6-2005

Heterogeneity in returns to schooling: Econometric evidence from Ethiopia

Sourafel Girma
University of Leicester, U.K., sg116@le.ac.uk

Abbi M. Kedir
University of Leicester, U.K.

Heterogeneity in returns to schooling: Econometric evidence from Ethiopia

Sourafel Girma[*] *and Abbi Kedir*
Department of Economics
University of Leicester
University Road
Leicester LE1 7RH

Abstract

The paper investigates whether returns to schooling in Ethiopia vary across the wages distribution of individuals. To do so it adopts an instrumental variables quantile regression framework that allows for both endogeneity of schooling resulting from unmeasured ability, and possible heterogeneity in the impact of schooling. The empirical estimates indicate that education contributes more to the earnings of the individuals at a lower end of the income distribution. Under the assumption that the wage and ability distributions are related, this result is consistent with the notion that education and ability are substitutes.

Key Words: returns to schooling; quantile regression
JEL Classification: I2; J3

---

[*] *Corresponding author: e-mail: sg116@le.ac.uk.* *The authors would like to thank the Economic Department of Addis Ababa University for giving access to the data used in this study. Kedir gratefully acknowledges financial assistance from the Experian Centre for Economic Modelling at the University of Nottingham.*

The empirical literature on the returns to education focuses mainly on developed countries[1], and much of the literature in developing countries compares the returns to vocational and academic education (Psacharopoulos, 1994; Bennell, 1996), or seeks to identify the impact of completing a given schooling cycle on earnings (Appleton, 2001). The aim of this study is to contribute to the literature by conducting a systematic investigation on the returns to education in Ethiopia. In particular it asks to what extent returns to education vary across the wages distribution. It also examines the empirical implications of neglecting the possible endogeneity of schooling in the wages determination equations.

To simultaneously address the two issues of heterogeneity in returns and endogeneity of schooling, we adopt an instrumental variable quantile regression framework. Our empirical estimates indicate that education contributes more to the earnings of the individuals at the lower end of the income distribution. The relatively low (but still economically significant) returns to education at the higher end of the conditional earnings distribution is indicative of the importance of inherent ability or personal connections in securing high paying jobs..

The rest of the paper is organised as follows. Section II presents a selected review the literature. Section III outlines the econometric methodology, and this is followed by the data description in Section IV. Section V discusses the empirical results, and Section VI concludes.

[1] For an excellent summary of the literature see Card (1999).

II. Literature review

It is widely argued that any investment in human capital has a pure productivity element. But there are criticisms levelled against this argument. The main criticism centres on the idea that the effect of education is simply to enhance the productivity of the individual undertaking the specific education. This is the pure human capital hypothesis. The alternative hypothesis suggests that education is not productivity enhancing but simply acts as a screen to identify highly productive individuals. The signalling/screening hypothesis states that individuals have an inherent ability and education raises their earnings. It is the attainment of specific levels of education that is used to command higher earnings, and as such highly intelligent individuals will choose to make human capital investments. However, the primary role of education is to signal to employers as to the inherent ability of individuals and not to enhance the productivity of an individual. The evidence for and against the screening hypothesis has been sought by providing the presence/absence of a diploma/sheepskin effect which is tested empirically by introducing dummy variables for various levels of completed schooling (Bauer et al 2002; Antelius, 2000).

Rosenzweig (1995) developed a framework for investigating the circumstances under which schooling improves productivity in the market and in the household, based on the notion that schooling enhances information acquisition. He focuses on two channels through which schooling may enhance productivity: i) by improving access to information sources such as newspapers or instruction manuals, which are found to be a major route in Thomas et al (1991) and ii) by improving the ability to decipher new information, whether from external sources or from own experience, as suggested by Schultz (1975).

Two important implications stand out of Rosenzweig's framework. The first implication of the model is that the returns to schooling should be higher in regimes or economies in which there is greater scope for misusing an input, or when tasks are sufficiently complex that substantial learning is required to execute them efficiently. Conversely, where tasks are simple and easy to master, schooling should have little influence on productivity. His model also implies that schooling returns are not necessarily augmented by the introduction of new technologies, if the new technology is relatively simple to use. This is corroborated by estimates from a reproduction function in relation to the contraceptive revolution (Rosenzweig and Schultz, 1989). Foster and Rosenzweig (1993) report that high-tech and high-schooling returns are correlated based on the Green Revolution data of India

Psacharopoulos' (1994) finds that returns to schooling (particularly for primary schooling) in least developed countries (LDCs) are high, but Bennell (1996) begs to differ. He argues that with chronically low internal and external efficiencies at all educational levels in most Sub-Saharan Africa SSA countries, it seems highly implausible that rates of return to education are higher than in the advanced countries. Looking at returns country by country, it is certainly not the case that returns to primary education is consistently higher than either secondary or higher education (e.g., Appleton, et al, 1999)

When it comes to the analysis of returns to schooling in Ethiopia, there is very little empirical evidence. Using Youth Employment Survey of 1990 from Ethiopia, Krishnan (1996) investigates the impact of family background on both entry into employment in the private and public sector and its effect on returns to education. She finds that family networks to be a key determinant of entry into public sector work. However, education seems to serve as a screening mechanism in finding productive employees in the private sector. In another study (Krishnan et al, 1998) asks whether returns to education have changed over time following recent economic reforms. The study shows that returns to education, as measured by the total percentage returns from completing a particular level of education, have remained largely unaffected by the structural reforms.

III. Econometric methodology

It is now well-understood that OLS fails to account for the heterogeneity in the effect of education on earnings as well as the bias introduced due to the endogeneity of schooling (Buchinsky, 1998; Card, 1999). It is therefore important to adopt an empirical strategy that fits the earnings model across different ability levels, while at the same time allows for endogeneity of schooling. To this end, we deploy quantile regression techniques due to Koenker and Bassett (1978) in the estimation of standard Mincerian earning functions. As is customary in the literature (cf. Buchinsky, 1998; Arias et al, 2001), we assume that the unobserved ability distribution can be approximated by the conditional earnings distribution.

Let y_i denote the log of hourly wage of worker i and let X be the vector of covariates which consists of year of schooling, experience, experience squared, and the full set of for gender, ethnicity (as proxy for personal connections), year and location dummies.

The θ^{th} quantile of the conditional distribution of y_i given X is specified as:

$$Q_\theta(y_i \mid X) = \alpha(\theta) + X_i'\beta(\theta), \quad \theta \in (0,1). \tag{1}$$

where $Q_\theta(y_i \mid X)$ denotes the quantile θ of log earnings conditional on the vector of covariates. Following Koenker and Basset (1978), the θ^{th} quantile estimator can be defined as the solution to the problem:

$$\min_{\beta}\frac{1}{n}[\sum_{i:y_i\geq x'\beta}\theta|y_i - X_i'\beta_\theta| + \sum_{i:y_i\prec x_i'\beta}(1-\theta)|y_i - X_i'\beta_\theta|] = \min_{\beta}\frac{1}{n}\sum_{i=1}^{n}\rho_\theta(u_{\theta i}) \quad (2)$$

where $\rho_\theta(.)$ is known as the 'check function' and is defined as $\rho_\theta(u_{\theta i}) = \theta u_{\theta i}$ if $u_{\theta i} \geq 0$ and $\rho_\theta(u_{\theta i}) = (1-\theta)u_{\theta i}$ if $u_{\theta i} < 0$. The minimisation problem can be solved by using linear programming methods (Buchinsky, 1998). Like standard OLS estimates, a quantile regression estimate can be interpreted as the partial derivative with respect to a particular regressor at the relevant quantile.

To allow for the potential endogeneity of schooling alluded to earlier, we follow a two-stage quantile regression approach in which the schooling variable is instrumented with the years of schooling completed by the parents of the individuals under investigation[2]. Here the underlying assumption is the plausible scenario in which children of relatively more educated parents are likely to have more education. Since instrumental variables estimation within a quantile framework this is a non-standard problem, the variance-covariance matrices of the resulting estimates are obtained using bootstrapping techniques[3].

IV. Data

The paper uses panel data drawn from the 1994, 1995 and 1996 Ethiopian Urban Household Survey, conducted in seven urban areas. Members of each household are asked to report their wages (monthly, weekly and hourly), years of schooling completed, age, gender, ethnic origin, marital status, work experience in years. Information on the number of years of schooling completed by the parents of individuals covered in the survey is also available. For our study, we selected

[2] See Arias et al. (2001) for a recent application of instrumental variables quantile regression .

[3] The estimations for this study have been conducted using the Stata Release 7, and further details are available from the authors.

individuals from this survey based on the following three criteria: i) individuals who are currently wage employed either in the public or private sector; ii) individuals who are not attending full time schooling during the survey period; and iii) individuals who are between the ages of 15 and 59.

Table 1 reports some basic descriptive statistics. The average hourly wage for 1.658 Ethiopian *Birr*. This is equivalent to an average monthly earnings of about 347 *Birr,* which is nearly 3 times the minimum wage.[4] The wage data exhibit quite a high variation, which suggests the prevalence of substantial wage inequality. Figures 1 and 2 display the relationship between years of schooling for females and males in the sample with wages and conditional wages[5].

[Table 1 here]

V. Empirical estimate

We first estimate the Mincerian earning functions by assuming that the schooling variable is exogenous, in order to indicate the bias that might be introduced by neglecting the endogeneity issue. Table 2 reports the panel random effect and the quantile regression estimates for five values of θ.

According to the panel estimate the average return to one extra year education is 15%. This rather high figure is consistent with findings elsewhere in the developing world. But it is obvious that panel estimate masks important heterogeneity in the impacts of education. For example, the quantile regressions show that at the lower end of the earnings distribution (the 10th quantile) the marginal effect of schooling is more than 22%, whereas at the upper end it is only 11%.

[Table 2 here]

[4] There is no minimum wage legislation in Ethiopia but a wage of 120 Birr (US $15) per month is currently acceptable as minimum rate payable for unskilled workers.
[5] Conditional wages are obtained as a residuals from the regression of wages on experience location time and ethnic dummies.

As suggested by theory there is an inverted U-shaped relationship between earnings and experience. Furthermore, females appear to be discriminated against in the Ethiopian labour market, especially at the higher end of the income distribution.

Our main empirical findings from the instrumental variables quantile regressions are reported in Table 3. The panel IV estimate shows that the endogeneity-corrected schooling effect is on average 13%. Thus it would appear that OLS overestimated the average effect of schooling by two percentage points (or by about 12%). This is consistent with the direction and magnitude of OLS biases reported elsewhere in the literature (Card, 1999, 2001; Griliches, 1977).

[Table 3 here]

In our analysis we were careful to check for the appropriateness of parents' years of schooling as instruments for our schooling variable. Firstly, we apply a Sargan test for the over-identifying restrictions implied by the instruments. We find that parents' schooling and the disturbance term of the conditional earnings function are uncorrelated, suggesting that the instruments we employed are valid. Second, we also examine whether the instruments and the potentially endogenous schooling variable exhibit sufficiently high correlation. It has been noted in the econometric literature (see, for example, Staiger and Stock, 1997) that when the partial correlation between the instrument and the instrumented variables is low, instrumental variables regression is biased in the direction of the OLS estimator. Staiger and Stock (1997) recommend that the F-statistics (or equivalently the p-values) from the first-stage regression be routinely reported in applied work. The F-statistic tests the hypothesis that the instruments should be excluded from the first-stage regressions (i.e. they are irrelevant instruments). If we this hypothesis cannot be rejected (the F-statistic is too small or the corresponding p-value is large), the instrumental variable estimates and the associated confidence interval would be unreliable. Reassuringly, we find that the parents' schooling variables are relevant instruments.

The endogeneity-corrected quantile regression estimates show that the impact of an additional year of education at the lower end of the wage distribution is an increase in wages of 14.7%. This is nearly 24% lower compared with the equivalent coefficient in Table 2, emphasising that the bias introduced by endogenous schooling could be serious.

It is interesting to note that that the impact of schooling at the 25^{th} quantile is more than 10 percentage points higher than the returns to education at the 90^{th} quantile. Our finding returns to schooling diminishes with the level of income can be interpreted education being more beneficial to the less able, under the widely used assumption that the distributions of the unobserved ability and wages are positively related. Our finding is in line with the results reported by Ashenfelter and Rouse (1998) based on a sample of genetically identical twins in the U.S, but in contrast to the finding by Bauer et al (2002) that returns are higher at the higher end of the income distribution in Japan. For South Africa, Mwabu and Schultz (1996) report that ability and returns are positively related among white South African who received higher education, whereas returns are homogenous amongst blacks with high education. But at the primary education level, they find that returns to education and ability are negatively related.

If following Mwabu and Schultz (1996), we interpret a negative ability-returns relationship as evidence that education is a substitute for ability, this means that maximising (private) returns to schooling requires the expansion of educational opportunities for the less able or the more disadvantaged. By contrast, the relatively low (but still economically significant) returns at the higher end of the earning spectrum is consistent with the notion that there are important factors leading to high-paying employment, which act independently of education-generated human capital. This may take the form of inherent ability, or family connections as argued by Krishnan (1996) using a Youth Employment Survey in Ethiopia (see also Krueger, 2000 for a similar argument).

By way of robustness analysis we investigate whether the returns to education are different for public and private sector workers. As reported in Table 4, the panel IV estimates suggest that on average that education is more beneficial to private sector workers. However the quantile regressions indicates that the returns to schooling at the lower end of the income distribution are higher in the public sector.

[Table 4 here]

VI. CONCLUSION

The paper uncovers evidence that returns to schooling in urban Ethiopia exhibit substantial heterogeneity across the income distribution. It also shows that controlling for the endogeneity of schooling that results from its association with unmeasured ability is important for the accurate identification of the impacts of education. The empirical estimates indicate that education is more beneficial to at the lower spectrum of the income distribution, suggesting that the expansion of educational opportunities to the disadvantaged members of society might contribute to the maximisation of the private rate of returns.

REFERENCES

Antelius, J. (2000). Sheepskin Effects in the Returns to Education: Evidence on Swedish Data. FIEF Working Paper Series, NO. 158, Stockholm.

Appleton, S. (2001). Education, Incomes and Poverty in Uganda in the 1990s. School of Economics, University of Nottingham, CREDIT Research Paper, No. 01/22.

Appleton, S., A. Bigsten and D. K. Manda (1999). Educational Expansion and Economic Decline: Returns to Education in Kenya, 1978-1995. Centre for the Study of African Economies, Working Paper Series/99-6.

Arias, O; Hallock, K.F and Sosa-Escudero,W (2001). Individual heterogeneity in the returns to schooling: instrumental variables quantile regression using twins' data, *Empirical Economics*, 26, 7-40.

Ashenfelter, O. and Rouse, C (1998). Income, schooling, and ability: Evidence from a new sample of identical twins. *Quarterly Journal of Economics,* 113: 253-84.

Bauer, T. K., P. J. Dross and J. P. Haisken-DeNew (2002). Sheepskin Effects in Japan. IZA, Discussion Paper No. 593.

Bennell, P. (1996). Rates of Return to Education: does the Conventional Pattern Prevail in Sub-Saharan Africa? *World Development*, 24(1). : 183-200.

Buchinski, M.(1998). Recent advances in quantile regression models, *Journal of Human Resources*, 33: 88-126.

Card, D. (2001). Estimating the return to schooling: progress on some persistent econometric problems, *Econometrica*, 69(5): 1127-1160.

______ (1999). The Causal Effect of Education on Earnings, in (O. Ashenfelter and D. Card, eds). , *Handbook of Labour Economics*, vol. 3A. Amsterdam: Elsevier Science, North-Holland, and pp.-1801– 1863.

Foster, A. D., and M. R. Rosenzweig (1993). Information, Learning and Wage Returns in Low-Income Rural Areas. *Journal of Human Resources*, **28**(4). : 759-790.

Griliches, Z. (1977). Estimating the Returns to Schooling: Some Econometric Problems, *Econometrica*, **45**: 1-22.

Koenker, R. and G. Bassett (1978). Regression quantiles. *Econometrica*, Vol. 50: 43-61.

Krishnan, P. (1996). Family background, education and employment in urban Ethiopia, *Oxford Bulletin of Economics and Statistics*, Vol. 58 (1). , 1996.

Krishnan, P., T. G/Selassie and S. Dercon (1998). The Urban Labour Market During Structural Adjustment: Ethiopia 1990-1997. Centre for the Study of African Economies, WPS 98-9, University of Oxford.

Krueger, A. (2000). Education Matters: Selected Essays by A. B. Krueger. Edward Elgar, Cheltenham, UK and Northampton, MA, USA.

Mwabu, G. and Schultz, T.P.(1996). Education returns across quantiles of the wage function: Aternative explanations for returns to education by race in South Africa. *American Economic Review*, 8692): 335-339.

Psacharopoulos, G. (1985). Returns to Education: A Further International Update and Implications. *The Journal of Human Resources*, 20(4). : 583 – 605.

____________ (1994). Returns to Investment in Education: a global update. *World Development* 22(9). : 1325-1344.

Rosenzweig, M.R. (1995). Why are There Returns to Schooling? *American Economic Review*, Papers and Proceedings, 85(2). : 153-158.

Rosenzweig, M.R. and T.P. Schultz (1989). Schooling, Information and Non-market Productivity: Contraceptive Use and Its Effectiveness. *International Economic Review*, 30(2). : 457- 77.

Schultz, T. P. (1988). Education Investments and Returns, in H. Chenery and T. N. Srinivasan (eds.). , *Handbook of Development Economics*, i. Amsterdam: North-Holland.

Staiger, D. and J. H. Stock (1997). Instrumental Variables Regression with Weak Instruments, *Econometrica*, 65: 557-586.

Thomas, D., J. Strauss and M-H. Henriques (1991). How Does Mother's Education Affect Child Height? *Journal of Human Resources* 26 (2). : 183-211.

Table 1: Descriptive statistics

Variable	
Hourly wage in	
Mean	1.658
St. deviation	2.045
Quantiles	
Q10	0.104
Q25	0.469
Q50	1.194
Q75	2.343
Q90	3.703
Gender (% of females)	36.8
Public sector (%)	62.9%
Mean years of Schooling	9.00
(St. deviation)	4.389
Mean years of experience	10.60
(St. deviation)	(12.436)

Note: Wages are expressed in real Ethiopian currency- Birr.

Table 2
Returns to education in Urban Ethiopia:
Panel data and quantile regression estimates

Variable	Panel data estimates	10th quantile	25th quantile	50th quantile	75th quantile	90th quantile
Years of schooling	0.147	0.193	0.189	0.152	0.121	0.109
	(17.91)***	(15.61)***	(22.48)***	(25.85)***	(21.11)***	(9.93)***
Female dummy	-0.079	0.508	-0.019	-0.157	-0.214	-0.249
	(1.08)	(4.06)***	(0.25)	(2.99)***	(4.60)***	(3.47)***
Experience	0.078	0.157	0.099	0.068	0.049	0.044
	(15.05)***	(15.07)***	(15.47)***	(16.89)***	(14.20)***	(8.58)***
Experience squared	-0.001	-0.002	-0.001	-0.001	-0.0008	-0.0008
	(13.02)***	(11.71)***	(15.65)***	(15.28)***	(10.24)***	(6.89)***
Constant	-2.099	-4.581	-3.194	-1.886	-0.939	-0.264
	(15.36)***	(19.57)***	(20.77)***	(18.82)***	(10.56)***	(1.90)*
Observations	1476	1476	1476	1476	1476	1476

Notes:

(i) t-statistics are reported in parentheses;
(ii) * significant at 10%; **significant at 5%; *** significant at 1%;
(iii) The full set of time, ethnic and location dummies are included in the regressions.

Table 3
Returns to education in Urban Ethiopia:
Instrumental variable panel and quantile regression estimates

Variable	Panel IV estimates	10th quantile	25th quantile	50th quantile	75th quantile	90th quantile
Years of schooling	0.137	0.126	0.196	0.186	0.107	0.092
	(4.13)***	(2.40)**	(7.84)***	(7.66)***	(5.40)***	(1.95)*
Female dummy	-0.049	0.425	0.061	-0.146	-0.174	-0.179
	(0.88)	(2.38)**	(0.78)	(1.97)**	(2.81)***	(3.15)***
Experience	0.086	0.159	0.119	0.086	0.060	0.030
	(16.48)***	(10.44)***	(17.72)***	(14.39)***	(12.70)***	(6.97)***
Experience squared	-0.001	-0.002	-0.001	-0.001	-0.001	-0.0008
	(9.60)***	(10.54)***	(15.18)***	(11.55)***	(8.82)***	(4.66)***
Constant	-2.074	-4.299	-3.568	-2.442	-0.658	0.725
	(6.19)***	(7.43)***	(12.89)***	(9.16)***	(2.96)***	(3.29)***
Observations	1476	1476	1476	1476	1476	1476

Notes:

(i) t-statistics are reported in parentheses;
(ii) * significant at 10%; **significant at 5%; *** significant at 1%;
(iii) The full set of time, ethnic and location dummies are included in the regressions
(iv) The Sargan test for the validity of instruments conducted within the panel IV GMM framework gives a p-value of 0.247, validating the use of parents education as instruments
(v) We also checked the quality (relevance) of instruments by examining the joint significance in the first stage regressions. The resulting F statistic which is 10.28 (p-values =0) indicates a strong correlation between parents and offspring's education.

Table 4
Are the returns to education for public sector workers different?
Instrumental variable panel and quantile regression estimates

Variable	Panel IV estimates	10th quantile	25th quantile	50th quantile	75th quantile	90th quantile
schooling *public	0.118	0.178	0.236	0.199	0.131	0.101
	(13.27)***	(20.54)***	(27.15)***	(23.92)***	(19.96)***	(12.71)***
Schooling*private	0.163	0.117	0.159	0.153	0.144	0.128
	(7.73)***	(7.00)***	(14.74)***	(14.52)***	(14.42)***	(10.97)***
Female dummy	-0.124	0.144	0.023	-0.247	-0.205	-0.161
	(2.33)**	(1.54)	(0.36)	(3.86)***	(3.56)***	(2.23)**
Experience	0.066	0.065	0.057	0.054	0.043	0.038
	(12.67)***	(8.90)***	(11.59)***	(10.46)***	(9.19)***	(6.65)***
Experience squared	-0.001	-0.001	-0.001	-0.0001	-0.0001	-0.0001
	(7.77)***	(6.35)***	(8.84)***	(7.48)***	(6.05)***	(4.76)***
Constant	-1.893	-3.900	-2.851	-1.905	-0.730	0.012
	(13.48)***	(22.33)***	(25.16)***	(16.78)***	(7.34)***	(0.09)
Observations	1476	1476	1476	1476	1476	1476

Notes:

(i) t-statistics are reported in parentheses;
(ii) *significant at 10%; **significant at 5%; *** significant at 1%;
(iii) The full set of time, ethnic and location dummies are included in the regressions

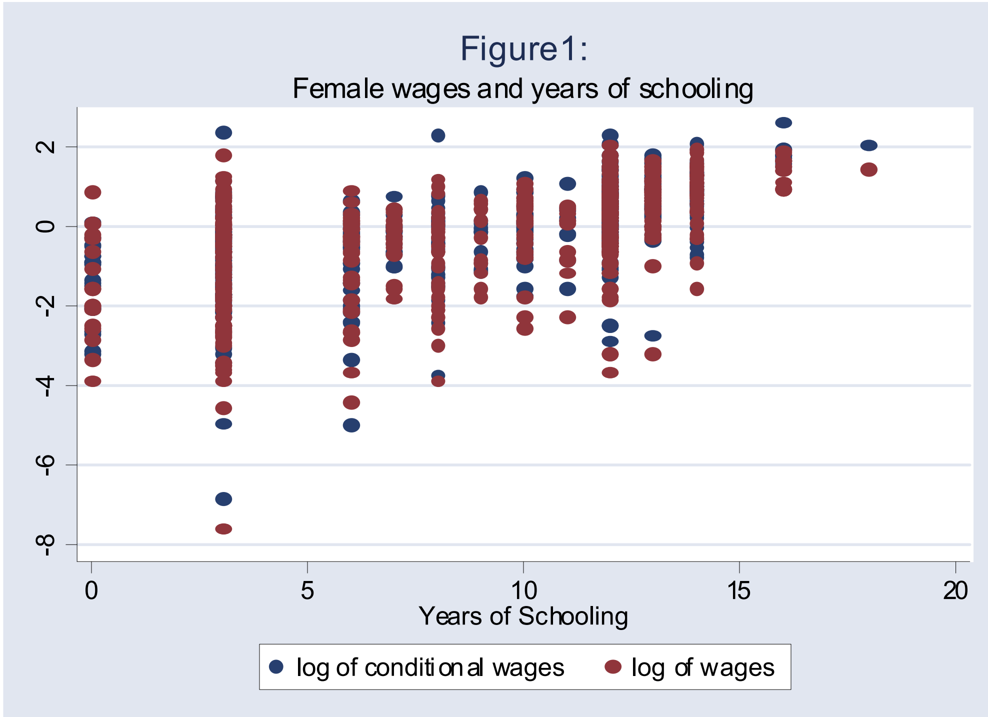
Figure1:
Female wages and years of schooling
2
0
-2
-4
-6
-8
0
5
10
15
20
Years of Schooling
log of conditional wages
log of wages


Figure2:
Male wages and years of schooling
5
0
-5
-10
0
5
10
15
20
Years of Schooling
log of conditional wages
log of wages