Evaluating the Impact of Working While in School on School Progress: The Case of Rural Bangladesh

Cristóbal Ridao-Cano [1]

Department of Economics and Population Program at IBS

University of Colorado at Boulder, Campus Box 484, Boulder, CO 80309

E-mail: ridaocan@colorado.edu

August 2001

Abstract

This paper examines the causes underlying the low levels of school progress in rural Bangladesh, while focusing on the effect of working while in school. To this end, a switching probit model for the transition to secondary school is estimated, where the switching is determined by endogenous employment status. This approach allow us to recover a variety of mean "treatment" effects, and to assess the relative importance of observables and unobservables in understanding the selection into work and school outcome processes. The results show that overall those children who are more likely to work are those who suffer the most from working in terms of making the transition to secondary school or, alternatively, who would benefit the most from not working. This result suggests that policies aimed at increasing the rate of transition to secondary school through reductions in child labor are both most relevant and most effective. A set of policies in this regard are suggested that aim at changing the environment in which child work and schooling decisions are made.

JEL CLASSIFICATION: D12, I21, J13, J22, O15

KEYWORDS: Child work, school progress, selection, switching model.

[1] This research was made possible in part thanks to financial support provided by Jane Menken (University of Colorado at Boulder) through the National Institute of Aging grants 5P01 AG/HD11952-04 (LP075) and 5P01 AG/HD11952-05 (LP115). I am grateful to Jeffrey Zax, Jose Canals, Terra McKinnish, Robert McNown, Jane Menken and participants at the 2001 Annual Meetings of the European Society for Population Economics (Athens, Greece) for helpful comments and suggestions.

1 Introduction

Although labor force participation rates for school-age children (i.e. aged 5-14) have been declining over time, recent ILO estimates (1996) show that child labor continues to be a very pervasive phenomenon, particularly in the developing world, where it is generally accompanied by low levels of educational achievement.

Bangladesh is a typical example of this pattern. Despite the remarkable improvements in school enrolment rates and educational attainment in recent years, the school performance of children in Bangladesh continues to be very poor, particularly in rural areas. In particular, recent estimates (Filmer, 1999) indicate that among children aged 15-19, 27.5 percent have never attended school, of those who ever attended school 36 percent started school later than 6 years of age (the official starting age), 73.8 percent had completed primary school, and 69.2 percent had survived to secondary.

On the other hand, estimates based on the Child Labor Survey 1995-96 (BBS, 1996) indicate that child labor is still a very pervasive phenomenon in Bangladesh, especially in rural areas. In particular, 19 percent of children aged 5-14 are in the labor force, and child labor constitutes about 12 percent of the total labor force of Bangladesh. Rahman (1997), found that about 48 percent of the working children had never attended school, 43 percent had attended school earlier, and 9 percent were still in school.

What has increased in contemporary times is the awareness of, and concern for, working children, both at national and supranational levels. The perception of child labor as a "problem" stems from its believed harmful effects on the health and intellectual development of children, although it is recognized that not all work is necessarily detrimental. There is, however, wide disagreement on how to tackle the "problem" of child labor, which stems from lack of awareness of the causes and consequences of child labor.

The literature on child labor and schooling in developing countries has been rapidly expanding in recent years.[2] Fewer studies, however, have analyzed the decisions regarding school and work simultaneously. Most of these studies look at the determinants of child labor and schooling and then conclude that child work and schooling are close substitutes based on the negative correlation between observables and/or unobservables in the work and school equations.[3]

Some studies have also attempted to examine indirectly the impact of work on schooling by considering the response of work and schooling to exogenous changes in the price of schooling caused by school incentive programs such as the Food-for-Education scheme in rural Bangladesh (Ravallion and Wodon, 2000; Arends-Kuenning and Amin, 2000) and Progresa in rural Mexico (Demombynes, 2001; Schultz, 2001). These studies generally find that enrolment subsidies reduce the incidence of child labor, but this effect only accounts for a small proportion of the increases in school enrolment and time spent on schooling activities.

Finally, a few papers have provided direct estimates of the effect of work on education outcomes, although in most cases the possible self-selection into work is not accounted for, thereby clouding the interpretation of such estimates as structural effects. In general these studies find negative effects of working on education "inputs" such as school attendance and hours of study (Akabayashi and Psacharopoulos, 1999), years of schooling and grade progression (Psacharopoulos, 1997), and education "outputs" such as cognitive achievement (Heady, 2000).[4]

This paper examines the causes underlying the low levels of school progress in rural Bangladesh, while focusing on the effect of working while in school. The analysis is focused on one dimension of school progress, namely, transition from primary to secondary school, although grade repetition in primary school is also considered. I focus on the transition to secondary school since the acquisition of basic skills needed in the job market and in life generally requires having at least

[2] See Grootaert and Kanbur (1995), and Basu (1999) for surveys of the literature.

[3] See, for example, Canagarajh and Coulumbe (1997), Grootaert (1998), Skoufias (1994), Duraysamy (2000), Assad *et al.* (2000), Duryea *et al.* (2001), and Ridao-Cano (2001).

[4] There is a growing parallel literature in the U.S. examining the effect of working while in high school on a variety of school outcomes. Since in general the environment in which work and schooling decisions are made in the U.S. is very different from that in developing countries, it is not surprising to find that the evidence from this literature is more mixed. See Donahoe and Tienda (2000) for a review of this literature.

some secondary education. Furthermore, the transition from the last grade in primary to the first grade in secondary is one of the single most important turning points in the Bangladeshi school system.

The contribution of this paper is twofold. First, this is one of the first papers in the context of developing countries that explicitly examines the direct effect of combining school and work on subsequent human capital accumulation, while accounting for the simultaneous nature of work and school decisions. Second, this paper proposes a robust and flexible econometric model, not previously used in the literature, for evaluating the effect of working while in school. In particular, the paper estimates a switching probit model for the transition to secondary school, where the switching is determined by endogenous employment status. This approach allows us to recover a variety of mean "treatment" effects, and to assess the relative importance of observables and unobservables in understanding the selection into work and school outcome processes.

The estimated model works well in characterizing the work and school outcome processes. The main result of this paper is, however, that working while in school has a negative impact on the transition to secondary school, but that this effect is significantly more negative for those children who actually select into work. In particular, those children who are more likely to work are those who suffer the most from working or, alternatively, who would benefit the most from not working.

This finding has important policy implications. In particular, the above result suggests that policies aimed at increasing the rate of transition to secondary school through reductions in the incidence of child labor are both most relevant, given the magnitude of the effect of child work on school progress, and most effective, since most working children are perfectly able to reach secondary school in the absence of work. Effective policies should be aimed at changing the environment in which work and schooling decisions are made. To this end, the reinforcement of the ongoing school incentive schemes offer a promising route, since they provide cash or in-kind subsidies to qualifying children that are conditional on school attendance and are aimed at compensating parents for the foregone income from their children's labor.

The rest of the paper is organized as follows. Section 2 describes the institutional setting in which work and school decisions are made. Section 3 presents the data used for the analysis. Section 4 presents the model and derives the expression for the alternative mean treatment parameters. Section 5 reports the general results of the model and the estimates of the mean treatment effects. Section 6 concludes and suggests some policy implications.

2 Institutional Setting

Bangladesh is a poor country by any measurable standard. Income per capita was $279 in 1999, and about 45 percent of its population of more than 128 million lives below the poverty line. Agriculture is by far the largest sector of the economy, accounting for 32 percent of GDP. Agriculture is still characterized by high labor intensity and low productivity. It is also constantly subject to flooding during the monsoon season, in a country where credit and insurance markets are still poorly developed. Although Bangladesh is a small country, multiple rivers and inadequate transportation systems make one third of the country inaccessible.

Formal education in Bangladesh is mainly delivered by the Government and is divided into a 5-year cycle of primary education, 5 years of secondary education (2 years of junior secondary and 3 years of secondary), and 2 years of higher secondary education. Higher education comprises courses of 2-5 years of duration and beyond. A non-formal education system also exists, delivered by the Government and NGOs, and targeting disadvantaged children and young adults. Finally, an Early Childhood Care and Development Program is also in place, targeting the children aged 3-5.

Primary school is run in two shifts: the first shift of two hours for grades 1 and 2, and the second shift of three and a half to four hours for grades 3 to 5. Primary education is compulsory as of 1992 and tuition-free, and textbooks are supplied free of cost. The official age of entry into primary school is 6, but as already mentioned children often enter at later ages.[5] Until 1995,

[5] This pattern is very systematic so as to be attributed only to age misreporting arising from the lack of official birth registry in many rural areas.

grade promotion to the next grade was based on student performance for all grades. Since 1995, however, all students in grades 1-2 get promoted to the next grade. No certificate is awarded after completion of primary school, but it is considered as the norm of attainment of basic literacy.

Secondary education is more examination- and labor market-oriented. The length of the school day increases with grade but it is still relatively short. The first public examination, known as the Secondary School Certificate (SSC) examination, is held at the end of grade 10, and it must be passed by the student in order to move to higher secondary. At the end of higher secondary school another public examination is held, leading to Higher Secondary Certificate (HSC). Tuition is free for girls in grades 6-8 in rural areas since 1990.

Although the Government of Bangladesh has been taken important steps toward extending and improving the education system since 1973, the changes implemented during the 1990-1995 development plan marked a turning point. In particular, major changes during this period include the introduction of compulsory primary education in 1992 (along with the creation of committees for social mobilization and campaign for education and against child labor at the grassroots level), construction of new schools and the improvement of existing ones.

In addition, several school incentive programs were introduced in 1994: the Food-for-Education program, which provides food assistance to parents of poor children aged 6-10 provided these children attend school 85 percent of the school year, and the secondary-school scholarship program for girls, which provides monthly stipends to girls enrolled in grades 6 and 9 (subsequently extended to include girls in grades 7 and 8), provided they attend 65 percent of the school year and maintain a passing grade.[6] The latter program, along with the 1990 law that made tuition in secondary school free for girls, was intended to counterbalance the important gender differences in education that still exist today in Bangladesh.

Despite the steady increases over the years, public expenditures in education as a percentage of GDP is among the lowest in South Asia. To make matters even worse, each year an ever increasing

[6] Parents must also sign an agreement that their daughters will not get married before the age of 18.

number of children enter school, creating bottlenecks in both the quantity and quality of schools.

3 Data Description

The data for the analysis comes from the 1996 Matlab Health and Socio Economic Survey (MHSS).[7] The survey covers 141 villages of Matlab, a region of rural Bangladesh where there is an ongoing prospective Demographic Surveillance System. The MHSS collected extensive current and retrospective information on multiple domains from approximately 38,000 individuals in a sample of over 7000 households, and conducted a detailed community survey. A distinctive feature of the MHSS is its multistage sampling procedure which takes into account the social structure in rural Bangladesh.[8]

The present analysis focuses on the education and work experiences of individuals who ever attended school and were aged 15-25 at the time of the survey. This sample contains both young adults living with their parents as well as young adults living separately, thus avoiding the typical sample selection bias arising from just considering those in the former group. I am selecting the sample of children who ever attended school, since the focus of the analysis is on investigating the impact of working while in school on school progress. Hence all inferences will refer to those children.[9]

The reason for setting the lower age limit at 15 is twofold. First, by using a sample of young adults (15 years of age or older), the information collected is as reported by the same individuals and not their parents, which is crucial in the reporting of the work history. Second, it minimizes issues of censoring arising from children who are still in primary school. The reason for setting

[7] MHSS was carried out as part of a Program Project funded primarily by the National Institute on Aging, with additional support from the National Institute of Child Health and Human Development and the Mellon Foundation. The Project is a collaboration among investigators in U.S. universities and the International Centre for Diarrhoeal Disease Research (ICDDR) in Bangladesh.

[8] As a result, appropriate weights are then needed in the analysis of household or individual data to correct for the non-random sample distribution. However, analysis carried out with and without weights yields similar results. For details on sample design, see Rahman *et al.* (1999), which can be found, along with other documentation and the data, at http://ftp.rand.org/software_and_data/FLS/mhss/.

[9] In future research, it would be interesting to simultaneously investigate the impact of working on the timing of entry into school and whether the child ever enters school at all. Alas, the MHSS has very limited information on child work outside school.

the upper age limit at 25 is fourfold. First, the older the individual the more likely he or she is to make recollection errors. Second, the parental and origin household information becomes more limited as we consider older individuals, since they are more likely to live apart from their parents, and in this case parental information is reported by the individual and not his or her parents. Third, extending the sample to older adults would require us to deal with children and their parents in the same model. Finally, I want to relate as much as possible my results to the current status quo of education in Bangladesh, so as to make the proposed policies more relevant and significant. Using these age cutoffs, and dropping the few observations with missing values in the key variables, the sample used for the empirical analysis contains 1954 individuals, 113 of which are still in primary school.[10]

The MHSS contains detailed information on education histories, including the school entry age, school drop out age, grades attended and completed, grade repetition, and scores in the SSC and HSC examinations.[11] Individuals report the school they attended at each school level, information that can be then used to collect the current and retrospective school quality information from the community survey. In practice, however, the school level survey is limited to a handful of schools, covering a small percentage of the schools attended by the sample of young adults. Information is available, however, on the presence in each village of schools at each education level, and the year in which these schools started operating.

Among the possible school outcomes, this paper focuses on the transition from primary to secondary school for two main reasons. First, as mentioned earlier, survival to secondary school among those who started primary school is only 69.2 percent, and in fact a significant proportion of early school drop outs occur after completing primary education, particularly girls. Second, in the present context I am particularly interested in the effect of work on the acquisition of basic

[10] These are mostly children who started school when they were 11 years of age or older, which is not such an uncommon phenomenon in rural Bangladesh.

[11] For those young adults who attended school in the previous year, information is also available on educational expenditures, school hours and distance to school.

skills needed in the job market and in life. The acquisition of these basic skills is generally believed to require having at least some secondary education.

To reinforce some of the findings using the transition to secondary school outcome, this paper also makes use of the information on whether the young adult repeated some grade in primary school. Grade repetition conveys information on another important dimension of school progress, namely, the rate at which individuals progress through school. Using grade repetition as the school outcome has the problem that schools may not apply uniform standards in enforcing grade repetition. Furthermore, repeating a grade implies persisting in school, and hence comparing children who repeat a grade with those who do not puts in the same category those who make the transition to the next grade and those who fail and drop out, which are clearly different groups. This calls for the adequate distinction between those who make the transition to the next grade, those who fail a grade and repeat it, and those who fail and drop out. In fact we could potentially combine the information on transition to secondary school and grade repetition in primary school into a single comprehensive school outcome indicator. In practice, however, the lack of sufficient variation in the grade repetition variable precludes such possibility.

It is widely recognized that not all work is necessarily detrimental for child schooling, although the question of how detrimental is an empirical one. In any case, the key is to identify the kind of work that can potentially interfere with a child's schooling. In the present context, I am interested in work that was consistently undertaken while the child was attending primary school, and not in sporadic and/or small amounts of part-time work or employment during school vacations.

The MHSS contains retrospective discrete information on whether a young adult consistently performed some kind of productive activity while he or she was attending each school level (i.e. primary, secondary, higher secondary, and higher education). This information is reported by the young adult, thus minimizing the typical under-reporting when it is the mother or the father who provides this information. Only children who consistently performed some kind of productive work while in primary school are likely to report to have been working, which is exactly the group of children that I am interested in.

This definition of work ignores, however, household chores such as caring for younger siblings. Levison *et al.* (2001) point out that this is likely to significantly underestimate the amount of work carried out by girls, and thus ignores a potential impediment for their schooling. However, if this were true in this context we should expect systematic differences in work status by gender. In fact, the proportion of boys and girls that report having worked during primary school is almost the same.

Table 1 shows the numbers and proportions of boys and girls by work status during primary school. Almost 28 percent of children performed productive work while in primary school, and there are not significant differences by gender. This high proportion of working students is not surprising since the short duration of the school day during primary school makes it easier for many children to combine school attendance with work responsibilities, particularly if they perform farm work. However working children may find themselves less able to learn as a result of exhaustion or insufficient time to complete homework, which increases their chances of failing and repeating a grade or dropping out of school all together. Furthermore, this ability to combine school and work diminishes as the child moves to higher grades, where the required schooling time is higher. This is particularly so in moving from the last grade in primary school (i.e. grade 5) to the first grade in secondary school (i.e. grade 6), where not only does the required schooling time increase but also the chances of having a secondary school in the village are lower.

Tables 2 reports the Kaplan-Meier estimates of the survival rates to grade 6 by gender and work status.[12] Almost 30 percent of the children never made it to secondary school, which is very close to the national average, and thus speaks for the representativeness of the data for the whole nation. There are, however, statistically significant differences by gender, with a smaller proportion of girls reaching grade 6. Although 76.4 percent of non-working children made it

[12] See Table A.2 for Kaplan-Meier estimates of survival rates for all grades. This table confirms that the transition from grade 5 (last grade in primary) to grade 6 (first grade in secondary) is the single most important turnning point, particularly for girls.

to secondary school, this figure is only 56.4 percent for working children, and the difference of 20 percentage points is statistically significant.[13] This difference is much greater for girls (24 percent) than for boys (15.6 percent).[14]

Table 3 shows the Kaplan-Meier estimates of the repetition rate in primary school by gender and work status,[15] and reveals that 11 percent of children repeated some grade in primary. Again, there are statistically significant differences by gender, with a higher proportion of girls repeating some grade in primary. Although 9 percent of non-working children repeated some grade in primary, this figure goes up to 16 percent for working children, a statistically significant difference of 7 percentage points. This difference is slightly higher for girls than for boys.

4 A Switching Probit Model for the Work and School Outcome Processes

There is large literature on sample selection in the context of continuous variables. An important strand of this literature has been concerned with evaluating the effectiveness of some social program or "treatment". Individuals participating in a program or receiving treatment often have different characteristics than the average person in the population. Thus evaluating the returns to a program requires accounting for the non-random assignment of individuals into the treated and untreated states. A popular approach in this context for dealing with selection bias has been the two-stage estimation method suggested by Heckman (1976).

There are, however, significantly fewer results on selection in the context of discrete outcomes. In this context, the standard two-stage instrumental variable method does not guarantee consistency under the assumptions of non-linear discrete choice models (Smith and Blundell, 1986).

[13] It is worth noting that in looking at grade-specific survival rates, the big difference between working and non-working children starts occurring at the beginning of the second cycle of primary (i.e. grade 3), which coincides with an increase in the required schooling time.

[14] This gender difference may be partly explained by selection on observables and unobservables, but it may also indicate that working girls are more likely to combine productive activities with household chores (which I am not considering here) than working boys, so that they end up working longer hours.

[15] Those still in primary school are assumed to be exposed to the risk of repeating a grade in primary.

Furthermore, the alternative linear probability model is incompatible with the observed data when dichotomous endogenous regressors are present (Carrasco, 1998).

An alternative approach, yielding consistent estimates, is a bivariate probit model, where we simultaneously estimate the program participation and outcome equations. Although we can interact the treatment variable with every observable in the outcome equation, this approach constrains the unobservables generating the outcome under the treated and untreated states to be the same. Thus, this approach does not allow us to distinguish between the potential effect of the program for an individual randomly selected from the population, and the potential effect for an individual randomly selected from the population of participants. This constraint also rules out some possible combinations of outcomes under the treated and untreated states.

This paper makes use of a flexible and robust method for evaluating the effect of working while in school on a discrete school outcome, based on the general framework developed by Aakvik, Heckman, and Vytlacil (2000).[16] In particular, the paper estimates a switching probit model for the transition to secondary school, where the switching is determined by endogenous employment status. This approach allows us to recover a variety of mean "treatment" effects, and to assess the relative importance of observables and unobservables in understanding the selection into work and school outcome processes.

4.1 The Model

For each child i, assume two potential dichotomous schooling outcomes (S_{1i}, S_{0i}) corresponding, respectively, to the potential school outcome in the working state (i.e. $W_i = 1$) and non-working state (i.e. $W_i = 0$). Let S_i be the observed school outcome so that

$$S_i = W_i S_{1i} + (1 - W_i) S_{0i}$$

It is assumed that S_{0i} and S_{1i} are defined for everyone and that these outcomes are independent across children. This paper assumes that W, S_1 and S_0 are generated by the following linear latent

[16] Referred to as AHV 2000 for the remainder of the paper.

index structure

$$W_i = \mathbf{1}(W_i^* \geq 0) = \mathbf{1}(Z_i\beta_w + U_{wi} \geq 0) \tag{1}$$

$$S_{1i} = \mathbf{1}(S_{1i}^* \geq 0) = \mathbf{1}(X_i\beta_1 + U_{1i} \geq 0) \ \text{ iff } \ W_i = 1 \tag{2}$$

$$S_{0i} = \mathbf{1}(S_{0i}^* \geq 0) = \mathbf{1}(X_i\beta_0 + U_{0i} \geq 0) \ \text{ iff } \ W_i = 0 \tag{3}$$

where W_i^*, S_{1i}^* and S_{1i}^* are the net utilities for child i (as perceived by his or her parents) from working, making it to secondary school under the working state, and making it to secondary under the non-working state, respectively; Z_i, X_i are vectors of observables generating the work and school outcomes, respectively; U_{wi}, U_{1i} and U_{0i} are unobservables (to the econometrician) generating the work outcome, school outcome in the working state, and school outcome in the non-working state, respectively; and $\mathbf{1}[.]$ is the indicator function.

Throughout the paper I make the following two main assumptions:[17]

1. $Z\beta_w$ is a non-degenerate random variable conditional on X
2. (U_w, U_1) and (U_w, U_0) are independent of (Z, X)

Assumption (1) requires an exclusion restriction: there exists a variable that determines the work outcome but does not directly affect the school outcome conditional on work status.

To complete the specification of the model (U_1, U_0, U_w) are assumed to be normally distributed with zero mean vector and covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho_{10} & \rho_{1w} \\ & 1 & \rho_{0w} \\ & & 1 \end{pmatrix} \tag{4}$$

[17] Other, more technical, assumptions are the absolute continuity of (U_w, U_1) and (U_w, U_0); S_1 and S_0 have finite sample moments (required if mean treatment parameters are to be well defined, which is satisfied trivially in our case since S_1 and S_0 are binomial); and the standard assumption that for each set of X variables, we observe children in both states.

Since S_1 and S_0 are never jointly observed, the joint distribution of (U_1, U_0) is not identified and, consequently, ρ_{10} is not identified. The log-likelihood function of the model, from which maximum likelihood estimates can be obtained, is as follows

$$L = \sum_{i:\ W_i=1, S_i=1} \log P_i^{11} + \sum_{i:\ W_i=1, S_i=0} \log P_i^{10} + \sum_{i:\ W=0, S=1} \log P_i^{01} + \sum_{i:\ W=0, S=0} \log P_i^{00} \tag{5}$$

where

$$\begin{aligned} P_i^{11} &= \Pr[W_i = 1, S_i = 1] = \Phi[Z_i\beta_w, X_i\beta_1, \rho_{1w}] \\ P_i^{10} &= \Pr[W_i = 1, S_i = 0] = \Phi[Z_i\beta_w, -X_i\beta_1, -\rho_{1w}] \\ P_i^{01} &= \Pr[W_i = 0, S_i = 1] = \Phi[-Z_i\beta_w, X_i\beta_0, -\rho_{0w}] \\ P_i^{00} &= \Pr[W_i = 0, S_i = 0] = \Phi[-Z_i\beta_w, -X_i\beta_0, \rho_{0w}] \end{aligned} \tag{6}$$

The probabilities in (6) refer to completed (i.e. uncensored) primary school spells. The contribution to the likelihood function of those children still attending primary and working is [18]

$$\Pr[W = 1, S = 1] + \Pr[W = 1, S = 0]$$

and the contribution of those children still attending primary school but not working is

$$\Pr[W = 0, S = 1] + \Pr[W = 0, S = 0]$$

Finally, those children who have not started school by age 15 (at least) are assumed to contribute nothing to the likelihood function.[19]

[18] Most of the children still attending primary school are in grades 4 and 5. Thus, since my definition of work refers to productive work consistently performed during primary school, if these children report no work up to the grade they are currently attending then they are treated as non-working observations for the analysis. If we also allowed for censoring in the work variable, the contribution of each censored observation to the likelihood function would be 1.

[19] Of the total sample of children aged 15-25, 540 (i.e. 21%) have never attended school.

4.2 Mean Treatment Parameters

An important advantage of the above latent variable model is that it can be used to generate mean treatment parameters from a common set of structural parameters. Identification of the bivariate distribution of $(W,\ S_1)$ and $(W,\ S_0)$ implies identification of the mean treatment parameters. Identification of the distributional treatment parameters requires knowledge of the full trivariate distribution of $(W,\ S_1,\ S_0)$, which we do not have in the context of the switching probit model developed above, but it can be achieved in the context of a factor loading switching probit (AHV 2000).

Let Δ denote the effect of working for a given child, where $\Delta = S_1 - S_0$. This person-specific effect is a counterfactual. For a given child, it answers the question of what would be his or her school outcome if he or she worked compared to the case where he or she had not worked. In our case, Δ can take three values

1. $\Delta = 1$ $(S_1 = 1,\ S_0 = 0)$ if the child would make it to secondary in the working state and would not make it in the non-working state.

2. $\Delta = 0$ if the child would make it to secondary in either state $(S_1 = 1,\ S_0 = 1)$, or if he or she would not reach secondary in either state $(S_1 = 0,\ S_0 = 0)$.

3. $\Delta = -1$ $(S_1 = 0,\ S_0 = 1)$ if the child would make it to secondary in the absence of work and would not make it if he or she were to work.

Clearly we expect most children to be of type 2 or 3. One can rarely estimate Δ for a given person.[20] Instead, it is more common to work with population means or distributions of these variables. In this paper, I consider three different mean treatment parameters: the average treatment effect (ATE), the effect of treatment on the treated (TT) and the marginal treatment effect (MTE). Each of these parameters corresponds to an average value of Δ but defined on different conditioning sets.

[20] Some panel data estimators indentify Δ for each person (see Heckman and Smith, 1998, for a discussion).

In the context of our model, the average treatment effect (ATE) is the average effect of working for a child randomly selected from the population of children aged 15-25 who ever attended school, and is given by

$$\begin{aligned}\Delta^{ATE}(x) &= E[\Delta \mid X = x] \\ &= \Pr[S_1 = 1 \mid X = x] - \Pr[S_0 = 1 \mid X = x] \\ &= \Phi[x\beta_1] - \Phi[x\beta_0] \qquad (7)\end{aligned}$$

The effect of treatment on the treated (TT) is the average effect of working for children who actually worked while in primary school, and is given by

$$\begin{aligned}\Delta^{TT}(x, W = 1) &= E[\Delta \mid X = x, W = 1] \\ &= \Pr[S_1 = 1 \mid X = x, W = 1] - \Pr[S_0 = 1 \mid X = x, W = 1] \\ &= \frac{\Phi[x\beta_1, z\beta_w, \rho_{1w}] - \Phi[x\beta_0, z\beta_w, \rho_{0w}]}{\Phi[z\beta_w]} \qquad (8)\end{aligned}$$

The marginal treatment effect (MTE) is the average effect of working for children with different values of $U_w = u_w$ or, alternatively, the average effect of working for children who are on the margin of indifference between working ($W = 1$) or not ($W = 0$), i.e. $u_w = -Z\beta_w$. Hence, MTE identifies the effect of an intervention to make children work on those children induced to change working states by this intervention. MTE is given by [21]

$$\begin{aligned}\Delta^{MTE}(x, u_w) &= E[\Delta \mid X = x, U_w = u_w] \\ &= \Pr[S_1 = 1 \mid X = x, U_w = u_w] - \Pr[S_0 = 1 \mid X = x, U_w = u_w] \\ &= \Phi\left[\frac{x\beta_1 + \rho_{1w}u_w}{\sqrt{1-\rho_{1w}^2}}\right] - \Phi\left[\frac{x\beta_0 + \rho_{0w}u_w}{\sqrt{1-\rho_{0w}^2}}\right] \qquad (9)\end{aligned}$$

[21] Note that under our distributional assumptions

$$(S_j^* \mid X = x, U_w = u_w) \sim N\left[x\beta_j + \rho_{jw}u_w, \left(1 - \rho_{jw}^2\right)\right] \ , \quad j = 1, 0$$

where $u_w = -z\beta_D$. In the results section, ATE and MTE are evaluated at the population means of the Xs (i.e. $x = \overline{x}$), and TT is evaluated at the means of the Xs and Zs for the working children population (i.e. $x = \overline{x}_w$ and $z = \overline{z}_w$). The standard errors for ATE and TT are calculated using the delta method.

Within this framework, one can also calculate the marginal effects of changes in elements of X on ATE and TT. The marginal effects on ATE are calculated as

$$\frac{\partial ATE(\overline{x})}{\partial x_k} = [\phi(\overline{x}\beta_1) - \phi(\overline{x}\beta_0)][\beta_{1k} - \beta_{0k}]$$

for a continuous variable x_k, and the difference between ATE evaluated at $x_k = 1$ and ATE evaluated at $x_k = 0$ for dummy variables. The marginal effect of a continuous variable x_k on TT is calculated as the derivative of the first conditional probability in (8) with respect to x_k minus the derivative of the second conditional probability with respect to x_k, where the expression for the former is [22]

$$\phi(\overline{x}_w\beta_1)\,\Phi\left[\frac{\overline{z}_w\beta_w - \rho_{1w}\overline{x}_w\beta_1}{\sqrt{1-\rho_{1w}^2}}\right]\frac{1}{\Phi[\overline{z}_w\beta_w]}\beta_{1k} + \tag{10}$$
$$\left[\frac{\Phi[\overline{z}_w\beta_w]\phi(\overline{z}_w\beta_w)\,\Phi\left[\frac{\overline{x}_w\beta_1 - \rho_{1w}\overline{z}_w\beta_w}{\sqrt{1-\rho_{1w}^2}}\right] - \Phi[\overline{x}_w\beta_1, \overline{z}_w\beta_w, \rho_{1w}]\phi(\overline{z}_w\beta_w)}{(\Phi[\overline{z}_w\beta_w])^2}\right]\beta_{wk}$$

For dummy variables, the marginal effect is the difference between TT evaluated at $x_k = 1$ and TT evaluated at $x_k = 0$.

ATE, TT and the MTE parameters are related through the MTE parameters. In particular, ATE can be expressed an average of the MTE parameters, whereas TT can be written as a weighted average of the MTE parameters, where those children with a higher propensity to work in terms of unobservables receive the most weight in the average.[23]

[22] The expression for the second term is the same as (9) but replacing β_1 for β_0 and ρ_{1w} for ρ_{0w}.

[23] See AHV 2000 for a derivation of these relationships.

The three mean treatment parameters define different average effects of working if U_w is not mean independent of $U_1 - U_0$ but they are identical if U_w is mean independent of $U_1 - U_0$ conditional on $X = x$. In particular if $Corr[U_w, U_1 - U_0] = 0$, or $\rho_{1w} = \rho_{0w}$, TT reduces to ATE in (7). In this case, unobservables would not have any role in jointly generating the work and school outcomes, and thus selection bias would not be an issue.

4.3 Model Covariates

The model is estimated for the whole sample of children since there is not sufficient variation in the data to estimate such a complex model by gender. In order to gain precision in the estimated mean treatment parameters, I choose a parsimonious specification for the observables determining the work and school outcome processes. The observables for the school outcome include a set of child, parental and community characteristics that are, to a large extent, relevant to the period when the child was in primary school.

Child characteristics include sex, school entry age, and a dummy for whether the child was still in school in or after 1992, the year that marked an important turning point in education policy. The policies changes that occurred after 1992 affected different dimensions of the shadow price of schooling, including preferences for school and work, as well as direct and indirect costs of schooling. To isolate the effect of this composite policy variable from the overall trend in the work and school outcome, the age of the young adult at the time of the survey is included.

The school entry age is a state variable reflecting a parental decision or outcome preceding our work and school outcome processes, and as such it is potentially endogenous to those.[24] Given the already complex nature of the model, a two-step approach for dealing with endogeneity in this context is adopted here, following the method developed in Smith and Blundell (1986), and Datt and Ravallion (1994). In particular, time until entry in school is modelled using an ordered probit. The residuals from this model are then included in the work and school outcome equations along

[24] This is particularly so if children entering late in school did so because they were working, and then they continued working once in school.

with the actual school entry age. Provided we have at least one instrument, the structural effect of school entry age is identified, and the significance of the coefficient on the residuals provides a test for weak exogeneity.[25] In all cases, however, the residuals were found to be highly insignificant, so the final model was reestimated without the residuals.[26]

Parental characteristics include years of schooling of the mother and the father of the child, as well as an indicator of whether the mother's father was wealthier than her father-in-law at the time of marriage. The latter variable is intended to investigate whether the mother and the father differ in their preferences over the work and school outcomes considered.[27] This variable, however, was not found to be significant in any of the equations, and thus was excluded from the final specification.[28]

Household level variables have several dimensions. Household demographics are summarized by the age and gender composition of the child's siblings at the time the child started primary school. Household productive assets are summarized by whether the household has farm land, and non-farm business assets. For children living with their parents this information refers to the time of survey, and for children living separately this information refers to the time of the survey if parents are alive and to the time of death if parents are dead. While the amount of land owned or the value of non-farm business assets are likely to change over time, it is less likely that whether the household owns some of these assets changes over time.[29] In any case, if the

[25] Table A.1 shows the results of estimating the model for school entry age. In this model, I use variables that were not significant in explaining the work and school outcomes as instruments, such as the presence of a primary school in the village, whether the father had died before the child reached 6 years of age, and an indicator of the intra-household distribution of power.

[26] It is worth noting that we may get a different result when estimating school entry age, work during primary, and transition to secondary simultaneously.

[27] The idea here is to find an exogenous distribution factor that affects the final outcome only through the bargaining power of the mother relative to the father's. The fact that the mother's father was richer than her father-in-law is likely to get translated into more assets brought to marriage by the mother than the father, which in turn is likely to get translated into a higher relative bargaining power of the mother. This type of analysis circumscribes in the newly developed literature on household collective models (see, for example, Browning and Chiapporti, 1998). For and application to child labor and schooling in developing countries see Ridao-Cano (2001).

[28] This result could either suggest that parents do not differ in their preferences over our particular work and school outcomes, or simply indicate that our variable is not a good proxy for the intra-household distribution of power.

[29] Particularly in the short period of time from completion of primary schooling that I am considering in most

cross-sectional pattern in these variables does significantly change over time, we should not expect any relationship between these variables and the work and school outcomes.

Household wealth is summarized by the current value of non-productive assets, such as homestead land, precious metals and savings. In this case, looking at whether the household owns any asset or a particular asset such homestead land is not applicable as all households own some kind of asset and most own homestead land. The final qualification made for household productive assets applies here as well. This asset information is not available for those children who are living away from their parents, so a dummy is included to control for those.

In order to supplement the household wealth information (particularly for those children for whom it is not available), an indicator for whether the household has a modern latrine (i.e. septic or slab latrine) is also used. This variable also proxies for the health environment that the child was exposed to during primary school. This information refers to the time of the survey for children living with their parents and to the time right before leaving the parental home for children living separately.[30]

Finally, a variety of village level variables were included the model, such as the presence of a tubewell for drinking water, presence of a modern health facility, presence of credit institutions, village electrification, road infrastructure, agriculture modernization (e.g. irrigation facilities, use of high yielding varieties), village economy diversification (i.e. presence in the village of any mill, factory, workshop or cottage industry), distance to market and the capital of Matlab, and the presence of primary and secondary schools. I also included an indicator for whether the village was part of the treatment or control areas of a unique experimental family planning that began in Matlab in 1978.[31] All village level variables refer to the period before the child completed his or her primary school spell.[32] Even though tabular analysis shows significant relations between

of the cases.

[30] This information is available from the migration history of each individual in the sample.

[31] See Fauveau (1994) for a detailed description of the program.

[32] The village where the child resided during primary school is obtained from the migration history of each child. In most cases, this village coincides with the current village of residence. In some cases the child migrated with his

many of these variables and the school outcome, only the presence of a tubewell for drinking water, and a modern health facility were found to be significant in explaining the school outcome after controlling for the other covariates.

The observables in the work equation include those in the school outcome equations plus two variables that could potentially affect the school outcome conditional on work status but were found to be highly insignificant, namely, the distance to the capital of Matlab, and the indicator of the diversification of the village economy. In addition, I use an indicator of whether the household cultivated land (own land, rented or sharecropped) when the child was in primary as the basic identifying instrument. This variable is constructed on the basis of the current cultivation status of the household, and the retrospective information on parental occupation. Conditional on working status, and controlling for land ownership, whether the household cultivates land should not affect the school outcome, and, in fact, it does not.

Table 4 reports means and standard deviations of the variables used in the empirical analysis.

5 Estimation Results

In this section, I first discuss the estimated coefficients in the work and school outcome equations. I then present the estimated mean treatment parameters, the relation between selection into work and school outcomes, and the heterogeneity in observable characteristics with respect to the mean treatment parameters. Finally, to reinforce some of main findings of this paper, I report the results of a bivariate probit model of grade repetition in primary school.

5.1 The Work and School Outcome Equations

Estimates of the parameters of the work equation (W^*), the school outcome equation for working children (S_1^*), and the school outcome equation for non-working children (S_0^*) are reported in Table 5.1 and Table 5.2. Table 5.1 reports, for each equation, the parameter estimates, t-values

or her family while the child was still in primary school. For those observations, the destination village is selected as the relevant village for the child.

for these estimates, estimated correlation coefficients, and the Wald for the overall significance of the model. Table 5.2 presents the marginal effects and their corresponding t-values.[33] The Wald test shows that the model fits the data quite well.

Children who delay school entry are more likely to work during primary school, and less likely to make it to secondary school in either working state. The effect is particularly strong in the case of making the transition to secondary school in the working state, where each additional year out of school above the official school entry age decreases the probability of making it to secondary school by about 7 percent. Having been unable to reject the hypothesis of weak exogeneity, we can interpret the effect of late entry as a structural effect. The effect on work is hardly surprising since older children are physically more able to do productive work, but the effect on the school outcome may indicate that older children may have a hard time fitting in classrooms with younger class mates, and may have fewer chances of continuing in school after failing a grade.[34]

If the child was in primary school after 1992, the year that marked an important turning point in education policy, his or her chances of working are lower while his or her chances of reaching secondary school in either working state are greater. These effects stand apart from the overall time trend, as captured by the effect of current age on the work and school outcomes. This effect is particularly strong for working children, for whom this composite of policy chances increased their likelihood of making it to secondary school by 20 percent.

Both the education of the mother and of the father significantly reduce child work and increase the child's chances of reaching secondary school, but the effect of mother's education is consistently greater across outcomes. Parental education can potentially influence the allocation of children's

[33] Marginal effects are evaluated at the means of the Xs and Zs for the whole population of children. Standard errors are calculated using the delta method. The marginal effect for a continuous variable x_k on S_1 is calculated as

$$\frac{\partial \Pr[S_1 = 1 \mid X = \overline{x}]}{\partial x_k} = \phi\left(\overline{x}\beta_1\right)\beta_{1k}$$

The corresponding expression for a dummy variable is

$$\Phi\left[\overline{x}\beta_1 + \beta_{1k}\right] - \Phi\left[\overline{x}\beta_1\right]$$

[34] See Ahlburg *et al.* (2001) for a similar finding in the context of Egypt.

time directly, mainly through income and preferences, and indirectly through its effect on the bargaining power of the mother relative to that of the father (Ridao-Cano, 2001). Even assuming equal effects on household income, this differential effect of the mother's and the father's education may suggest that women have a higher preference for child schooling than men. This prediction holds even if education does not affect the distribution of power within the household, but it gets reinforced if it does.[35] The effect of parental education is particularly strong for working children, for whom each additional year of schooling of the mother and of the father increases their likelihood of making it to secondary school by 4.1 percent and 1.3 percent, respectively.

Household wealth also plays a significant role in determining the work and school outcomes. In particular, a 1 percent increase in household assets reduces the likelihood of working by 2.1 percent, and increases a non-working child's chances of reaching secondary school by 1.7 percent, with no significant effect for working children. This result may indicate that, apart from its income effect, household wealth plays a very important role as a cushion against economic shocks in the absence of well-developed capital markets.[36] The presence of a modern latrine, which conveys further information about household wealth and the health environment of the child, has no significant role in explaining child work but it does significantly increase the likelihood of making it to secondary school, particularly among working children (by 15 percent).

Household productive assets increase both child work and the child's chances of reaching secondary school, but only among non-working children. Household productive assets, such as farm land and non-farm business assets, have both a positive income effect and a negative substitution effect on a child's schooling time.[37] The income effect is likely to be underplayed in this case,

[35] The variable indicating whether the mother's father was richer than her father-in-law provided a direct test of parental differences in preferences. As I already mentioned, however, this variable did not prove to be significant in explaining the work and school outcomes. Nevertheless, it does significantly decrease the chances of late school entry (see Table A.1).

[36] Jacobi and Skoufias (1997) present evidence on how child time is used as an insurance mechanism against economic shocks in the absence of well-functioning capital markets.

[37] The presence of a productive household asset increases the shadow price of schooling by increasing the value of the marginal product of children (see, for example, Rosenzweig and Evenson, 1977).

since I am just considering whether the productive asset is owned and not how much of it is owned. The positive effect of household productive assets on child work (about 13.9 percent) indicates that the substitution effect dominates the income effect for this outcome. Conditional on working status, all we have left is the income effect, which is only significantly positive for non-working children, for whom owning both types of assets increases their likelihood of making it to secondary school by about 8.8 percent.

Controlling for whether the household owns farm land, the fact that the household cultivates land (whether own land, or rented/sharecropped) increases the likelihood of child work by about 7 percent.

The age and sex composition of siblings play a mixed role in the allocation of children's time. In general, the presence of other siblings in the household increases competition for household resources for education (in the case of school-age siblings) and otherwise. The pressure on household resources is particularly intensified by the presence of younger siblings, since they are less likely to contribute to household income, thus increasing the need for child work. The presence of older siblings makes the sharing of work responsibilities more likely, reduces the value of the marginal product of the child, and potentially results in higher household income.

The number of older siblings has the expected negative effect on child work, but this effect is not statistically significant. Conditional on working, each additional older sibling increases the child's chances of reaching secondary school by 4 percent, which mainly reflects a net positive income effect. Conditional on not working, an additional older sibling has a negative but insignificant effect on the transition to secondary school. Surprisingly, the presence of younger siblings only has a significant effect on the probability of making it to secondary school among working children, and this effect is positive.[38] However, the proportion of younger sisters has a negative effect on the child's chances of reaching secondary school, although only significantly so among working

[38] This positive effect may be explained by the combination of younger siblings not competing for education resources and also contributing to household income.

children. This finding may indicate that younger sisters are less likely to contribute to household income than younger brothers, and/or be the result of lower direct and indirect costs of secondary schooling for girls relative to boys as a result of the policy changes that started in 1990.[39]

The presence in the village of a tubewell for drinking water and a modern health facility only have a significant effect on the school outcome for non-working children, where they increase the likelihood of reaching secondary school by 7.8 percent and 9.2 percent, respectively.

Villages further away from the capital of Matlab have a higher incidence of child work, whereas a more diversified village economy decreases the likelihood of child work. Child work is likely to be related with subsistence agriculture, and, in many cases, be the result of shocks which agriculture is prone to. In this context, the capital offers a variety of important services such as a big market and credit institutions. It is also the place to learn about new agricultural technologies, as well as a place with more (and more diversified) employment. More remote and isolated areas are harder to reach for the implementation and supervision of child labor standards and social mobilization against child labor. On the other hand, goods produced at a factory or workshop will generally be sold beyond the village boundaries, and thus are less likely to be as sensitive to shocks at the local level as agricultural products. Furthermore, child work in a factory is more open to reputation and social sanctioning problems than work on the family farm.

Finally, Wald tests show that the estimated ρ_{0w} is statistically significant, the estimated ρ_{1w} is not statistically significant,[40] and both are jointly significant. Furthermore, the hypothesis of $Cov[U_w, U_1 - U_0] = 0$ is rejected. Hence we cannot reject the hypothesis of selection bias due to unobservables, which gives a first indication of the importance of selection in these data.

---

[39] It is worth noting that the number of older and younger siblings both increase the school entry age of the child, whereas the proportion of younger siblings reduces it. The latter is consistent with girls entering school later than boys. See Table A.1 for these results.

[40] However, the estimated ρ_{1w} is large, even if imprecisely estimated.

5.2 Estimated Mean Treatment Parameters

Conditional on the maximum likelihood parameter estimates, we can compute the different mean treatment parameters. The average treatment effect and the treatment effect on the treated are found to be

$$\begin{aligned}\Delta^{ATE}(\overline{x}) &= E[\Delta \mid X = \overline{x}] = -0.1718\ (0.1854) \\ \Delta^{TT}(\overline{x}_w, W = 1) &= E[\Delta \mid X = \overline{x}_w, W = 1] = -0.4244\ (0.0479)\end{aligned}$$

where the numbers in brackets are standard errors. Even though ATE is imprecisely estimated, these results show that, for the entire population of children, working while in primary school has a negative effect on the probability of making the transition to secondary school, but that the loss is far greater for those children who actually select into work. This suggests that children whose combined U_w and Z values make them more likely to work are those who suffer the most from working.

In comparison, as reported in Table 2, the non-parametric estimate of the mean difference in survival rates by work status is 0.20. Furthermore, the estimated ATE and TT from the same model but imposing independence of the error terms across equation (i.e. $\rho_{1w} = \rho_{0w} = 0$) are -0.2357 (0.0476) and -0.2417 (0.0480), respectively.[41] Hence, controlling for observables slightly reduces ATE to -23.57 percent and TT to -24.17 percent. Further controlling for unobservables increases ATE to -17.18 percent and significantly reduces TT to -42.44 percent. Thus controlling for selection on observables and, particularly, unobservables appears to be very important in these data. We now examine more in detail the relation between selection on observables and unobservables and school outcomes.

[41] The estimated parameters of this model are available from the author upon request. In this model, ATE and TT are calculated as

$$\begin{aligned}ATE &= \Phi[\overline{x}\beta_1] - \Phi[\overline{x}\beta_0] \\ TT &= \Phi[\overline{x}_w\beta_1] - \Phi[\overline{x}_w\beta_0]\end{aligned}$$

5.2.1 The Relation Between Selection Into Work and School Outcomes

A central question in this paper is whether those children who suffer the most from working are those most likely to work. We have already noted that TT is more negative than ATE, that is, those children with a higher propensity to work in terms of observables and unobservables are those who have more to lose from working. This suggests that overall children with a higher propensity to work are more likely to be of type $\Delta = -1$ ($S_1 = 0$, $S_0 = 1$), that is, children who would make it to secondary school in the absence of work and would not make it working. In order to examine the contribution of observables and unobservables to this overall effect, it is necessary to relate Δ to $Z\beta_w$ and U_w.

One way to look at the relation between unobservables, U_w, and Δ is by plotting the estimated MTE parameter against different values of $u_w = -z\beta_w$ (see Figure 1). The MTE parameter is decreasing in u_w, that is, in terms of unobservables, those more likely to work are those who suffer the most from working. Another way to make the same point is to inspect the correlations among the unobservables

$$\begin{aligned} Corr[U_w, U_1] &= \rho_{1w} = -0.2683 \\ Corr[U_w, U_0] &= \rho_{0w} = 0.8624 \end{aligned}$$

Thus, in terms of unobservables, the higher the propensity to work the higher the propensity to make it to secondary school in the non-working state, and the lower the propensity to reach secondary school in the working state.[42] Hence, in terms of unobservables, those children with higher values of u_w are more likely to have lower values of Δ, holding constant X and Z, that is, working children are those who would benefit the most from not working. Although it is difficult

[42] This latter effect is, however, smaller and less statistically significant than the former. The finding that working children have unobservable characteristics that are positively associated with the school outcome in the non-working state is very robust to changes in model specification. This relationship could indicate that parents may select for work children whom they judge to be more able to combine school and work responsabilities, because of their greater physical and mental fitness and/or greater self-discipline. This point deserves further future research.

to generalize to rural areas in other parts of the developing world, this result contradicts previous views suggesting that working children may be innately less interested in academic achievement. There is nothing inherently wrong with these working children, they are perfectly able to make it to secondary school in the absence of work.

Determining the effect of $Z\beta_w$ on Δ is a more difficult matter (see AHV 2000). We can, however, try to give a first look at this issue by examining the correlations among the indices ($Z\beta_w$, $X\beta_1$, $X\beta_0$)

$$
\begin{aligned}
Corr[Z\beta_w, X\beta_1] &= -0.4429 \\
Corr[Z\beta_w, X\beta_0] &= -0.4495 \\
Corr[X\beta_1, X\beta_0] &= 0.7608 \\
Corr[Z\beta_w, (X\beta_1 - X\beta_0)] &= -0.1280
\end{aligned}
$$

Thus, unlike the case that arises in the analysis of unobservables, in terms of observables, a higher propensity to work is associated with lower school outcomes in either state. Furthermore, the fourth correlation is negative which suggests that, contrary to the effect of unobservables on Δ, those children who are more likely to work in terms of observables are those who would benefit the least from not working. However, given the size of this correlation, this effect is likely to be small.

Hence, since selection on unobservables is stronger than selection on observables, overall children who are more likely to work are those who suffer the most from working or, alternatively, who would benefit the most from not working. The fact that working children are more sensitive in terms of school progress to a marginal change in work status may reflect the relatively more adverse environment to which these children are exposed.

This overall effect of selection gets reflected in TT being more negative than ATE. This result is also consistent with the comparison of the unconditional difference in survival rates by work status and the mean treatment parameters of the switching probit model.

5.2.2 Heterogeneity in Observables

The degree to which treatment effects vary with observable characteristics can be seen by studying the marginal effect of each observable characteristic on the expected treatment effect. This exercise allows us to identify which groups of children are more likely to suffer from working. Table 6 shows these marginal effects on both ATE and TT.

Given the imprecision with which ATE is estimated, and the finding that work is more harmful for working children, I focus on the marginal effects on TT. Overall the results show that TT varies considerably with observed characteristics. In particular, the prototypical working child who suffers the most from working is a boy who started school late, completed/quit primary before 1992, whose parents are uneducated, and who lived in a household with no modern latrine, no non-business assets, few older siblings, and a high proportion of younger sisters.[43]

5.3 Grade Repetition in Primary

To reinforce the idea that working children have unobservable characteristics that are positively associated with the school outcome, I further consider the effect of working while in school on grade repetition in primary school. However, due to the lack of sufficient variation in the repetition variable,[44] a switching probit model could not be estimated. Instead, a standard bivariate probit model is estimated.[45] As already mentioned, this approach is less flexible than our switching model. In particular, the process generating the school outcome for working and non-working children is only allowed to differ in the constant term. More importantly, the correlation between unobservables in the work equation and unobservables in the school outcome equations for each

[43] It is worth noting that, as it refers to the gender of this prototypical child, the above result is in sharp contrast with the unconditional gender difference in survival rates by work status. Thus, after controlling for selection on observables and unobservables the relative disadvantage of girls turns into relative advantage.

[44] Only 171 children (i.e. 8.75% of the sample) report having repeated some grade in primary school.

[45] This model also takes into account censoring of those children who are still attending primary school.

working state are constrained to be the same.[46]

Furthermore, as noted in the data section, using grade repetition as the school outcome has the problem of treating equally those who failed a grade and quit school and those who did not repeat a grade in primary. Despite the drawbacks associated with the methodology and school outcome used, it is interesting to investigate whether some of the results for the transition to secondary school still hold in this setting.

Table 7.1 and Table 7.2 report, respectively, the coefficient estimates and marginal effects for this model. Being a girl, starting school late, and living in a household with no productive assets and few older siblings are all associated with a higher probability of repeating a grade. Whereas the effect of the father's education has the expected sign, mother's education appears to increase the chances of repeating. The latter is probably a reflection of the problem, mentioned above, with this school outcome variable. In particular, it is likely that mother's education reduces the chances of failing a grade. However, conditional on failing a grade, the education of the mother is likely to increase the probability of persisting in school (i.e. repeating) versus dropping out. If the latter effect is stronger than the former, we then observe an overall positive effect on repetition.

From this model we can only estimate a single treatment parameter, which controls for both selection on observables and selection on unobservables. This mean treatment parameter is just the marginal effect of the work variable on repetition. The estimated treatment effect and its standard error are 0.4118 (0.1892). Thus working while in primary school increases the probability of repeating a grade by 41.18 percent.

In comparison, as reported in Table 3, the non-parametric estimate of the difference in repetition rates by work status is 0.07. Furthermore, the estimated treatment effect from the same model, but imposing independence of the error terms across the repetition and work equations, is

[46] As a result of that, this model rules out the possibility that a child would repeat a grade when not working and would not repeat when working. In any case, this model is not in contradiction with the data, since the above outcome is very unlikely.

0.06354 (0.01715).[47] Thus controlling for observable characteristics does not significantly change the effect of working, but once we control for selection on unobservables the effect of work on repetition increases dramatically. This result indicates that working children have unobservable characteristics that are negatively related with the repetition outcome. The negative estimated correlation coefficient is also consistent with this finding. Again, there is nothing inherently wrong with these working kids, they are likely to progress in school without repetition in the absence of work.

6 Conclusions and Policy Implications

Bangladesh is a country characterized by very poor school performance and a high incidence of child labor. This paper investigates the causes underlying the low levels of school progress in rural Bangladesh, while focusing on the effect of working while in school. To this end, this paper develops a flexible and robust econometric model for evaluating the effect of working while in school on school progress. In particular, the paper estimates a switching probit model for the transition to secondary school, where the switching is determined by endogenous employment status. This approach allows us to recover a variety of mean "treatment" effects, and to assess the relative importance of observables and unobservables in understanding the selection into work and school outcome processes.

The estimated model works well in characterizing the work and school outcome processes. The main result of this paper, however, is that working while in school has a negative impact on the transition to secondary school, but that this effect is significantly more negative for those children who actually select into work. In particular, those children who are more likely to work are those who suffer the most from working or, alternatively, those who would benefit the most from not working. The fact that working children are more sensitive in terms of school progress to a marginal change in work status may reflect the relatively more adverse environment to which

[47] The treatment effect in this model is also the marginal effect of the work variable on repetition. The estimated parameters of this model are available from the author upon request.

these children are exposed.

This finding has important policy implications. In particular, the above result suggests that policies aimed at increasing the rate of transition to secondary school through reductions in the incidence of child labor are both most relevant, given the magnitude of the effect of child work on school progress, and most effective, since most working children are perfectly able to make it to secondary school in the absence of work.

In most cases, the parental decision to send their children to work is the rational response to the environment in which work and schooling decisions are made, and hence effective policies should be aimed at changing this environment. In this context, banning child labor is not likely to produce the intended effects and may even be counterproductive. On the other hand, the reinforcement of the ongoing school incentive schemes, such as the food-for-education program and the secondary school scholarship program for girls, offer a promising route. These programs provide cash or in-kind subsidies to qualifying children that are conditional on school attendance, and are aimed at compensating parents for the foregone income from their children's labor.

Ravallion and Wodon (2000), and Arends-Kunning and Amin (2000) have found, however, that while these programs have reduced the incidence of child labor, this effect only accounts for a small proportion of the total effect of the programs on children's time spent in schooling activities. These results suggest that while these programs are an important step in the right direction, there is a lot of room for improvement in their design and targeting.

The effect of these programs is likely to be limited by supply constraints, so efforts should continue to increase the number of schools and classrooms. These programs should also be accompanied by other policy measures such as improving the quality of education, diversification of the rural economy, and further development of credit and insurance markets. Finally, getting children to start school at an early age is likely to have a big payoff since, on the top of its direct impact on school progress, early schooling reduces the incidence of child labor.

References

Aakvik, A., J.J. Heckman, and E.J. Vytlacil, 2000. Treatment Effects for Discrete Oucomes when Responses to Treatment Vary Among Observationally Identical Persons: An Application to Norwegian Vocational Rehabilitation Programs. *NBER Technical Working Paper* 262, NBER, Cambridge, MA.

Ahlburg, D., B.P. McCall , and R. Assad, 2001. Schooling in Egypt: The Impact of Delayed Entry and Working While in School. Paper presented at the poster session of the 2001 annual meetings of the Population Association of America, March 2001, Washington D.C.

Akabayashi, H., and G. Psacharopoulos, 1999. The Trade-Off Between Child Labor and Human Capital Formation: A Tanzanian Case Study. *Journal of Development Studies* 35, 120-140.

Arends-Kuenning, M., and S. Amin, 2000. The Effects of Schooling Incentive Programs on Household Resource Allocation in Bangladesh. *Policy Research Division Working Paper* 133, Population Council, New York.

Assad, R., D. Levison, and N. Zibani, 2000. Child Work and Schooling in Egypt. Mimeo, Humphrey Institute of Public Affairs, University of Minnesota.

Bangladesh Bureau of Statistics, 1996. *Report on National Sample Survey of Child Labor in Bangladesh 1995-1996*. BBS, Dhaka.

Basu, K., 1999. Child Labor, Cause, Consequence and Cure, with Remarks on International Labor Standards. *Journal of Economic Literature* 37, 1083-1119.

Browning, M., and P.A. Chiappori, 1998. Efficient Intra-Household Allocations: A General Characterization and Empirical Tests. *Econometrica* 66, 1241-1278.

Canagarajah, S., and H. Coulumbe, 1997. Child Labor and Schooling in Ghana. Mimeo, Africa Technical Department, the World Bank, Washington, D.C.

Carrasco, R., 1998. Binary Choice with Binary Endogenous Regressors in Panel Data: Estimating the Effect of Fertility on Female Labor Participation. *FEEM Note di Laboro* 1999.3, Fondazione Eni Enrico Mattei, Italy.

Cigno, A., and F.C. Rosati, 2000. Why do Indian Children Work, and is it Bad for Them?. *IZA Discussion Paper* 115, Institute for the Study of Labor, Bonn, Germany.

Datt, G., and M. Ravallion, 1994. Transfer Benefits from Public-Works Employment: Evidence for Rural India. *Economic Journal 104*, 1346-1369.

Demombynes, G., 2001. Incentive-Oriented Programs and Child Time Allocation: The Case of Progresa. Paper presented at the 2001 annual meetings of the Population Association of America, March 2001, Washington D.C.

Donahoe, D., and M. Tienda, 2000. The Transition from School to Work: Is There a Crisis and What Can Be Done?. In Sheldon Danziger, Jane Waldfogel (eds.), *Securing the Future: Investing in Children from Birth to College*, Russel Sage Foundation, New York.

Duraisamy, M., 2000. Child Labor and Child Work in India. Paper presented at the 2000 World Congress of the Econometric Society, August 2000, Seattle, WA.

Duryea, S., D. Lam, and D. Levison, 2001. Effects of Economic Shocks on Children's Employment and Schooling in Brazil. Paper presented at the 2001 annual meetings of the Population Association of America, March 2001, Washington D.C.

Fauveau, V., 1994. Matlab: Women, Children and Health. ICDDR, Bangladesh.

Grootaert, C., and R. Kanbur, 1995. Child Labor: A review. *World Bank Policy Research Working Paper* No. 1454, the World Bank, Washington, DC.

Grootaert, C., 1998. Child Labor in Cote d'Ivoire: Incident and Determinants., Mimeo, Social Development Department, the World Bank, Washington, DC.

Heady, C., 2000. What is the Effect of Child Labor on Learning Achievement? Evidence from Ghana. *Innocenti Working Paper* 79, UNICEF, Innocenti Research Center, Florence, Italy.

Heckman, J., 1976. The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement* 5, 475-492.

Heckman, J., and J. Smith, 1998. Evaluating the Welfare State. In S. Strom (eds.), *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial*, Econometric Society Monograph Series, Cambridge University Press, Cambridge, MA.

ILO, 1996. *Child Labor: Targeting the Intolerable*. Report VI, Part 1, ILO, 85th Session, Geneva.

Jacobi, H.G., and E. Skoufias, 1997. Risk, Financial Markets, and Human Capital in a Developing Country. *Review of Economic Studies* 64, 311-335.

Khandker, S.R., 1996. Education Achievements and School Efficiency in Rural Bangladesh. *World Bank Discussion Paper* 319, the World Bank, Washington, D.C.

Levison, D., K. Moe, and F. Knaul, 2001. Youth Education and Work in Mexico. *World Development* 29, 167-188.

Patrinos, H., and G. Psacharopoulos, 1997. Familiy Size, Schooling and Child Labor in Peru: An Empirical Analysis. *Journal of Population Economics* 10, 387-405.

Rahman, O., J. Menken, A. Foster, M.N. Khan, C.E. Peterson, R. Khun, P. Gertler, ICD-DRB, Mitra and Associates, 1999. The Matlab Health and Socioeconomic Survey: Overview and User's Guide", DRU-2018/1, RAND, Santa Monica CA.

Rahman, W., 1997. *Child Labor Situation in Bangladesh: A Rapid Assessment*. ILO-UNICEF, Dhaka, Bangladesh.

Ravallion, M., and Q. Wodon, 2000. Does Child Labor Displace Schooling? Evidence on Behavioral Responses to an Enrollment Subsidy. *Economic Journal* 110, C158-C175.

Ridao-Cano, C., 2001. Child Labor and Schooling in a Low Income Rural Economy. Paper presented at the 2001 annual meetings of the Population Association of America, March 2001, Washington D.C.

Rosenzweig, M.R., and R. Evenson, 1977. Fertility, Schooling, and the Economic Contribution of Children in Rural India: An Econometric Analysis. *Econometrica* 45, 1065-1079.

Schultz, T.P., 2001. What are the Returns to Poverty Reduction: Evaluation of Progresa in Mexico. Paper presented at the 2001 annual meetings of the Population Association of America, March 2001, Washington D.C.

Skoufias, E., 1994. Market Wages, Family Composition and the Time Allocation of Children in Agricultural Households. *Journal of Development Studies* 30, 335-360.

Smith, R. J., and R.W. Blundell, 1986. An Exogeneity Test for a Simultaneous Equation Tobit Model with Application to Labor Supply. *Econometrica* 54, 679-686.

Table 1. Children by Work Status

	All		Girls		Boys	
	Count	Percent	Count	Percent	Count	Percent
Working	560	27.69	288	27.38	272	27.97
Not working	1394	72.31	697	72.62	697	72.03
Total	1954	100	985	100	969	100

Notes: Based on weighted data.

Table 2. Kaplan-Meier Estimates of Survival Rates to Secondary School

	All		Girls		Boys	
	Survival	Std.Error	Survival	Std.Error	Survival	Std.Error
All	0.7078	0.0105	0.6774	0.0152	0.7386	0.0143
Working	0.5636	0.0216	0.5059	0.0305	0.6251	0.0303
Not working	0.7638	0.0115	0.7462	0.0167	0.7812	0.0158

Notes: Based on unweighted data. The mean difference in survival rates, all children, by work status is 0.20, which is significantly different from zero according to the likelihood-ratio test ($\chi^2(1) = 56.80$ (0.000)). The corresponding mean differences for girls and boys are 0.2403 ($\chi^2(1) = 38.53$ (0.000)) and 0.1561 ($\chi^2(1) = 18.71$ (0.000)), respectively.

Table 3. Kaplan-Meier Estimates of Repetition Rates

	All		Girls		Boys	
	Repetition	Std.Error	Repetition	Std.Error	Repetition	Std.Error
All	0.1102	0.0072	0.1334	0.0110	0.0868	0.0091
Working	0.1613	0.0158	0.1864	0.0233	0.1347	0.0210
Not working	0.0900	0.0077	0.1118	0.0120	0.0684	0.0096

Notes: Based on unweighted data. The mean difference in repetition rates, all children, by work status is 0.0713, which is significantly different from zero according to the likelihood-ratio test ($\chi^2(1) = 135.52$ (0.000)). The corresponding mean differences for girls and boys are 0.0746 ($\chi^2(1) = 54.31$ (0.000)) and 0.0663 ($\chi^2(1) = 88.28$ (0.000)), respectively.

Table 4. Means and Standard Deviations of Characteristics

	All children		Working children		Non-working children	
	Mean	Std.Dev	Mean	Std.Dev	Mean	Std.Dev
Girl	0.467	0.499	0.462	0.499	0.469	0.499
School entry age	7.719	1.611	7.999	1.634	7.607	1.588
Current age	19.20	3.045	19.27	3.099	19.17	3.012
In school in or after 1992	0.300	0.458	0.295	0.453	0.302	0.458
Mother's education	1.481	2.424	1.067	2.059	1.647	2.536
Father's education	3.907	3.747	3.318	3.536	4.143	3.801
Modern latrine	0.309	0.462	0.295	0.456	0.314	0.464
Log(household assets)[1]	10.95	1.428	10.84	1.557	11.00	1.374
Household assets missing	0.151	0.358	0.162	0.369	0.146	0.349
Cultivating household	0.757	0.429	0.808	0.394	0.736	0.440
Owns farm land	0.800	0.400	0.833	0.372	0.787	0.407
Owns non-farm business	0.439	0.496	0.508	0.495	0.412	0.485
Older siblings (number)	2.514	2.131	2.469	1.937	2.532	2.205
Younger siblings (number)	1.258	1.012	1.309	1.047	1.237	0.982
Younger sisters (%)[2]	0.470	0.489	0.467	0.466	0.471	0.498
Health facility in village[3]	0.697	0.488	0.702	0.488	0.695	0.488
Tubewell in village[4]	0.827	0.402	0.847	0.384	0.818	0.409
Distance to Matlab capital[5]	5.727	4.810	6.500	5.628	5.414	4.395
Industry in village	0.639	0.510	0.633	0.514	0.642	0.509
Village outside Matlab	0.087	0.273	0.081	0.273	0.089	0.283

Notes: Based on weighted data. (1) Mean defined for non-missing observations. (2) Mean defined for observation with positive number of younger siblings. (3) Means of village characteristics defined for children whose village of residence during primary was in Matlab, since only those villages were surveyed. (4) Out of the villages that currently report tubewell as the main source of drinking water. (5) In Kms.

Table 5.1. Switching Probit Model for Child Work and the Transition to Secondary School (Coefficient Estimates)

	Work equation		Transition to secondary equations			
			Working state		Non-working state	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Intercept	-0.706	-1.647	0.246	0.240	-2.115	-3.796
Girl	0.022	0.317	0.276	1.761	-0.080	-0.964
School entry age	0.105	4.506	-0.211	-3.931	-0.121	-3.992
Current age	-0.018	-1.248	0.046	1.622	0.077	4.198
In school in or after 1992	-0.243	-2.352	0.710	3.316	0.793	5.826
Mother's education	-0.041	-2.645	0.126	3.628	0.102	4.498
Father's education	-0.017	-1.759	0.040	2.008	0.041	3.232
Modern latrine	0.004	0.061	0.484	2.880	0.369	4.104
Log(household assets)	-0.067	-2.287	-0.010	-0.175	0.102	2.394
Household assets missing	-0.459	-1.349	-0.963	-1.369	0.688	1.461
Cultivating household	0.230	2.737				
Owns farm land	0.192	2.084	0.191	0.926	0.286	2.867
Owns non-farm business	0.258	3.920	0.267	1.466	0.202	2.488
Older siblings (number)	-0.026	-1.582	0.122	3.185	-0.016	-0.841
Younger siblings (number)	0.009	0.278	0.245	3.204	0.069	1.611
Younger sisters (%)	-0.059	-0.729	-0.482	-2.652	-0.142	-1.475
Health facility in village	0.082	1.107	-0.002	-0.016	0.482	5.571
Tubewell in village	0.068	0.784	-0.068	-0.381	0.393	3.971
Distance to Matlab capital	0.033	4.778				
Industry in village	-0.201	-2.983				
Village outside Matlab	0.064	0.384	-0.796	-2.269	0.875	4.720
ρ_{1w}	-0.268	-0.586				
ρ_{0w}	0.862	9.063				
Log-L	-1952.38					
Wald $\chi^2(20)$	118.31 (0.000)					

Notes: Based on weighted data. Weights are appropriately scaled so that they sum of observations in each working state. $Cov[U_w,U_1-U_0] = -1.1307$. The Wald test rejects the null hypothesis $Cov[U_w,U_1-U_0] = 0$ ($\chi^2(1) = 5.86$ (0.015)). The Wald test also shows that ρ_{1w} and ρ_{0w} are jointly significant ($\chi^2(2) = 82.59$ (0.000)).

Table 5.2. Switching Probit Model for Child Work and the Transition to Secondary School (Marginal Effects)

	Work equation		Transition to secondary equations			
			Working state		Non-working state	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Girl	0.007	0.316	0.090	1.411	-0.014	-0.951
School entry age	0.033	4.443	-0.069	-2.832	-0.020	-3.824
Current age	-0.006	-1.242	0.015	1.450	0.013	3.815
In school in or after 1992	-0.073	-2.436	0.210	2.259	0.111	5.903
Mother's education	-0.013	-2.622	0.041	2.314	0.017	4.131
Father's education	-0.005	-1.742	0.013	1.762	0.007	3.012
Modern latrine	0.001	0.061	0.149	1.665	0.057	3.969
Log(household assets)	-0.021	-2.479	-0.003	-0.170	0.017	3.017
Household assets missing	-0.126	-1.666	-0.356	-1.207	0.085	2.613
Cultivating household	0.069	2.862				
Owns farm land	0.058	2.222	0.065	0.759	0.054	2.373
Owns non-farm business	0.081	3.929	0.087	1.051	0.034	2.394
Older siblings (number)	-0.008	-1.573	0.040	1.999	-0.003	-0.844
Younger siblings (number)	0.003	0.279	0.080	1.816	0.012	1.573
Younger sisters (%)	-0.018	-0.734	-0.158	-1.653	-0.024	-1.422
Health facility in village	0.025	1.119	-0.001	-0.016	0.093	4.653
Tubewell in village	0.021	0.795	-0.022	-0.391	0.078	3.336
Distance to Matlab capital	0.010	4.642				
Industry in village	-0.064	-2.931				
Village outside Matlab	0.020	0.377	-0.297	-1.916	0.091	6.058

Notes: See text for details on the calculation of marginal effects. Marginal effects measure the percentage increase in the probability that the appropriate outcome is equal to 1 resulting from a 1 unit increase in a variable (for continuous variables) and a change from 0 to 1 in the value of the variable (for dummy variables). In the case of log(household assets), the marginal effect measures the change in probability resulting from a 1 percent increase in household assets.

Figure 1. Estimated Marginal Treatment Effects ($MTE(\bar{x}, u_w = -z\beta_w)$ **)**

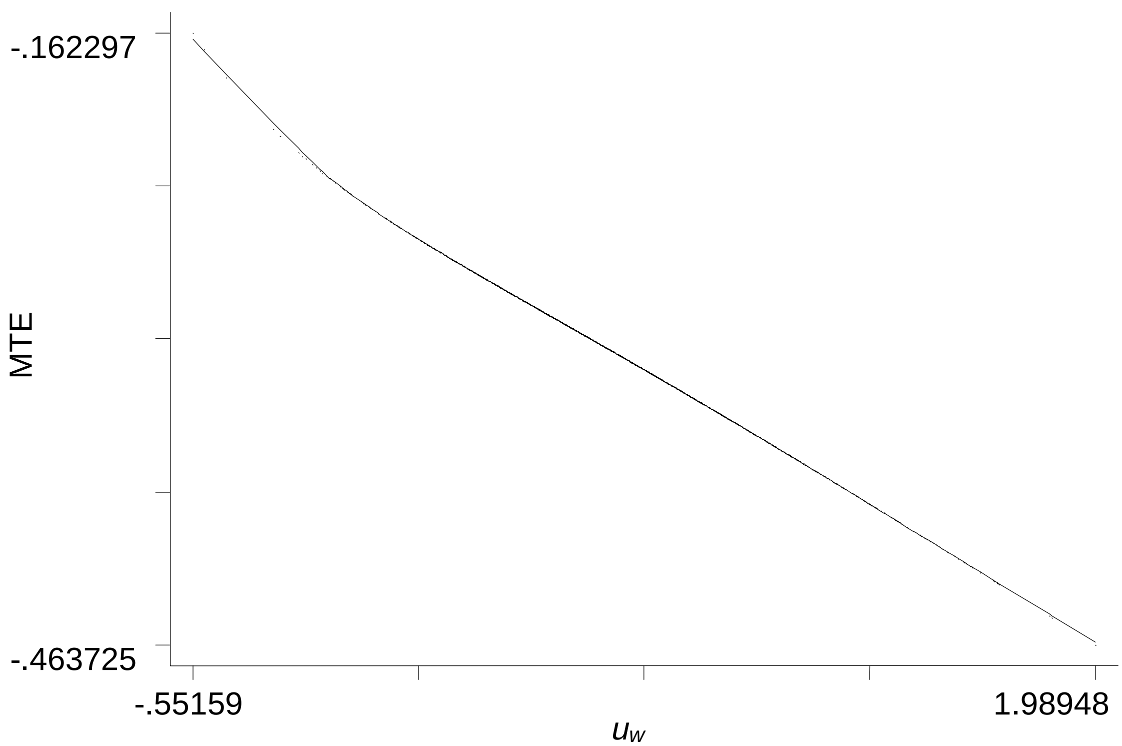

Table 6. Marginal Effects of Model Covariates on the Average Treatment Effect (ATE) and the Treatment Effect on the Treated (TT)

	ATE		TT	
	Coeff.	t-ratio	Coeff.	t-ratio
Girl	0.104	1.579	0.113	1.810
School entry age[1]	-0.014	-1.190	-0.049	-1.195
Current age	-0.005	-0.725	0.001	0.037
In school in or after 1992	0.099	1.039	0.252	3.408
Mother's education[2]	0.004	0.544	0.025	0.791
Father's education[3]	-0.000	-0.009	0.006	0.441
Modern latrine	0.092	1.011	0.189	3.264
Log(household assets)	-0.018	-0.829	-0.032	-0.831
Household assets missing	-0.441	-1.481	-0.408	-1.956
Cultivating household			0.019	0.549
Owns farm land	0.011	0.124	0.094	1.291
Owns non-farm business	0.053	0.631	0.128	2.337
Older siblings (number)	0.022	1.205	0.050	3.131
Younger siblings (number)	0.028	0.985	0.085	2.406
Younger sisters (%)	-0.054	-0.908	-0.171	-2.137
Health facility in village	-0.093	-1.688	0.005	0.084
Tubewell in village	-0.101	-1.622	-0.022	-0.317
Distance to Matlab capital			0.003	0.580
Industry in village			-0.016	-0.535
Village outside Matlab	-0.388	-2.491	-0.309	-2.561

Notes: See text for details on the calculation of these marginal effects. Because of the way the mean parameters are defined, a positive marginal effect means a reduction in the negative effect of working while in school. (1) Starting 1 year late versus starting on time decreases TT by –0.015 (t-ratio = -1.989). (2) Having a mother with one year of education versus none increases TT by 0.046 (t-ratio = 3.335). (3) Having a father with 1 year of education versus none increases TT by 0.015 (t-ratio = 1.795).

Table 7.1. Bivariate Probit Model for Work and Grade Repetition (Coefficient Estimates)

	Work equation		Repetition eq.	
	Coeff.	t-ratio	Coeff.	t-ratio
Intercept	-0.664	-1.607	-1.692	-3.053
Work while in school			1.548	4.463
Girl	0.010	0.878	0.195	2.225
School entry age	0.108	4.682	0.111	3.094
Current age	-0.019	-1.305	-0.042	-1.981
In school in or after 1992	-0.256	-2.461	0.139	1.043
Mother's education	-0.042	-2.704	0.043	2.253
Father's education	-0.017	-1.735	-0.023	-1.837
Log(household assets)	-0.073	-2.493	0.039	0.999
Household assets missing	-0.532	-1.547	0.457	1.015
Cultivating household	0.194	2.291		
Owns farm land	0.250	2.695	-0.222	-2.154
Owns non-farm business	0.267	4.046	-0.170	-1.946
Older siblings (number)	-0.028	-1.819	-0.042	-2.051
Tubewell in village	0.095	1.098	-0.310	-2.968
Distance to Matlab capital	0.025	3.384		
Industry in village	-0.115	-1.669		
Village outside Matlab	0.008	0.052	-0.134	-0.721
ρ	-0.677	-3.764		
Log-L	-1610.05			
Wald $\chi^2(16)$	103.15 (0.000)			

Notes: Based on weighted data. ρ is the estimated correlation between unobservables in the work equation and unobservables in the grade repetition equation.

Table 7.2. Bivariate Probit Model for Work and Grade Repetition (Marginal Effects)

	Work equation		Repetition eq.	
	Coeff.	t-ratio	Coeff.	t-ratio
Work while in school			0.412	2.177
Girl	0.003	0.153	0.067	1.871
School entry age	0.033	4.610	0.039	6.718
Current age	-0.006	-1.299	-0.015	-2.902
In school in or after 1992	-0.075	-2.559	0.048	0.859
Mother's education	-0.013	-2.676	0.015	1.968
Father's education	-0.005	-1.716	-0.010	-1.696
Log(household assets)	-0.022	-2.735	0.014	1.002
Household assets missing	-0.139	-2.015	0.144	1.158
Cultivating household	0.057	2.386		
Owns farm land	0.072	2.942	-0.074	-1.817
Owns non-farm business	0.082	4.047	-0.060	-1.672
Older siblings (number)	-0.010	-1.801	-0.014	-1.860
Tubewell in village	0.028	1.121	-0.101	-2.269
Distance to Matlab capital	0.010	3.338		
Industry in village	-0.035	-1.651		
Village outside Matlab	0.002	0.051	-0.048	-0.697

APPENDIX

Table A.1. Ordered Probit Model for School Entry Age

	School entry age equation	
	Coeff.	t-ratio
Girl	0.333	6.270
Current age	-0.063	-6.949
Mother's education	-0.091	-7.654
Father's education	-0.022	-2.958
Mother's father richer	-0.107	-1.928
Father dead by age six	0.692	5.956
Latrine	-0.249	-4.530
Log(household assets)	0.005	0.220
Household assets missing	-0.107	-0.389
Cultivating household	0.201	3.033
Owns farm land	-0.081	-1.153
Owns non-farm business	-0.003	-0.068
Older siblings (number)	0.038	3.067
Younger siblings (number)	0.221	7.972
Younger sisters (%)	-0.137	-2.170
Health facility in village	-0.365	-6.527
Tubewell in village	-0.074	-1.214
Distance to Matlab capital	0.006	0.987
Industry in village	-0.252	-4.720
Primary school in village	-0.224	-3.648
Village outside Matlab	-0.298	-2.263
Cutoff 1	-2.198	-7.393
Cutoff 2	-1.410	-4.758
Cutoff 3	0.707	2.389
Cutoff 4	-0.204	-0.691
Cutoff 5	0.514	1.734
Log-L	-3033.65	
Likelihood ratio $\chi^2(20)$	446.57 (0.000)	

Notes: Based on weighted data. The number of observations in some of the school entry ages was very small, so the dependent variable was redefined and rescaled. It is equal to 0 for children who started at age 5 or 6 (478), then 1 for those who started at age 7 (499), 2 for starting age 8 (439), 3 for starting age 9 (228), 4 for starting age 10 (206), and 5 for starting ages 11 and beyond (104). Village-level information refers to the time until the child was six years of age (the official school entry age).

Table A.2. Kaplan-Meier Estimates of Survival Rates to Each Grade

	All		Girls		Boys	
	Survival	Std.Error	Survival	Std.Error	Survival	Std.Error
Primary						
Grade 2	0.9882	0.0024	0.9848	0.0039	0.9917	0.9546
Grade 3	0.9601	0.0044	0.9655	0.0058	0.9546	0.0067
Grade 4	0.9091	0.0065	0.9134	0.0090	0.9047	0.0094
Grade 5	0.8621	0.0078	0.8668	0.0152	0.8573	0.0113
Junior secondary						
Grade 6	0.7078	0.0105	0.6774	0.0152	0.7386	0.0143
Grade 7	0.6703	0.0109	0.6409	0.0157	0.7000	0.0150
Secondary						
Grade 8	0.6240	0.0114	0.5949	0.0163	0.6533	0.0158
Grade 9	0.5625	0.0120	0.5078	0.0172	0.6171	0.0165
Grade 10	0.5223	0.0125	0.4553	0.0179	0.5888	0.0171
Higher secondary						
Grade 11	0.4025	0.0141	0.3282	0.0194	0.4760	0.0199
Grade 12	0.3220	0.0162	0.2639	0.0219	0.3796	0.0237
Higher education	0.2505	0.0222	0.1783	0.0323	0.3155	0.0309

Notes: Based on unweighted data. The likelihood ratio test for the null hypothesis of no gender differences in survivor functions is rejected ($\chi^2(1) = 13.75$ (0.000)).