JOURNAL OF ECONOMIC DEVELOPMENT
Volume 33, Number 2, December 2008

MEASUREMENT ERROR BIAS IN RETURNS TO EDUCATION: EVIDENCE FROM A DEVELOPING COUNTRY - SRI LANKA

RASIKA RANASINGHE AND THOMAS HERTZ *

Connecticut College and American University

There is a continuing debate about the size and direction of the bias in estimates of returns to education. Evidence from developing countries is particularly scarce. This paper addresses the problem of measurement error bias in returns to schooling for Sri Lanka, by exploiting dual measurements of reported schooling for a sub-sample of the data and deriving a reliability estimate of schooling. This is used to obtain measurement error corrected fixed effects estimates of the proportionate increase associated with an additional level of schooling. This corrected measure is 5.5%, which is less than the OLS estimate of 7.8% for two person households.

Keywords: Rate of Return, Economic Development, Educational Economics, Human Capital
JEL classification: O12, O53

1. INTRODUCTION

The importance of education in the process of economic growth and development has been increasingly recognized, leading to the resurgence of interest in studying the relationship between education and labor market outcomes and earnings over the past decade (Card (2001), Behrman (1999)). Obtaining accurate and credible measures of returns to schooling involves minimizing the upward bias caused by omitted variables, and the downward attenuation bias caused by measurement errors in schooling. A main drawback of Ordinary Least Squares (OLS) estimates of returns to education is that they suffer from omitted variables bias. But correcting for omitted variable bias using household fixed-effects greatly exacerbates the measurement-error bias.

Thus, in order to obtain credible fixed-effects estimates, they must be corrected for measurement errors in reported schooling. While there are many studies that address these issues in estimating returns to education in developed countries (especially in the

* The authors are grateful to an anonymous referee for very helpful comments and suggestions.

US and OECD countries),[1] there is a paucity of studies that take into account measurement error and omitted variables in estimating returns to schooling in developing countries.[2] Existing evidence from Sri Lanka indicate that there are high and increasing returns to schooling at each additional level of education.

Private rates of return to education ranging between 10 to 15 percent for schooling above the secondary level are reported by Aturupane (1997), and Ranasinghe & Hartog (1997) find a 9 percent return to an extra level of schooling on average. However, these studies do not address issues of measurement error and omitted variable bias, and hence these results should be interpreted accordingly.

To our knowledge, there are no publicly available studies that address measurement error bias in returns to schooling in Sri Lanka. A main reason for this could be that up to now, there were no appropriate data available for this purpose. The present study utilizes data from the newly available Sri Lanka Integrated Survey of 1999/2000, to obtain measurement error corrected estimates of returns to schooling.[3] Exploiting the existence of dual measures of schooling for a subset of the sample, reliability of reported schooling is estimated, which is then used to correct for measurement error in schooling. The resulting within-family measurement error corrected estimate is 5.5 percent, while the initial OLS estimates are 7.6 and 7.8 percent for the full sample and two-member households respectively, which are considerably lower than other currently available estimates for Sri Lanka.

These results raise some important policy issues concerning Sri Lanka's education system, where the provision of free education up to university level has been in existence since the early 1940's. If the returns to education are much lower than currently supposed, this casts doubts on the effectiveness of free education as a tool in raising earnings of individuals. As will be evident from the analysis, there is still a considerably strong positive relationship between family background and schooling, despite free education. Hence, there is still reason to anticipate that OLS estimates of the returns to education will be upwardly biased by omitting family background.

The paper is organized as follows. The next section provides a brief overview of the education system in Sri Lanka, in order to place the issue at hand in context. Section three describes some selected approaches to modeling of returns to education, and related issues. Data and definitions, and model specifications are given in section four, followed by the analysis of empirical results in section five. The final section concludes

[1] For example, Behrman and Rosenzweig (1999), Ashenfelter and Krueger (1994), and Ashenfelter and Zimmerman (1993).

[2] An exception is Duflo (2001) who provides instrumental variables estimates for Indonesia.

[3] It has been shown by Mincer (1974) that if the return to schooling is independent of schooling level, and if the only costs of schooling are foregone earnings, then the proportional increase in earnings for each level of schooling is the return to education. Following conventional practice, the proportional increase per level of schooling is simply referred to as the rate of return to schooling.

and discusses the policy implications of the results.

2. OVERVIEW OF EDUCATION IN SRI LANKA

Sri Lanka has had a record of outstanding performance in education, making it an outlier among developing countries. As far back as the 1950's, Sri Lanka had surpassed today's low-income country averages in education indicators. It has since then preserved its edge over other developing countries in these traditional indicators. The estimated adult literacy rate is 90 percent, and gross enrollment rates in primary and secondary schooling are 105 and 74 percent respectively.[4] Moreover, there are no stark gender disparities in Sri Lanka's education outcomes, and this has positive implications on health and population outcomes. The introduction of free education in 1945 has had considerable long-term impacts on the socio-economic outcomes of Sri Lanka. By the time independence from nearly a century of British colonial rule was achieved in 1948, Sri Lanka already had a well established school system throughout the country.

Some significant policy changes that took place over the next few decades are worth noting. Until about the early 1950's two types of schools were in operation: fee-levying Christian missionary schools whose medium of instruction was English; and free vernacular schools. The former had preferential treatment in terms of educational facilities, employment and higher education opportunities. Hence in 1954, a new "language policy" was introduced which made the primary language of instruction Sinhala (the vernacular language). This policy was expected to provide the poor and the underprivileged equal access to educational facilities. The free provision of text books was introduced in the late 1970's as a means of lowering the direct costs of schooling. These policy measures, whether well intentioned or politically motivated-have had long-term effects on the education system, and on the overall social development in Sri Lanka.

Provision of subsidized education is expected to fulfill both *equity* and *efficiency* concerns. The former deals with minimizing family background effects on schooling decisions, thus enabling inter-generational mobility. Subsidized education is efficient in that it allows for optimal resource allocation within the family, since capital market constraints are a major obstacle to financing education. Even with the free education system that is in place in Sri Lanka, direct costs of schooling are not negligible.[5] In fact, Ranasinghe (2004) shows that a household on average spends nearly 3 percent of their total annual income on school related expenses. Low income households spend a substantially larger proportion on educational expenses. For example, households below the 20th income percentile spend up to 10 percent of their annual income on schooling

[4] The World Bank (1998).

[5] Free education covers tuition fees, text books, and school uniforms for low-income students.

related expenses. In contrast households at the 90th income percentile spend only about 1.5 percent of their income on schooling. Therefore, it seems that the costs of schooling are significant, despite "free" access to educational facilities. There are considerable out of pocket costs involved in schooling that are likely to constrain low-income households in their schooling decision. This view is supported by Ranasinghe & Hartog (2002) who conclude that despite the free education system, family background plays an important role in children's schooling choice and attainment.

The school structure for formal education begins with the primary school, followed by six years of secondary school that terminates with the General Certificate of Education-Advanced Level [G.C.E. (A/L)], which is the national university entrance examination. At present, about 10 percent of the children in the country do not attend school. School drop out begins by about the (junior) secondary level, and only a small fraction completes the full cycle. For instance, in 1996 only 21 percent of the students qualified for G.C.E. (A/L)'s, which implies that 79 percent of the students left school after G.C.E. (Ordinary Level), or the 10th grade. Of the students who do qualify for G.C.E. (A/L), only about 2 percent will actually gain admission to University. At present there are 13 State run universities in Sri Lanka, and nearly 20 percent of the university graduates remain unemployed, which raises questions about the tertiary education system.

Currently the entire education system is under pressure from international donors (the World Bank and IMF in particular) on the one hand who insist on major educational reforms including a fee-levying educational system; and from the general public on the other hand who now considers free public education as their right, and oppose any reforms. Government spending on education has been declining over the past few decades. From about 5 percent of GNP in the 1960's and the seventies, public expenditure on education has fallen to about 3 percent in the 1980's. In 1999 the total expenditure on education was 2.6 percent of GNP. In addition to insufficient financing, a number of other persistent issues have been identified regarding the nature and outcomes of the schooling system in Sri Lanka. The quality of education provided has been a cause for concern in the recent past. Contributory factors are out dated and irrelevant curricula, poor teacher quality, poor educational infrastructure such as dilapidated school buildings, and lack of teaching material and equipment. Additionally, there are increasing concerns regarding high educated unemployment, unsatisfied demand for university education, and low quality and relevance of education to labor market and social needs (World Bank (1998)).

3. SOME ASPECTS OF RETURNS TO EDUCATION

The widely used conventional Least Squares estimates of returns to education are subject to several possible sources of bias. First, income and schooling may be endogenous and jointly determined, leading to an upward bias of the OLS estimate (for

example, Card (2000)). This upward bias may occur when schooling attainment covaries with unobserved ability, family background, parental education, school quality, and socio-economic and locational factors (Card (1999), Griliches (1977)). Second, and the main focus of this study, is the attenuation caused by measurement errors in schooling that bias OLS estimates downwards. This attenuation bias can be partially offset by any upward bias from ability and unobserved omitted variables (Griliches (1977)).

Surveys of recent studies on returns to education indicate that least squares estimates are only slightly upward biased, or that the upward bias of the least squares estimate due to omitted variables is more or less offset by the attenuation bias caused by measurement error (Card (2000), Ashenfelter, Harmon & Oosterbreek (1999)). Ashenfelter & Zimmerman (1993) in their study of returns to education among siblings in the United States, find that the OLS estimates for brothers may be biased upward by as much as 25 percent due to omitted family background variables. However, when corrected for measurement error, their intrafamily estimate of returns to education is biased downward by about 25 percent, thus offsetting any bias in the least squares estimate. Ashenfelter & Rouse (1998), and Rouse (1999) also find similar results in studies of returns to education among twins in the US.

This result however may not hold true for developing countries, where the bias may be larger. In these countries, family and community background and capital constraints are likely to be important factors in determining both earnings and education. For instance, Hertz (2003) finds that measurement error corrected estimates in South Africa to be considerably lower than the OLS estimates. Yet Duflo (2001) shows that ability bias and measurement error bias approximately offset each other in Indonesia. Hence the debate on the size and the direction of the bias in the returns estimate still continues, with mixed evidence from different countries at different times.

4. THE DATA

The Sri Lanka Integrated Survey (SLIS) of 1999/2000 is a nationally representative household survey, carried out across all provinces of the country between October 1999 and the third quarter of 2000. A total of 7,500 households and 34,330 individuals were surveyed in 500 urban, rural and estate communities.[6] The SLIS is also the first household survey of this kind where the war-torn North and the Eastern provinces are included. The World Bank, in collaboration with the Sri Lanka Department of Census and Statistics (DCS) constructed the survey and sample design. The SLIS is organized under the World Bank Living Standards Measurement Survey (LSMS) framework, and consists of three modules: household, community, and, price questionnaires. Households

[6] The estate sector in Sri Lanka consists mainly of tea plantations where the majority of workers belong to the minority Tamil ethnic group (whose origins are in South India).

can be linked to the community they belong to, via a community serial number.

The SLIS is a cross sectional data set with multi-stage stratified random sampling where sectors (urban/rural) of each district are the domains for stratification. The master sample frame of the Demographic Survey (1994)[7] was used as the basis of population estimates for this present survey. Out of the 4000 Primary Sampling Units (PSU's) in this frame of the Demographic Survey, a sample of 375 PSUs was drawn at random for the SLIS. From each sampled PSU, 20 households (the final sampling unit) were drawn to construct a sample of 7500 households.

This study makes use of the household module of the survey, which contains detailed information on household members including demographic, income/expenditure, health, employment/unemployment, and migration. A unique feature of the household module is the inclusion of a separate section on the parents of household head and spouse. The purpose of this section is to gather intergenerational educational and mobility information. The head/spouse is asked to furnish their parents' level of education, in addition to other questions relating to landownership and migration. Additional sections on household socio-economic and demographic factors provide information on past and current educational attainment of household members, including that of the parents of head/spouse. The survey therefore provides two measures of the reported level of schooling for parents[8] — a feature exploited in this paper in order to obtain a reliability estimator of schooling, which is then used to correct for measurement error in returns to schooling for the entire sample.

The survey provides information on activities of household members 10 years and older, which enables the identification of labor market participants and employment activities. Earnings are reported for all economic activities stated for each household member. Monthly cash and in-kind receipts are reported (in Sri Lankan rupees, Rs.) for casual wage and estate workers, while monthly take-home pay, other payments (tips, bonuses), and the value of other benefits received are reported for regular salaried workers. Using this information, total monthly earnings are calculated for all activities.

Literacy and schooling are reported for household members 5 years and older. Follow up questions to those currently in school and those who attended school in the past provide detailed information not only on the highest level of schooling, but also on reasons for leaving school (if applicable), number of attempts at standardized exams, expenses on education, and type of school attended/attending. Schooling is classified into 16 levels, ranging from no-schooling to post-graduate qualifications. Reported level

[7] A Census of the population is conducted in Sri Lanka every ten years since 1871, except in 1991 due to political instability in certain parts of the country. A one time Demographic Survey was done in 1994 to make up for this, which covered 100,000 households. Information on demographic characteristics, housing, disability, education, employment, income, expenditure and migration are collected in the survey.

[8] "Parents" refer to the parents of household head or spouse throughout the rest of the paper, unless otherwise stated.

of education, which is the primary variable of interest, was carefully reconstructed using the available information on literacy, and on past and current enrollment levels, in order to minimize the frequency of missing values. Information on literacy provides a starting point to obtaining individual schooling: those with no schooling are categorized as illiterate, while the rest are literate with some level of education. This latter group is further subdivided into those who are no longer in school, and currently in school. This information is then used to obtain the relevant educational attainment of individuals who have reported some level of schooling.

For the purpose of this study, 6428 individuals aged 16 years and over, from the urban and rural sectors with non-missing incomes were retained as the final sample. Males constitute nearly 75 percent of the sample. Approximately 77 percent of the people are from the rural sector, which is not surprising given that Sri Lanka is mainly a rural agricultural economy.

Table 1. Descriptive Statistics

Overall Sample: All Household Members Aged 15 years and Above						
	Female		Male		All	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Sample size	1586		4842		6428	
Age	34	11.2	36.2	12.9	36	12.01
Education	9.6	4.1	8.6	3.5	8.9	3.7
Monthly income (R_S)	3647	2975	4378	3329	4198	3261
Log monthly income	7.9	0.7	8.2	0.6	8.1	0.6
Parents sub-sample: Parents of household head and spouse						
	Female		Male		All	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Sample size	443		161		604	
Age	65.8	11.4	70.6	12.4	67.3	11.9
Education	4.4	3.6	5.2	3.6	4.7	3.6
Monthly income (R_S)	2093	1195	2900	2798	2604	2347
Log monthly income	7.5	0.6	8.1	0.7	7.7	0.7

As reported in Table 1, the average level of schooling for the entire sample is about 9 years, with females reporting higher educational attainment than males. The average monthly earnings are about Rs. 4198.00 (nominal value at 1999 prices) for the overall sample, with males earning nearly 20 percent more than the females. Nearly 5 percent of the final sample has had no schooling, while the majority (39 percent) has had up to lower secondary level of schooling. Only 4.4 percent has university education and above.

Table 2. Matrix of Reported Schooling for Parents of Household Head/Spouse: Parents' Roster Versus Household Roster

Reported Schooling in the Parents' Roster (number of years)	**Reported Schooling in the Household Roster (number of years)**																	
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**Total**
0	**102**	0	6	4	4	5	1	0	2	0	1	3	0	0	0	0	0	**128**
1	1	**2**	2	0	0	0	1	0	0	0	0	0	0	1	0	0	0	**7**
2	7	2	**25**	4	3	1	1	0	0	0	0	0	0	0	0	0	0	**43**
3	6	0	4	**38**	3	2	2	1	0	0	0	1	0	0	0	0	0	**57**
4	6	0	1	4	**25**	7	3	3	1	1	0	1	0	0	0	0	0	**52**
5	5	0	3	7	4	**62**	5	1	3	1	1	1	0	0	0	0	0	**93**
6	2	0	0	0	1	7	**34**	0	3	0	3	0	1	1	0	0	0	**52**
7	1	0	0	0	1	2	3	**18**	4	1	3	0	0	0	0	0	0	**33**
8	1	0	0	1	3	4	3	1	**40**	5	4	3	0	0	0	0	0	**65**
9	1	0	0	2	0	2	1	0	3	**14**	2	1	0	0	0	0	0	**26**
10	0	0	0	0	1	0	1	0	2	2	**13**	5	0	0	0	0	0	**24**
11	0	0	0	0	0	0	1	1	0	0	2	**12**	0	0	0	0	0	**16**
12	0	0	1	0	0	0	0	0	0	0	0	0	**1**	0	0	0	1	**3**
13	0	0	0	0	0	0	0	0	0	0	0	0	0	**2**	0	0	0	**2**
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	**1**
15	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	**1**	0	**2**
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**0**
Total	**132**	**4**	**42**	**61**	**45**	**92**	**56**	**25**	**58**	**24**	**29**	**27**	**2**	**4**	**1**	**1**	**1**	**604**

4.1. Dual Education Measures

Of the total sample, 604 observations with dual educational measures for parents of household head/spouse were retained. These were obtained by comparing the two reported levels for them — one from the parents' information module (which contains only parental information), and the other from the household roster (which contains information on all household members, including the parents). For each parent, the reported schooling in the household roster was compared with the entry for the schooling variable in the parents' roster, in order to obtain the dual schooling measure.

There are 421 households with only one parent of head/spouse, while 87 households have two parents and three households with three parents. Thus there are 511 households in the parents sub-set. Approximately 22 percent of the parents have had no schooling, while the average level of their education is about 6 years of schooling. The lower panel of Table 1 describes the parents (of head/spouse) sub-sample. Of the 604 observations, 443 are females (mother of head/spouse). Mean schooling for the overall sample is about 9 years of schooling, indicating that the average schooling for parents is not the same as that for the entire sample.

Table 2 presents the matrix of reported schooling in the household roster versus schooling reported in the parents' roster. This gives an idea of the nature of the reliability of schooling data. If there were no measurement errors, the reported schooling for parents in both modules should be identical, and the off-diagonal elements should be zero. However, as Table 2 clearly shows there are considerable measurement errors in schooling data, which again emphasizes the importance of correcting for such errors when estimating returns to schooling.

5. MODELING RETURNS TO SCHOOLING

If schooling is measured perfectly, with no measurement errors, and there is no endogeneity between earnings, schooling and other covariates, then the returns to education will be consistent. That is, its probability limit will be equal to the average marginal return to schooling $[p\lim b = E(b)]$ However, if this is not the case, and an individual's education level is measured erroneously, it will be correlated with the regression error and the OLS estimate of the schooling coefficient will be biased. To further explore this, consider the Mincerian (log) earnings function for individual *i*;

$$\log y_i = \alpha_i + \beta_i S_i + u_i , \tag{1}$$

where $y =$ total monthly earnings, $S =$ schooling, and $u =$ person-specific random error term. If schooling is measured error-free, and there is no endogeneity in earnings and schooling and other covariates, β_i represents the unbiased person-specific returns

to schooling [thus $E[\beta_i] = \overline{\beta}$].

However, unobserved factors such as ability, family and individual characteristics, school and teacher quality, lead to endogeneity in schooling and earnings. This implies that y_i and S_i are correlated, resulting in biased OLS estimators of β_i. In order to capture this correlation, the linear projections of α_i and ($\beta_i - \overline{\beta}$) on schooling (S_i) can be written as follows:

$$\alpha_i - \alpha_0 = \lambda(S_i - \overline{S}) + u_i , \tag{2a}$$

$$\beta_i - \overline{\beta} = \varphi(S_i - \overline{S}) + v_i . \tag{2b}$$

$\overline{S}$ represents the mean schooling across all individuals, and $E[S_i u_i] = E[S_i v_i] = 0$, and λ and φ are theoretical regression coefficients with $\lambda = \text{cov}(\alpha_i, S_i)/\text{var}(S_i)$ and $\varphi = \text{cov}(\beta_i, S_i)/\text{var}(S_i)$.

Assuming that β and S are jointly symmetrically distributed, Card shows that the probability limit of the OLS estimate of β (from Equation (1)) is;

$$p\lim(\beta_{ols}) = \overline{\beta} + \lambda + \varphi\overline{S} , \tag{3}$$

where $\overline{\beta} = E[\beta_i]$. The magnitude and direction of the bias will depend on the effect of omitted variables-such as ability (cognitive skills, availability of funds, willingness), quality of education (including school/teacher quality) and family background, on schooling.

In order to derive the corresponding within-family estimator which controls for family effects, consider a household with two individuals with reported incomes.

$$\log y_1 = \alpha_1 + \beta_1 S_1 , \tag{4a}$$

$$\log y_2 = \alpha_2 + \beta_2 S_2 . \tag{4b}$$

The linear projections onto observed schooling are;

$$\alpha_1 - \overline{\alpha} = \lambda_{11}(S_1 - \overline{S}_1) + \lambda_{12}(S_2 - \overline{S}_2) + u_1 , \tag{5a}$$

$$\beta_1 - \overline{\beta} = \varphi_{11}(S_1 - \overline{S}_1) + \varphi_{12}(S_2 - \overline{S}_2) + v_1 , \tag{5b}$$

$$\alpha_2 - \overline{\alpha} = \lambda_{21}(S_1 - \overline{S}_1) + \lambda_{22}(S_2 - \overline{S}_2) + u_2 , \tag{5c}$$

$$\beta_2 - \overline{\beta} = \varphi_{21}(S_1 - \overline{S}_1) + \varphi_{22}(S_2 - \overline{S}_2) + v_2, \tag{5d}$$

where u_i and v_i are orthogonal to S_1 and S_2. $\overline{\alpha}, \overline{\beta}$ and $\overline{S}$ are expected values across households. λ_{11} reflects the correlation between the first person's own education and α, and λ_{12} is the connection between the first person's α and the second person's schooling. Similarly φ_{11} and φ_{12} reflect the correlation between the first person's own education and β, and the connection between the first person's β and the second person's schooling. Substituting (5a) through (5d) to the income equations (4a) and (4b) and rearranging terms lead to the following income equations:

$$\log y_1 = c_1 + S_1\tau_{11} + S_2\tau_{12} + e_1, \tag{6a}$$

$$\log y_2 = c_2 + S_1\tau_{21} + S_2\tau_{22} + e_2, \tag{6b}$$

where (the own-schooling effect and the other-family-member's schooling effects)

$$\tau_{11} = \overline{\beta} + \lambda_{11} + \phi_{11}\overline{S}_1,$$

$$\tau_{12} = \lambda_{12} + \phi_{12}\overline{S}_1,$$

$$\tau_{21} = \lambda_{21} + \phi_{21}\overline{S}_2,$$

$$\tau_{22} = \overline{\beta} + \lambda_{22} + \phi_{22}\overline{S}_2.$$

Assuming symmetry ($\lambda_{11} = \lambda_{22}, \lambda_{12} = \lambda_{21}, \varphi_{11} = \varphi_{22}, \varphi_{12} = \varphi_{21}$), the within family (fixed-effects) estimator can be obtained as follows:

$$\begin{aligned} p\lim(\tau_{11} - \tau_{12}) &= \overline{\beta} + \lambda_{11} + \varphi_{11}\overline{S} - \lambda_{12} - \varphi_{12}\overline{S}, \\ p\lim(\beta_{\Delta ols}) &= \overline{\beta} + (\lambda_{11} - \lambda_{12}) + (\varphi_{11} - \varphi_{12})\overline{S}. \end{aligned} \tag{7}$$

5.1. Measurement Errors in Schooling

The result of measurement errors in schooling is to cause a downward bias in the OLS estimate of the schooling coefficient (Griliches (1977)). To incorporate the effect of measurement errors on the returns to education in the SLIS data, two alternative measures of (non-missing) reported schooling for each parent of the household head and spouse are used. The first measure comes from the household roster where the education level of all household members is recorded, and the second measure of schooling is from

the parents' module, which specifically targets the parents of head/spouse.

Let reported schooling of parent *i* in the household roster be S_i^1, and the level of schooling reported in the parents' module be S_i^2. If there is no measurement error in observed schooling, $S_i^1 = S_i^2 = S_i$, where S_i = true schooling for the i^{th} parent. Assuming that true schooling and measured schooling differ by an additive error, and dropping the individual subscript *i* for clarity, true schooling for each parent in the household roster and parents module can be expressed as follows:

$$S^1 = S + e_1, \tag{8}$$

$$S^2 = S + e_2, \tag{9}$$

where $E[e_1] = E[e_2] = E[e_1 e_2] = E[e_j s_i] = 0$ and $E[e_i^2] = \sigma^2$.

If Equation (3) represents the probability limit of the OLS estimator of β based on *true* schooling, the OLS regression of log earnings in the household roster on *reported* schooling will produce the following estimator of β:

$$p\lim(\beta_{ols}) = R(\overline{\beta} + \lambda + \varphi\overline{S}), \tag{10}$$

where R is the reliability (i.e., the ratio of signal-to-total-variance) of reported schooling in the household roster (S^1), which can be expressed as:

$$R \equiv \frac{\text{cov}[S^1, S]}{\text{var}[S^1]} = \frac{\text{var}(S)}{\text{var}(S) + \text{var}(e_1)}.$$

Thus when measurement errors are orthogonal to true schooling, the reliability ratio (R) will be less than 1, and the least squares estimates of returns to education using observed schooling will be biased towards zero (relative to when there is no measurement error).

The estimated reliability of reported schooling in the household roster, obtained as described above is 79 percent, implying that 21 percent of observed schooling variance is error. This figure is comparable with that of Hertz (2003) who reports a 77 percent reliability for South Africa.

The variance and the covariance of the measurement error term in the household roster (e_1) can be expressed as follows:

$$\text{var}(e_1) = (1 - R)\,\text{var}(S^1),$$

$$Cov(e_i^1, e_j^1) = (1-\Omega)\,\mathrm{cov}(S_i^1, S_j^1),$$

where $\Omega \equiv \dfrac{\mathrm{cov}(S^1, S^2)}{\mathrm{cov}(S^1, S^2) + \mathrm{cov}(e_i^1, e_j^1)}$ reflects the correlated errors.

The measurement error corrected with-in family differenced (fixed-effects) estimator can be obtained similarly from Equation (7):

$$p\lim(\beta_{\Delta FE}) = R_\Delta[\overline{\beta} + (\lambda_{11} - \lambda_{12}) + (\varphi_{11} - \varphi_{12})\overline{S}], \tag{11}$$

where R_Δ is the reliability of the difference in schooling which can be written as follows:

$$R_\Delta = (R - \Omega\rho_s)/(1-\rho_s).$$

5.2. Estimating R

The reliability of reported schooling in the household roster (R) can be estimated by regressing S^2 on S^1. Thus R can be expressed as:

$$\hat{R} \equiv \left(\frac{\mathrm{cov}(S^2, S^1)}{\mathrm{var}(S^1)} \right).$$

Empirically, R can be obtained by a correlation of the two schooling measures, for the sample of parents. The reliability estimate for within-family differences can be obtained in a similar fashion from the corresponding regression of the within-family differences:

$$\hat{R}_\Delta \equiv \left(\frac{\mathrm{cov}(\Delta S_2, \Delta S_1)}{\mathrm{var}(\Delta S_1)} \right) = \frac{R - \Omega\rho_S}{1-\rho_S}, \tag{12}$$

where ρ is the correlation of the two schooling measures, $Corr(S_i^1, S_j^1)$. The classical model assumes independent measurement errors, $Cov(e_i^1, e_j^1) = 0$ which implies that $R_\Delta < R$. That is, the attenuation caused by measurement error in schooling is higher in the fixed-effects estimator than the standard cross-sectional result. The within-family differenced estimator is especially susceptible to measurement error since differencing within families removes much of the true signal in education as well. Generally the fixed-effects (within family) estimator will be smaller when the measurement errors are

contemporaneously correlated. As before, R_Δ is empirically obtained by a correlation of the two schooling measures-differenced within households, for the sample of two-member households.

6. RESULTS

Reliability estimates of parent's reported schooling, calculated as described above in section 4 are reported in Table 3. Estimated reliability of parents' schooling is 79 percent, for the entire sample of parents (N=604). The reliability estimate is 76 percent for the within-family differences (R_Δ), which controls for family effects. This latter estimate is based on the sub-sample of households with two reported measures of schooling for parents (87 households with 174 individuals). These estimates are comparable with those for South Africa reported by Hertz (2003) which fall between 0.72 and 0.61, but are considerably lower than the 0.90 rule-of-thumb reliability for the United States. This implies that schooling data from developing countries are not as reliable as those from developed countries. The within-household correlation coefficient of the two schooling measures (ρ) is 0.56, and the imputed value of Ω which reflects the correlated errors, is 0.81.

Table 3. Reliability Estimates of Parents' Reported Schooling: 1999/2000

$\hat{R} \equiv \left(\frac{\text{cov}(S^2, S^1)}{\text{var}(S^1)} \right)$	0.79 (N=604)
$\hat{R}_\Delta \equiv \left(\frac{\text{cov}(\Delta S^2, \Delta S^1)}{\text{var}(\Delta S^1)} \right)$	0.76 (N=174)
$\rho = Corr(S_i^1, S_j^1)$	0.56 (0.065)
$\Omega \equiv \frac{\hat{R} - \hat{R}_\Delta(1-\rho)}{\rho}$	0.81

Notes: S^1 = Reported schooling from the household roster. S^2 = Reported schooling from the parents' module.

Table 4 presents the OLS, fixed-effects and measurement error corrected estimates for the full sample of 6428 wage-earning individuals, and for two-member household, which has a sample of 3421 people. The least squares estimate of the return to schooling for the full sample is 0.076. This estimate increases to 0.078 when only two-member households are considered, but falls considerably to 0.037 with household fixed-effects. Correcting for measurement error raises the estimate to 0.055, but it is still lower than

the OLS estimate.

Table 4. OLS, Fixed Effects and Measurement Error Corrected Estimates

	Full Sample	Two-Member Families		
	OLS	OLS	Household Fixed-Effects	Measurement Error Corrected Fixed Effects (Reliability=0.76)
Level of Schooling	0.076	0.078	0.037	0.055
	(0.002)	(0.0029)	(0.005)	(0.005)
Age	0.039	0.040	0.028	0.030
	(0.004)	(0.0048)	(0.005)	(0.004)
Age^2	-0.0004	-0.00044	-0.0003	-0.00034
	(0.00004)	(0.00006)	(0.00006)	(0.00005)
Male	0.287	0.258	0.275	0.277
	(0.0175)	(0.0226)	(0.0214)	(0.016)
R^2	0.20	0.21	0.11	0.12
Sample Size	6428	3421	3421	3421

Since R_Δ is itself an estimate, subject to sampling error, in order to obtain a credible range of estimates of returns to education, the reliability estimate for the differenced data is calculated for the 95 percent confidence interval, which results in a range of values for R_Δ (the lower 95 percent confidence interval estimate of R_Δ, the central value, and the upper 95 percent confidence interval estimate of R_Δ). These reliability estimates are then used to correct for measurement error in the household-fixed effects model. Therefore, it is possible to obtain a range of estimates for the measurement error corrected returns to schooling, in addition to the central estimate of returns to education.

The values for R_Δ, estimated as above, range from 0.62 at the lower 95 percent confidence interval, to 0.89 at the upper bound of the 95 percent confidence interval. The corresponding measurement error corrected returns to schooling in the fixed-effects model now range from a lower bound of 0.044 (for $R_\Delta = 0.89$) up to 0.077 (for $R_\Delta = 0.62$), which is approximately close to the OLS estimate of 0.078 for two-member households. Therefore, it appears that depending on the reliability estimate used, it is possible to obtain a range of measurement error corrected estimates of returns to schooling that could be nearly as low as the household fixed-effects estimate and as high as the least squares estimate.

7. CONCLUSIONS

In this paper, measurement error corrected estimates of returns to education were obtained using a reliability estimator from dual measures of education for a sub-sample of parents of household heads and spouses. The results indicate that the standard OLS estimate of returns to schooling is biased upward. The measurement error corrected return to a year of schooling is 5.5 percent for the household fixed-effects model, which is about 20 percent lower than the standard OLS estimate.

The reliability estimate for household differences can be obtained as a range of estimates in the 95 percent confidence interval, which will produce a corresponding range of measurement error corrected returns to schooling. These estimates lie between 4.4 percent and 7.7 percent. Hence it seems that there is considerable bias due to measurement error in schooling in Sri Lanka, which points to the issue of reliability of education data from developing countries, and the importance of correcting for such biases in order to obtain better estimates of returns to schooling.

The results of the preceding analysis would suggest that returns to education in Sri Lanka are not as high as one would like to think at first glance. The policy implications of such a finding are non-trivial given that Sri Lanka has had a free education system in place for the past six decades. While the provision of free education has increased access to schooling, and has brought about improvements in the general socio-economic conditions of the people, the results of this study pose some important questions on the overall effectiveness of the free education system in Sri Lanka as a means of raising individual earnings.

REFERENCES

Arias, O., Hallock, K., and W. Sosa-Escudero (2001), "Individual Heterogeneity in the Returns to Schooling; Instrumental Variables Quantile Regression Using Twins Data," *Empirical Economics*, 26(1), 7-40.

Ashenfelter, O., and D. Zimmerman (1993), "Estimates of the Returns to Schooling from Sibling Data; Fathers, Sons and Brothers," Working Paper, 318, Princeton: Industrial Relations Section, University of Princeton.

Ashenfelter, O., and A.B. Krueger (1994), "Estimates of the Economic Return to Schooling from a New Sample of Twins," *American Economic Review*, 84(5), 1157-1173.

Ashenfleter, O., and C. Rouse (1998), "Income, Schooling & Ability; Evidence from a New Sample of Twins," *Quarterly Journal of Economics*, 113(1), 253-284.

Aturupane, H. (1997), "Earnings Functions and Rates of Return to Education in Sri Lanka," UC-ISS Working Paper, 970, Colombo: University of Colombo, Sri Lanka.

_____ (1998), *Education and Poverty in Sri Lanka*, Colombo: Ministry of Planning, Sri Lanka. (Background Paper for the Sri Lanka Poverty Assessment, 1999).

Behrman, J.R., and M.R. Rosenzweig (1999), "Ability Biases in Schooling Returns and Twins; A Test and New Estimates," *Economics of Education Review*, 18(2), 159-167.

Card, D. (2000), "Estimating the Returns to Schooling; Progress on Some Persistent Econometric Problems," Working Paper, 7769, Cambridge, MA: National Bureau of Economic Research, Cambridge.

_____ (1999), "The Causal Effect of Education on Earnings," in Ashenfelter, O., and D. Card eds., *The Handbook of Labor Economics - Vol.3*, Elsevier Science.

_____ (1995), "Earnings, Schooling and Ability Revisited," in S. Polacheck ed., *Research in Labor Economics - Vol. 14*, JAI Press Inc.

Chesher, A. (1998), "Measurement Error Bias Reduction," Discussion Paper, 98/449, Bristol: Department of Economics, University of Bristol.

Dufflo, 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 Returns to Schooling; Some Econometric Problems," *Econometrica* , 45(1), 1-22.

Hertz, T. (2003), "Upward Bias in the Estimated Returns to Education; Evidence from South Africa," *American Economic Review*, 93(4), 1354-1368.

Kane, T., Rouse, C., and D. Staiger (1999), "Estimating Returns to Schooling When Schooling is Misreported," Working Paper, 7235, Cambridge, MA: National Bureau of Economic Research, Cambridge.

Kling, J. (2000), "Interpreting Instrumental Variables Estimates of the Returns to Schooling," Working Paper, 7989, Cambridge, MA: National Bureau of Economic Research, Cambridge.

Mincer, J. (1974), *Schooling, Experience and Earnings*, New York: Columbia University Press.

Ranasinghe, R. (2004), "Returns to Schooling, Mobility and the Gender Wage Gap; Evidence from Sri Lanka and India," unpublished Ph.D. dissertation, Department of Economics, American University, Washington, DC.

Ranasinghe, A., and J. Hartog (1997), *Investment in Post-compulsory Education in Sri Lanka*, Tinbergen Institute, Amsterdam, The Netherlands.

_____ (2002), "Free Education in Sri Lanka; Does it Eliminate the Family Effect?" *Economics of Education Review*, 21(6), 623-633.

Rouse, C. (1998), "Further Estimates of the Economic Return to Schooling from a New Sample of Twins," Working Paper, 388, Princeton, NJ: Industrial Relations Section, Princeton University.

The World Bank (1998), "Sri Lanka Social Services; A Review of Recent Trends and Issues," Report, 17748-CE, Washington, DC: Poverty Reduction and Economic Management, South Asia Region, The World Bank.

_____ (2000), *Sri Lanka Poverty Assessment*, Washington, DC: The World Bank.

Mailing Address: Rasika Ranasinghe, Department of Economics,Box 5552, Conneticut College, 270 Mohegan Avenue, New London, CT 06320, U.S.A. Tel: 860-447-1911. E-mail: rasika.ranasinghe@conncoll.edu

Manuscript received July 2007; final revision received September 2008.