Average and Marginal Returns to Upper Secondary Schooling in Indonesia

Pedro Carneiro

University College London, IFS and Cemmap

Michael Lokshin

Development Research Group
World Bank

Cristobal Ridao-Cano

Human Development East Asia Pacific
World Bank

Nithin Umapathi[†]

University College London, IFS and Cemmap

May 2008

[PRELIMINARY AND INCOMPLETE]

Abstract:

We estimate heterogeneity in returns to schooling in Indonesia using a non-parametric method of local instrumental variables. In this paper we estimate average and marginal returns to upper secondary schooling in Indonesia, using data from the Indonesia Family Life Survey (IFLS). In order to account for the fact that schooling is unlikely to be exogenous in the population, we instrument the schooling choice with the distance to the nearest secondary school, measured in traveling time. We do not focus on a single instrumental variable parameter, but instead we estimate a range of average returns for different individuals. In particular, we estimate the marginal treatment effect. We estimate that the return to upper secondary schooling varies widely across individuals with unobserved characteristics: it can be as high as 200% or as low as 0%. Our results can be informative about the distributional consequences of extending the upper secondary schooling system.

[†]Nithin Umapathi gratefully acknowledges the financial support from the Economic and Social Research Council for the ESRC Centre for Microdata Methods and Practice (grant reference RES-589-28-0001).
These are the views of the authors and do not reflect those of the World Bank, its Executive Directors, or the countries they represent.
Correspondence by email to n.umapathi@ucl.ac.uk.

1. Introduction

Indonesia has an impressive record of educational expansion since 1970's. The gross enrollment rates for elementary schooling are universal and are around 75% for the secondary education (Beegle and Newhouse, 2006). There is ongoing effort to extend universal education attainment to the secondary level through higher government expenditure on this sector. We argue that it is important to analyze the distributional consequences of such an educational expansion arising from heterogeneity in treatment effects. The returns for the marginal person attracted by decreased costs of upper secondary schooling may not be the same as the returns for the average person in the population. Hence, average returns estimated using standard instrumental variables approaches may mask considerable inequality in returns. Our finding of large heterogeneity in returns for unobservable characteristics raises an important policy relevant question, whether educational expansion at the secondary level can be combined with efforts to bridge the gap in the inequality in returns. There is growing evidence that some skills are not substitutable over life-time. Therefore, efforts such as improvement in quality of education at primary level or even earlier pre-school interventions could remove the disadvantages that hinder some of the students in secondary school.

Rates of return to schooling have been estimated for most countries in the world for which micro data is available (see, e.g., Psacharopoulos and Patrinos, 2004). However, many of these estimates, especially for developing countries, are based on correlations between education and wages, and fail to account for the non-random allocation of schooling in the population (e.g., Card, 2001). Recently, some studies have emerged for developing countries where the endogeneity of schooling is accounted for through the use of instrumental variables. For example Maluccio (1998) uses distance to school as an instrument for estimating return to schooling in the Philipines, Duflo (2001) uses a school construction program as the instrument in Indonesia, and Patrinos and Sakellariou (2005) instrument account for endogeneity of schooling by the change in the compulsory schooling law in Venezuela.

Instrumental variables estimates (IVE) could be good guides for education policy if the return to schooling is assumed to be common for all individuals in the economy. In that case IVE identifies the average treatment effect and a single parameter is sufficient to characterize the mean impact of schooling on wages. However, when the return to schooling varies across population and individuals decide to participate in schooling based on their idiosyncratic gains, the standard features of IVE break down and mean impact parameters are not generally recoverable. In this setting IVE estimates are identified but not easily interpretable (see Imbens and Angrist, 1994, Card, 2001, Carneiro, Heckman and Vytlacil, 2006).

For instance, consider a hypothetical comparison of two educational programs that assumes absence of general equilibrium effects: compulsory schooling or tuition subsidies. The two programs may attract different groups of individuals:a tuition subsidy attracts people with positive net benefit, while compulsory schooling policy affects the entire population. Hence, compulsory education may even have a lower average individual impact on wages, by virtue of who is attracted, than a policy of tuition. Therefore, it is useful to move beyond instrumental variables estimates, and analyze how the returns to schooling vary across individuals, and who are the individuals most likely to be induced to go to school by a specific policy or a reform (e.g., Heckman and Vytlacil, 2001).

In this paper we estimate average and marginal returns to upper secondary schooling in Indonesia, using data from the Indonesia Family Life Survey (IFLS). Our estimates of the returns to schooling are slightly higher than the returns to schooling for Indonesia found in Duflo (2000) with the qualification that the dataset, the instrumental variable, and the time period are not the same. However, unlike Duflo (2000) we allow for a more general empirical model that relaxes the common impact assumption, and explicitly accounts for the unobserved heterogeneity by using the method of local instrumental variables (Heckman and Vytlacil, 2005). Instead of focusing on a single instrumental variable parameter, we estimate a range of returns for different individuals. In particular, we estimate the Marginal Treatment Effect, a parameter introduced by Bjorklund and Moffitt (1987). Our results indicate that the return to upper secondary schooling varies significantly across individuals: it can be as high as 200% or as low as 0% depending on unobservable characteristics. Our methods allow us to estimate interpretable population estimates of returns such as average treatment effect (ATE = 86.3%), treatment on treated

(ATT=102.3%) and treatment on untreated (ATU = 77.85%.). We find that despite the heterogeneity, the return for the average individual (ATE) with upper secondary schooling is very similar to the IV estimate, however this is not guaranteed to happen.

The next secion discusses the data, Section 3 presents the econometric framework which is used in section 4 to understand the importance of heterogeneity and selection in the returns to schooling. Section 5 concludes.

2. Data

We use the third wave of the Indonesia Family Life Survey (IFLS3) fielded in 2000[1]. Our sample consists of 2445 adults aged 25-60 who are employed in the labor market, reported non-missing wages and provide information on schooling and their place of residence at age of 12. **Table 1** presents descriptive statistics on the main variables used in our analysis. It shows that individuals with upper secondary or higher levels of education have, on average, 72% higher wages than those with lower education. The respondents with an upper secondary education are younger, more likely to have better-educated parents, live in towns or cities at age 12, and live closer to upper secondary schools, compared to those with less than an upper secondary education.

3. A Semi-Parametric Selection Model

Our empirical strategy is based on the framework developed by Heckman and Vytlacil (2000, 2005) and Carneiro, Heckman and Vytlacil (2005). In order to understand the patterns of heterogeneity and selection in the returns to schooling in Indonesia we estimate a selection model with two schooling levels:

$$\begin{aligned} Y_1 &= \alpha_1 + X\beta_1 + U_1 \\ Y_0 &= \alpha_0 + X\beta_0 + U_0 \end{aligned} \quad (1)$$

$$S = 1 \text{ if } Z\gamma - U_s > 0 \quad (2)$$

Where, Y_1 are log wages of individuals with upper secondary education and above, Y_0 are log wages of individuals without upper secondary education, X is a vector of observable

[1] For a description of the survey see Strauss, J., K. Beegle, B. Sikoki, A. Dwiyanto, Y. Herawati and F. Witoelar. "The Third Wave of the Indonesia Family Life Survey (IFLS): Overview and Field Report", March 2004. WR-144/1-NIA/NICHD.

characteristics that might affect wages, and U_1 and U_0 are the error terms. Z is a vector of characteristics affecting the schooling decision. It will also be convenient to rewrite the selection equation as:

$$S = 1 \text{ if } P(Z) > V \quad (3)$$

where $P(Z) = F_{U_S}(Z\gamma)$ and $V = F_{U_S}(U_S)$ and F_{U_S} is cumulative distribution function of U_S, V is distributed uniformly, which is an innocuous assumption given that U_S can have any density. Finally, observed wages are:

$$Y = SY_1 + (1-S)Y_0 \quad (4)$$

Notice that the return to schooling is

$$Y_1 - Y_0 = X(\beta_1 - \beta_0) + U_1 - U_0 \quad (5)$$

and it varies across individuals with different X's and different U_1, U_0. For identification Z must be independent of U_1 and U_0. Furthermore, Z needs to be continuous and with large support, and it has to be a strong determinant of the schooling decision. In practice we will use a stronger assumption: X,Z is independent of U_1, U_0, U_S. As a result, instead of requiring continuity and large support in Z, we only require these for $P(Z)$.[2] The main parameter of interest in our analysis is the marginal treatment effect (or MTE, a parameter introduced in Bjorklund and Moffitt, 1987, and developed in Heckman and Vytlacil, 1999, 2000, 2005) that in our notations can be expressed as:

$$MTE(x,v) = E\left(Y_1 - Y_0 \mid X = x, V = v\right) = \left(\alpha_1 - \alpha_0\right) + x\left(\beta_1 - \beta_0\right) + \mathrm{E}\left(U_1 - U_0 \mid V = v\right) \quad (6)$$

The MTE measures how the average return to schooling varies for individuals with different levels of observed (X) and unobserved (V) heterogeneity.

Assuming that the unobservables in the wage and selection equations are jointly normally distributed the MTE can be estimated a standard switching regression model (see Heckman, Tobias and Vytlacil, 2001). That model relies on strong assumptions about the distribution of the error terms in equation (1). Instead, we use the method of local instrumental variables (Heckman and Vytlacil, 2000) that imposes no distributional assumptions on the unobservables of the model. In particular, Heckman and Vytlacil (2000) show that:

[2] See Heckman and Vytlacil (2005), and Carneiro, Heckman and Vytlacil (2006), for a formal statement of the hypotheses of the model, and for details of its implementation.

$$MTE(x,v) = \frac{\partial E(Y \mid X, P)}{\partial P}\Big|_{X=x,P=v} \quad (7)$$

Where,

$$\begin{aligned} E(Y \mid X, P) &= E\left[\alpha_0 + X\beta_0 + S(\alpha_1 - \alpha_0) + SX(\beta_1 - \beta_0) + U_0 + S(U_1 - U_0) \mid X, P\right] \\ &= \alpha_0 + X\beta_0 + PX(\beta_1 - \beta_0) + E(U_1 - U_0 \mid S = 1, X, P)P \\ &= \alpha_0 + X\beta_0 + PX(\beta_1 - \beta_0) + K(P) \end{aligned} \quad (8)$$

(*K(P)* is a function of *P*). Therefore, taking the derivative of (8) with respect to P, (7) becomes:

$$MTE(x,v) = \frac{\partial E(Y \mid X, P)}{\partial P}\Big|_{X=x,P=v} = X(\beta_1 - \beta_0) + K'(P) \quad (9)$$

Our estimation procedure has several steps. First, we estimate the propensity score *P* assuming a probit model for the selection equation. Then we examine the support of *P* for individuals with $S = 0$ and $S = 1$. The MTE can only be estimated over the points of *P* in the intersection of these two sets (common support). We regress log wages on a set of explanatory variables (*X*), interactions of explanatory variables and the propensity score, and on the third degree polynomial of propensity score. We exclude from our sample observations for which there is little overlapping support. Next, from (8) we compute the residual that only retains the unobserved heterogeneity:

$$R = Y - \left[\alpha_0 + X\beta_0 + PX(\beta_1 - \beta_0)\right] \quad (10)$$

and run a nonparametric (local quadratic) regression of *R* on *P* to obtain *K(P)* and to derive *K'(P)*. A simple test of heterogeneity in the impact by unobserved characteristics is a test of whether *K'(P)* is flat, or if *E(Y |X, P)* is nonlinear in *P*. If the derivative is flat heterogeneity is not important.

Conventional treatment parameters such as the Average Treatment Effect (ATE), Average Treatment on the Treated (ATT) and Average Treatment on the Untreated (ATU) can be estimated as weighted averages of the MTE for the corresponding sub-populations, as proposed by Heckman and Vytlacil (2000, 2005). In particular:

$$
\begin{aligned}
&ATE(x) = E\left(Y_1 - Y_0 \mid x\right) = \int_0^1 MTE\left(x,v\right)dv \\
&ATT(x) = E\left(Y_1 - Y_0 \mid x, D=1\right) = \int_0^1 MTE\left(x,v\right)h_{TT}\left(x,v\right)dv \\
&ATU(x) = E\left(Y_1 - Y_0 \mid x, D=0\right) = \int_0^1 MTE\left(x,v\right)h_{TUT}\left(x,v\right)dv \\
&where \\
&h_{TT}\left(x,v\right) = \left[\int_v^1 f\left(p \mid X = x\right)dp\right]\frac{1}{E\left(P \mid X = x\right)} \\
&h_{TUT}\left(x,v\right) = \left[\int_0^v f\left(p \mid X = x\right)dp\right]\frac{1}{E\left(1-P \mid X = x\right)}
\end{aligned}
\tag{11}
$$

4. Empirical Results

4.1 Validity of IV

To account for the fact that schooling is unlikely to be exogenous in the population we instrumented the schooling choice with the distance to the nearest secondary school, measured in traveling time[3]. We argue that the distance to school at should affect the probability of schooling and should have no effect on the adult wages. Distance to the nearest school has been used as an instrument for schooling by Card (1995), Kane and Rouse (1995), Kling (2001), Currie and Moretti (2003), Cameron and Taber (2004) and Carneiro, Heckman and Vytlacil (2006). However, its use is not uncontroversial, because families and even schools may not randomly locate across Indonesia. For example, Carneiro and Heckman (2002) and Cameron and Taber (2004) show that individuals living closer to universities have higher levels of cognitive ability and come from better family backgrounds. Therefore, it is important to control for characteristics capturing the location decision of the household. In our paper we include both household characteristics, such as parental education, and village characteristics, such as an indicator for whether the individual was living in a city or village at age 12, and dummies for the province, district and sub-district of residence. Therefore our identification comes from variation in school availability or distance to school between

[3] Similar instruments have been used in the literature by Card (1995), Kling (2001), Kane and Rouse (1995), Maluccio (1998), Currie and Moretti (2003), Cameron and Taber (2004), and Carneiro, Heckman and Vytlacil (2006).

villages located within the same sub-district; our assumption is that, conditional on these detailed household and location controls, the location of schools is exogenous.

First, in **Table 2** we examine whether distance to upper secondary school is correlated with the number of repeated grades as a measure of early school success, and whether the individual worked while in primary school, a measure of early educational environment. If our instrument is valid it should not be correlated with such early measures. Our results show no apparent correlation between distance to school and these measures, increasing our confidence in the validity of the instrument. This is true even after excluding several village and family controls from the regression.

4.2 Standard estimates of returns

In Table 3 we also show that distance to the nearest secondary school is a strong predictor of enrolment in secondary school. We include distance to health post as a proxy for location characteristics and reassuringly unlike distance to school, distance to health post does not predict school enrollment. We include all the Xs and the distance to upper secondary school, measured in time (minutes) at the time of the survey. We use log hourly wage in 2000 as our dependent variable. In order to compare our estimates with the rest of the literature, we present the results of OLS and IV linear regressions of log wages on years of schooling. We also report the coefficients on our instruments in the first stage regression, as well as an F-test for their joint significance. The first column in Table 3 shows an OLS estimate of the return to schooling of 10%. The second column shows first stage of the 2SLS, that the distance to school is negatively related to school attainment: the coefficient is negative (the F-statistic is 7.98). Finally, column 4 shows the IV estimate of the return to schooling, which is about 14.7%. Our estimates of return to schooling are higher than those by Duflo(2001), although our wage data is slightly more recent. As in Duflo (2001) and the rest of the literature, IV estimates of the return to education are larger than OLS estimates of this parameter. As argued in Griliches (1977) and Card (2001), at first sight we would not expect the IV estimate to be above the OLS estimate. However, if the return to schooling varies across individuals this may well happen. Card (2001) suggests that such a finding indicates that the marginal individual induced to enroll in school by the change in the instrument has a higher return than the average individual. Carneiro, Heckman and Vytlacil (2006) show that IV estimates can be above OLS estimates even if the marginal individual has a lower return than the average, and show this to be the case for college attendance in the US. Another reason why IV can exceed OLS is measurement error in schooling. Although schooling is relatively well measured in the US (Card, 1999) and other developed countries, that is not necessarily the case in Indonesia. In the remaining of this paper we ignore measurement error, and focus instead on heterogeneity.

Since the remaining of the paper deals with the case of binary schooling, it is convenient to replicate **table 3** using as independent variable a dummy for upper secondary school attendance. This is what we do in **table 4**, which shows basically the same patterns as in **table 3**. Most notably, the OLS estimate of the return to upper secondary schooling is 69.4%, while the IV estimate[4] is 161.4%. Given that our instrument seems to be acceptable both in terms of strength and arguably in validity, and that our estimates of the returns to schooling are quite reasonable. For example Duflo(2001) finds a 10.6% economic return on education in Indonesia. We proceed to investigate the extent to which returns are heterogeneous, and we estimate several parameters of potential interest.

4.3 Average and Marginal Treatment effect estimates

In order to calculate the average and marginal treatment effect we define two schooling categories: i) completed lower secondary or below, and ii) attendance of upper secondary or above. This division ignores the fact that in reality there are many more levels of schooling, but it simplifies the model and is standard in many studies of the returns to schooling (e.g., Willis and Rosen, 1979, Taber, 2001, Carneiro, Heckman and Vytlacil, 2005). Furthermore, we lack exogenous variation for every schooling transition to estimate a model with more than one level of schooling. Therefore, the estimated return parameters in the case of two schooling categories should be interpreted as the return to upper secondary schooling or above. We construct *P* as a predicted probability

[4] Any person attending upper secondary school would on average have attended 10-12 years of school. Therefore using back of the envelope calculations we can back out the return for additional year of education, for example if a student attended 11th grade the our IV estimate of 1.614 implies 14.6% return which squares well with our IV estimate for years of schooling in Table 3

of a probit regression of schooling on the X and Z variables of section 3. **Table 5** reports the coefficients of our model estimated by the method of local instrumental varialbes. In addition to the variables reported in the table, this model includes dummy variables for province, district and sub-district of residence. All coefficients have the expected sign. The coefficient on the distance variable is strong and negative. The coefficients on the Xs have the expected signs, parental education is positively related to school attendance and individuals who lived in the village at age 12 are less likely to go to upper secondary school. **Table 6** shows the bootstrapped results for a simple test of essential heterogeneity, described in section 2, where the coefficients on the non-linear terms of P are statistically significant. Therefore the MTE is not flat, which means that heterogeneity in returns is important empirically.

Figure 1 presents the distributions of the predicted P for individuals in $S = 0$ and in $S = 1$. The supports for these distributions overlap almost completely, although the support at the tails is thin, especially for those individuals for whom $S = 0$. Because of lack of support at the extremes of the distribution of P, in the rest of the analysis we only use individuals for whom P is between 0.05 and 0.90 (2445 observations). This means that all our results are restricted to this population (see Heckman and Vytlacil, 2005, Carneiro, Heckman and Vytlacil, 2005). In particular, we cannot estimate ATE, ATT AND ATU for individuals with extreme values of V. We interpret our estimates of such parameters as conditional on the support of the data. Finally we compute $R = Y - [\alpha_0 + X\beta_0]$ as described in section 2, and to estimate *K(P)* we run a local quadratic regression of R on P, using a Gaussian kernel and a bandwidth of 0.2. The implied *MTE*(v) is computed by calculating the slope on the linear term of the local quadratic regression as in (9).

Figure 2 shows that the MTE is monotonically decreasing, taking positive and high values for low V's (individuals who are likely to enroll in upper secondary school or facing high costs), and low values for individuals who are unlikely to enroll in upper secondary school (and therefore have high V's). The bootstrapped 5-95% confidence interval is also plotted. The figure demonstrates a large heterogeneity in the returns to schooling that range from 200% for individuals with V around 0.2 to 0% for those with V close to 0.7. The fact that returns are the lowest for individuals who are least likely to go to school is consistent with a simple economic model where agents sort themselves based on their comparative advantage.

The main treatment parameters are presented in **table XXX** (5%-95% bootstrapped confidence intervals are presented in parenthesis). ATE is the average treatment effect; TT is treatment on the treated and TUT is treatment on the untreated. In reality these parameters are not identified in our data because we do not have full support. As mentioned above, the correct way to interpret these numbers is conditional on the observed support of the data ($0.05 < P < 0.90$). The return to upper secondary school for a random person is 86.35%. The returns for the individuals who was enrolled in the upper secondary schooling is considerably higher, at 102.26%.(SE) If individuals who did not go to upper secondary school would have gone there, they could expect the returns of 77.85%.(SE).

5. Conclusion

In this paper we estimate the returns to upper secondary schooling in Indonesia using data from the IFLS. We estimate that ATE = 86.3%, ATT = 102.3% and ATU = 77.85%. We find that there is a considerable degree of heterogeneity in returns to schooling, with returns being as high as 150% for individuals very likely to attend upper secondary schooling and as low as 0% or even negative for those who are not likely to attend.

Indonesia has an impressive record of educational expansion since 1970's. The enrollment rates are nearly universal for elementary schooling and are around 75% for the secondary education. There is ongoing effort to extend universal education attainment to the secondary level. And although enrollment in secondary education continues to rise we find striking inequality in returns to education. The individuals who are most likely to be attracted by educational expansion at the upper secondary level are also the ones with the lowest returns. The policy relevant question arises, whether efforts at educational expansion can also bridge the gap in the inequality in returns. There is a growing body of literature that argues that human capital outcomes later in life are largely determined early in life (Cunha and Heckman, 2007; Heckman and Masterov, 2007). It is therefore important for the design of schooling policy to determine whether the inequality in secondary schooling outcomes can be remedied at earlier stages for example during elementary education or even earlier. In an impressive drive to increase the quantity of education, there should be a renewed emphasis on quality of education that ensures a more relevant learning environment for the disadvantaged children that reduces the inequity in lifetime outcomes. This can be achieved by raising the return for the marginal student, and therefore the increased equity need not cost in terms of efficiency.

References

Bjorklund, A. and R. Moffitt (1987) , ``The Estimation of Wage Gains and Welfare Gains in Self-Selection Models," *Review of Economics and Statistics* , 69:42-49.

Cameron, S. and C. Taber (2004), ``Estimation of Educational Borrowing Constraints Using Returns to Schooling", *Journal of Political Economy* , part 1, 112(1): 132-82.

Card, D. (1995), ``Using Geographic Variation in College Proximity to Estimate the Return to Schooling ", *Aspects of Labour Economics: Essays in Honour of John Vanderkamp* , edited by Louis Christofides, E. Kenneth Grant and Robert Swindinsky. University of Toronto Press.

Card, D. (1999), ``The Causal Effect of Education on Earnings," Orley Ashenfelter and David Card, (editors), Vol. 3A, *Handbook of Labor Economics,* Amsterdam: North-Holland.

Card, D. (2001), ``Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica* , 69(5): 1127-60.

Card, D. and T. Lemieux (2001), ``Can Falling Supply Explain the Rising Return to College For Younger Men? A Cohort Based Analysis," *Quarterly Journal of Economics* 116: 705-46.

Carneiro, P. and J. Heckman (2002), ``The Evidence on Credit Constraints in Post-secondary Schooling," *Economic Journal* 112(482): 705-34.

Carneiro, P., J. Heckman and E. Vytlacil (2006), ``Understanding what Instrumental Variables estimate: estimating average and marginal returns to schooling", working paper, The University of Chicago.

Currie, J. and E. Moretti (2003), Mother's Education and the Intergenerational Transimission of Human Capital: Evidence from College Openings, *Quartely Journal of Economics* , 118:4.

Duflo, E. (2001), ``Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment", *American Economic Review*, 91(4), 795-813.

Griliches, Z. (1977), ``Estimating the Return to Schooling: Some Persistent Econometric Problems", *Econometrica.*

Heckman, J., S. Urzua and E. Vytlacil (2005), ``Understanding What Instrumental Variables Really Estimate in a Model with Essential Heterogeneity", working paper, The University of Chicago.

Heckman, J. and E. Vytlacil (1999), Local Instrumental Variable and Latent Variable Models for Identifying and Bounding Treatment Effects, *Proceedings of the National Academy of Sciences*, 96, 4730-4734.

Heckman, J. and E. Vytlacil (2000), ``Local Instrumental Variables," in C. Hsiao, K. Morimune, and J. Powells, (eds.), *Nonlinear Statistical Modeling: Proceedings of the Thirteenth International Symposium in Economic Theory and Econometrics: Essays in Honorof Takeshi Amemiya*, (Cambridge: Cambridge University Press, 2000), 1-46.

Heckman, J. and E. Vytlacil (2001), ``Policy Relevant Treatment Effects", *American Economic Review Papers and Proceedings*.

Heckman, J. and E. Vytlacil (2005), ``Structural Equations, Treatment, Effects and Econometric Policy Evaluation," *Econometrica*, 73(3):669-738.

Imbens, G. and J. Angrist (1994), ``Identification and Estimation of Local Average Treatment Effects,' *Econometrica*, 62(2):467-475.

Kane, T. and C. Rouse (1995), ``Labor-Market Returns to Two- and Four-Year College", *American Economic Review*, 85(3):600-614.

Kling, J. (2001), ``Interpreting Instrumental Variables Estimates of the Returns to Schooling", *Journal of Business and EconomicStatistics*, 19(3), 358-364.

Maluccio, J. (1998), ``Endogeneity of Schooling in the Wage Function: Evidence from Rural Philippines", Food Consumption and Nutrition Division Discussion Paper No. 54, International Food Policy Research Institute, Washington, DC.

Psacharopoulos, G. and H. Patrinos (2002), ``Returns to Investment in Education: a Further Update",*Education Economics*, 12(2), 111-134.

Willis, R. and S. Rosen (1979), ``Education and Self-Selection,"*Journal of Political Economy*, 87(5):Pt2:S7-36.

Table 1: Sample statistics

	More than upper secondary	*Less than upper secondary School*	*Difference*	*SE of difference*
	N=1027	*N= 1418*		
Log hourly wages	8.207	7.491	0.716	0.035
Age	37.024	38.877	-1.852	0.378
Religion Protestant	0.050	0.020	0.029	0.007
Catholic	0.026	0.008	0.018	0.005
Other	0.063	0.044	0.019	0.009
Father's education elementary	0.515	0.513	0.002	0.020
secondary and higher	0.337	0.060	0.277	0.015
missing	0.010	0.025	-0.016	0.006
Mother's education elementary	0.493	0.410	0.082	0.020
secondary and higher	0.206	0.022	0.185	0.012
missing	0.080	0.112	-0.032	0.012
Type of place at age 12, village	0.532	0.733	-0.202	0.019
Town	0.301	0.206	0.095	0.018
City	0.167	0.061	0.107	0.012
Moved since age 12	0.352	0.241	0.110	0.018
Rural	0.242	0.476	-0.234	0.019
Distance in minutes	9.031	9.805	-0.774	0.252

Note: With the exception Father's elementary education all differences are significant at 5

Table 2: Regression of junior secondary education experience on distance to school

	Failed grade		**No of repeats**		**Worked**	
	coef	**se**	**coef**	**se**	**coef**	**se**
Distance to school in minutes	-0.002	*0.006*	-0.005	*0.006*	0.005	*0.008*
Age	0.002	*0.006*	0.003	*0.006*	-0.002	*0.009*
Age Squared	0.000	*0.000*	-0.000	*0.000*	0.000	*0.000*
Protestant	0.001	*0.021*	0.011	*0.027*	0.026	*0.044*
Catholic	0.017	*0.045*	0.019	*0.045*	0.007	*0.060*
Other religions	0.048	*0.077*	0.049	*0.078*	-0.052*	*0.031*
Father's education elementary	-0.004	*0.020*	-0.007	*0.021*	-0.037	*0.026*
Junior secondary or higher	-0.008	*0.022*	-0.005	*0.023*	-0.045	*0.029*
Fathers education missing	0.009	*0.045*	0.009	*0.046*	0.057	*0.103*
Mothers education elementary	-0.015	*0.022*	-0.011	*0.021*	0.016	*0.027*
Junior secondary or higher	-0.025	*0.026*	-0.026	*0.026*	-0.057*	*0.033*
Mothers education missing	-0.037*	*0.019*	-0.034*	*0.019*	-0.023	*0.039*
Lived in a village at age 12	-0.004	*0.013*	-0.001	*0.013*	-0.017	*0.020*
Ever moved since age 12	-0.001	*0.010*	-0.005	*0.010*	0.048**	*0.019*
Rural household	0.000	*0.016*	0.004	*0.016*	0.055**	*0.024*
Province, district and sub-district dummies suppressed						
Distance to health post in min	-0.009*	*0.004*	-0.008*	0.004	-0.012*	0.007
Number of observations		1,433		1,431		1,431
R2		0.064		0.063		0.088

note: .01 - ***; .05 - **; .1 - *;

Table 3: OLS and IV estimates of the return to a year of schooling

	OLS		2SLS			
			First stage		IV estimate	
	coef	se	coef	se	coef	se
Years of education	0.100***	0.005			0.147***	0.054
Protestant	0.117	0.072	1.424***	0.406	0.050	0.107
Catholic	0.033	0.131	1.599*	0.827	-0.041	0.147
Other religions	0.028	0.117	1.015	0.653	-0.025	0.133
Age	0.054***	0.018	-0.004	0.077	0.055***	0.017
Age Squared	-0.000*	0.000	-0.000	0.001	-0.000*	0.000
Father elementary	0.060	0.050	1.763***	0.233	-0.024	0.105
Father junior high	0.081	0.066	3.311***	0.315	-0.077	0.189
Missing	-0.032	0.122	-0.202	0.478	-0.025	0.123
Mother elementary	-0.061	0.047	0.725***	0.210	-0.096	0.059
Mother junior high	-0.094	0.079	2.523***	0.323	-0.215	0.156
Missing	-0.170***	0.056	0.258	0.305	-0.185***	0.061
Lived in a village at age 12	0.022	0.037	-0.405**	0.172	0.041	0.044
Ever moved since age 12	0.054	0.035	0.737***	0.179	0.018	0.058
Rural household	0.145***	0.051	-0.300	0.304	0.166***	0.057
Province, district and sub-district dummies suppressed						
Distance to health post in min	0.001	0.014	0.010	0.094	0.003	0.013
Distance to school in minutes			-0.064***	0.023		
_cons	5.333***	0.349	7.128***	1.897	5.029***	0.497
Number of observations		2,445		2,445		2,445
F-test instruments				7.98		
R2		0.295		0.360		0.260
note: .01 - ***; .05 - **; .1 - *;						

Table 4: OLS and IV estimates of the return to upper secondary schooling

	OLS		First Stage		IV estimates	
	coef	**se**	**coef**	**se**	**coef**	**se**
Years of education	0.694***	0.044			1.614**	0.653
Protestant	0.200***	0.075	0.084*	0.047	0.124	0.106
Catholic	0.100	0.165	0.132*	0.072	-0.019	0.186
Other religions	0.109	0.128	0.035	0.072	0.068	0.139
Age	0.049**	0.019	0.008	0.009	0.042**	0.020
Age Squared	-0.000	0.000	-0.000	0.000	-0.000	0.000
Father elementary	0.130**	0.051	0.153***	0.025	-0.012	0.113
Father junior high	0.163**	0.068	0.360***	0.036	-0.171	0.248
Missing	0.010	0.126	-0.086	0.063	0.084	0.152
Mother elementary	-0.037	0.049	0.069***	0.025	-0.100	0.066
Mother junior high	-0.016	0.080	0.254***	0.036	-0.254	0.181
Missing	-0.169***	0.059	0.041	0.034	-0.213***	0.074
Lived in a village at age 12	0.012	0.038	-0.045**	0.021	0.054	0.053
Ever moved since age 12	0.089**	0.037	0.057***	0.021	0.035	0.059
Rural household	0.128**	0.054	-0.027	0.031	0.165***	0.062
Province, district and sub-district dummies suppressed						
Distance to health post in min	0.001	0.018	0.000	0.009	0.004	0.017
Distance to school in minutes			-0.006**	0.002		
_cons	5.939***	0.374	0.116	0.203	5.891***	0.418
Number of observations		2,445		2,445		2,445
F-test instruments				5.63		
R2		0.234		0.286		0.060
note: .01 - ***; .05 - **; .1 - *;						

Table 5: Probit regression of schooling choice

	Schooling equation	
	coef	**se**
Distance to school in minutes	-0.019***	0.007
Age	0.033	0.028
Age Squared	-0.000	0.000
Protestant	0.292*	0.165
Catholic	0.483*	0.255
Other religions	0.164	0.244
Father elementary	0.526***	0.080
Father junior high or higher	1.150***	0.112
Missing	-0.272	0.232
Mother elementary	0.189**	0.078
Mother junior high or higher	0.923***	0.141
Missing	0.135	0.108
Lived in a village at age 12	-0.132*	0.069
Ever moved since age 12	0.149**	0.065
Distance to health post in min	0.001	0.028
Rural household	-0.086	0.092
Province, district and sub-district dummies are suppressed		
_cons	-1.355**	0.635
Number of observations		2,561

Table 6: Test for presence of heterogeneity

	Observed Coef.	Bootstrap Std. Err.
Propensity score	5.101**	2.230
Squared	-7.431***	2.806
Cubed	4.449**	1.921
Age	0.039	0.032
Age Squared	-0.000	0.000
Protestant	0.358	0.262
Catholic	-0.129	0.454
Other religions	0.141	0.343
Fathers education elementary	-0.146	0.160
Father junior secondary	-0.172	0.421
Father's education missing	0.007	0.231
Mother education elementary	-0.171	0.113
Mother junior secondary	-0.922**	0.466
Mothers education missing	-0.206*	0.120
Location at 12 village	0.118	0.093
Ever moved since 12	0.058	0.100
kmsd	-0.020	0.022
Rural household	0.163	0.120
Province, district and subdistrict dummies are suppressed		
Interactions of Xs with Propensity score suppressed		
_cons	5.727***	0.698
Number of observations		2,446
R2		0.156

Note: Standard errors are bootstrapped using 100 replications

Figure 1: Propensity score support for each schooling group

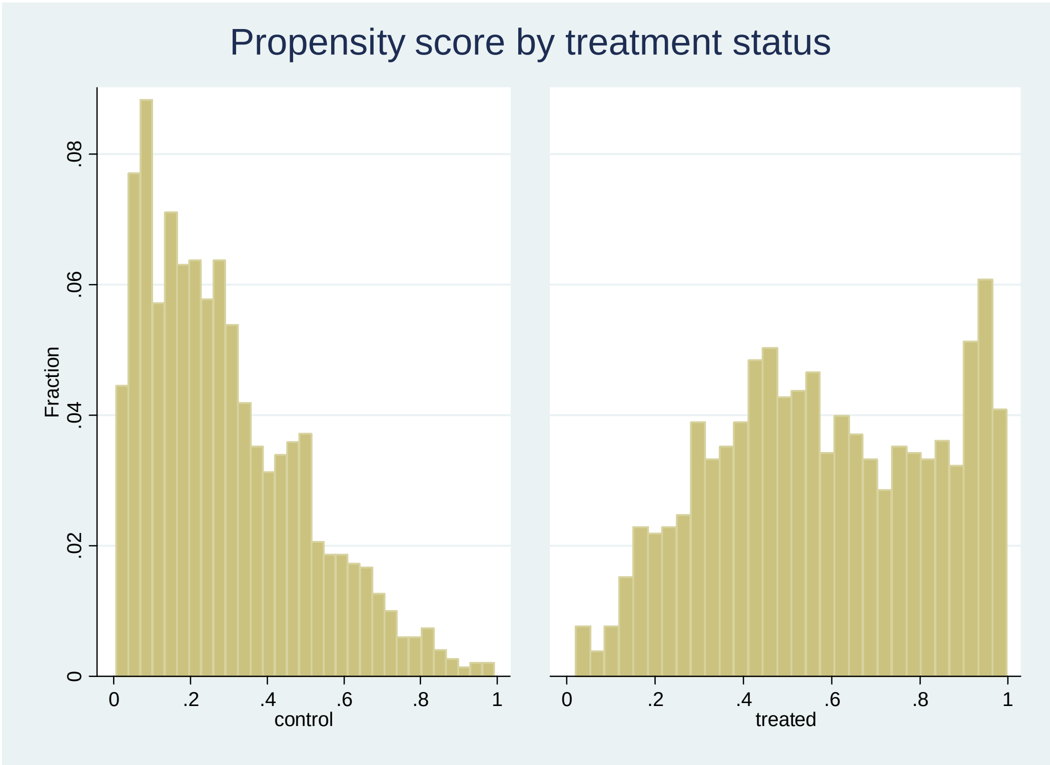

Figure 2: Marginal treatment effect