

This is the accepted version of the manuscript Rozana Himaz, "Intrahousehold Allocation of Education Expenditure: The Case of Sri Lanka," Economic Development and Cultural Change 58, no. 2 (January 2010): 231-258. <https://doi.org/10.1086/648187>

Intrahousehold Allocation of Education Expenditure: The Case of Sri Lanka

Rozana Himaz
University of Oxford
August 8, 2008

Short-Title: Intrahousehold Allocation of Education expenditure

Author Name: Rozana Himaz

Mailing Address: Queen Elizabeth House,

3, Mansfield Road,

University of Oxford

Oxford OX1 3TB

Telephone number (mobile): 44-(0)7986872517 (work)-44-(0)1865289971

E-mail: rozana.himaz@qeh.ox.ac.uk

Keywords: Intrahousehold Allocation, Sri Lanka, Gender

JEL Classification: D13, I21, J16

* I would like to thank Sonia Bhalotra, Sara Horrel, Pramila Krishnan and 3 anonymous referees for helpful comments on an earlier version of this article. The usual disclaimer applies. E-mail: rozana.himaz@qeh.ox.ac.uk.

Abstract

This paper uses demand conventional Engel Curve analysis as well as a double-hurdle model to explore whether there are intrahousehold differentials in the allocation of education expenditure between boys and girls in rural Sri Lanka. Contrary to most developing countries there is a significant bias favouring girls for 1990/91 and 1995/6 for age groups categories 8-9, 14-16 and 17-19 and in 2000 for age category 14-16 and 17-19. Significant differences in enrolment favouring girls for age groups 17-19 explain part of the girl-bias observed in 1990/91 and 2000/01 but most of the bias is driven by positive expenditure given enrolment. The biases favouring girls are observed at critical stages of the schooling career in the run-up to key national exams. The 8-9 age group captures the run-up to the Year 5 scholarship exams that are used to gain entry to better performing secondary schools. The 14-16 and 17-19 age groups capture those who read for important National level qualifications vital for the job market. This paper also looks at various possible explanations for the bias.

1. Introduction

Developing countries have been a testing ground for the investigation of gender bias in the allocation of educational resources -mainly schooling expenditure- in a household (Burgess and Zhuang 2000, Rudd 1993, Kingdon 2005). The rationale for the exercise is often found in the hypothesis that girls may be less favoured than boys (or even discriminated against) in terms of parents spending towards their education. Often macro data lends support to this contention with female schooling rates being significantly less than those for men

Sri Lanka is a particularly interesting case to look at in this regard. This is because it is a developing country in South Asia well known for her high achievements in male and female literacy and gender 'equality' in terms of school enrolment and exam completion. In recent years, however, this success seems to have translated into a very interesting phenomena: School enrolment rates are *higher* for girls compared to boys especially at more *senior* grades in school-grades that run up to important national level examinations such as the Ordinary Levels (ages 14-16) and Advanced Levels (ages 17-19). Table 1 shows that this gap in enrolment is significant throughout the 1990s, for the age 14-16 category. It is interesting examine whether this gap observed at the aggregate level is reflected in household level investment in education and if there is a significant favouring of boys or girls at the margin, in terms of both enrolment and expenditure on education once enrolled.

It is also interesting to attempt to explain why a household-level gender bias exists, if at all, in the allocation of education expenditure. Could it be due to gender differences in returns to schooling in the labour market? Could it be indicative of an insidious form of 'discrimination' that happens in the labour market where girls have to be more educated than their counterparts in the opposite gender when competing

for the same job? Alternatively, could the answer lie outside the labour market reflecting perhaps differences in ‘cultural’ attitudes to the education of boys and girls?

This paper attempts, primarily, to establish whether there are gender differences in expenditure towards education in rural Sri Lanka over the decade beginning in the 1990s. It is an issue that has not been investigated so far. It also looks at some possible explanations for the results observed. I use both demand analysis (i.e., the conventional Engel Curve method) as well as a double-hurdle model. The study uses 3 household income and expenditure surveys for 1990/91, 1995/96 and 2000/01. Section 2 discusses the Engel curve and double-hurdle models. Section 3 discusses the data and section 4 looks at the results and checks for robustness. Section 5 attempts to explain the results. Section 6 concludes.

2. Intrahousehold Allocation of Education Expenditure

2.1. Engel Framework

If individual level data was available, then we could directly compare expenditure on education for males and females. However, given the lack of such individual level data, intrahousehold allocational differences have to be estimated indirectly. Data is available at the level of the household and therefore I try to detect gender-biases in education expenditure by investigating how the presence of individuals of similar ages but opposite sexes affect household expenditure on education. The Working-Lesser Engel form for demand analysis is used with a linear relationship assumed between the share of the budget on each good and the log of total household expenditure. Deaton and Muellbauer (1980, p.75) argue that such a relationship has the theoretical advantage of being consistent with a utility function

and conforms to data 'in a wide range of circumstances'. As discussed in Deaton (1997:231), Working's Engel curve can be extended to include household demographic composition where age classes are denoted by n_j and are broken down by gender. Separate γ_{ij} coefficients can be calculated for males and females:

$$w_i = \alpha_i + \beta_i \ln(x/n) + \eta_i \ln n + \sum_{j=1}^{J-1} \gamma_{ij} (n_j/n) \tau_i \tilde{z} + u_i \quad (1)$$

where w_i is the share of the household budget devoted to the i^{th} good (education expenditure in this paper) calculated as $p_i q_i / X$ with p_i and q_i denoting the price and quantity of good i (education) and x is total expenditure per household, n is household size, n_j is the number of people in age-sex class j (there are J such classes in total)¹.

The age categories adopted for children are important because each of it is the run up to an important national exam that qualifies a student to enter the next stage and even make the choice of entering an institution that is reputed for better performance than the one which he or she leaves. The vector z contains other socio-economic variables such as the education of the household head, ethnic group, location (district) dummies. Finally, u_i is the error-term for good i (education). The coefficient β determines whether the good is a luxury or a necessity. If $\beta > 0$ then the good is a luxury with the budget share increasing with total outlay making the total expenditure elasticity greater than 1. The good is a necessity if $\beta < 0$ with an

¹ The potential endogeneity of household expenditure per capita is checked for with the use of the instrumental variable approach. I use unearned income and its square as instruments. Unearned income comprise dividends, interest and rents. Roughly 13 per cent of the rural households in the HIES have some form of positive unearned income. The instruments are relevant with an F test on the joint significance of the instruments in an equation predicting the potentially endogenous variable being significant at the 5 per cent level. An over identification test asserts that the instruments are valid. However, the Hausman-Wu test performed fails to reject the endogeneity of the log of expenditure per capita for all years 1990/91, 1995/6 and 2000/01 and I have therefore retained this variable in the wage equation.

expenditure elasticity less than 1. Gender bias in the allocation of good i can be detected through a straight forward F test checking whether the coefficients $\gamma_{ij} = \gamma_{ik}$ where j and k reflect boys and girls in the same age group.

The model is fitted, as is conventional, on the sample of all households with a child aged 0-19 years, regardless of whether the households incur a zero or positive budget share of a particular expenditure^{2,3}.

Kingdon (2005) argues that fitting one model to capture the school enrolment decision as well as the expenditure given enrolment decision may be one of the reasons as to why the conventional Engel curve analysis may fail to pick up a gender-bias in schooling expenditure in India as in Subramanian and Deaton (1991). She argues that a gender-bias in schooling can work through both the above channels and that averaging across them may lead to the conclusion of no gender bias if the bias works through just one channel. She therefore proposes a hurdle model that separates the households decision whether to incur any expenditure (i.e., enrol the child in school or not) from how much is actually spent given that it is decided to incur expenditure (i.e., expenditure given enrolment) In her sample, zero purchases equate to the child not being sent to school while in nearly 98% percent of those sent to

² Another issue to consider is whether the Engel curve is indeed linear, as assumed, or if it is non-linear, with households considering education a luxury at lower levels of income and a necessity at higher levels of income. In the regression analysis, this would be reflected by the coefficient on the log of household expenditure being positive for the lower income groups and negative for the higher income groups when the analysis accounts for household socio-economic status. This is not the case, as seen in results discussed later, education is a luxury across all income groups with an elasticity of 1.02, 0.85, 0.88 and 0.67 in 1990 across the poorest to the richest quintile; 1.09, 1.08, 1.07 and 1.06 in 1995 and an almost equal elasticity of 0.8 across all groups in 2000/01. Even simple descriptive statistics (unreported) of the budget share of education expenditure by expenditure quartile show a positive correlation. This is a sign that education is a luxury with the budget share devoted towards it increasing with income(expenditure). I therefore work with the assumption that the Engel curve is linear.

³ Note also that issues such as those raised in Jensen (2000) whether gender differences occur within or across households if girls tend to be in larger households due to parents 'son preferring, differential stopping behaviour' is not applicable in the Sri Lankan case. Household size tends to be the same with larger households not being those with more girls than boys, in Sri Lanka, with no evidence to back a claim of son/daughter preferring behaviour.

school (i.e., enrolled) incur some positive expenditure. She therefore uses the terms positive expenditure and enrolment interchangeably. Kingdon finds that the basic discriminatory mechanism is via differential enrolment rates for boys and girls.

A hurdle model, however, as discussed above, is inappropriate for the Sri Lankan case. This is because enrolment in school does not necessarily mean that a positive expenditure is incurred on education. Schooling in Sri Lanka in State schools (as do a majority of children in rural areas) has been largely 'free' since the 1930s. Text books and uniform material is provided free of charge by the government as well as a highly subsidised bus-pass scheme. In some schools free midday meals were provided. This means that around 25 per cent of our sample with a child enrolled in school do not report a positive education expenditure, during the period covered in the relevant question: The survey asks for expenditure on school books, stationery, fees, travelling etc. during the month before the survey and clothing expenditure (uniforms, tailoring) for the six months before the survey.

This means that in the Sri Lankan case, zero expenditure on education may reflect non-enrolment or that no expenditure is incurred even if the child is enrolled in school. What is more appropriate for the Sri Lankan case is a double-hurdle model explained below.

2.2. Double-hurdle Model

The double-hurdle model proposed originally by Cragg(1971), assumes that a household makes two independent decisions regarding the purchase of an item, with each decision explained by a different set of explanatory variables. A positive level of expenditure is observed only when two hurdles are crossed. In our case the first hurdle is whether or not to enrol a child in school (a binary decision that can be

estimated using a probit) and the second being incurring a positive expenditure given the decision to school (that can be modelled using a truncated regression). Each decision process is modelled using a different latent variable, where the latent variable may or may not be observed. For each household with a child in age group 1 to 19,

$$w_1^* = \mu a + \nu \quad \text{school enrolment decision} \quad (2a)$$

$$w_2^* = \delta b + \varepsilon \quad \text{expenditure decision} \quad (2b)$$

$$w = \delta b + \varepsilon \quad \text{if } w_{i1}^* > 0 \text{ and } w_{i2}^* > 0 \quad (3)$$

$$w = 0 \quad \text{otherwise}$$

where w_1^* is a latent variable describing whether the household has a child enrolled in school or not, w_2^* is a latent variable describing expenditure on education and w is the observed dependent variable (household budget share of education expenditure), a a vector explaining the school enrolment decision and b a vector explaining the school enrolment decision, ν and ε are the respective independent error terms distributed as $\nu \sim N(0,1)$ and $u \sim N(0,1)$.

The double-hurdle model with independent error terms can be estimated by the following log-likelihood function (Moffatt, 2005; Aristei et al, 2007):

$$LL = \sum_0 \ln \left[1 - \Phi(\mu a) \phi \left(\frac{\delta b}{\sigma} \right) \right] + \sum_+ \ln \left[\Phi(\mu a) \frac{1}{\sigma} \phi \left(\frac{w - \delta b}{\sigma} \right) \right] \quad (4)$$

The first term reflects the contribution of all the zero observations. It shows zeros occurring due to non-enrolment $1 - \Phi(\mu a_i)$, as well as those that might possibly occur

even in the second stage decision $\phi\left(\frac{\delta b}{\sigma}\right)$. The second term corresponds to the contribution of all the values indicating positive education expenditure⁴.

Under the assumption of independence between the two error terms, the log-likelihood function of the double-hurdle model is equivalent to the sum of the log-likelihoods of a probit model and a truncated regression model. Thus maximising a probit over all observations followed by a regression on just the non-zero observations on the non-zero observations is equivalent to maximising the log likelihood of the double-hurdle model (McDowell 2003), since the likelihood function is separable⁵.

To identify whether significant within household differentials exist in the allocation of education expenditure the coefficients on the demographic controls are compared using a F-test, as for the case of the conventional Engle Curve analysis, The most important thing about this analysis is that since the enrolment and expenditure decisions are now estimated separately, we can unpack the effects picked up by the Engel curve analysis and see whether the bias occurs in the enrolment or expenditure decision.

3. Data

⁴ Note that fitting a double-hurdle model is quite different to fitting a Tobit model on all the observations of w_i . Fitting a Tobit implies that the zero values come about as a corner solution, not as a part of a separate decision process that governs two different outcomes (decision to participate versus the level or intensity of participation). The double hurdle model also differs from a Heckman selection model because it allows for the possibility of zero observations in the second stage. Put differently, the log-likelihood function with independent error terms using the Heckman (1979) two-step method is

almost identical to (4) above, apart from $\phi\left(\frac{\delta b}{\sigma}\right)$ in the first term that reflects the contribution of zero values occurring in the second stage.

⁵ This paper adapts Julian Fennema's user-written code in STATA available at <http://www.sml.hw.ac.uk/sojaf/Stata/> for the estimation of the double-hurdle model.

The data comes from three cross-section Household Income and Expenditure Surveys (HIES) for 1990/91, 1995/6 and 2000/01 carried out by the Department of Census and Statistics (DCS), Sri Lanka. The DCS conducts the HIES once every 5 years. Data collection is done in twelve equal monthly rounds to capture seasonal variations in income and expenditure. A two stage stratified random sample design is used with urban, rural and estate sectors as the domains for stratification. The primary sampling unit is a census block and the secondary sampling unit are the housing units within the selected census blocks.

The overall quality of the HIES is quite good with high response rates and a coverage that is consistent with other independent surveys carried out on the same population. Let us discuss these two survey quality indicators separately. The 1990/91 and 1995/6 surveys have a 95 per cent response rate while the 2000/01 survey has a 91 per cent response rate. Non-response is due mainly to respondents being unable to complete the schedule, refusing to do so, being temporarily away or due to some unspecified 'other' reason. The incidence of non-response showed no significant seasonal variation--i.e., the amount of non-response was roughly the same during all 12 months of the year. These response rates compare quite favourably with those of several other countries. For example, six popular US government household surveys conducted during the 1990-1999 period indicate an initial response rate between 84 per cent to 95 per cent while UK's General household survey, which is also based on face-to-face interviews, records response rates averaging 80 per cent over the decade.

The coverage rate compares the estimated number of people from the HIES in a specific demographic group to the same estimate from an independent population total--usually Census estimates. For example, the under 17 population by race and

gender. The DCS carried out an all-island Census in 1981 and 2001. The HIES 2000/01 results are consistent with the Census 2001 results in terms of demographic composition with the main limitation that the HIES excludes several areas in the North and East of the country, due data collection problems in these war-torn areas. The broad composition of the HIES is also similar to that of the Consumer Finance Survey carried out by the Central Bank of Sri Lanka during the 1990s. The two surveys gather similar information and have roughly the same demographic and ethnic composition. No groups are noticeably under or over-represented compared to the Consumer Finance Survey that also excludes areas in the North and East in their work.

Table 2 contains summary statistics for the variables used to estimate the education Engel curves and double-hurdle model for rural areas. Education expenditure as a share of total expenditure for all households with at least one child aged 0 to 19 is not large at around 2 per cent in the 1990s, with the share growing, albeit slightly, over the decade. In rupee terms, the average rural household education expenditure was around Rs. 50 in 1990 and Rs. 324 a month by the year 2000. The rather small share of education expenditure is unsurprising because education is basically 'free' in Sri Lanka. However, there is still a cost and specially an opportunity cost to education. If a child attends a state school (as do a majority of children in rural areas), text books and uniform material as well as a mid-day meal in the case of a few schools was provided during the decade beginning in the 1990s. However costs of private tuition fees, exercise books, travelling, equipment and additional costs of uniforms (such as shoes, socks, etc.), a nominal school fee etc. was still to be borne by households. The various categories of education expenditure have been aggregated and it is assumed that reparability is not an issue since the

composition of various expenditure categories over time has remained roughly the same.

4. Results

4.1. Engel Curve estimations

Equation 1 is used to run OLS regressions for the budget share of education in rural areas for 1990/91, 1995/6 and 2000/01 (Table 3), for households with at least one child aged 0 to 19. F-tests for the equality of γ coefficients are presented at the bottom of the table.

The goodness of fit of the linear Engel curve is around 0.13 for all 3 years. The coefficients on log expenditure are positive and close to unity in some cases showing that education is treated as a luxury. The elasticity is highest in 1995 at 1.1 when the country's poverty rates were the highest during the decade at 33 per cent (see DCS 2005). The lower elasticity of 0.87 for 2000 suggests that education has become to be treated as less of a luxury towards the end of the decade with the country's economy picking up and poverty rates dropping to 25 per cent in rural areas⁶.

The coefficient on household size is significant and positive for all three years. This matches theoretical arguments that suggest that larger households will be better off due to economies of scale that accrue from shared public goods, at any given level of per capita resources. However, note also that household size may be endogenous because parents with a higher taste for schooling may choose to have smaller families and a higher education budget share. Unfortunately we do not have data on

⁶ Unfortunately no previous study exists for Sri Lanka pertaining to a similar analysis to compare these education expenditure elasticities. Kingdon(2005)'s estimates for India for 1994 show that for most States the elasticity is close to or above unity. Subramanian (1995) reports elasticities ranging between 1.3 and 2.75 for some of India's poorest States for the mid 1980s. In comparison to these figures, Sri Lanka's rural sector seems to treat education as less of a luxury than most Indian States.

households across time in order to capture household level fixed effects and so address the issue of the potential endogeneity of household size.

The education of the household head is also positively significant for all three years. This indicates a higher demand for schooling among households with more educated heads. Ethnic group is also significant with being a Tamil household affecting the budget share negatively.

The most important result for this paper is to note the coefficients against the age-cohort variables and the F tests at the bottom of the table that compare these coefficients between boys and girls of the same age group. Compared to the omitted category of females aged 30 to 54, children in age categories between 5 and 19 exert a significant positive impact on a household's budget share on education. This is not surprising. What is particularly interesting to note, however, is that the coefficients are higher for girls than for boys for most age groups especially in 1990 and 1995. In 1990, for instance, if a child had been a girl rather than a boy in the 14-16 age group within the same household, 0.9 percent more would have been spent on her towards education, once controlled for other factors such as household size, etc. The corresponding figure for 1995 was 1.1 per cent. Similarly, in 1990, adding an extra girl in the 17-19 age group increased the household education budget by 1.8 per cent more than adding a boy of the same age group. The F-tests at the bottom of the table summarise these results by highlighting the significant biases. In both 1990/91 and 1995/6, statistically significant biases favouring girls are indicated for the age cohorts 5 to 9, 14-16 and 17 to 19. The Engel curve estimates do not indicate any significant biases for the year 2000/01.

3.2. Double-hurdle Model

The results for the double-hurdle model estimations are presented in Table 4. For each year, a probit looking at participation (i.e., household has a child enrolled in school) and a truncated regression based on the level of expenditure is estimated. At the bottom of the table, F-statistics and their significance are reported for the comparison of coefficients between boys and girls in the various age categories.

For 1990/91, the F-statistics for the participation equation shows that significant differences between girls and boys are found in terms of school enrolment for the 17-19 category, favouring girls. This matches the school enrolment statistics presented in Table 1, where it was clear that 9 per cent more girls in that age group are enrolled at school than boys⁷. The F-tests for the expenditure results show that for age-groups 5-9, 14-16 and 17-19, adding an extra girl rather than a boy increases household education expenses significantly. Thus the results in the Engel curve analysis in the previous section seem to be driven mainly by education expenditure once children are enrolled in school. The same holds for the year 1995/6. For the year 2000, the double-hurdle model picks up biases in the enrolment decision for age category 17-19 and expenditure decision for age category 14-16, both favouring girls. This was not picked up in the Engel curve analysis. This is most probably because the two channels through which ‘discrimination’ works (enrolment and expenditure given enrolment) cancel each other out, when aggregated under the conventional Engel curve method or are weakly significant on their own (as in the case of the girl

⁷ Table 1 was based on individual-level data. The estimations in Table 4 are based on household level data. The F-statistics tell us that adding an extra girl in a household in the age group 17-19 means that there is a significantly higher chance that she will be enrolled in school than adding an extra boy. The household-level estimations do not show a significant difference in the enrolment rates of boys and girls in the 14-16 age category. This may be because the control variables in the regression explain all of this difference or because household level aggregation of enrolment (and indeed education expenditure) under-estimates the true bias. Unfortunately we do not have individual level data on education expenditure to test this hypothesis.

bias in the expenditure for 14-16 year olds in 2000 that is exactly significant at the 10 per cent level), which cancels out when combined with the enrolment decision.

Robustness Checks

As a robustness check, I re-run the estimations for households with only boys or girls (with the observations pooled) to see if the results are similar to that of pooling all households together. Around 20 per cent of the households have only boys and 20 per cent only girls. Descriptive statistics for these two groups show that their mean values for budget share of education expenditure, household expenditure, household size and siblings are statistically the same, as the t-value testing the means to be different is rejected in all cases. The regression analysis for this sub-sample of households indicate results that match that of all households. In other words, age categories 5-9, 14-16 and 17-19 still indicate a significant girl bias, as the F-tests reveal (results unreported).

Do the baseline results of a significant bias favouring girls in 1990/1 and 1995/6 result hold across expenditure groups or is it something that it driven by the poor (non poor)? For all three years, the budget share of education increases as the household group becomes richer. For example, in 1990, the poorest quartile spend 1.47 of their budget share on education, the second poorest 1.5, the third poorest 1.5 and the richest 1.9. In 1995 the corresponding shares are 1.34, 1.43, 1.66 and 2.25. In 2000 it is 1.55, 1.81, 1.93 and 2.22. Household size decreases by expenditure group. None of the other regressors vary notably between income quartiles. To see if baseline results hold across expenditure groups, I interact the age cohort variable with a dummy indicating whether the household belongs to the poorest quartile and re-regress the education Engel curve. I then interact the age cohort variable with a dummy

indicating whether the household belongs to the richest quartile. The results of the F-test for this exercise (unreported) match those of the baseline case. The bias in the 5-9 category is indicated in the two middle-income groups. In 1995 girls in the 14 to 16 age group among the poorest quartile were favoured at a 1 per cent level of significance while it was around 12 per cent for the richest quartile. The age group where the bias arises varies with the income group but what is common is that the biases always favour the girls.

Let us now discuss the findings regarding a girl-bias, for the 'baseline' results in more detail.

Each of the age cohorts 5 to 9, 14 to 16 and 17-19 run up to and culminate at important State examinations. The age cohort 5 to 9 culminates in the Grade 4 scholarship exams, a competitive national exam, the results of which can be used to gain entrance to better-performing secondary schools. The age group 14-16 is the senior secondary level ending with the Ordinary Level (O/L) examinations. It is an important educational milestone that completes secondary education and is needed to gain entrance to better performing schools and to qualify for high-school education. High school education is captured by age group 17-19 that culminates with students reading for the Advanced Level (A/L) examinations qualifying them to enter university⁸.

In order to support the argument that the differential expenditure favouring girls is observed mainly in the run-up to key state exams, I split the age groups much more narrowly and use age in years rather than age group (i.e., age 5, 6, 7 and so on until 19, together with other groups such as 0 to 4, 20-29, 30-54 and over 55 with females aged 30 to 54 omitted) and re-estimate the baseline regression. I report just the F-tests

⁸ Note that there is no grade repetition in Sri Lankan state schools.

of the comparison of coefficients for boys and girls in Table 5. The results provide strong support for the baseline results. In 1990/91 a difference in expenditure (driven mainly by expenditure once enrolled) is observed at ages 8, 9, 18 and 19 (with the difference in age groups 15 and 16 high and significant at the 13 per cent), in 1995/6 at ages 8, 9 and 16 and in 2000/01 at ages 15 and 18. In 2000/01, the participation equation also indicates a significant difference favouring girls at age 17, implying that boys tend to drop out more from school at this age, just after their O/Ls, instead of choosing to read for the A/Ls⁹. No differences are seen at the other ages.

Thus the results show that according the Engel curve methodology, rural households allocate the extra rupee towards daughters at age cohort 5-9 (primary school, especially as the children are aged 8 and 9), 14-16 (Ordinary levels) and 17-19 (A/Ls) in 1990/91 and 1995/6. A bias is also seen during in the 14-16 and 17-18 category for 2000/01. Higher investment at the primary school level may mean daughters can gain entry to better performing secondary schools. Higher investment at the senior secondary and high-school level will, most probably, bring about better performance at the State examinations.

So why does an extra girl increase the household education budget more than a boy does?

5. Explaining the bias

Is it simply noise in the data?

It could be noise in the data that gives us these results. We have, however, noted in the discussion in section 2.2 that the overall quality of the data is quite good as judged

⁹ I have chosen to use results for groupings 5-9, 10-13, 14-16 and 17-19 as the baseline rather than the using age at levels because I think they better capture the 'run-up' to state exams, especially at more senior levels. For example, the O/L curriculum for the exams at age 16 begin from age 14/15, while the A/L curriculum spans roughly 2.5 years from ages 17 to 19.

by the coverage rate and non-response rate and the care with which it has been gathered and cleaned. Moreover, the results are not sensitive to the way we break up the age-cohorts. The results remain robust even when the age-composition of children is changed in order to make sure that the regression outcomes are not due to simply to the way I split the age categories up. For example, the age categories were changed to (a) 5-9, 10-14 and 15-19 and (b) 5-14, 15-19 and yet again as (c) 5-9, 10-16 and 17-19. For both 1990/91 and 1995/96 it was (a) 5-9, 15-19 (c) 5-9, 10-16 and 17/19 with the most significant ages being 8/9, 15/16 and 18/19. Thus it is not noisy data that explains our results. It seems that the addition of a girl child in certain age-cohorts genuinely increased the household education budget share more than adding a boy of the same age group.

Is it just a matter of taste?

In order to get a better understanding of what component of education expenditure actually causes the bias, I disaggregate education expenditure into expenditure on books, fees, travelling and other expenses, and calculate the budget share of each of these components. I then replace the left-hand-side of equation (1), i.e., w_i with each of these budget shares separately and re-estimate (1) and carry the F test to check whether the gender bias is more obvious in any one component of education expenditure. The unambiguous result is that in it is expenditure towards school books that cause a significant bias. Expenditure in other categories such as fees are often higher for girls than boys but is not statistically significant. Is this indicative that there is no ‘favouring’ that happens but that it is simply a matter of ‘taste’ with girls preferring more books and input to education? If this is the case, it seems strange that this ‘taste’ is not something found in every age group and that this taste is not

consistent across the years. Further, a cohort-based analysis (i.e., splitting the age group as 0-5, 5-10, 10-15 and 15-20) shows that the ‘taste’ factor does not move along with the cohort—the significant biases appear only when girls reach a particular age category.

The above disaggregated analysis also shows that the girl bias is not due to girls being more ‘expensive’ in that parents are more keen to send them to school using safer modes of transport or that their uniforms cost more. It is *school books* (supplementary text books, equipment and exercise books) that cause the most significant difference between boys and girls even though expenditure other categories also differ between boys and girls, favouring girls.

Is there a ‘cultural’ preference for girls?

It is possible that the bias we notice favouring girls is a reflection of a wider ‘cultural’ attitude or preference for girls. This argument, however, can easily be refuted statistically but looking at similar analysis conducted with reference health expenditure or education expenditure in urban areas. The girl bias does not occur in Engel curves estimated for food and health shares separately (unreported). The bias exists only in terms of education expenditure. Notice also, that this expenditure bias is found only in the rural sector data. We do not find such biases for urban areas. So it seems that in rural areas, there is indeed a bias favouring girls in terms of education expenditure.

Education is a substitute for the ‘dowry’ at marriage, so parents chose to invest in education rather than the dowry

Perhaps the girl-bias arises because education and subsequent employment are used by parents as a substitute for the dowry at the time of a daughter's marriage. Dowries at marriage is still considered important by some Sri Lankan parents and youth- roughly 30 percent according to a survey of Sri Lankan youth by Malhotra and Tsui (1996). The trend of education and consequent earning potential being a substitute for the dowry is seen in South India as well, where higher education is often a dowry substitute and a means of improving a women's value in the marriage market. In some cases, dowries are often collected by the girl herself through employment as the age at marriage is pushed forwards¹⁰. However, in Sri Lanka, the age of marriage being pushed forwards is not indicative that it is purely for the sake of collecting a dowry. The higher age at first marriage in Sri Lanka seems to be driven by poor, rather than improved economic and political conditions (Caldwell et al 1989). For most females, employment before marriage is more for the income it provides, than purely for the dowry they can collect through 'saving' or the career they can build as a substitute for the dowry.

Female earnings, or rather, the earnings of both the man and the woman in the family has grown important over time due to slow economic growth and political unrest and thus, economic returns to education play an important role when a household decides to invest in a girl-child, rather than dowry or other 'cultural' or normative concerns. A majority of Sri Lankan women participate actively in the labour market and are not considered a burden. In 1995 and 2000, labour force

¹⁰ The median age at first marriage in Sri Lanka, for women in rural areas, has risen from around 20 in the 1960s to 25-27 in 2000 in rural areas (DHS 2000; Dissanayake 2000). Sri Lanka has historically had ages at first marriage that are higher than those of the rest of the region. For example, while her counterparts married on reaching puberty at ages as young as 13 or 14 in Northern India, Pakistan and Bangladesh, the Sri Lankan female has been marrying at around 18 or 19 even during as far back as the 1940s. During the last couple of decades, these median ages have increased further, matching other Asian countries such as Taiwan or Thailand, as have educational levels and social welfare for women (De Silva 1990, Thornton and Lin 1994).

participation of women was around 40 per cent (with this being higher for younger women) while for men it was around 80 per cent. Moreover, unlike in neighbouring countries such as India or Bangladesh, a woman often does not move in with her in-laws after marriage nor do her earnings after marriage accrue mainly to the family of her in-laws (Malhotra and Tsui 1996). This means that her natal family can benefit as much from a daughters economic returns after marriage as they could a son's. An implication of this is that parents do not necessarily choose to invest more in sons under the implicit assumption that it will be sons who will look after them financially in old age rather than daughters.

The reason for the observed bias, therefore, cannot be explained convincingly by noise in the data, tastes, cultural attitudes or preferences or arguments regarding the dowry. We shall now move to looking at the labour market for possible explanations.

Are households responding to labour market returns being higher for women?

The girl-bias may be due to returns to education being higher for girls than boys backed by the fact that the opportunity cost of leaving school was higher for a girl than a boy, given higher youth unemployment rates among girls with longer waiting periods (Salih 2001).

Several studies done using Sri Lankan data for 1980/81, 1985 and 1990 (Gutkind (1984), Glewwe (1985), Sahn and Alderman (1988), Aturupane (1993), Gunawardane (2002)) suggest that returns to female education is higher than that for males those completing secondary education (Ordinary Levels) or higher. In rural areas, returns are particularly high for completing Ordinary Level and Advanced

Level examinations¹¹. This means that if a girl stays on at school for an extra year at the secondary level or stays on to complete her education at the high-school level (A/Ls), the extra amount that she earns is higher than that for a boy.

Did higher returns to education for girls persist into the 1990s? If it did, the household level girl-bias may be seen to be an efficient allocation of resources and the decision itself may probably be a response to market outcomes in an economy, where females are economically active, assuming that a household's decision to invest in their children's education is guided by returns to education for adults at any given time.

In order to verify what happens in the 1990s, I estimate Mincerian wage functions for males and females by level of education for 1995 and 2000. I estimate returns to schooling in two ways. First using earnings as the dependent variable (i.e., including both wage earners and non-wage earners) and correcting the estimates for sample selection of females and secondly, a fixed effects estimation by fine-tuning the Mincerian function to account for household level unobserved heterogeneity following Behrman and Wolfe (1984), Khandker (1990), Behrman and Deolalikar (1993, 1995), Gunawardane (2002)).

I correct for sample selection indirectly through correcting for unobserved heterogeneity at the household level by using the fixed effects estimation to a cluster sample, where the well-defined cluster in this case is the household in each cross-section data set. I use deviations from household means for all households where there are two or more males (females) who are wage earners, to investigate whether fixed effects are important, assuming that differences are across households (if they

¹¹ Most of these works use the Mincerian earnings or wage function and some (i.e., Aturupane, Sahn and Alderman) correct the estimates for selectivity using the Heckman procedure to find that returns for females for secondary and A/L qualifications are higher than those for males and the direction of the effects remain unchanged when corrected for sample selection. The trends remain the same whether one uses earnings or wages as the dependent variable.

exist). If there are such unobservable fixed effects, and they are significant, the constant term of the fixed effects regression would be significant and the OLS estimations would be biased. Controlling for 'fixed effects' at the household level controls for the sample selection problem as observed by Pitt and Rosenzweig (1990), Heckman and McCurdy (1980). This is because the household-level observed variables used to control for selectivity in the paid labour force are those such as wealth, unearned income, assets etc., and controlling for household fixed effects should control for the selectivity in the paid labour force (Behrman and Deolalikar 1995:106). Apart from this, the fixed effects method addresses the issue of omitted variable bias arising from unobserved heterogeneity at the household and community level. However, this procedure only addresses the issue of heterogeneity bias arising from unmeasured attributes that are common to individuals in the same household, since the fixed effects estimation is limited to households that have more than one male or female earning.

If the unobserved household and community effects are random instead of being fixed, they would bias the error term and invalidate standard statistical tests. In order to test for this possibility, I estimate a random effects model using the same subsample. I then use the Hausman test to compare between the fixed or random effects estimations.

The data sources for both the earnings and wage functions are the same as for the previous analysis on intrahousehold allocation.

Due to space considerations, I have not reported the regression results, but calculated the returns to education on schooling for both the earnings and wage estimations. The trends in private rates of return to education estimated using both the above methods (returns to education with earnings as the dependent variable

corrected for sample selection and returns to education with wages as the dependent variable, corrected for unobserved heterogeneity at the household level using fixed effects) are very similar.

Table 6 reports fixed effects estimation based returns to education calculated, using the reported results for rural areas and unreported results for all areas¹². The table shows the extra return to an extra level of schooling. The results show that women have an unambiguously higher return to secondary and O/L education than do men in terms of both rural areas and all-sectors. This matches trends observed in the 1980s and 1990, reported by Aturupane, Gunawardane and other studies. The trends in all the estimations, regardless of econometric refinement and differences in the dependent variables, are broadly the same. Note also that the higher returns seem to be concentrated at the lower levels of education especially in 1995. Overall, the size of the return has fallen over the two decades beginning in 1981. This is partly because the earlier estimates are 'inflated' in not accounting for household level heterogeneity, and partly, because true returns have indeed fallen, at least from 1985 onwards when the estimates discussed become more comparable in terms of methodology adopted.

The calculations in this paper show, in addition, that the clear advantage females had in terms of returns in the 1980s and even 1995, diminishes slightly

¹² For the wage-based estimations the Breusch -Pagan Lagrange Multiplier test on the random effects model in all the cases show that unobserved household effects are indeed significant. The Hausman test, used to compare the fixed and random effects models, rejects the null hypothesis that the difference in coefficients is not systematic in all cases at a 1 per cent level of significance apart from females in 1995, where it is rejected at the 10 per cent level of significance. Assuming that the specification of the model is correct, I interpret this result to mean that unobserved household effects and the implicit correction for selection is indeed important, and that the fixed effects estimations are superior to the OLS or random effects estimations. I therefore use only the fixed effects-based estimations for the wage-based regressions. Note, however, that the earnings-based results corrected for sample-selection also reveal the same type of trend in terms of returns to education.

especially for rural areas by the year 2000. Secondary school educated women still have an advantage but the male-female gap has fallen.^{13,14}.

Thus in spite of returns for females continuing to be higher than for males, the analysis as it stands does not seem to support the hypothesis that the household level female bias of education expenditure is clearly a response higher returns. First, higher returns are observed in both rural and urban areas but gender based education expenditure differences are observed only for rural areas. Moreover, the higher returns seem to be strongest at lower levels of schooling. As such, the challenge of explaining gender differences in schooling in rural areas remains.

Is it actually discrimination in the labour market?

The data shows that for all three years 1990/91, 1995/6 and 2000/01, women are significantly more educated than men, for every occupation category apart from for clerical work in government service (see Table 7, where summary information for the year 2000 are presented). This can mean two things: First, that women are actually *discriminated* in the labour market with employers expecting women to be more qualified than men for the *same* job. Alternatively, that females have a higher taste for education and are therefore on average more educated than males, and this is reflected in the labour market—the latter argument, however leaves a puzzle behind: if women were more educated than men due simply to personal taste, then employers should prefer to employ more women rather than men assuming that education level is indicative of ability and productivity, and that both male and female education is

¹³ It is beyond the scope of this paper to discuss why the returns have changed as they have between 1995 and 2000, and indeed why returns overall have been going down since 1981. Suffice it to note that there have been many exogenous shocks the economy faced during the post-liberalisation period.

¹⁴ It is possible that returns to schooling and schooling attainment are determined simultaneously, which is not addressed in this paper.

valued equally in the labour market. The fact that more females tend to be unemployed than men—specially amongst the youth—along with the fact that more men are employed than women (as a proportion of those of each gender participating in the labour market) seems to suggest just the opposite. Being male seems to have some extra advantage attached to it than being female. Both these contentions need to be investigated more fully.

6. Conclusion

This paper showed that contrary to most developing countries, there is a bias favouring girls in rural Sri Lanka in the allocation of education expenditure within the household. This was seen using demand analysis assuming a linear Engel curve as well as a double-hurdle model that unpacks the enrolment and expenditure given enrolment decisions allowing for the fact that a household can report zero expenditure even when a child is enrolled in school due mostly to infrequency of purchase of items under the education category.

Sri Lankan rural households seem to be allocating more educational resources towards girls in 1990/91 and 1995/6 for age groups 5-9, 14-16 and 17-19 and in 2000 for age groups 14-16 and 17-19. This is driven mainly by the expenditure on girls versus boys given a household decides to spend on education. In 1990/91 and 2000/01 the double-hurdle model also showed that there was a bias towards enrolling more girls in school to read for their Advanced Levels (age group 17-19).

The age groups where the biases are obvious both in the enrolment and expenditure equations are important: The significant differentials in age group 5 to 9 happen at ages 8 and 9, corresponding to the run up to year 5 scholarship exams where children can gain entry into better performing state schools. The 14-16 age

group captures the culmination of secondary education with the reading for the Ordinary Level examination and gaining entrance to read particular subjects for the Advanced Levels. The 17-19 age group captures the run up to and culmination in Advanced Level examination, by far the most competitive exam in Sri Lanka that allows a student to gain university entrance. The fact that biases are all observed for these age groups and not for the 10 to 13 age group for instance or for the ages that fall outside 8/9, 14/15/16 or 18/19 in the robustness checks that split-up the age categories in different ways (as well as various us checks for robustness) show that the observed biases in expenditure are not random. Rural households seems to be indicating a deliberate preference for spending more on education when an extra girl is added to the household rather than a boy close to important national-level examinations.

The paper then went onto explore various non-labour market and labour market oriented explanations that are possible for this observed bias. Cultural biases favouring girls were ruled out as no bias was found in terms of food or health share expenditure or education expenditure for non-rural areas. Arguments based on education being a possible substitute for the dowry were also considered and argued to be weak. Several other taste and attitude based explanations were also looked at. I also looked at labour market returns to education between males and females. Returns continue to be higher for females in the Sri Lankan labour market continuing trends in the early 1980s, even though the gap between males and females is reducing in recent years. The higher returns would make it more efficient for a household to invest in an extra year of education for a girl rather than a boy, at the margin. However, we cannot conclude that this explains household-level gender differences in education expenditure because higher returns are observed both in rural and urban

areas, strongest at lower levels of education. Another possible explanation is that there is an insidious form of discrimination of females going on in the Sri Lankan labour market, with females having to be more educated than their male counterparts in order to compete for the same job. All these explanations contribute perhaps in varying degrees to the female-bias in household expenditure observed in rural Sri Lanka and need further investigation.

References

Aristei, D, F. Perali and L. Pieroni. 2008 “Cohort, age and time effects in alcohol consumption by Italian households: A double-hurdle approach.” *Empirical Economics*, 35, no.1:29-61.

Aturupane, H. 1993. “Is Education Beneficial? A Microeconomic Analysis of the Impact of Education on Economic Welfare in a Developing Country.” Ph.D. diss., Cambridge University, UK.

Behrman, J, and A. Deolalikar. 1993. “Unobserved Household and Community Heterogeneity and the Labour Market Impact of Schooling.” *Economic Development and Cultural Change* 41(April):461-488.

Behrman, J, and A. Deolalikar. 1995. “Are there Differential Returns to Schooling by Gender? The Case of Indonesian Labour Markets.” *Oxford Bulletin of Economics and Statistic*. 57:97-117.

Behrman, J. R, and B.L.Wolfe.1984. “The Socioeconomic Impact of Schooling in a Developing Country.” *The Review of Economics and Statistics* 66, no.2:296-303.

Burgess, R, and J. Zhuang.2000. “Modernisation and Son Preference.” *STICERD - Development Economics Papers* 29, Suntory and Toyota International Centres for Economics and Related Disciplines, London School of Economics.

Caldwell, J. 1989. “Is marriage delay a response to pressures for fertility decline? The case of Sri Lanka.” *Journal of Marriage and Family* 51:327--335.

Cragg, J. 1971. "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods." *Econometrica* 39, no.5:829-844

Deaton, A. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Johns Hopkins Press for the World Bank.

Deaton, A, and J. Muellbauer. 1980. *Economics and Consumer Behavior*. Cambridge University Press.

Department of Census and Statistics (DCS) .2005. *Statistical Abstract 2005*, Department of Census and Statistics, Colombo, Sri Lanka.

De Silva, W. I. 1990. "Age at marriage in Sri Lanka: Stabilizing or declining?" *Journal of Biosocial Science* 22, no.4:395-404

DHS. 2000. *Demographic and Health Survey Sri Lanka 2000, Final Report*. Department of Census and Statistics, Sri Lanka.

Dissanayake, L. 2000. "Factors influencing stabilization of women's age at marriage." In *Demography of Sri Lanka*. Department of Demography, University of Colombo: 45-58

Glewwe, P. 1985. *An Analysis of income distribution and labour markets in Sri Lanka*, Ph.D. diss., Stanford University.

Gunawardane, D. 2002. "Reducing the gender wage gap in Sri Lanka: Is education enough?" *Sri Lanka Economic Journal* 3, no. 2:57-103

Gutkind, E. 1984 *Earnings Functions and Returns to Education in Sri Lanka*. International Labour Organization, Geneva.

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

Heckman, J. J, and T. E. McCurdy. 1980. "A Life Cycle Model of Female Labour Supply." *Review of Economic Studies* 47 , vol.1: 47-74

Jensen, R. 2000. "Agricultural Volatility and Investments in Children." *American Economic Review* 90, no. 2: 399-404

Khandker, S.R. 1990. "Labour market participation, returns to education and male-female wage differences in Peru." *Policy Research Working Paper 461.*, World Bank, Washington D.C.

Kingdon, G.G. 2005. "Where has all the bias gone?: Detecting gender-bias in the household allocation of educational expenditure." *Economic Development and Cultural Change* 53, no. 2:409–451

Malhotra, A, and A. O. Tsui. 1996. "Marriage Timing in Sri Lanka: The Role of Modern Norms and Ideas." *Journal of Marriage and Family* 58, no. 2: 476-490

McDowell, A. 2003. "From the help desk:hurdle models." *The Stata Journal* 3, no. 2:178-184

Moffat, P. 2005. "Hurdle models on loan default." *Journal of the Operational Research Society* 56:1063-1071

Pitt, M, and M.R. Rosenzweig. 1990. "Estimating the behavioural consequences of health in a family context: the intrafamily incidence of infant illness in Indonesia." *International Economic Review* 31:969-89

Rudd. J. B. 1993. "Boy-girl discrimination in Taiwan:evidence from expenditure data." Research program in development studies, Princeton University, processed.

Sahn, D. E, and H. Alderman. 1988. "The Effects of Human Capital on Wages, and the Determinants of Labor Supply in a Developing Country." *Journal of Development Economics* 29, no.2:157-183

Salih, R. 2001. *Youth employment in Sri Lanka: A review of the current labour market situation, policy and programs*, The ILO/Japan Tripartite Regional Meeting on Youth Employment in Asia and the Pacific.

Subramanian, S. 1995. "The demand for food and calories." *Journal of Political Economy* 104, no.4: 133-62

Subramanian, S, and A. S. Deaton.1991. "Gender effects in Indian consumption patterns." *Sarveksnana* 14:1-12

Thornton, A. and Lin. H. 1994. *Social Change and the Family in Taiwan*. University of Chicago Press.

Table 1: School Enrolment Rates by Gender in Rural Areas¹

	1990/91	1995/96	2000/01
Males 5-9 years	n/a	n/a	n/a
Females 5-9 years	n/a	n/a	n/a
GAP			
Males 10-13 years	94.6	95.7	94.7
Females 10-13 years	95.2	96.8	95.6
GAP	(0.6)	(1.1)	(0.9)
Males 14-16 years	78.3	84.8	86.1
Females 14-16 years	83.2	87.4	88.3
GAP	(4.9)***	(2.6)**	(1.85)*
Males 17-19 years	42.1	46.5	46.3
Females 17-19 years (n=)	51.1	48.9	49.0
GAP	(9.0)***	(2.4)	(2.7)

Note: Boys (girls) enrolled as a percentage of all boys (girls) in that age cohort. For age group 5-9, enrolment information is not available. Enrolment information is extracted from the question that asks each member of the household what their usual activity is. Unfortunately the question is asked only for persons 10 years and over. Positive household education expenditure is not a good proxy for enrolment as children can be enrolled in school but the household may not have incurred any positive expenditure on education during the reference period (roughly 20 per cent of the sample for children over 10). Other sources indicate, however, that primary school enrolment rates in Sri Lanka are quite high at over 97% for both boys and girls during 1990-2000 (UNICEF).

Source: Own calculations using HIES 1990/91, 1995/6 and 2000/01.

Table 2: Summary Statistics for Budget Share of Education and Contributory variables

Variable name	Description of variable	1990/91	1995/96	2000/01
Educ_share	Budget share of education: household education expenditure/Total expenditure on food and non-food items X 100	1.84 (0.04)	2.11 (0.05)	2.56 (0.08)
Exp_TOT	Total expenditure (in rupees) on food and non-food items per member of the household	791.03 (9.96)	1303.36 (22.23)	2808.4 (49.12)
hsize	household size	5.41	4.96	4.59
M0_4	males aged 0 to 4 as a proportion of all household members	.04	.04	.04
f0_4	females aged 0 to 4 as a proportion of all household members	.04	.04	.03
M5_9	males aged 5 to 9 as a proportion of all household members	.05	.03	.04
f5_9	females aged 5 to 9 as a proportion of all household members	.05	.04	.03
M10_13	males aged 10 to 13 as a proportion of all household members	.05	.04	.03
f10_13	females aged 10-13 as a proportion of all household members	.05	.04	.03
M14_16	males aged 14-16 as a proportion of all household members	.03	.03	.03
f14_16	females aged 14-16 as a proportion of all household members	.03	.03	.03
M17_19	males aged 17-19 as a proportion of all household members	.03	.03	.03
f17_19	females aged 17-19 as a proportion of all household members	.03	.03	.03
M20_29	males aged 20-29 as a proportion of all household members	.06	.08	.07
f20_29	females aged 20-29 as a proportion of all household members	.07	.08	.08
M30_54	males aged 30-54 as a proportion of all household members	.14	.16	.16
f30_54**	females aged 30-54 as a proportion of all household members	.14	.16	.17
M55	males above age 55 as a proportion of all excluding boarders and lodges)	.05	.05	.06
f55	females above age 55 as a proportion of all household members (excluding boarders and lodges)	.04	.04	.04
hhead_educ	Education level (in years) of the household head	6.15	7.20	8.72
Tamil	Ethnicity of the household head is Tamil	.01	.04	.02
Minority	Ethnicity of the household head is Muslim, burgers or other	.04	.04	.09
Dependents (double hurdle model)	Members in the family who are unable to work, unemployed or engaged in household work and stay at home as a proportion of all household	0.25	0.24	0.25
observations		9103	11597	9689

Note: ** Reference group for education Engel-curve estimations. Standard errors in parenthesis for continuous variables.

Table 3: Engel curves for education expenditure

Dependent Variable: Budget share of education

	1990/91	1995/6	2000/01
Log expenditure	0.988 (13.30)**	1.103 (16.07)**	0.879 (13.75)**
Log household size	0.722 (7.52)**	0.765 (8.28)**	1.094 (9.02)**
M0_4	-2.554 (5.93)**	-2.943 (7.16)**	-3.002 (6.79)**
f0_4	-2.384 (5.45)**	-2.705 (6.63)**	-2.838 (6.36)**
M5_9	0.312 (0.73)	0.453 (1.13)	1.194 (2.76)**
f5_9	1.126 (2.64)**	1.203 (2.95)**	1.029 (2.35)*
M10_13	1.937 (4.33)**	2.075 (4.98)**	3.092 (6.85)**
f10_13	2.331 (5.22)**	2.633 (6.14)**	2.485 (5.55)**
M14_16	2.420 (4.95)**	2.535 (5.60)**	3.168 (6.42)**
F14_16	3.317 (6.51)**	3.639 (8.11)**	4.085 (8.00)**
M17_19	0.975 (1.95)	1.110 (2.42)*	3.036 (6.26)**
F17_19	2.717 (5.56)**	1.907 (4.21)**	2.561 (5.25)**
M20_29	-1.720 (4.31)**	-1.991 (5.51)**	-3.200 (7.74)**
F20_29	-0.457 (1.35)	-0.356 (1.12)	-1.225 (3.42)**
M30_54	-1.304 (3.05)**	-1.486 (3.92)**	-2.236 (5.26)**
M55	-0.994 (2.19)*	-1.535 (3.66)**	-2.011 (3.89)**
F55	-0.808 (2.12)*	-0.464 (1.31)	-1.855 (4.31)**
hhead_educ	0.103 (11.27)**	0.123 (13.85)**	0.038 (4.25)**
Tamil	-0.603 (2.18)*	-0.634 (4.20)**	-0.495 (2.18)*
Minority	-0.310 (1.96)*	-0.252 (1.54)	-0.186 (1.16)
Constant	-6.408 (10.23)**	-7.111 (11.41)**	-6.209 (10.75)**
Observations	9103	11597	9689
r-squared	0.13	0.13	0.12
F-Tests: 5-9	4.23(0.039)**	3.52(0.060)*	0.21(0.731)
10-13	0.88(0.347)	1.80(0.179)	1.45(0.228)
14-16	3.14(0.076)**	5.70(0.017)	2.41(0.121)
17-19	12.1(0.000)***	2.72(0.098)	0.72(0.396)

Note: F-Tests refer to a testing for the equality of coefficients of, for example, m5_9 and f5_9, with the corresponding F-statistic and significance reported within brackets. Absolute value of t statistics in parentheses; District dummies included in regressions but not reported. * Significant at 5%; ** significant at 1%.

Table 4: Double hurdle model estimations for education expenditure

Dependent Variable: Budget share of education

	1990/91		1995/6		2000/01	
	Participation Equation	Expenditure Equation	Participation Equation	Expenditure Equation	Participation Equation	Expenditure Equation
Log expenditure	0.272	1.635	0.287	2.062	0.376	1.362
	(2.03)*	(13.83)**	(3.02)**	(16.85)**	(5.18)**	(12.87)**
Log house-hold size	1.062	1.197	0.761	1.422	0.865	1.828
	(5.60)**	(6.65)**	(5.83)**	(7.35)**	(5.97)**	(8.36)**
M0_4	-2.109	-4.673	-2.560	-5.446	-0.520	-6.302
	(3.64)**	(5.57)**	(4.97)**	(6.02)**	(1.26)	(6.98)**
F0_4	-2.673	-4.544	-2.678	-5.466	0.716	-6.910
	(4.34)**	(5.35)**	(5.25)**	(5.94)**	(1.61)	(7.65)**
M5_9	37.633	-1.286	27.625	-1.341	18.878	-1.387
	(0.05)	(1.78)	(0.69)	(1.76)	(2.80)**	(1.84)
F5_9	32.801	-0.198	16.785	0.126	36.655	-1.744
	(0.18)	(0.27)	(2.13)*	(0.16)	(0.15)	(2.30)*
M10_13	26.915	1.296	41.417	1.685	11.642	1.103
	(0.53)	(1.74)	(0.05)	(2.16)*	(5.56)**	(1.45)
F10_13	40.288	1.820	14.970	2.387	20.312	0.420
	(0.06)	(2.46)*	(2.22)*	(3.00)**	(1.27)	(0.56)
M14_16	7.715	0.789	7.902	1.226	8.195	0.202
	(3.60)**	(0.96)	(4.72)**	(1.45)	(6.41)**	(0.24)
F14_16	9.178	2.123	9.261	2.759	9.866	1.591
	(3.79)**	(2.54)*	(3.96)**	(3.34)**	(6.85)**	(1.91)
M17_19	2.762	-0.392	1.260	0.700	1.892	3.163
	(3.34)**	(0.45)	(2.17)*	(0.76)	(4.58)**	(3.44)**
F17_19	1.423	3.765	0.960	2.444	2.698	1.793
	(2.18)*	(4.16)**	(1.85)	(2.62)**	(5.80)**	(1.84)
M20_29	-0.267	-3.761	-1.107	-4.702	0.240	-7.178
	(0.52)	(5.07)**	(2.69)**	(5.79)**	(0.63)	(8.48)**
F20_29	1.054	-1.740	0.157	-1.309	0.422	-3.325
	(2.32)*	(2.91)**	(0.43)	(2.01)*	(1.26)	(4.97)**
M30_54	0.329	-2.642	-0.928	-2.955	0.863	-4.864
	(0.62)	(3.41)**	(2.08)*	(3.80)**	(2.02)*	(6.33)**
M55	0.474	-2.354	-0.874	-2.994	0.218	-3.221
	(0.75)	(2.65)**	(1.82)	(3.19)**	(0.47)	(3.45)**
F55	0.772	-2.306	-0.482	-0.183	-0.047	-3.219
	(1.52)	(3.21)**	(1.22)	(0.24)	(0.12)	(4.10)**
hhead_educ	0.009	0.149	0.038	0.191	0.022	0.034
	(0.57)	(10.32)**	(2.91)**	(12.15)**	(2.15)*	(2.36)*
Dependents	-2.785		-1.890		-2.276	
	(8.55)**		(7.84)**		(9.91)**	
Constant	-2.822	-11.635	-2.844	-14.979		-9.588
	(2.54)*	(11.22)**	(3.27)**	(12.89)**		(9.35)**
Observations: censored	3444		5279			
uncensored	5659		6318			
F Tests: 5-9	0.00(0.994)	3.90(0.048)**	0.07(0.791)	5.72(0.016)*	0.01(0.942)	0.30(0.581)
10-13	0.00(0.984)	0.84(3.358)	0.00(0.976)	1.27(0.260)	0.29(0.590)	1.02(0.312)
14-16	0.27(0.603)	3.22(0.072)*	0.27(0.604)	4.38(0.036)**	0.86(0.354)	2.69(0.100)*
17-19	3.37(0.066)*	22.5(0.000)***	0.33(0.563)	3.48(0.062)*	3.09(0.078)*	1.76(0.185)

Note: Absolute value of t statistics in parentheses; * Significant at 5%; ** significant at 1%. Twenty three district and 2

ethnicity (Tamil, minority) dummies included in regressions but not reported

Table 5: F-Tests based on age in levels (rather than age group)

	1990/91				1995/6		2000/01		
	Engel Curve	Double Hurdle		Engel Curve	Double Hurdle		Engel Curve	Double Hurdle	
		Participation Equation	Expenditure Equation		Participation Equation	Expenditure Equation		Participation Equation	Expenditure Equation
Age 5	0.64(0.42)	0.17(0.67)	0.11(0.73)	0.79(0.37)	0.06(0.80)	0.11(0.74)	2.80(0.11)	0.21(0.64)	1.95(0.16)
Age 6	0.06(0.81)	0.00(0.99)	0.04(0.83)	0.20(0.65)	0.00(0.99)	0.25(0.61)	2.58(0.12)	0.00(0.99)	2.21(0.13)
Age 7	0.61(0.43)	0.00(0.99)	0.76(0.38)	0.02(0.87)	0.00(0.99)	0.02(0.89)	0.00(0.95)	0.00(0.99)	0.04(0.84)
Age 8	4.43(0.03)*	0.00(0.99)	4.15(0.04)*	6.28(0.01)*	0.00(0.99)	6.70(0.00)**	1.31(0.25)	0.00(0.99)	0.55(0.45)
Age 9	1.95(0.16)	0.00(0.99)	3.45(0.06)*	2.74(0.09)*	0.00(0.99)	4.45(0.03)*	0.53(0.46)	0.00(0.98)	0.38(0.53)
Age10	0.00(0.97)	0.00(0.99)	0.01(0.92)	0.86(0.35)	0.02(0.89)	1.18(0.27)	1.10(0.29)	0.02(0.88)	0.37(0.54)
Age 11	1.70(0.19)	0.00(0.99)	1.35(0.24)	0.73(0.39)	0.00(0.99)	1.16(0.28)	1.23(0.26)	0.00(0.99)	0.59(0.44)
Age 12	1.85(0.17)	0.01(0.93)	1.58(0.20)	2.21(0.13)	0.00(0.99)	1.39(0.23)	0.15(0.70)	0.00(0.99)	0.15(0.70)
Age 13	0.05(0.82)	0.21(0.64)	0.00(0.98)	1.26(0.26)	0.00(0.99)	1.20(0.27)	0.36(0.54)	0.00(0.98)	0.94(0.33)
Age 14	0.02(0.89)	0.27(0.60)	0.22(0.63)	0.11(0.74)	0.00(0.99)	0.10(0.75)	0.26(0.61)	0.00(0.96)	0.11(0.74)
Age 15	2.10(0.14)	0.00(0.96)	1.26(0.26)	2.13(0.14)	0.08(0.77)	1.23(0.26)	3.11(0.07)*	0.14(0.70)	3.30(0.06)*
Age 16	2.10(0.14)	0.15(0.73)	2.25(0.12)	9.69(0.00)**	0.17(0.68)	8.49(0.00)**	1.83(0.17)	2.85(0.09)	1.26(0.26)
Age 17	0.05(0.81)	1.59(0.20)	1.62(0.20)	1.50(0.22)	1.96(0.16)	1.96(0.16)	0.58(0.44)	7.55(0.00)**	0.17(0.68)
Age 18	10.14(0.00)**	0.64(0.42)	14.54(0.00)**	0.85(0.35)	1.65(0.19)	1.65(0.19)	4.45(0.03)*	0.07(0.78)	3.60(0.05)*
Age 19	6.03(0.01)*	0.10(0.75)	7.29(0.00)**	1.09(0.29)	0.83(0.36)	0.83(0.36)	0.00(0.99)	0.15(0.69)	0.08(0.71)

Note: ** significant at the 1 per cent level, * significant at the 5 percent level

Table 6: Private rates of return to education in Sri Lanka 1995 and 2000*

	Level of Education			
	Secondary (6-10 years of schooling)	Completed O/L (11 years of schooling)	Completed A/L (13/14 years of schooling)	University Graduate (15-17 years of education)
1995				
Male- Rural	2.5	3.6	3.5	8.6
Female-Rural	6.6	8.2	2.4	4.4
Male-All sectors	2.4	4.2	3.1	7.5
Female-All Sectors	5.4	9.6	3.2	6.4
2000				
Male- Rural	1.7	1.1	3.2	2.5
Female-Rural	4.8	1.6	2.6	4.1
Male-All sectors	2.4	2.7	3.9	3.9
Female-All Sectors	6.3	5.8	5.0	8.4

* Returns for rural sector estimated using results reported in Table 10.1 and 10.2. Returns for all sectors calculated using unreported results. Calculations based on restricted sample fixed effects estimates.

Table 7: Average education levels by occupation category for men and women in 2000.

Occupation Category	All Sectors-Average education level in years				Rural areas only-Average education level in years			
	men	women	Mean compari son	Ratio women: men	men	Women	Mean compari son	Ratio women: men
Legislators, senior officials and managers	12.48	12.94	-2.18*	1 : 3.13	12.34	12.48	-0.48	1 : 2.80
Professionals and technicians	12.27	12.94	-8.32*	1 : 1.20	12.09	12.93	-8.43*	1 : 1.15
Clerks	11.51	12.30	-7.73*	1 : 1	11.63	12.36	-6.19*	1 : 1.06
Service workers and shop and market sales workers	9.83	10.36	-2.21*	1 : 5.64	9.81	10.46	-2.24*	1 : 5.88
Skilled agricultural workers/fishermen	8.16	8.13	-0.03	1 : 2.86	8.19	7.05	-0.84	1 : 4.57
Craft related	8.45	9.36	-7.29*	1 : 2.55	8.4	9.33	-7.04*	1 : 2.36
Plant and machine operators and assemblers	9.16	9.93	-5.04*	1 : 3.09	9.18	9.98	-4.59*	1 : 3.88
Elementary occupation	7.84	8.69	-6.39*	1 : 2.31	7.88	8.55	-4.09*	1 : 2.94
Security Forces	10.76	12.14	-1.64	1:24.9	10.65	12	-1.5	1:22.6