection eqn	Wage eqn
Wage observed	0.596 (0.491)		0.184 (0.388)	
ln (hourly wage)		0.519 (0.565)		0.447 (0.500)
ln (wage+subsidies)		0.679 (0.666)		0.500 (0.489)
Years of schooling	6.641 (3.947)	7.477 (3.979)	5.183 (3.861)	9.003 (2.838)
Years of potential labor market experience[a]	16.873 (12.894)	17.561 (11.250)	17.138 (12.569)	11.448 (7.638)
Extended family live in household	0.730 (0.444)		0.731 (0.444)	
Non-labour income[b] (household total)	0.065 (0.269)		0.065 (0.266)	
Married	0.604 (0.489)	0.734 (0.442)	0.753 (0.431)	0.710 (0.454)
Tenure (years)		6.292 (6.220)		4.716 (4.769)
Public sector worker[c]		0.450 (0.498)		0.500 (0.501)
Union member		0.217 (0.413)		0.159 (0.366)
Received on-the-job training		0.390 (0.488)		0.338 (0.474)
No. of observations	2048	1220	1577	290

Note: Standard deviation in ( ).
[a] Defined as age minus completed school years minus seven.
[b] K'000 per fortnight.
[c] Includes national and provincial governments, statutory bodies and local government.

Table 3
Basic Mincerian earnings functions

Variable	All	Males	Females
Years of schooling	0.111	0.109	0.128
	(24.80)**	(22.71)**	(10.16)**
Years of experience	0.036	0.037	0.034
	(10.37)**	(9.53)**	(3.89)**
$(\text{Experience})^2 \div 100$	-0.043	-0.044	-0.032
	(5.58)**	(5.38)**	(1.33)
Male	0.126		
	(4.73)**		
Intercept	-0.886	-0.748	-1.031
	(14.01)**	(11.37)**	(6.59)**
R^2	0.39	0.39	0.39
F-statistic	$F_{(4,1505)}$=188**	$F_{(3,1216)}$=209**	$F_{(3,286)}$=43.1**
No. of observations	1510	1220	290

Note: The dependent variable is (log) hourly, after-tax wages. Heteroscedastically-robust *t*-statistics in parentheses; + significant at 10%; * significant at 5%; ** significant at 1%

Table 4
Extended earnings functions including non-wage benefits in remuneration measure

Dependent variable:	ln (wage)		ln (wage + subsidies)	
	Basic	Extended	Basic	Extended
Years of schooling	0.111	0.075	0.125	0.076
	(24.80)**	(8.23)**	(25.06)**	(7.88)**
Years × Secondary		0.023		0.031
		(4.21)**		(5.16)**
Years × Higher		0.031		0.042
		(5.33)**		(6.68)**
Years of experience	0.036	0.039	0.040	0.043
	(10.37)**	(11.08)**	(10.04)**	(10.91)**
$(\text{Experience})^2 \div 100$	-0.043	-0.050	-0.050	-0.059
	(5.58)**	(6.42)**	(5.88)**	(6.97)**
Male	0.126	0.125	0.249	0.248
	(4.73)**	(4.74)**	(8.90)**	(8.78)**
Intercept	-0.886	-0.799	-0.988	-0.870
	(14.01)**	(11.90)**	(14.10)**	(11.81)**
R^2	0.39	0.40	0.39	0.41
F-statistic	$F_{(4,1505)}$=188	$F_{(6,1503)}$=153	$F_{(4,1505)}$=200	$F_{(6,1503)}$=159

Note: Dependent variable is (log) hourly after-tax wage (with employer-provided subsidies where noted). Heteroscedastically-robust t-statistics in parentheses; + significant at 10%; * significant at 5%; ** significant at 1%. N=1510.

Table 5
The effect of adding covariates to the extended earnings functions

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Years of schooling	0.065 (6.95)**	0.074 (7.77)**	0.075 (7.81)**	0.073 (7.60)**	0.071 (7.54)**	0.060 (6.64)**
Years × Secondary	0.027 (4.79)**	0.028 (4.69)**	0.030 (4.97)**	0.027 (4.56)**	0.028 (4.84)**	0.022 (3.93)**
Years × Higher	0.036 (6.01)**	0.040 (6.30)**	0.040 (6.26)**	0.037 (5.76)**	0.039 (6.43)**	0.032 (5.18)**
Years of experience	0.025 (6.01)**	0.035 (7.72)**	0.043 (10.77)**	0.041 (10.59)**	0.037 (9.25)**	0.017 (3.72)**
$(\text{Experience})^2 \div 100$	-0.038 (4.42)**	-0.046 (5.18)**	-0.059 (7.05)**	-0.055 (6.61)**	-0.048 (5.75)**	-0.023 (2.57)*
Male	0.235 (8.67)**	0.257 (8.98)**	0.250 (8.95)**	0.228 (8.18)**	0.231 (8.28)**	0.216 (7.91)**
Tenure	0.029 (12.93)**					0.024 (10.07)**
Married		0.146 (4.48)**				0.103 (3.35)**
Public sector			0.063 (2.35)*			-0.050 (1.85)+
On-the-job training				0.193 (6.93)**		0.135 (5.06)**
Union member					0.296 (9.44)**	0.196 (6.42)**
Intercept	-0.700 (9.61)**	-0.854 (11.76)**	-0.874 (11.95)**	-0.853 (11.74)**	-0.795 (10.91)**	-0.652 (9.20)**
R^2	0.46	0.41	0.41	0.43	0.44	0.49
F-statistic	170**	140**	142**	146**	159**	121**

Note: Dependent variable is (log) hourly after-tax wage (including employer-provided subsidies). N=1510. Heteroscedastically-robust t-statistics in parentheses; + significant at 10%; * significant at 5%; ** significant at 1%.

Table 6
Maximum likelihood selection and wage equations

	All		Males		Females	
	Selection	Wage	Selection	Wage	Selection	Wage
Years of schooling	0.084 (6.43)**	0.084 (8.51)**	0.067 (3.89)**	0.074 (7.11)**	0.120 (4.34)**	0.076 (2.43)*
Years × Secondary	0.060 (6.03)**	0.038 (5.66)**	0.044 (3.57)**	0.030 (4.13)**	0.080 (4.48)**	0.024 (1.43)
Years × Tertiary	0.122 (10.64)**	0.053 (6.61)**	0.108 (7.62)**	0.044 (4.99)**	0.131 (6.46)**	0.021 (1.04)
Years of experience	0.112 (15.29)**	0.054 (9.41)**	0.138 (15.87)**	0.043 (5.31)**	0.073 (4.61)**	0.031 (3.18)**
$(\text{Experience})^2 \div 100$	-0.188 (12.67)**	-0.077 (6.85)**	-0.232 (13.19)**	-0.059 (3.95)**	-0.152 (3.40)**	-0.038 (1.38)
Male	1.149 (22.12)**	0.370 (7.02)**				
Extended household	-0.048 (0.89)		-0.216 (2.85)**		0.230 (2.10)*	
Non-labour income (household total)	-0.209 (1.70)+		-0.317 (1.73)+		0.042 (0.34)	
Intercept	-2.796 (22.23)**	-1.309 (7.43)**	-1.604 (10.99)**	-0.594 (3.30)**	-2.867 (12.76)**	-0.575 (2.17)*
λ (Mills ratio)		0.192 (2.47)**		-0.037 (0.37)		-0.069 (0.98)
χ^2 (slopes=zero)		$\chi^2_{(6)}$=254		$\chi^2_{(5)}$=241		$\chi^2_{(5)}$=35.1
No. of observations	3625	1510	2048	1220	1577	290
Selectivity test: r = 0	$\chi^2_{(1)}$=5.60		$\chi^2_{(1)}$=0.14		$\chi^2_{(1)}$=0.97	

Note: Dependent variable of selection equation is indicator for whether wages are observed. Dependent variable of wage equation is (log) hourly after-tax wage (including employer-provided subsidies). Heteroscedastically-robust t-statistics in parentheses; + significant at 10%; * significant at 5%; ** significant at 1%.

Table 7
Maximum likelihood selection and wage equations with full set of covariates

	All		Males		Females	
	Selection	Wage	Selection	Wage	Selection	Wage
Years of schooling	0.082 (6.19)**	0.067 (7.18)**	0.057 (3.24)**	0.059 (6.22)**	0.124 (4.49)**	0.066 (2.12)*
Years × Secondary	0.058 (5.84)**	0.028 (4.62)**	0.038 (3.04)**	0.021 (3.13)**	0.085 (4.76)**	0.020 (1.22)
Years × Tertiary	0.120 (10.51)**	0.041 (5.74)**	0.104 (7.19)**	0.035 (4.23)**	0.137 (6.82)**	0.015 (0.77)
Years of experience	0.105 (12.51)**	0.026 (4.55)**	0.119 (11.80)**	0.019 (2.42)*	0.104 (5.00)**	0.012 (1.13)
$(\text{Experience})^2 \div 100$	-0.179 (11.22)**	-0.038 (3.54)**	-0.206 (11.11)**	-0.027 (1.88)+	-0.213 (3.73)**	-0.007 (0.23)
Married	0.100 (1.46)	0.115 (3.64)**	0.315 (3.94)**	0.119 (3.14)**	-0.431 (3.35)**	0.060 (1.01)
Male	1.167 (21.39)**	0.327 (7.19)**				
Tenure		0.024 (10.04)**		0.025 (9.49)**		0.020 (3.99)**
Public sector		-0.044 (1.63)		-0.063 (2.05)*		0.031 (0.59)
Union member		0.198 (6.56)**		0.225 (6.56)**		0.042 (0.81)
On-the-job-training		0.137 (5.21)**		0.139 (4.53)**		0.116 (2.51)*
Extended household	-0.047 (0.87)		-0.205 (2.64)**		0.248 (2.18)*	
Non-labour income (household total)	-0.212 (1.64)		-0.319 (1.64)		0.049 (0.39)	
Intercept	-2.786 (22.13)**	-1.047 (7.09)**	-1.491 (9.93)**	-0.480 (2.90)**	-2.881 (12.77)**	-0.510 (1.86)+
λ (Mills ratio)		0.173 (2.77)**		0.020 (0.21)		-0.050 (0.67)
χ^2 (slopes=zero)		$\chi^2_{(11)}$=747		$\chi^2_{(10)}$=530		$\chi^2_{(10)}$=96.3
No. of observations	3625	1510	2048	1220	1577	290
Selectivity test: r = 0	$\chi^2_{(1)}$=7.05		$\chi^2_{(1)}$=0.04		$\chi^2_{(1)}$=0.45	

Note: Dependent variable of selection equation is indicator for whether wages are observed. Dependent variable of wage equation is (log) hourly after-tax wage (including employer-provided subsidies). Heteroscedastically-robust t-statistics in parentheses; + significant at 10%; * significant at 5%; ** significant at 1%.

Appendix Table 1
Testing for endogenous treatment effects in augmented wage equations

	Public Sector		On-the-Job-Training		Union Member	
	Selection	Wage	Selection	Wage	Selection	Wage
Years of schooling	0.041 (1.75)+	0.087 (6.65)**	0.055 (2.26)*	0.068 (6.43)**	0.063 (2.35)*	0.073 (7.47)**
Years × Secondary	0.047 (2.86)**	0.042 (4.01)**	0.053 (3.16)**	0.021 (2.73)**	0.033 (1.73)+	0.029 (4.91)**
Years × Higher	0.092 (5.37)**	0.065 (3.53)**	0.071 (4.02)**	0.027 (2.98)**	0.032 (1.60)	0.040 (6.53)**
Years of experience	0.026 (2.46)*	0.049 (8.28)**	0.034 (3.07)**	0.038 (8.10)**	0.091 (7.13)**	0.040 (7.90)**
$(\text{Experience})^2 \div 100$	0.003 (0.12)	-0.058 (6.71)**	-0.066 (2.61)**	-0.049 (5.09)**	-0.163 (5.50)**	-0.053 (5.38)**
Male	-0.090 (1.03)	0.224 (6.50)**	0.366 (4.09)**	0.192 (4.99)**	0.335 (3.17)**	0.239 (8.01)**
Public sector share (other householders)	0.035 (1.62)					
Public sector		-0.691 (1.30)				
Job training rate (other householders)			0.379 (3.16)**			
On-the-job training				0.528 (2.15)*		
Unionisation rate (other householders)					0.769 (6.11)**	
Union member						0.161 (1.10)
λ (Mills ratio)		0.460 (1.41)		-0.204 (1.37)		0.080 (0.95)
Intercept	-1.305 (7.12)**	-0.835 (10.81)**	-1.856 (9.92)**	-0.822 (10.80)**	-2.834 (13.33)**	-0.829 (10.01)**
$\chi^2_{(df\ 7)}$ (slopes=zero)	185**		744**		136**	
F-statistic $_{(df\ 8,1501)}$		125**		128**		139**
R^2		0.41		0.43		0.44

Note: Dependent variable of selection equation is indicator for whether the worker is in the public sector (or received on-the-job-training or is a union member). Dependent variable of wage equation is (log) hourly after-tax wage (including employer-provided subsidies). Heteroscedastically-robust t-statistics in parentheses; + significant at 10%; * significant at 5%; ** significant at 1%.

Acknowledgements

We are grateful to Chris Bennett, Boe Douna and Nick Suvulo of the National Statistical Office in Papua New Guinea for assistance with obtaining the data. Helpful comments have been received by Geua Boe-Gibson, Sholeh Maani and Begum Zaman. The financial support of the Institute of National Affairs in Papua New Guinea is gratefully acknowledged.

Notes

[1] Tannen (1991) estimates wage equations for Brazil that include the value of payments-in-kind in the dependent variable but he does not present any comparisons when this component of remuneration is excluded.

[2] For example, see Kugler and Psacharopoulos (1989).

[3] If expatriates are included in the estimating sample, the basic Mincerian estimate of the private returns to education (pooled across males and females) rises from 11.1% to 12.7%, while the pattern of returns across schooling levels is the same as when the expatriates are excluded.

[4] Potential years of labour market experience is derived as age minus school years completed minus seven.

[5] Using pre-tax wages, the estimated rate of return would be 11.6%.

[6] For a dummy variable in a semi-logarithmic regression, this is calculated as $100\times[\exp(0.126) - 1]$.

[7] The sample had only 18 university graduates, so higher education is defined as post-Provincial high school rather than as university completion. This reflects the reality of admission to university after Provincial high school and

the parallel treatment of National high school and university students. The higher education category also includes those with technical education achieved after provincial High school.

[8] The $\bar{R}^2$ is 0.36 for the intercept dummy model and 0.38 for the model with separate slopes. For the 'J' test, when the predicted wages from the intercept dummy model were added to the right-hand side of the separate slopes model, the estimated coefficient was statistically insignificant (t=1.49). But when the predicted wages from the separate slopes model were added to the intercept dummy model, the coefficient on the added variable was statistically significant (t=9.56). This suggests that the separate slopes model explains some of the variation in wages that the intercept dummy model does not.

[9] For example, the probability of receiving on-the-job training rises by 4.8% for each additional year of schooling.

[10] The significance of the coefficients for these household-level effects in the three selection equations in Appendix Table 1 appear to support this assumption (the significance of the public sector share of household employment is just outside conventional levels, at p=0.11).

[11] Specifically, the coefficient on training becomes 0.172 (t=2.31), there is a positive interaction between training and secondary years of schooling (0.030

with t=2.16) and no significant interaction between training and the other educational variables.

[12] The incidence of training is 48% in the public sector and 29% in the private sector, and this difference is statistically significant at $p<0.01$.

[13] On average, each year of completed education raises the probability of female wage earning by 10%, and of male wage earning by 6%, where this difference is statistically significant (t=5.59).