

The Dynamics of School and Work in Rural Bangladesh

José Canals and Cristóbal Ridao-Cano¹

Department of Economics

University of Colorado at Boulder

December 2002

¹ Corresponding author: Jose J. Canals-Cerda, Department of Economics, University of Colorado at Boulder, Campus Box 256, Boulder, Colorado 80309-0256. Phone: (303) 492-7869, Fax: (303) 492-8960. E-mail: jose.canals-cerda@colorado.edu.

The Dynamics of School and Work in Rural Bangladesh

Abstract

This paper investigates the causes underlying the poor school performance of children in rural Bangladesh, while focusing on the effect of work on school progress. To this end, a dynamic switching model is presented for the sequence of school and work outcomes up to the end of secondary school, where the switching in each school level considered is determined by the endogenous work sequence up to that level. This approach allow us to characterize the full sequence of school and work choices of children, and to evaluate the dynamic effects of work on schooling. We find that work has a negative and sizable effect on school progress for the entire population, as well as for all the subpopulations considered, including the different groups of working children. We are also able to identify at each school level the observable and unobservable characteristics of working children relative to those of non-working children. This characterization and the magnitude of the estimated effects of work have important policy implications.

JEL CLASSIFICATION: C35, D12, O15

KEYWORDS: Dynamics of school and work, dynamic evaluation, selection

1 Introduction

Although labor force participation rates for school-age children (i.e. aged 5-14) have been declining over time, recent International Labor Organization 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, particularly in rural areas. Recent estimates (Filmer, 1999) indicate that among children aged 15-19, 27.5 percent have never attended school. Among those who attended school, 36 percent started school later than 6 years of age (the official school entry age), 69.2 percent had reached secondary school, and less than 20 percent had completed secondary school. Furthermore, estimates based on the Child Labor Survey 1995-96 (Bangladesh Bureau of Statistics, 1996) indicate that 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.

The literature on child labor and schooling in developing countries has been rapidly expanding in recent years.² Most studies either look at child labor and schooling separately or focus on one or the other, particularly schooling. Most of the studies looking at child schooling focus on static measures (e.g. school enrolment in a particular year) with only a few papers looking at the dynamics of schooling (Lillard and Willis, 1994; Sawada and Lokshin, 2001). To our knowledge, all of the studies looking at child labor focus on static measures, such as work participation or hours of work in a particular year, or, at most,

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

monthly variations within a year in these measures.

A number of 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 make statements about the degree of substitution between child work and schooling based on the correlation between observables and/or unobservables in the work and school equations.³ Some studies have also examined the impact of work on schooling indirectly by considering the response of work and schooling to exogenous changes in the price of schooling caused by school incentive programs.⁴

Finally, a few papers have provided direct estimates of the effect of work 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). However, in most cases the possible self-selection into work is not accounted for, thereby clouding the interpretation of such estimates as structural effects. Furthermore, the studies that account for the endogeneity of work do so within a restrictive framework that allows them to recover a single average estimate of the effect of work.⁵ From a policy perspective, the important question is not so much to find out what the average effect of work is but to identify the children who are most affected by it in terms of school progress, so that informed policies can be developed.

³ See, for example, Canagarajh and Coulumbe (1997), Grootaert (1998), Skoufias (1994), Duraysamy (2000), and Ridao-Cano (2001).

⁴ Examples are the Food-for-Education scheme in rural Bangladesh (Ravallion and Wodon, 2000) and Progreso in rural Mexico (Schultz, 2001).

⁵ This is also true for the growing number of studies examining the effect of working while in high school on a variety of school outcomes in the U.S. See, for example, Eckstein and Wolpin (1999) and Dagenais *et al.* (2001).

This paper investigates the causes underlying the poor school performance of children in rural Bangladesh, while focusing on the effect of work on school progress. To this end, a dynamic switching model is presented for the sequence of school and work outcomes up to the end of secondary school, where the switching in each school level is determined by the endogenous work sequence up to that level. This approach allow us to characterize the full sequence of school and work choices of children, and to evaluate the dynamic effects of work on schooling.

We extend the existing literature in a number of ways. First, to our knowledge this is the first paper in the context of developing countries to analyze the joint dynamics of school and work. Second, the dynamic structure of our model allows us to extend some of the static concepts from the program evaluation literature to a dynamic context.

The model provides a good description of the dynamics of school and work for the children in the sample. The model is able to capture two types of interrelated selection processes, namely selection into work and selection into school level. The main result of this paper is that work has a negative and sizable effect on school progress for the entire population, as well as for all the subpopulations considered, including the different groups of working children. Furthermore, the effect of work becomes more negative the earlier in life an individual is exposed to work. We are also able identify at each school level the observable and unobservable characteristics of working children relative to those of non-working children. This characterization and the magnitude of the estimated effects of work have important policy implications.

The rest of the paper is organized as follows. Section 2 describes the data used in this study. Section 3 presents the dynamic model of school and work. Section 4 develops the framework for evaluating the dynamic effects of work on school progress. Section 5 presents the general results of the model, analyzes the dynamic effects of work and discusses the effects of a variety policies. Section 6 concludes and suggests some policy implications.

2 Data Description

The data for the analysis come from the 1996 Matlab Health and Socio-Economic Survey (MHSS). 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.⁶

The present analysis focuses on the school and work experiences, up to the end of secondary school, of individuals who 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 on their own, thus avoiding the typical sample selection bias arising from just considering those in the former group.

⁶ As a result, weights are needed in the analysis of these data to correct for the non-random sample distribution. 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/.

The reason for setting the lower age limit at 15 is twofold. First, by using the sample of young adults (15 years of age or older), the information on the individuals is as reported by the same individuals and not their parents. Second, starting primary school at the official age, all the individuals in the sample except those aged 15 would have had the chance to complete secondary school. The reason for setting the upper age limit at 25 is threefold. 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. Finally, we want to relate as much as possible our 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 after dropping a few observations with missing values in the key variables, the sample used for the empirical analysis contains 2489 individuals, 113/684 of which were still in primary/secondary school at the time of the survey.

The MHSS contains detailed information on education histories including the school entry age, school exit age, grades attended and completed, grade repetition. This paper looks at the school and work outcomes in three “school” levels: school entry, primary school and secondary school. The school outcome in the entry level looks at whether school entry occurred at each age, starting from age 5 (the first reported entry age) up to age 14, beyond which the child is no longer of primary-school age and thus he or she is assumed to be no

longer at risk of entering school.⁷ The school outcome in the primary level looks at whether a child reached secondary school in each possible time interval, starting from 5 years, which is the minimum number of years required to reach secondary from school entry. For those individuals still attending primary school, the school outcome looks at each of the years in which the child could have not reached secondary school, being this observation censored after the last year in which the child could have not reached secondary school. The school outcome in the secondary school level is constructed the same way, but for secondary school completion.

A few points are worth noting. First, the information on the timing of the school events in the primary and secondary school levels is based on the number of repetitions in each level. Second, the maximum time to reach and complete secondary school in the sample is 10 years. Third, the school outcome in the primary/secondary school level for individuals who dropped out of school before reaching/completing secondary school is zero for all years considered.

Fourth, the consideration of the timing of schooling events allows us to make use of the information on censored observations that otherwise could not be used. This is particularly important when we estimate the effect of work on schooling, as part of the reason why some children are still in school may lie in the lack of adequate school progress, which may in turn be affected by their work status. If this is the case, then the estimated effects of work on the probability of reaching/completing secondary school may be biased downward to the extent

⁷ We do not observe any children entering school after age 14.

that a significant proportion of these children will never reach/complete secondary school.

Fifth, for the school outcome in the primary school level we choose to focus on whether the child who entered primary school reached secondary school, instead of whether he or she completed primary school, because the transition from the last grade in primary to the first grade in secondary is the single most important turning point in the Bangladeshi school system, particularly for girls.

It is widely recognized that not all work is necessarily detrimental for child schooling, although the question of how detrimental is an empirical one. The key is to identify the kind of work that can potentially interfere with a child's schooling. The MHSS contains retrospective information on the age at which each individual started performing productive work, which we use to construct the work status variable in the entry level (i.e. work before school entry age). In particular, an individual is considered to be in the working state at a particular age at which school entry is considered if he or she was working at the prior age.⁸ The MHSS also contains retrospective information on whether an young adult performed productive work while he or she was attending each school level (i.e. primary, secondary, higher secondary, and higher education), which we use to define the work status variable in the primary and secondary school levels.

All these pieces of information are reported by the young adult, thus minimizing the typical under-reporting when it is the mother or the father who provides this information. Individuals who consistently performed some kind of productive work before school entry or

⁸ Since the school year starts in January, if the individual started working at the same age he or she entered school, he or she is considered not to have worked before school entry.

while in primary or secondary school are those most likely to report work according to the above definitions, which are exactly the individuals that we are interested in. However, the definition of work as productive work ignores household chores such as caring for younger siblings which, as Levison et al. (2001) point out, is likely to underestimate the amount of work carried out by girls, and thus the role of work as a potential impediment for their schooling. In our sample, all the children who start working in a given level continue working in subsequent levels provided they reach those levels. Finally, it is worth noting that the nature and motivation for work in each level is likely to be different. For example, while the child may have little to say in the decision to work before or during primary school, he or she is likely to have a greater role in the decision to work during secondary school.

We choose to focus on the sequence school and work outcomes up to the end of secondary school for two main reasons. First, we are mainly interested in the school and work experiences of children. To this extent, a child starting school at the official age of 6 would, in the absence of school delay, complete secondary school by age 15, which defines the beginning of adulthood in Bangladesh. Second, we are mainly interested in the effect of work on the acquisition of basic skills needed in the labor market and in life, skills that are provided by basic education which is, in turn, delivered by primary and secondary education.⁹

Table 1 shows the numbers and proportions of children by work sequence in each level, where the work sequence in a given level is defined by the work status in that level and the work status in previous levels. The first two columns show these figures for the row data,

⁹ In addition, the number of individuals in the sample pursuing higher education levels is very small.

while the figures in the last two columns are adjusted for censoring. As regards censoring, it is assumed that if the individual is attending the last grade of the school level considered and reports no work, then he or she is no longer at risk of working during that level, while he or she is considered to be at risk if attending some other grade in that level.

The figures in Table 1 show that the proportion of working children is high and increasing by school level. This is not surprising since, other things being equal, older children are not only more able to do work but they are also expected to contribute more to household income. Very few of those who worked in the entry level entered school. The high proportion of working children in the primary school level may be partly explained by the short duration of the school day during primary school, which allows children to combine school and work, particularly 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 altogether. Furthermore, this ability to combine school and work diminishes as the child moves to higher grades, where the required schooling time is greater.¹⁰

Table 2 reports the transition rates associated to the school outcome in each level by work sequence. Individuals still attending a school level only contribute to the estimation sample for that level if they could not have experienced the schooling event in 5 years, and do so in the work sequence reported at the time of the survey. As expected, the greater the difference in work intensity between two work sequences the greater the difference in the

¹⁰ This is particularly so in moving from primary school to secondary school where not only does the required schooling time increase, but also the chances of having a secondary school nearby are lower.

probability of experiencing the school event between them.

2.1 Model Covariates

For the model to be presented in the next section, we choose a parsimonious specification for the observables determining the work and school outcome processes. These observable characteristics include a set of child, parental and community characteristics that are, to a large extent, relevant to the period when the child was in each school level.

Child characteristics that are common to the three levels include sex and age at the time of the survey.¹¹ We also include endogenous predetermined variables in the primary and secondary school levels: school entry age in both levels, and number of grade repetitions in the secondary school level. In addition, the primary and secondary school levels include policy variables. In the primary school level we include an indicator for whether the child was in primary school in or after 1992, the year in which compulsory primary education was introduced,¹² and a gender-specific indicator for whether the child was in primary school when free tuition for girls in secondary school grades 6-8 was implemented (i.e. 1990).¹³ In the secondary school level we include an indicator for whether the child was in grades 6-8 when free tuition for girls in secondary school grades 6-8 was in place.

Parental characteristics include years of schooling of the mother and the father of the

¹¹ The age at the time of the survey captures the overall time trend in the work and school outcomes.

¹² More generally, 1992 marked an important turning point in education policy. The policies changes affected different dimensions of the shadow price of schooling, including preferences for school and work, as well as direct and indirect costs of schooling.

¹³ The idea is that the expectation of lower direct cost in secondary school increases the incentive to make the transition to secondary school. This policy change may have also affected the primary school outcome of boys, as the cost of educating girls relative to boys decreases, and household resources are released.

child. Household level variables have several dimensions. Household demographics are summarized by the number of younger siblings and the number of older siblings when the child was 6 years of age, for the entry level, and at the time the child started primary and secondary school, for the primary and secondary levels, respectively. 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 either the time of the survey if parents are alive or 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. In any case, if the 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 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 on their own.¹⁴

Finally, a variety of village level variables are included, such as the presence of a tubewell for drinking water, presence of a modern health facility, village economy diversification (i.e. presence in the village of any mill, factory or workshop), distance to the capital of Matlab, and the presence of primary and secondary schools. Village level variables refer to the period when the child was 6 years of age (for the entry level), and the period prior to completion of or drop out from primary and secondary school (for the primary and secondary levels, respectively).¹⁵ The indicators for the presence of a primary and secondary school in the village are included in the entry and primary levels, but only the latter is included in the secondary level.

The work equations include, in addition, an indicator of whether the household cultivated land (own land, rented or sharecropped) around the time the child was 6 years of age (for the entry level), and around the time the child was in primary and secondary school (for the primary and secondary levels, respectively). This variable is constructed on the basis of the current cultivation status of the household, and the retrospective information on parental occupation.¹⁶ Table 3 reports means and standard deviations of the above variables by

¹⁴ This information is available from the migration history of each individual in the sample.

¹⁵ The village where the child resided during each level is obtained from the migration history of each child.

¹⁶ This variable proved to have a significant effect on work but not on school progress conditional on work

school level and work sequence.

3 A Dynamic Model of School and Work

The empirical framework is based on a traditional schooling-transition model (Mare 1980) augmented to take into account the decision to work at each period.¹⁷ We focus on the most important school transitions in the education system of Bangladesh: school entry, transition from primary to secondary school and completion of secondary school. At each period the household decides on the schooling and work status of the child conditional on current and past information.¹⁸ Since the information set at each period includes the work history of the child up to that period, the potential school outcomes at each period are not only defined by the working state in that period but also by the working states in previous periods. Furthermore, the model specification allows for the existence of observable and unobservable, to the econometrician, individual specific heterogeneity.

Heckman and Cameron (1998) have shown that the schooling-transition model can be rationalized based on an economic model of choice. Their main criticism of the model is that it requires myopic behavior on the part of the individuals. As they indicate, the assumption of myopic decision making implies that: “individuals ignore the potential value of future

status. As it is shown in the next section, this variable is not required for identification. However, we gain in identification power to the extent that it is a valid exclusionary restriction.

¹⁷ As mentioned in section 2, a period in the entry level is measured in years of age, starting at age 5 and ending at age 14. In the primary and secondary levels, a period is measured in years to reach secondary school and years to complete secondary school, respectively, starting from 5 years.

¹⁸ A household is viewed as a setting in which individuals with different preferences and bargaining power interact.

shocks and act as though today’s shock determines the value of all future shocks.” We agree that this is an important criticism but consider that it is an acceptable restriction given the limitations of the data.

Based on the schooling-transition model, we develop a dynamic switching model for the sequence of school and work outcomes described in section 3, where the switching in each school level is determined by the endogenous work sequence up to that level. In particular, the working state in the entry level is defined by the age prior to the age at which school entry is considered, and defined by level in the primary and secondary levels.¹⁹

Let t represent the period in which the school event in level $k = \{e, p, s\}$ is considered, where e represents school entry, p represents primary school, and s represents secondary school. Also, let the working status in each (t, k) be represented by $W_{tk} \in \{0, 1\}$ and denote the work sequence up to k (inclusive) by H_k . As noted in section 2, in our sample once a child starts working he or she works thereafter. Hence the set of possible work sequences is

$$H_e \in \{0, 1\}; H_p \in \{(0, 0), (0, 1), (1, 1)\}$$

$$H_s \in \{(0, 0, 0), (0, 0, 1), (0, 1, 1), (1, 1, 1)\}$$

For each child i who has reached level k but has not experienced the school event by period t , we consider the set of potential school outcomes associated with each work sequence:

¹⁹ However, to keep notation consistent throughout the model, we index work outcomes by the period in which the school outcome in a given level is considered.

$S_{tk}(H_k) \in \{0, 1\}$. Thus the potential school outcomes at (t, k) are not only dependent on the working state at (t, k) but also on the work history prior to k . We assume that W_{tk} and $S_{tk}(H_k)$ are generated by the following latent index structure

$$\begin{aligned} W_{tki} &= \mathbf{1}[W_{tki}^* \geq 0] = \mathbf{1}[W_{tki}^*(t, Z_{ki}, \theta_i^w, \varepsilon_{tki}^w) \geq 0] \\ S_{tki}(H_{ki}) &= \mathbf{1}[S_{tki}^*(H_{ki}) \geq 0] = \mathbf{1}[S_{tki}^*(t, X_{ki}, \theta_i^s, \varepsilon_{tki}^s; H_{ki}) \geq 0], \end{aligned} \quad (1)$$

where W_{tki}^* and $S_{tki}^*(H_{ki})$ are the net utilities associated with W_{tki} and $S_{tki}(H_{ki})$, respectively;²⁰ Z_{ki}/X_{ki} represents the vector of observed (by the econometrician) characteristics affecting the school/work outcomes at k ; θ_i^w/θ_i^s represents the unobserved (by the econometrician), individual-specific, propensity for work/schooling that is constant over time and common across work/school outcomes; (θ_i^w, θ_i^s) i.i.d. $F_\theta(\bullet)$; ε_{tki}^w represent level/period/individual i.i.d. shocks to work; and $\varepsilon_{tki}^s(H_{ki})$ represent state/level/period/individual i.i.d. shocks to schooling.

In particular, we consider the following convenient specification:

$$\begin{aligned} W_{tki}^*(t, Z_{ki}, \theta_{wi}, \varepsilon_{tki}^w) &= \lambda_w(t; \beta_{k\lambda}^w) + \beta_{kz}^w Z_{ki} + \beta_{k\theta}^w \theta_i^w + \varepsilon_{tki}^w \\ S_{tki}^*(t, X_{ki}, \theta_{si}, \varepsilon_{tki}^s; H_{ki}) &= \lambda_s(t; \beta_{k\lambda}^s(H_k)) + \beta_{kx}^s(H_k) X_{ki} + \beta_{k\theta}^s(H_k) \theta_i^s + \varepsilon_{tki}^s(H_k), \end{aligned} \quad (2)$$

²⁰ Since in our sample once a child starts working he or she works thereafter, the work outcome at (t, k) is deterministic (i.e. it is 1 with probability 1) for those who worked prior to (t, k) .

where $\lambda_w(\cdot)$ and $\lambda_s(\cdot)$ are the baseline hazard functions for the work and school equations, respectively.²¹ We assume (i) $(\varepsilon_{tki}^w, \varepsilon_{tki}^s(H_k))$ are mutually independent and identically distributed extreme value random variables. The extreme value assumption produces the logistic probabilities

$$\Pr(W_{tki} = \mathbf{1}) = \Lambda(W_{tki}^*) \text{ and } \Pr(S_{tki}(H_k) = \mathbf{1}) = \Lambda(S_{tki}^*(H_k)) \quad (3)$$

$$\text{with } \Lambda(z) = \exp(z)(1 + \exp(x))^{-1}$$

In addition, we assume (ii) (θ_i^w, θ_i^s) are independent of $\{(\varepsilon_{tki}^w, \varepsilon_{tki}^s(H_k)); k = e, p, s\}$ and both independent of $\{(X_{ki}, Z_{ki}); k = e, p, s\}$; (iii) $E(\theta^w) = E(\theta^s) = 0$, $Var(\theta^w) = Var(\theta^s) = 1$, $\beta_{k\theta}^l$ is finite for all k and $l \in \{w, s\}$ and $(\beta_{e\theta}^w, \beta_{e\theta}^s(0)) = (1, 1)$; (iv) $Supp(\beta_{k\theta}^w \theta_i^w + \varepsilon_{tki}^w, \beta_{k\theta}^s(H_k) \theta_i^s + \varepsilon_{tki}^s(H_k)) \subseteq Supp(\beta_{kz}^w Z_{ki}, \beta_{kx}^s(H_k) X_{ki})$, and each component of $(\beta_{kz}^w Z_{ki}, \beta_{kx}^s(H_k) X_{ki})$ assume either arbitrarily large or arbitrarily small values or both; (v) $F_\theta(\bullet)$ is a discrete distribution with a finite and known number of mass points $\{\theta_m\}_{m=1}^M$, $\pi_m \geq 0$ is the probability associated with mass point $\theta_m = (\theta_m^w, \theta_m^s)$, and $\sum_{m=1}^M \pi_m = 1$ (Heckman and Singer, 1984); and (vi) (X_{ki}, Z_{ki}) vary across levels, and X_{ki} varies across work sequences within each level. Under these conditions identification of the model is a result of Theorem 4 and Theorem 5 in Heckman and Cameron (1998).²²

While the original distribution of θ is reasonably flexible, because no functional form

²¹ $\lambda_w(\cdot)$ only applies to the work equation in the entry level.

²² Although condition (v) is not an assumption required for identification, it is satisfied by our data and in practice should aid with the identification of the model.

assumptions are made about $F_\theta(\bullet)$, it still assumes that the initial distribution of θ^s in the working and non-working states is the same. Thus, we also estimate the model under an alternative specification $\theta = (\theta^w, \theta_0^s, \theta_1^s)$ and assume that θ is distributed a la Heckman-Singer. In this case, identification requires that $\beta_{e\theta} = (1, 1, 1)$. Finally, we also estimate the model for the case of $(\theta^w, \theta^s) \sim N(0, [1, 1, \rho])$.

As a result of the factor structure of the model, dependence between work/school outcomes occurs through θ^w/θ^s , while dependence between school and work outcomes arises from the correlation between θ^w and θ^s .²³ In particular, the structure of the model allows the school and work decisions in a given level to be correlated and subject to selectivity with respect to school and work decisions in previous levels.

3.1 The Likelihood Function

A sample observation i consists of a set of work and school outcomes $y_i = \{(W_{tki}, S_{tki}) : t \in T_k, k \in \{e, p, s\}\}$, and a set of instruments $x_i = \{(X_{ki}, Z_{ki}) : k \in \{e, p, s\}\}$, where T_k represents the set of possible periods in k , with $T_e = \{5, \dots, 14\}$, $T_p = \{5, \dots, 10\}$ and $T_s = \{5, \dots, 10\}$.²⁴ In this framework, the contribution to the sample likelihood of observation i

²³ Dependence also occurs through observable characteristics.

²⁴ As noted in section 2, a child is assumed to be no longer at risk of entering school beyond 14 years of age. The treatment of censoring of the work and school outcomes in the primary and secondary school levels is outlined in section 3.

conditional on θ_i , is

$$\begin{aligned}
\Pr(y_i, x_i; \theta_i) &= \Pr(\{(W_{tki}, S_{tki}) : t \in T_k, k \in \{e, p, s\}\} | x_i, \theta_i) \\
&= \prod_{k \in \{e, p, s\}} \prod_{t \in T_k} \Pr(W_{tki}, S_{tki} | H_k, x_i, \theta_i) \\
&= \prod_{k \in \{e, p, s\}} \prod_{t \in T_k} \Pr(S_{tki}(H_k) | x_i, \theta_i) \Pr(W_{tki} | x_i, \theta_i)
\end{aligned} \tag{4}$$

Finally, after integrating out the unobserved heterogeneity component, we obtain the following expression for the unconditional likelihood function for a sample of N individuals

$$L = \prod_{i=1}^N \int \Pr(y_i, x_i; \theta_i) dF_\theta. \tag{5}$$

When the Heckman-Singer approach is considered, the integration term is substituted by a sum over the space of unobserved heterogeneity types. In this case, the probabilities associated with the mass points and the mass points themselves are estimated along with the other model parameters by maximum likelihood.

4 Dynamic Evaluation of the Effect of Work on School Progress

This section presents a dynamic extension of the static program evaluation framework based on selection models.²⁵ The crucial feature of this framework is the presence of heterogenous

²⁵ See Heckman et al. (1999) for a review of the program evaluation literature.

responses to treatment among observationally equivalent individuals upon which individuals act. In this context, the treatment effect is a random variable that cannot be summarized by a single parameter. Thus a variety of treatment effects can be defined depending on the conditioning sets and the summary statistics desired. This framework allows us to identify which groups are most affected by treatment, and to assess the relative importance of observables and unobservables in understanding the selection into treatment and outcome processes.

An important advantage of the dynamic model of school and work developed in section 4 is that it can be used to generate a rich set of dynamic mean treatment parameters from a common set of structural parameters.

In the present context, we are ultimately interested in evaluating the effect of work on the probability that a child ever experiences the school event in each level. In particular, let \bar{t}_k represent the maximum survival time in the sample for level k . Thus, for each possible work sequence H_k in level k , the school outcome we are ultimately interested in is

$$S_k(H_k) = \mathbf{1} [t \leq \bar{t}_k \mid H_k]$$

Let $\Delta_k(H_k^+, H_k) = S_k(H_k^+) - S_k(H_k)$ denote the effect on the school outcome in level k of the work sequence H_k^+ compared to a “lower” work sequence H_k for a given child, and for any pair of work sequences belonging to the space of possible work sequences in level k . 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 had the work sequence H_k^+ compared to the case where he or she had the work sequence H_k . In our case, $\Delta_k(H_k^+, H_k)$ can take

three values

1. $\Delta_k (H_k^+, H_k) = 1$ ($S_k (H_k^+) = 1, S_k (H_k) = 0$) if the child would experience the school event in level k under the work sequence H_k^+ and would not experience it under the lower work sequence H_k .
2. $\Delta_k (H_k^+, H_k) = 0$ if the child would experience the school event under both work sequence ($S_k (H_k^+) = 1, S_k (H_k) = 1$), or if he or she would not experience the school event under any work sequence ($S_k (H_k^+) = 0, S_k (H_k) = 0$).
3. $\Delta_k (H_k^+, H_k) = -1$ ($S_k (H_k^+) = 0, S_k (H_k) = 1$) if the child would experience the school event under the lower work sequence H_k and would not experience the school under the work sequence H_k^+ .

In the present context, we cannot estimate $\Delta_k (H_k^+, H_k)$ for a given person since we never observe the same individual under both work sequences. Instead, we can work with population means or distributions of these variables. In this paper, we focus on a variety of dynamic mean treatment parameters that differ in the conditioning set on which they are defined. In particular, we consider the dynamic versions of three familiar static mean treatment parameters, namely the dynamic average treatment effect (DATE), the dynamic average treatment effect on the treated (DTT) and the dynamic average treatment effect on the untreated (DTU).

Let us start by defining the short-run effects. The first one we consider is the dynamic average treatment effect on the school outcome in level k of the work sequence H_k^+ compared

to a “lower” work sequence H_k . This parameter is defined as the average effect of the work sequence H_k^+ over the work sequence H_k for a child randomly selected from the initial population of children with characteristics x and θ , and is given by

$$DATE_k(H_k^+, H_k) = E(\Delta_k(H_k^+, H_k) | x, \theta) \quad (6)$$

If we define (6) for the work status during level k only we obtain the static ATE of working during level k . The dynamic average treatment effect on the treated is defined as (6) but for a child randomly selected from the population of children who actually experienced the work sequence H_k^+ and has characteristics x and θ . This parameter is given by

$$DTT_k(H_k^+, H_k) = E(\Delta_k(H_k^+, H_k) | x, \theta, H_k^+) \quad (7)$$

Note that (7) controls for selection into the work sequence H_k^+ on the basis of both observable and unobservable characteristics. Also, if we define (7) for the work status during level k and for the working population in that level we obtain the static TT of working during level k . Finally, if instead we condition (7) on H_k , or any other possible work sequence at k different from H_k^+ , we obtain the dynamic average treatment effect on the untreated (DTU).

The fact that we find ex-post heterogeneity in $\Delta_k(H_k^+, H_k)$ among observationally equivalent individuals is a necessary but not sufficient condition for (6) and (7) to be different. In particular, unless individuals act on this unobserved heterogeneity (i.e. select into a particular work sequence on the basis of unobservable gains/losses associated with that sequence

compared to another), the two dynamic treatment parameters are identical.

We compute the above dynamic mean treatment effects using simulation techniques. In particular, we use the estimated model to get a sufficiently large simulated sample.²⁶ Then we use the model to simulate the sequence of school outcomes under alternative work sequences. The outcomes under any given pair of work sequences are then compared either for full sample (DATE) or subsamples of individuals. In each level, these subsamples can be defined by the groups of individuals for whom the unrestricted model predicts would select into each possible work sequence in that level (DTT and DTU). Standard errors for the dynamic mean treatment effects are computed using the parametric bootstrap method.²⁷

5 Model Results

We estimate the dynamic switching model under the three alternative specifications of θ described in section 3. For the two specifications involving a non-parametric distribution of θ , we find, after experimentation, that three types for each element of θ describes the data quite well. The lack of sufficient variation in the data prevented us from estimating a separate equation for $S_{tp}(1, 1)$. Instead we constrain the difference between $S_{tp}(1, 1)$ and $S_{tp}(0, 1)$ to a constant. Likewise, the only individual with work sequence $(1, 1, 1)$ contributing to the estimation sample in the secondary school level is aggregated with those with work sequence

²⁶ For each individual in the sample we generate a θ type based on the estimated distribution $F_\theta(\bullet)$. Likewise, we generate i.i.d. shocks to work, ε_{tki}^w , and schooling, $\varepsilon_{tki}^s(H_{ki})$.

²⁷ In particular, we first use the estimated asymptotic normal distribution of the vector of parameters to generate vectors of parameter values. We then compute the dynamic treatment effects under each of these vectors and calculate their standard deviation.

$(0, 1, 1)$ in equation $S_{t_s}(0, 1, 1)$.

To discriminate among competing models (including the one without unobserved heterogeneity), we use the likelihood ratio tests (LR) and the Bayesian Information Criterion (BIC).²⁸ Appendix D shows that the two models with a non-parametric specification of unobserved heterogeneity compare favorably against the model with normal heterogeneity and the model without unobserved heterogeneity. Between the two non-parametric specifications, the LR test tends to favor somewhat the more flexible three-factor model. However, using BIC (which accounts for the difference in the number of parameters) the more parsimonious two-factor model is preferred. In addition, in contrast with the two-factor model, all of the probabilities associated with mass points are estimated very imprecisely in the three-factor model. Thus, the two-factor model appears to be the specification that best describes the dynamic sequence of work and school outcomes given the variation in observable characteristics. For the remaining of the paper we focus the discussion on the results of the two-factor model.²⁹

Estimates of the parameters associated with the work and school outcome equations in the three school levels are reported in Table 4.1 through Table 4.3. Table 4.1 also reports the estimated correlation between θ^w and θ^s .

In addition to the tests in Appendix D, another way of evaluating the importance of controlling for unobserved heterogeneity is to examine its impact on the estimated coefficients.

²⁸ The p-values of the LR tests are meant as heuristic guides only, and cannot be interpreted using the standard Chi-square tables as the models being compared are non-nested. BIC is, however, valid for discriminating between non-nested models.

²⁹ The results for the other models are available upon request from the authors.

In particular, all the parameters associated with the unobserved heterogeneity components are sizable and statistically significant, which indicates the importance of accounting for unobserved heterogeneity when analyzing selection into work and selection into school level. The presence of selection on unobservable characteristics in these data has a significant impact on the parameters associated with exogenous covariates, which are generally larger in magnitude and more statistically significant than in model without unobserved heterogeneity. This gives an indication of the extent to which the parameters associated with exogenous covariates in the model without unobserved heterogeneity are subject to dynamic selection bias. Likewise, once we account for selection on unobservables the parameter estimates on predetermined endogenous variables change significantly with respect to those in the model without unobserved heterogeneity, which indicates evidence of endogeneity bias.

As regards the baseline hazards, the results show that both the probability of working before school entry increases with age, while the probability entering school follows an inverted u-curve. In the primary and secondary levels, the probability of experiencing the school event decreases as the child accumulates more school delay.

In general, the observable characteristics that make an individual more likely to work also make him or her less likely to experience the schooling event under any work sequence. This offers insight into the presence of selection into work on the basis of observable characteristics. Likewise, the observable characteristics that make an individual more likely to experience the school event in one level also make him or her more likely to succeed in subsequent levels. This indicates evidence of school level selectivity on the basis of observable characteristics.

Child characteristics

Delaying school entry increases the probability of working during primary school and, specially, secondary school. Conditional on work sequence, however, school entry age only has a significantly negative effect on the probability of reaching secondary for those who did not work during primary school. To the extent that work reduces the likelihood of experiencing the school event, these results show that school entry age has an indirect negative effect on schooling. The effect on work may indicate that older children are physically more able to do productive work, and they are expected to contribute more to household income than younger children. The effect on schooling indicates that older children may have a hard time fitting in classrooms with younger classmates, and may have fewer chances of continuing in school after failing a grade.

Grade repetition in primary school increases the probability of working during secondary school. Conditional on work sequence, grade repetition reduces the probability of completing secondary school, but only significantly so for those who started working during secondary school. For those who were working before entering secondary school, grade repetition has an indirect negative effect on secondary school completion.

Girls are significantly less likely to work in all levels except in primary school. This gender difference in work propensities is particularly marked in the secondary level. As noted in section 2, our definition of work refers to productive work only, and thus ignores household chores, which are more likely to be performed by girls. Girls are significantly less likely to experience the school event in all three levels but only under the non-working sequences.

However, these are the sequences where girls are more likely to be. Hence the gender gap in education persists as we move to higher school levels among non-working children but seems to disappear after school entry among working children.

If the child was in primary school when the compulsory schooling law was in place, his or her chances of working during primary school are significantly lower, while his or her chances of reaching secondary school are greater, although only significantly so in the non-working state. These effects stand apart from the overall time trends in school and work, as captured by the effect of age at the time of the survey.

If the child was in primary school when free tuition for girls in secondary school was introduced, his or her chances of working in primary school are unaffected. However, the effect of this policy on schooling is equally positive for both boys and girls in the working state, but only significantly positive for girls in the non-working state. These results suggest that the anticipation of this policy had a positive income effect on the probability of reaching secondary school for both boys and girls, and a substitution effect in favor of girls among non-working children. If the child was in secondary school grades 6-8 when the free tuition policy was active, his or her chances of working are lower, specially for boys. This mainly reflects an income effect whereby the need for a child to contribute to household income is reduced, and since boys are more likely to contribute to household income than girls the effect is far greater for them. The effect of this policy on secondary school completion is significantly positive for girls only among non-working children, and significantly positive for both genders among those who started working in secondary school, specially for boys. In

contrast with the effect of this policy in primary school, here there seems to be a surprising substitution effect in favor of boys among children who start working in secondary school. However, this result must be interpreted with caution as girls are much less likely to work in secondary school.

Parental and household characteristics

With regard to parental education, only the education of the father significantly reduces the probability of working in primary school. The education of either parent increases the probability of entering school and reaching secondary school under all work sequences, although the effect of mother's education is larger in all cases. The positive effect of mother's education persists into secondary school in all work sequences, but that of the father only has a significant effect among non-working children, and even then the effect is smaller than that of mother's education. Parental education can potentially influence the allocation of children's 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 income effects, this differential effect of education by gender of the parent may suggest that women have a higher preference for child schooling than men.

Household wealth, as indicated by household assets or the ownership of a modern latrine, significantly reduces the probability of working in all but the primary level, while it significantly increases the probability of experiencing the school event in each level under all possible work sequences. These results indicate that, apart from its income effect, household wealth may play an important role as a cushion against economic shocks in the absence of

well-developed capital markets.³⁰

In all but the entry level, the ownership of at least one of two household productive assets significantly increases the probability of work, particularly the ownership of a non-farm business. Likewise, in all but the secondary school outcome for children who were working before entering secondary school, the ownership of at least one of two household productive assets significantly increases the probability of work.

Household productive assets have both a positive income effect and a negative substitution effect on a child's schooling.³¹ The positive effect of household productive assets on child work indicates that the substitution effect dominates the income effect for this outcome. However, conditional on working status all we have left is the income effect, which is significantly positive in all but one school outcome equation. Controlling for whether the household owns farm land, the fact that the household cultivates land (whether own land or rented/sharecropped) significantly increases the probability of work in all but the entry level.

In theory, the age composition of siblings plays a mixed role in the allocation of children's time. The presence of other siblings in the household increases competition for household resources for education 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

³⁰ 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.

³¹ 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).

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 work, but this effect is only significant in the secondary level, while the number of younger siblings increases the probability of work before school entry but decreases the probability of work in secondary school. Conditional on work sequence, the effect of the number of older siblings on schooling is always positive whenever significant, while the effect of the number of younger siblings is significantly negative in the secondary level among those who start working in that level, but significantly positive among working children in the primary level.³²

Village characteristics

The presence of a secondary school in the village has a significant effect on work, but only in the secondary level where the presence of a secondary school nearby seems to facilitate the combination of school and work activities. School availability should in theory have a significant effect on schooling. The results show that this does generally seem to be case for secondary schools but not for primary schools, which is not surprising given the fact that most children in the sample lived in a village with a primary school by the time they were 6. The presence of health-related infrastructure in the village has a significantly negative effect on work in the secondary level but significantly positive in the primary level. The benefits on schooling seem to be particularly strong in the entry and primary levels.

Child work is likely to be related with subsistence agriculture. In the absence of capital

³² This unexpected positive effect may be explained by the combination of younger siblings not competing for education resources and also contributing to household income.

markets, reliance on agriculture makes children's time more likely to be used as an insurance mechanism against shocks, which are common in the agriculture sector. A more diversified village economy (as indicated by the presence of some form of industry) has no significant effect on work but it does generally have a consistently positive effect on schooling throughout the three levels. The capital of Matlab provides access to a big market, health facilities, credit institutions, schools and employment opportunities. Villages further away from the capital of Matlab have a higher incidence of child work in the primary level, and consistently lower school transition and completion rates.

5.1 Simulated Effects of Work on School Progress

Table 5 reports the simulated mean effects of work by school outcome, pair of work sequences and population. As work in the entry level may refer to any age prior to the school entry age, we define the working state in the entry level as having a probability one of working at each age.

The dynamic average treatment effects (DATE) are found to be negative and sizable in all levels and for all possible pairs of work sequences. The greater the difference in work intensity between two work sequences the larger the difference in school outcomes. Thus, for a random individual work reduces his or her probability of experiencing the school event at each level, and the younger this individual is exposed to work the more negative the effect of work. For example, if a child starts working in primary school his or her probability of reaching secondary school is reduced by 10.73 percent, while if he or she starts working

before school entry this probability is reduced by 33.20 percent.

In all three school levels, the dynamic treatment effect on the treated (DTT) is less negative than the DATE for all pairs of work sequences involving work in the entry level, except the one in which the sequence being compared is no work up to secondary school (inclusive). More precisely, the difference in the entry and primary school outcomes between a work sequence involving work before school entry and a less work-intensive sequence is generally less negative for those who worked in the entry level than for the entire population, particularly in the entry and primary school levels. The opposite is true for all pairs involving work starting in primary school or work starting in secondary school. Furthermore, the effect of work on the school outcomes of each treated population becomes more negative the greater the difference in work intensity between the work sequence of the treated population and the comparison sequence.

In a dynamic context, however, it is probably more informative to compare, the treatment effect on the treated with the treatment effects for those who select into other possible work sequences in that level. To focus the discussion below, we choose to report the dynamic treatment effect on the untreated (DTU) for the group of children who select into the work sequence being compared. The difference in all three school outcomes between a more work-intensive sequence and a less work-intensive sequence is always more negative for the untreated population than for the treated population or the entire population, particularly when the treated population started working before school entry. The effect of work starting in secondary school is, however, more negative for the treated population than for the

untreated population.

What these comparisons indicate is that in most cases treated populations have a combination of observable and unobservable characteristics that make them less likely to be affected by the work sequence they have selected in than untreated populations. This is particularly so the more work-intensive the treatment sequence is. In particular, the effect of work appears to be smaller for working children than for non-working children and, within working children, the longer a child has been exposed to work the smaller the impact of work. The exception to this pattern involves children who started working in secondary school.

Comparing the simulated treatment effects with the non-parametric mean differences of Table 2, we observe that although the basic picture of more work leading to worse school outcomes remains the same, there are systematic differences in the magnitudes. In particular, the treatment effects on the treated are significantly less negative than the corresponding mean differences for the primary and, particularly, entry school outcomes. However, the three pairs of sequences in the secondary level for which a comparison can be made, the reverse is true.³³ Thus controlling for selection appears to be very important in these data. The next step in the analysis is then to examine the contribution of observables and unobservables to this pattern.

³³ Note that none of these pairs includes the sequence (1, 1, 1).

5.2 The Relation Between Selection and School Outcomes

A central question in this paper is to identify the observable and unobservable characteristics of the different groups of working children that make them more or less sensitive to the work regime they have selected in than other groups of children. As it will be noted in the concluding section, this characterization has major implications for policies aimed at increasing school progress through reductions in child work.

To this extent, this section examines the contribution of observables and unobservables to the differences between DTT and DTU. For this task, we need to relate selection on observables and unobservables with the counterfactual $\Delta_k(H_k^+, H_k)$. In doing so, we need to distinguish between two types of interrelated selection processes: selection into work and selection into school level. While selection into school level occurs until the last school level an individual attends, selection into work (i.e. the decision of whether to work or not) only occurs up to the level the individual starts working (inclusive) or up to the secondary level for those who never work (inclusive). This is because the work decision is deterministic once an individual starts working, and thus his or her work sequence in future levels is determined by the work sequence up to the level he or she started working (inclusive).

Hence, for a given level, selection into work can help us understand the difference between DTT and DTU for work sequences that only differ in the work status in that level. For example, we use the selection into work argument when comparing the effect of $(0, 1)$ versus $(0, 0)$ for those who select into $(0, 1)$ and those who select into $(0, 0)$. However, when comparing the treatment effects in the secondary level for children in sequence $(0, 1, 1)$ with those in

sequences $(0, 0, 1)$ and $(0, 0, 0)$ we must invoke both the selection into work effect in the primary level and the selection into secondary school effect.

Let us start with a general description of the selection into school level process. Table 6 shows the distribution of θ^s by level and work sequence. For all groups, the distribution of θ^s shifts to the right across school levels as low θ^s individuals are screened out. Overall, the extent of cream-screening is larger among working children than among non-working children. This is because, relative to non-working children, working children are faced with lower values of $\beta_x^s X$, so they tend to continue schooling only if they also have a high θ^s .

By comparing the average values of $\beta_x^s X$ across levels for each work sequence, we observe that, first, there is a very significant selection on observables in the transition to primary school among those who worked in the entry level and, second, there is also a significant selection on unobservables in the transition to secondary school among children who attended primary school. Not surprisingly, this selection is more accentuated among non-working children in primary school who, relative to working children, need a higher value of $\beta^s X$ to compensate for their lower θ^s .

Let us now consider selection into work. To analyze selection into work on the basis of unobservables, and thus relate U^w to the counterfactual Δ , Table 7 reports, for each level, the correlations between the values of the unobservables in the work equation, $U^w = \beta_\theta^w \theta^w + \varepsilon^w$, and the values of the unobservables in the school equations in each working state, $U^s = \beta_\theta^s \theta^s + \varepsilon^s$. To examine selection into work on the basis of observables, and thus relate $\beta_z^w Z$ to Δ , Table 7 also reports, for each level, the correlations between $\beta_z^w Z$ and $\beta_x^s X$ in

each working state.

In the entry level, the higher the unobserved propensity to work the lower the unobserved propensity to enter school in either state, but particularly so in the non-working state. Thus, in terms of unobservables, working children would benefit less from not working. However, in terms of observables a higher propensity to work is associated with lower probability of school entry in either state, but particularly in the working state. Thus, in terms of observables, working children would benefit more from not working. Altogether, it appears that selection on unobservables is stronger than selection on observables in explaining why the effect of work in the entry level is less negative for those who work than for those who do not. In particular, children who work in the entry level are those who would benefit the least from not working. Furthermore, in terms of our typology, they are more likely to be of type $\Delta_e = 0$ ($S_e(1) = 0, S_e(0) = 0$).

Let us now consider selection into work in the primary level among those who did not work before school entry. In terms of unobservables, the higher the propensity to work the higher the propensity to reach secondary school in either state, but particularly in the working state. In terms of observables a higher propensity to work is associated with a lower probability of school entry in either state, but particularly so in the non-working state. Hence selection on observables and unobservables reinforce each other in explaining why those who start working in primary school are less affected by work than non-working children. Thus, as for those who work in the entry level, children who start working in primary school would benefit less from not working than non-working children, but for opposite reasons. In

particular, these working children are more likely to be of type ($S_p(0,1) = 1, S_p(0,0) = 1$) than those who do not work in primary school.

Despite the significant selection of those in the sequence (1, 1) out of the initial population in (1), they continue to have lower values of $\beta_x^s X$ and θ^s relative to the children in other primary work sequences. This explains the smaller effect of sequence (1, 1) versus sequences (0, 1) and (0, 0) for the treated than for the untreated populations. Thus, those in (1, 1) are less likely to benefit from not working before school entry than those in (0, 1) and (0, 0) in terms of reaching secondary school. This is because, relative to other children in primary school, they have characteristics that make them less likely to reach secondary school under any regime.

Let us now consider selection into work in the secondary level among those who did not work in primary school. In terms of unobservables, a high propensity to work is associated with an equally high propensity to reach secondary school in either state. This indicates that children who start working in secondary school have unobservable characteristics that make them more likely to complete secondary school than those who do not work in either state. However, these characteristics make them equally likely to perform well in both states. Hence, in terms of unobservables working children are equally likely to benefit from not working than non-working children. However, in terms of observables, working children have observable characteristics that make them less likely to complete secondary school than non-working children in either state, but particularly in the working state. This explains why the effect of starting to work in secondary school is actually less negative for those who

do not work than for those who started working in secondary school. Thus, working children are more likely to benefit from not working in secondary school than non-working children.

The difference in the effects for children in sequence (1, 1, 1) relative to those in sequences (0, 1, 1) and (0, 0, 1) can still be explained in terms of a lower θ^s . In the case of comparing children in (1, 1, 1) with those in (0, 0, 1), a lower θ^s is also accompanied by lower values of $\beta_x^s X$. The difference in the effect with respect to the non-working group cannot be explained in terms of θ^s since children in the sequence (1, 1, 1) do actually have a higher θ^s . However, this effect is more than offset by their much lower values of $\beta_x^s X$. Again, children in (1, 1, 1) are less affected by work than children in the other sequences because, relative to them, they have characteristics that make them less likely to complete secondary school under any regime.

The effect of the work sequence (0, 1, 1) versus the sequence (0, 0, 1) is not significantly different for the treated and untreated children, which indicates that the higher θ^s of the treated gets offset by their lower values of $\beta_x^s X$. Thus both groups of children are equally likely to benefit from not working in primary school relative to working in secondary school only in terms of completing secondary school. The effect of (0, 1, 1) versus (0, 0, 0) is less negative for the treated than for the untreated, which indicates that the lower values of $\beta_x^s X$ of the treated population more than offsets their higher θ^s . Thus, because of their relative higher values of unobservable characteristics, children in (0, 1, 1) would benefit more than non-working children from not having to work at any level.

Summarizing, the extent of cream-screening across school transitions is larger for working

children than for non-working children. Thus, working children that remain in school tend to have a higher unobserved propensity for schooling than non-working children. This tends to make the negative effect of work smaller for those who start working in primary school relative to non-working children in the primary level, and for all working children relative to non-working children in the secondary level. However, working children are subject to greater selection than non-working children because they face a worse environment in terms of observable characteristics.

Hence, by simply looking at the unobservable characteristics of working children in a particular school level, without actually understanding the nature of the selection process, we would tend to underestimate the benefits of policies to reduce child work, as we would observe many working children are able to make it through school without such a policy. However, if we were to expose working children to the same environment that non-working children have, we would not observe the systematic differences in selection patterns outlined above, and thus we would most likely observe similar negative effects of work for both groups of children.

These findings have important policy implications. First, any assessment of policies to reduce child work in primary school or secondary school must be based on a full understanding of the nature of the selection process across school transitions. Second, the magnitude of the effects of work on school progress makes policies aimed at increasing school progress through reductions in child work most relevant. Third, the effectiveness of these policies would be greatly increased if they were accompanied by efforts to improve the adverse environment

that working children face.

5.3 Policy Simulations

Motivated by the previous policy implications, in this section we simulate the dynamic effects of three policies. We first evaluate the effect of the compulsory primary schooling policy introduced in 1992. In particular, we compare the current environment with the simulated outcomes in the absence of this policy. Second, we consider a policy that makes school entry compulsory at age 6 and compare its effects with the simulated outcomes in the unrestricted model. Third, we consider a policy that makes access to secondary school universal against the simulated outcomes in the unrestricted model for different populations of working children. The first two policies are aimed at increasing school progress indirectly by reducing child work as well as directly. The third policy is aimed at increasing school progress directly by improving the schooling environment.

As Table 8 shows, the compulsory primary schooling policy had a significant impact on the probability of reaching secondary school, which is partly explained by its sizable effect on work during primary school. Making school entry compulsory at age 6 would significantly decrease the probability of working in primary school and secondary school, although its effect on schooling would only be sizable in the primary school level. The availability of secondary schools would have a beneficial effect on schooling, particularly in the entry level. In all three school levels, the benefits of this policy would be mainly accrued by working children, particularly those who started working younger, which highlights the importance

of policies aimed at improving the adverse environment that working children face.

6 Conclusions

This paper investigates the causes underlying the poor school performance of children in rural Bangladesh, while focusing on the effect of work on school progress. To this end, we specify an econometric model that is able to capture two types of interrelated selection processes in the data, namely selection into work and selection into school level.

We find that work has a negative and sizable effect on the school outcome in each level for the entire population, as well as for all the subpopulations considered. Furthermore, the effect of work becomes more negative the earlier in life an individual is exposed to work.

We also find that in most cases work appears to have a smaller effect for working children than for non-working children and, among working children, the younger a child started to work the smaller the impact of work. Thus, in terms of school progress, it appears that working children would benefit less from not working than non working children. The exception to this pattern involves children who started working in secondary school. In analyzing the contribution of observables and unobservables to this result, we find that working children in the primary and secondary school levels tend to have a higher unobserved propensity for schooling than non-working children, but that is because they have been subject to greater selection than non-working children which is, in turn, explained by the relatively more adverse environment that working children face. In particular, in all cases working children face a worse environment in terms of observable characteristics than non-

working children, more so the younger the child started working. Hence, by simply looking at the unobservable characteristics of working children in a particular school level, without actually understanding the nature of the selection process, we would tend to underestimate the benefits of policies to reduce child work, as we would observe many working children are able to make it through school without such a policy. However, if we were to expose working children to the same environment that non-working children have, we would not observe the systematic differences in selection patterns outlined above, and thus we would most likely observe similar negative effects of work for both groups of children.

These findings have important policy implications. First, any assessment of policies to reduce child work in primary school or secondary school must be based on a full understanding of the nature of the selection process across school transitions. Second, the magnitude of the effects of work on school progress makes policies aimed at increasing school progress through reductions in child work most relevant. Third, the effectiveness of these policies would be greatly increased if they were accompanied by efforts to improve the adverse environment that working children face.

References

- Aakvik, A., J.J. Heckman, and E.J. Vytlacil, 2000. Treatment Effects for Discrete Outcomes when Responses to Treatment Vary Among Observationally Identical Persons: An Application to Norwegian Vocational Rehabilitation Programs. *NBER Technical Working Paper* 262, NBER, Cambridge, MA.
- 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.
- 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.
- Canagarajah, S., and H. Coulumbe, 1997. Child Labor and Schooling in Ghana. Mimeo, Africa Technical Department, the World Bank, Washington, D.C.
- Dagenais, M., C. Montmarquette, and N. Viennot-Briot, 2001. Dropout, School Performance and Working while in School : An Econometric Model with Heterogeneous Groups. Working Paper 2001s-63, CIRANO, Montréal, Canada.
- Duraisamy, M., 2000. Child Schooling and Child Work in India. Paper presented at the 2000 World Congress of the Econometric Society, August 2000, Seattle, WA.
- Eckstein, Z., and K.I. Wolpin, 1999. Why Youths Drop out of High School: The Impact of Preferences, Opportunities, and Abilities. *Econometrica* 67, 1295-1339.
- Filmer, D., 1999. Educational Attainment and Enrollment Profiles: A Resource Book based on an Analysis of Demographic and Health Survey Data. Mimeo, Development Research Group, The World Bank.
- Grootaert, C., and R. Kanbur, 1995. Child Labor: A Review. *World Bank Policy Research Working Paper* 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., R. LaLonde, and J. Smith, 1999. The Economics and Econometrics of Training Programs. Forthcoming in: O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Volume III, North Holland, Amsterdam.

- 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.
- Levison, D., K. Moe, and F. Knaul, 2001. Youth Education and Work in Mexico. *World Development* 29, 167-188.
- Lillard, L.A., and R.J. Willis, 1994. Intergenerational Educational Mobility: Effects of Family and State in Malaysia. *Journal of Human Resources* 29, 1126-66.
- Mare, R.D., 1980. Social Background and School Continuation Decisions. *Journal of the American Statistical Association* 75, 295-305.
- Psacharopoulos, G., 1997. Child Labor Versus Educational Attainment: Some Evidence from Latin America. *Journal of Population Economics* 10, 377-386.
- Rahman, O., J. Menken, A. Foster, M.N. Khan, C.E. Peterson, R. Khun, P. Gertler, ICDDR, Mitra and Associates, 1999. The Matlab Health and Socioeconomic Survey: Overview and User's Guide", DRU-2018/1, RAND, Santa Monica CA.
- 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. Working Paper 01-16, Department of Economics, University of Colorado at Boulder.
- 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.
- Sawada, Y., and M. Lokshin, 2001. Household Schooling Decisions in Rural Pakistan. Working Paper 2541, the World Bank, Washington, DC.
- Schultz, T.P., 2001. School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program. Discussion Paper 834, Economic Growth Center, Yale University.
- Skoufias, E., 1994. Market Wages, Family Composition and the Time Allocation of Children in Agricultural Households. *Journal of Development Studies* 30, 335-360.

Table 1. Children by Work Sequence in Each Level

	Unadjusted for censoring		Adjusted for censoring	
	Count	Percent	Count	Percent
School entry				
$H_e = 1$	329	13.22	329	13.22
$H_e = 0$	2160	86.78	2160	86.78
Total	2489	100.00	2489	100.00
Primary school				
$H_p = (1,1)$	24	1.23	24	1.24
$H_p = (0,1)$	546	27.94	546	28.33
$H_p = (0,0)$	1384	70.83	1357	70.42
Total	1954	100.00	1927	100.00
Secondary school				
$H_s = (1,1,1)$	5	0.39	5	0.48
$H_s = (0,1,1)$	272	21.13	272	26.08
$H_s = (0,0,1)$	351	27.27	351	33.65
$H_s = (0,0,0)$	659	51.20	415	39.79
Total	1287	100.00	1043	100.00

Notes: H_e , H_p , and H_s represent the work sequence in the entry, primary and secondary school levels, respectively. The work sequence in each level includes the work status in that level as well as the work status in previous levels. For example, $H_p = (0,1)$ represents the work sequence in the primary school level of those children who did not work before school entry but worked during primary school.

Table 2. Estimated Transition Rates Associated to Each School Outcome by Work Sequence

	Estimate	Std. Err.	Comparison	Difference	LR test: $\chi^2(1)$
School entry					
(1) $H_e = 1$	0.073	0.014	(1) Vs (2)	-0.821	723.58 (0.000)
(2) $H_e = 0$	0.893	0.007			
Transition to secondary school					
(3) $H_p = (1,1)$	0.222	0.089	(3) Vs (4)	-0.283	3.21 (0.073)
(4) $H_p = (0,1)$	0.505	0.022	(3) Vs (5)	-0.513	14.86 (0.000)
(5) $H_p = (0,0)$	0.735	0.012	(4) Vs (5)	-0.231	96.20 (0.000)
Secondary school completion					
(6) $H_s = (0,1,1)$			(6) Vs (7)	-0.076	3.85 (0.050)
(7) $H_s = (0,0,1)$			(6) Vs (8)	-0.139	22.00 (0.000)
(8) $H_s = (0,0,0)$			(7) Vs (8)	-0.063	7.71 (0.005)

Notes: See notes to Table 1. There is only one case with $H_s = (1,1,1)$ contributing to the estimation sample (see text for details), which we aggregate with $H_s = (0,1,1)$. LR test corresponds to the likelihood ratio test for equality of transition rates (p values in parentheses).

Table 3. Means and Standard Deviations of Characteristics by Work Sequence

	$H_e = 1$		$H_e = 0$		$H_p = (0,1)$		$H_p = (0,0)$		$H_s = (0,1,1)$		$H_s = (0,0,1)$		$H_s = (0,0,0)$	
	Mean	(Stdv)	Mean	(Stdv)	Mean	(Stdv)	Mean	(Stdv)	Mean	(Stdv)	Mean	(Stdv)	Mean	(Stdv)
Girl	0.39	(0.49)	0.49	(0.50)	0.45	(0.50)	0.47	(0.50)	0.33	(0.47)	0.43	(0.50)	0.44	(0.50)
Age	19.99	(3.21)	19.29	(3.05)	19.42	(3.07)	19.32	(2.98)	21.19	(2.52)	21.15	(2.70)	21.07	(2.40)
School entry age			8.01	(1.67)	8.01	(1.67)	7.50	(1.49)	7.57	(1.54)	7.00	(1.32)	7.05	(1.40)
Grade repetitions in primary					0.27	(0.45)	0.27	(0.45)	0.25	(0.85)	0.06	(0.33)	0.06	(0.31)
Compulsory school policy					0.45	(0.50)	0.43	(0.49)	0.63	(0.48)	0.53	(0.50)	0.62	(0.49)
Free tuition policy					0.20	(0.40)	0.21	(0.41)	0.22	(0.41)	0.21	(0.41)	0.26	(0.44)
Free tuition policy * girl														
Mother's education	0.48	(1.29)	1.35	(2.35)	1.08	(2.05)	1.69	(2.56)	1.59	(2.40)	2.21	(2.62)	2.13	(2.94)
Father's education	1.37	(2.65)	3.58	(3.73)	3.35	(3.56)	4.15	(3.80)	4.17	(3.63)	4.93	(3.43)	5.23	(4.10)
Household assets missing	0.22	(0.41)	0.16	(0.37)	0.17	(0.37)	0.15	(0.36)	0.13	(0.34)	0.25	(0.43)	0.18	(0.38)
Log(household assets) ¹	10.24	(1.05)	10.92	(1.12)	10.95	(1.05)	11.05	(1.11)	11.34	(0.91)	11.13	(1.14)	11.45	(1.06)
Modern latrine	0.10	(0.30)	0.28	(0.45)	0.30	(0.46)	0.32	(0.47)	0.37	(0.49)	0.36	(0.48)	0.42	(0.49)
Cultivating household	0.71	(0.45)	0.74	(0.44)	0.82	(0.39)	0.73	(0.44)	0.80	(0.40)	0.71	(0.46)	0.68	(0.46)
Owens farm land	0.61	(0.49)	0.78	(0.42)	0.84	(0.37)	0.80	(0.40)	0.90	(0.30)	0.87	(0.33)	0.86	(0.35)
Owens non-farm business	0.38	(0.48)	0.43	(0.49)	0.51	(0.50)	0.42	(0.49)	0.59	(0.49)	0.46	(0.50)	0.41	(0.49)
Older siblings	2.07	(1.86)	2.50	(2.09)	2.53	(1.95)	2.51	(2.21)	2.65	(1.83)	2.28	(2.17)	2.66	(2.37)
Younger siblings	1.14	(0.93)	1.00	(0.83)	1.28	(1.05)	1.23	(0.97)	1.66	(1.42)	1.80	(1.31)	1.78	(1.50)
Village outside Matlab ²	0.11	(0.32)	0.09	(0.28)	0.08	(0.28)	0.09	(0.29)	0.05	(0.21)	0.17	(0.38)	0.13	(0.34)
Tubewell in village	0.77	(0.42)	0.78	(0.42)	0.84	(0.37)	0.82	(0.39)	0.84	(0.37)	0.85	(0.35)	0.85	(0.36)
Health facility in village	0.006	(0.08)	0.40	(0.49)	0.70	(0.46)	0.70	(0.46)	0.77	(0.42)	0.77	(0.42)	0.84	(0.36)
Industry in village	0.13	(0.34)	0.15	(0.35)	0.40	(0.49)	0.38	(0.49)	0.54	(0.50)	0.49	(0.50)	0.50	(0.50)
Primary school in village	0.64	(0.48)	0.75	(0.43)	0.74	(0.44)	0.77	(0.42)						
Secondary school in village	0.01	(0.10)	0.25	(0.43)	0.27	(0.45)	0.33	(0.47)	0.26	(0.44)	0.32	(0.47)	0.32	(0.47)
Distance to Matlab capital	6.38	(5.36)	5.82	(4.49)	6.48	(5.23)	5.42	(4.19)	7.53	(6.52)	5.03	(3.88)	5.36	(4.31)
Number of individuals	329		2160		533		1339		138		183		333	

Notes: Based on weighted data. Samples account for censoring (see text for details) so that the individuals in each sample are those used in the estimation. (1) Mean defined for non-missing observations. (2) Means of village characteristics defined for children whose village of residence in each level was in Matlab, since only those villages were surveyed. $H_p = (0,1)$ also includes the 19 individuals with $H_p = (1,1)$, while $H_s = (0,1,1)$ also includes the only individual with $H_s = (1,1,1)$.

Table 4.1. Dynamic Switching Model of School and Work: School Entry Level

	Work equation		School equations			
	W_{te}		$S_{te}(0)$		$S_{te}(1)$	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept	-6.259	-4.606	-6.465	-13.466	-1.518	-0.062
Baseline hazard (1)	2.868	3.085	1.688	13.547	9.040	1.892
Baseline hazard (2)	3.314	3.610	2.883	23.174		
Baseline hazard (3)	3.878	4.318	3.613	27.967		
Baseline hazard (4)	6.975	8.216	3.595	25.613		
Baseline hazard (5)	5.368	6.137	4.011	27.003		
Baseline hazard (6)	7.588	8.876	3.048	14.854		
Baseline hazard (7)	7.053	8.145	2.859	12.310		
Baseline hazard (8)			1.053	2.470		
Girl	-1.022	-5.046	-0.401	-5.187	1.017	0.489
Age	0.010	0.359	0.021	1.600	-1.486	-2.879
Mother's education	0.027	0.385	0.151	8.644	1.781	2.387
Father's education	0.038	1.032	0.086	7.848	1.082	2.105
Household assets missing	-2.207	-2.253	0.658	1.482	14.386	0.779
Log(household assets)	-0.244	-2.630	0.063	1.603	0.967	0.545
Modern latrine	-0.053	-0.198	0.599	7.441	7.984	2.392
Cultivating household	0.241	1.176				
Owns farm land	0.183	0.966	0.544	6.062	-2.458	-1.105
Owns non-farm business	0.015	0.089	0.120	1.665	5.595	2.621
Older siblings	-0.061	-1.279	-0.014	-0.758	1.031	2.331
Younger siblings	0.170	1.890	-0.010	-0.232	-0.409	-0.278
Village outside Matlab	0.211	0.566	0.466	2.617	-7.980	-1.089
Tubewell in village	-0.112	-0.532	0.141	1.593	-0.400	-0.186
Health facility in village	-0.628	-0.754	1.113	13.028	5.740	1.281
Industry in village	-0.133	-0.513	0.231	2.162	6.681	3.180
Primary school in village	-0.186	-1.026	0.121	1.368	-1.520	-0.722
Secondary school in village	-0.695	-1.026	0.701	7.395	7.732	1.466
Distance to Matlab capital	-0.007	-0.365	-0.044	-5.244	-0.349	-0.679
θ_w	1.000					
θ_s			9.602	3.432	1.000	
$Corr(\theta_w, \theta_s)$	-0.315					
Log-L	-8527.370					

Notes: The specification of the baseline hazard in each equation is given by the variation in the data. For equation W_{te} , baseline(1) through baseline(7) refer to work ages 7 through 13, while the reference is ages < 7. For equation $S_{te}(0)$, baseline(1) through baseline(7) refer to school entry ages 6 through 12, baseline(7) refers to ages 13 and 14, and the reference is age = 5. For equation $S_{te}(1)$, baseline(1) refers to school entry ages between 10 and 14, while the reference is ages < 10.

Table 4.2. Dynamic Switching Model of School and Work: Primary School Level

	Work equation				School equations			
	W_{ip}		$S_{ip}(0,0)$		$S_{ip}(0,1)$		$S_{ip}(1,1)$	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept	-3.064	-2.479	-8.188	-5.788	-11.728	-3.646	-13.628	-3.649
Baseline hazard (1)			-4.077	-22.866	-3.037	-11.709		
Girl	-0.046	-0.249	-0.667	-3.114	-0.073	-0.159		
Age	-0.022	-0.662	0.212	5.118	0.255	2.782		
School entry age	0.290	4.357	-0.176	-2.655	-0.006	-0.053		
Compulsory school policy	-0.511	-2.425	1.363	5.509	0.489	1.095		
Free tuition policy	-0.014	-0.063	0.363	1.380	1.262	2.595		
Free tuition policy * girl	-0.019	-0.075	0.822	2.802	0.245	0.425		
Mother's education	-0.041	-1.331	0.164	4.413	0.401	4.282		
Father's education	-0.040	-2.115	0.137	5.459	0.072	1.872		
Household assets missing	-0.552	-0.684	2.928	3.096	1.085	0.617		
Log(household assets)	-0.101	-1.444	0.342	4.061	0.228	1.517		
Modern latrine	0.087	0.639	0.590	3.659	1.630	4.372		
Cultivating household	0.434	2.494						
Owns farm land	0.390	2.116	0.541	2.818	0.741	1.662		
Owns non-farm business	0.510	3.864	0.294	1.982	0.896	2.658		
Older siblings	-0.034	-1.070	-0.013	-0.356	0.238	2.763		
Younger siblings	-0.004	-0.059	0.028	0.369	0.427	3.078		
Village outside Matlab	0.424	1.226	1.451	3.597	-1.844	-2.151		
Tubewell in village	0.061	0.351	0.653	3.406	-0.319	-0.803		
Health facility in village	0.406	2.568	1.014	5.570	0.596	1.689		
Industry in village	-0.075	-0.539	0.339	2.056	0.369	1.207		
Primary school in village	0.152	0.910	-0.262	-1.398	-0.069	-0.191		
Secondary school in village	-0.167	-1.101	-0.094	-0.545	1.097	2.819		
Distance to Matlab capital	0.049	3.317	-0.041	-2.197	-0.040	-1.280		
θ_w	-0.823	-2.699						
θ_s			0.836	3.723	1.901	4.408		

Notes: For all school outcome equations, baseline(1) refers to 6 or more years to reach secondary school, while the reference is 5 years.

Table 4.3. Dynamic Switching Model of School and Work: Secondary School Level

	Work equation		School equations					
	W_{ts}		$S_{ts}(0,0,0)$		$S_{ts}(0,0,1)$		$S_{ts}(0,1,1)/S_{ts}(1,1,1)$	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept	-4.354	-0.296	-37.013	-5.096	-50.748	-4.011	-36.176	-3.323
Baseline hazard (1)			-2.765	-5.174	-1.294	-2.265	-3.389	-4.767
Girl	-9.689	-2.563	-4.354	-4.026	-0.752	-0.533	-1.316	-1.059
Age	-0.697	-1.593	0.503	2.424	0.794	2.695	0.731	2.332
School entry age	6.351	3.487	0.326	1.299	-0.364	-1.007	0.091	0.274
Grade repetitions in primary	6.384	3.597	-0.374	-0.501	-4.411	-2.092	-1.714	-1.625
Free tuition policy	-10.921	-2.437	-1.100	-0.944	6.688	4.202	1.322	0.899
Free tuition policy * girl	10.239	2.448	5.988	4.217	-6.281	-3.302	0.909	0.626
Mother's education	-0.033	-0.174	1.057	6.031	1.438	4.373	0.448	1.825
Father's education	-0.078	-0.388	0.550	4.597	-0.070	-0.443	0.008	0.069
Household assets missing	-10.366	-0.898	12.162	2.604	14.424	1.635	10.265	1.679
Log(household assets)	-3.330	-2.411	1.433	3.728	1.929	2.387	1.103	2.058
Modern latrine	1.651	1.450	-0.704	-1.138	1.732	1.750	-0.707	-0.856
Cultivating household	24.740	3.887						
Owns farm land	2.762	1.577	2.540	2.397	2.638	2.146	-0.489	-0.441
Owns non-farm business	13.604	4.022	1.615	2.567	0.281	0.353	0.193	0.241
Older siblings	-2.997	-3.338	0.398	2.529	0.283	1.156	-0.406	-1.552
Younger siblings	-1.248	-2.036	0.216	0.970	-1.016	-2.076	0.250	0.964
Village outside Matlab	-8.244	-1.529	-0.611	-0.384	4.517	2.186	4.220	1.410
Tubewell in village	-6.867	-1.875	-0.953	-0.989	1.675	1.239	1.493	1.272
Health facility in village	-7.986	-2.821	-0.864	-1.009	1.818	1.569	2.517	2.050
Industry in village	4.436	1.742	6.149	5.177	-1.587	-1.534	1.927	2.289
Secondary school in village	7.856	2.945	3.548	3.812	-0.018	-0.013	2.417	2.344
Distance to Matlab capital	-0.439	-1.423	-0.360	-3.744	-0.005	-0.053	-0.234	-2.097
θ_w	-34.930	-3.490						
θ_s			5.923	6.690	7.123	4.314	2.810	2.903

Notes: For all school outcome equations, baseline(1) refers to 6 or more years to complete secondary school, while the reference is 5 years. There is only one case in $S_{ts}(1,1,1)$ that is thus aggregated with $S_{ts}(0,1,1)$.

Table 5. Dynamic Effects of Work by School Outcome, Work Sequence Pair and Population

	DATE	DTT	DTU
School entry			
$H_e = 1$ Vs $H_e = 0$	-0.2592 (0.042)	-0.1080 (0.046)	-0.2818 (0.046)
Transition to secondary school			
$H_p = (1,1)$ Vs $H_p = (0,1)$	-0.2247 (0.060)	-0.1347 (0.093)	-0.2984 (0.071)
$H_p = (1,1)$ Vs $H_p = (0,0)$	-0.3320 (0.060)	-0.1899 (0.093)	-0.4165 (0.070)
$H_p = (0,1)$ Vs $H_p = (0,0)$	-0.1073 (0.037)	-0.1197 (0.051)	-0.1369 (0.043)
Secondary school completion			
$H_s = (1,1,1)$ Vs $H_s = (0,1,1)$	-0.0395 (0.017)	-0.0289 (0.034)	-0.0966 (0.039)
$H_s = (1,1,1)$ Vs $H_s = (0,0,1)$	-0.0961 (0.048)	-0.0798 (0.066)	-0.1789 (0.088)
$H_s = (1,1,1)$ Vs $H_s = (0,0,0)$	-0.2474 (0.046)	-0.2781 (0.073)	-0.4746 (0.072)
$H_s = (0,1,1)$ Vs $H_s = (0,0,1)$	-0.0566 (0.046)	-0.1037 (0.086)	-0.1113 (0.081)
$H_s = (0,1,1)$ Vs $H_s = (0,0,0)$	-0.2079 (0.048)	-0.3458 (0.081)	-0.4149 (0.072)
$H_s = (0,0,1)$ Vs $H_s = (0,0,0)$	-0.1513 (0.038)	-0.3137 (0.083)	-0.2764 (0.066)

Notes: Standard errors in parentheses.

Table 6. Distribution of θ_s Across School Levels by Work Sequence

School entry				
	$H_e = 1$	$H_e = 0$		
Low type	27.61	4.89		
Middle type	68.80	65.53		
High type	3.60	29.58		
Primary school				
	$H_p = (1,1)$	$H_p = (0,1)$	$H_p = (0,0)$	
Low type	6.97	1.69	4.84	
Middle type	70.29	59.51	63.71	
High type	22.74	38.80	31.45	
Secondary school				
	$H_s = (1,1,1)$	$H_s = (0,1,1)$	$H_s = (0,0,1)$	$H_s = (0,0,0)$
Low type	3.87	0.49	0.22	5.45
Middle type	52.23	43.42	54.02	62.29
High type	43.90	56.09	45.75	32.27

Note: Numbers are percentages.

Table 7. Selection into Work by School Level

Unobservables		Observables	
School entry			
$Corr[U_e^w, U_e^s(1)]$	-0.0735	$Corr[\beta_{ez}^w Z_e, \beta_{ex}^s X_e(1)]$	-0.4201
$Corr[U_e^w, U_e^s(0)]$	-0.1495	$Corr[\beta_{ez}^w Z_e, \beta_{ex}^s X_e(0)]$	-0.3100
Primary school			
$Corr[U_p^w, U_p^s(0,1)]$	0.094	$Corr[\beta_{pz}^w Z_p, \beta_{px}^s X_p(0,1)]$	-0.2320
$Corr[U_p^w, U_p^s(0,0)]$	0.054	$Corr[\beta_{pz}^w Z_p, \beta_{px}^s X_p(0,0)]$	-0.3786
Secondary school			
$Corr[U_s^w, U_s^s(0,0,1)]$	0.3050	$Corr[\beta_{sz}^w Z_s, \beta_{sx}^s X_s(0,0,1)]$	-0.3592
$Corr[U_s^w, U_s^s(0,0,0)]$	0.3010	$Corr[\beta_{sz}^w Z_s, \beta_{sx}^s X_s(0,0,0)]$	-0.1026

Table 8. Policy Simulations

Policy 1: No compulsory primary schooling									
School level	Primary			Secondary					
School effect	-0.1934			-0.0229					
Work effect	0.0909			0.0499					
Policy 2: School entry at 6 years of age									
School level	Primary			Secondary					
School effect	0.0776			0.0054					
Work effect	-0.1609			-0.1876					
Policy 3: Secondary school availability									
School level	Entry		Primary				Secondary		
Work sequence	(1)	(0)	(1,1)	(0,1)	(0,0)	(1,1,1)	(0,1,1)	(0,0,1)	(0,0,0)
School effect	0.3989	0.0531	0.2558	0.1022	-0.0149	0.1903	0.1264	0.0059	0.0657

APPENDIX A
Model Comparison

Table A. Likelihood Ratio Tests and Bayesian Information Criterion

Model	Log-L	Parameters	BIC rank
(1) No heterogeneity	-8586.15	239	3
(2) Normal heterogeneity	-8601.82	248	4
(3) Non-parametric (2 factors)	-8527.37	257	1
(4) Non-parametric (3 factors)	-8512.37	275	2
Likelihood ratio tests			
(1) Vs (2)	NA		
(1) Vs (3)	$\chi^2(18) = 117.56 (0.000)$		
(1) Vs (4)	$\chi^2(36) = 147.57 (0.000)$		
(2) Vs (3)	$\chi^2(9) = 148.91 (0.000)$		
(2) Vs (4)	$\chi^2(27) = 178.91 (0.000)$		
(3) Vs (4)	$\chi^2(18) = 30.00 (0.000)$		

Notes: BIC rank goes from best to worst. Values in parentheses are p-values.

(NOT FOR PUBLICATION)

APPENDIX B
Model Without Unobserved Heterogeneity

Table B.1. School Entry Level

	Work equation		School equations			
	W_{te}		$S_{te}(0)$		$S_{te}(1)$	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept	-6.353	-5.205	-5.348	-13.823	-5.782	-0.998
Baseline hazard (1)	2.916	3.151	1.421	12.145	1.516	0.725
Baseline hazard (2)	3.420	3.743	2.304	19.984		
Baseline hazard (3)	4.006	4.479	2.773	23.238		
Baseline hazard (4)	6.983	8.275	2.578	19.839		
Baseline hazard (5)	5.375	6.184	2.844	20.562		
Baseline hazard (6)	7.410	8.741	1.877	9.642		
Baseline hazard (7)	6.766	7.892	1.643	7.377		
Baseline hazard (8)			-0.132	-0.315		
Girl	-1.005	-5.873	-0.298	-4.834	0.672	0.904
Age	0.005	0.193	0.024	2.317	-0.383	-2.932
Mother's education	0.077	1.299	0.116	7.942	0.207	0.979
Father's education	0.039	1.306	0.069	7.673	0.126	0.977
Household assets missing	-1.620	-1.999	0.293	0.823	6.982	1.237
Log(household assets)	-0.191	-2.480	0.033	1.039	0.552	1.100
Modern latrine	0.058	0.254	0.513	7.812	1.787	2.078
Cultivating household	0.176	1.030				
Owns farm land	0.234	1.497	0.404	5.670	0.011	0.010
Owns non-farm business	0.031	0.207	0.082	1.399	0.479	0.586
Older siblings	-0.063	-1.560	-0.008	-0.514	0.214	1.322
Younger siblings	0.156	2.067	-0.003	-0.089	0.373	1.139
Village outside Matlab	0.190	0.611	0.392	2.698	-2.399	-1.286
Tubewell in village	-0.101	-0.571	0.177	2.455	-0.512	-0.676
Health facility in village	-0.716	-0.865	0.935	13.379	2.838	1.279
Industry in village	-0.111	-0.502	0.257	3.030	1.581	2.273
Primary school in village	-0.125	-0.831	0.010	0.144	-0.072	-0.093
Secondary school in village	-0.777	-1.201	0.523	6.686	2.144	1.246
Distance to Matlab capital	-0.008	-0.518	-0.033	-4.914	-0.031	-0.385
Log-L	-8586.153					

Notes: See notes to table 4.1.

Table B.2. Primary School Level

	Work equation		School equations					
	W_{ip}		$S_{ip}(0,0)$		$S_{ip}(0,1)$		$S_{ip}(1,1)$	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept	-1.458	-1.678	-5.281	-5.221	-2.944	-1.973	-3.137	-1.926
Baseline hazard (1)			-4.217	-25.905	-3.407	-16.024		
Girl	-0.015	-0.086	-0.507	-2.791	0.365	1.212		
Age	-0.029	-0.965	0.200	5.583	0.131	2.582		
School entry age	0.196	4.386	-0.350	-6.683	-0.356	-4.408		
Compulsory school policy	-0.497	-2.633	1.232	5.557	0.385	1.215		
Free tuition policy	-0.040	-0.198	0.332	1.403	0.667	1.977		
Free tuition policy * girl	-0.025	-0.109	0.742	2.807	-0.076	-0.193		
Mother's education	-0.055	-2.013	0.128	3.848	0.186	3.723		
Father's education	-0.038	-2.211	0.101	4.944	0.030	1.049		
Household assets missing	-0.759	-1.046	2.355	2.833	0.164	0.132		
Log(household assets)	-0.111	-1.761	0.281	3.837	0.121	1.143		
Modern latrine	0.036	0.295	0.469	3.292	0.814	3.631		
Cultivating household	0.321	2.108						
Owns farm land	0.313	1.924	0.336	2.104	0.125	0.473		
Owns non-farm business	0.433	3.790	0.257	1.923	0.459	2.318		
Older siblings	-0.034	-1.178	-0.010	-0.303	0.166	2.971		
Younger siblings	0.002	0.027	0.011	0.152	0.289	3.039		
Village outside Matlab	0.242	0.792	1.128	3.348	-1.353	-2.232		
Tubewell in village	0.085	0.539	0.614	3.568	0.119	0.455		
Health facility in village	0.289	2.156	0.776	5.189	-0.115	-0.505		
Industry in village	-0.064	-0.508	0.277	1.875	0.101	0.494		
Primary school in village	0.030	0.206	-0.350	-2.093	-0.124	-0.485		
Secondary school in village	-0.226	-1.658	-0.271	-1.770	0.451	1.939		
Distance to Matlab capital	0.041	3.224	-0.027	-1.715	-0.013	-0.587		

Notes: See notes to table 4.2.

Table B.3. Secondary School Level

	Work equation				School equations			
	W_{ts}		$S_{ts}(0,0,0)$		$S_{ts}(0,0,1)$		$S_{ts}(0,1,1)/S_{ts}(1,1,1)$	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept	5.665	3.478	-7.302	-2.924	-10.644	-2.949	-22.623	-3.405
Baseline hazard (1)			-4.401	-11.189	-3.373	-8.629	-3.658	-5.900
Girl	-0.389	-1.296	-1.334	-3.041	0.051	0.091	-0.046	-0.053
Age	-0.099	-1.766	0.299	3.166	0.303	2.561	0.556	2.633
School entry age	0.078	1.060	-0.339	-2.984	-0.626	-3.593	-0.577	-2.658
Grade repetitions in primary	0.343	1.486	-0.236	-0.568	-2.675	-1.845	-1.792	-1.917
Free tuition policy	-0.922	-2.678	-0.438	-0.865	2.207	3.237	1.291	1.161
Free tuition policy * girl	0.288	0.819	2.456	4.526	-1.732	-2.327	0.124	0.112
Mother's education	-0.042	-1.214	0.328	5.688	0.238	3.110	-0.035	-0.288
Father's education	0.011	0.419	0.014	0.372	0.028	0.475	-0.023	-0.300
Household assets missing	-2.698	-2.322	0.091	0.052	4.545	1.791	9.879	2.470
Log(household assets)	-0.315	-3.112	0.155	1.061	0.559	2.530	1.013	2.911
Modern latrine	0.076	0.418	0.034	0.124	0.199	0.507	-0.699	-1.249
Cultivating household	0.223	0.995						
Owens farm land	-0.232	-0.855	0.247	0.655	-0.373	-0.722	-0.100	-0.135
Owens non-farm business	0.534	2.924	0.144	0.516	-0.141	-0.394	0.220	0.409
Older siblings	-0.077	-1.616	0.167	2.432	0.132	1.467	-0.294	-1.849
Younger siblings	0.005	0.076	0.024	0.226	0.177	1.237	0.383	1.934
Village outside Matlab	-0.677	-1.494	0.818	1.047	0.494	0.538	2.763	1.506
Tubewell in village	-0.130	-0.503	0.300	0.734	0.132	0.212	1.795	2.117
Health facility in village	-0.626	-2.589	-0.364	-0.945	0.129	0.289	0.323	0.512
Industry in village	0.279	1.376	1.111	3.423	1.050	2.393	1.193	2.083
Secondary school in village	0.014	0.073	0.648	2.116	-1.678	-3.389	1.304	2.133
Distance to Matlab capital	-0.030	-1.334	0.040	1.143	0.050	1.083	-0.094	-1.657

Notes: See notes to table 4.3.

(NOT FOR PUBLICATION)

APPENDIX C
Model with Normal Heterogeneity

Table C.1. School Entry Level

	Work equation		School equations			
	W_{te}		$S_{te}(0)$		$S_{te}(1)$	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept	-6.338	-4.621	-5.463	-11.643	-5.109	-0.902
Baseline hazard (1)	2.751	2.971	1.669	13.447	1.421	0.704
Baseline hazard (2)	3.189	3.479	2.824	22.857		
Baseline hazard (3)	3.728	4.141	3.521	27.378		
Baseline hazard (4)	6.858	8.016	3.456	24.642		
Baseline hazard (5)	5.277	5.970	3.881	25.946		
Baseline hazard (6)	7.489	8.609	2.910	14.005		
Baseline hazard (7)	6.946	7.861	2.696	11.408		
Baseline hazard (8)			0.831	1.921		
Girl	-1.025	-4.881	-0.387	-5.104	0.626	0.824
Age	0.005	0.165	0.015	1.196	-0.387	-2.910
Mother's education	0.041	0.574	0.158	8.961	0.253	1.138
Father's education	0.030	0.808	0.102	9.325	0.149	1.128
Household assets missing	-2.034	-2.021	-0.250	-0.561	6.693	1.240
Log(household assets)	-0.228	-2.378	-0.033	-0.830	0.520	1.078
Modern latrine	-0.020	-0.072	0.554	6.921	1.841	2.127
Cultivating household	0.240	1.160				
Owns farm land	0.151	0.765	0.575	6.481	0.138	0.176
Owns non-farm business	0.027	0.151	0.129	1.766	0.530	0.673
Older siblings	-0.079	-1.624	0.025	1.351	0.256	1.548
Younger siblings	0.161	1.756	0.018	0.429	0.350	1.046
Village outside Matlab	0.185	0.488	0.394	2.217	-2.341	-1.234
Tubewell in village	-0.149	-0.695	0.203	2.327	-0.431	-0.551
Health facility in village	-0.542	-0.642	0.998	12.000	3.063	1.394
Industry in village	-0.190	-0.716	0.227	2.201	1.707	2.352
Primary school in village	-0.159	-0.864	0.023	0.264	-0.081	-0.104
Secondary school in village	-0.730	-1.062	0.725	7.617	2.285	1.376
Distance to Matlab capital	-0.006	-0.304	-0.043	-5.014	-0.040	-0.499
θ_w	1.000					
θ_s			0.600	1.052	1.000	
$Corr(\theta_w, \theta_s)$	-0.964	-1.923				
Log-L	-8601.824					

Notes: See notes to table 4.1.

Table C.2. Primary School Level

	Work equation		School equations					
	W_{ip}		$S_{ip}(0,0)$		$S_{ip}(0,1)$		$S_{ip}(1,1)$	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept	-1.824	-2.020	-19.571	-6.995	-17.316	-4.314	-12.724	-2.833
Baseline hazard (1)			-2.486	-8.607	0.047	0.090		
Girl	-0.003	-0.015	-0.254	-0.570	2.865	1.733		
Age	-0.035	-1.109	0.311	3.869	0.457	2.250		
School entry age	0.247	4.846	-0.068	-0.616	-0.740	-2.329		
Compulsory school policy	-0.527	-2.729	3.515	4.768	-1.527	-1.067		
Free tuition policy	-0.029	-0.138	0.563	1.039	7.596	4.805		
Free tuition policy * girl	-0.058	-0.246	0.986	1.585	-7.804	-2.984		
Mother's education	-0.048	-1.722	0.341	4.749	2.663	5.403		
Father's education	-0.030	-1.696	0.544	6.048	0.518	3.833		
Household assets missing	-0.873	-1.179	5.230	2.085	-22.619	-3.117		
Log(household assets)	-0.125	-1.926	0.658	2.938	-1.488	-2.771		
Modern latrine	0.043	0.344	1.057	2.998	8.292	4.851		
Cultivating household	0.327	2.104						
Owens farm land	0.358	2.140	2.350	4.221	4.623	3.446		
Owens non-farm business	0.456	3.892	1.665	3.058	1.942	2.083		
Older siblings	-0.027	-0.924	0.048	0.493	3.688	4.875		
Younger siblings	0.012	0.211	0.144	1.028	4.773	4.853		
Village outside Matlab	0.262	0.840	4.323	3.963	-10.937	-4.083		
Tubewell in village	0.089	0.556	1.946	4.160	-0.574	-0.977		
Health facility in village	0.344	2.466	2.868	5.935	1.993	3.088		
Industry in village	-0.057	-0.444	0.626	1.717	8.110	5.104		
Primary school in village	0.016	0.104	-0.599	-1.843	-7.615	-4.165		
Secondary school in village	-0.176	-1.254	-0.219	-0.662	11.766	4.502		
Distance to Matlab capital	0.039	3.037	-0.184	-3.848	-0.258	-2.317		
θ_w	-0.326	-1.737						
θ_s			4.754	7.026	16.593	5.420		

Notes: See notes to table 4.2.

Table C.3. Secondary School Level

	Work equation		School equations					
	W_{ts}		$S_{ts}(0,0,0)$		$S_{ts}(0,0,1)$		$S_{ts}(0,1,1)/S_{ts}(1,1,1)$	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept	7.544	3.233	-30.685	-2.553	-38.910	-4.543	-28.497	-2.950
Baseline hazard (1)			-2.799	-4.635	-0.597	-0.860	-3.644	-5.743
Girl	-0.395	-1.136	-7.802	-2.764	1.525	1.789	-0.027	-0.030
Age	-0.088	-1.372	0.454	1.923	0.418	1.636	0.626	2.528
School entry age	-0.071	-0.665	0.764	1.823	0.339	1.454	-0.337	-1.154
Grade repetitions in primary	0.006	0.019	4.764	2.823	-0.321	-0.199	-1.641	-1.691
Free tuition policy	-0.917	-2.240	-2.841	-2.277	4.713	3.797	1.577	1.272
Free tuition policy * girl	0.257	0.640	10.199	2.830	-8.518	-5.888	-0.082	-0.071
Mother's education	-0.087	-1.864	1.728	3.197	2.072	4.230	0.098	0.568
Father's education	-0.024	-0.739	0.527	3.101	0.725	3.884	0.010	0.116
Household assets missing	-2.725	-2.002	-6.200	-1.470	-0.662	-0.399	9.165	2.075
Log(household assets)	-0.323	-2.655	-0.130	-0.356	0.471	3.028	0.990	2.566
Modern latrine	0.038	0.187	1.082	1.724	1.142	1.499	-0.457	-0.733
Cultivating household	0.216	0.863						
Owns farm land	-0.361	-1.133	1.526	1.574	1.063	1.825	0.163	0.205
Owns non-farm business	0.407	1.902	2.924	2.946	1.333	2.033	0.363	0.609
Older siblings	-0.106	-1.818	1.078	2.786	1.159	3.331	-0.216	-1.150
Younger siblings	-0.008	-0.098	0.218	0.973	0.669	2.684	0.446	1.939
Village outside Matlab	-0.798	-1.495	0.341	0.153	-0.192	-0.073	3.183	1.589
Tubewell in village	-0.155	-0.527	-0.258	-0.318	3.475	2.219	1.954	2.089
Health facility in village	-0.861	-2.615	2.156	1.905	-0.048	-0.044	0.511	0.721
Industry in village	0.308	1.322	2.302	3.118	2.517	2.582	1.531	2.222
Secondary school in village	-0.088	-0.382	2.732	3.274	0.986	0.866	1.773	2.187
Distance to Matlab capital	-0.021	-0.799	0.091	1.275	-0.448	-2.982	-0.093	-1.529
θ_w	0.928	1.461						
θ_s			8.041	3.463	10.760	4.526	1.219	1.141

Notes: See notes to table 4.3.

APPENDIX D
Model with Three-Factor Non-Parametric Heterogeneity

Table D.1. Entry Level

	Work equation		School equations			
	W_{te}		$S_{te}(0)$		$S_{te}(1)$	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept	-6.484	-4.556	-6.530	-12.244	-4.955	-0.847
Baseline hazard (1)	2.902	3.109	1.732	13.504	1.544	0.729
Baseline hazard (2)	3.383	3.620	2.938	22.990		
Baseline hazard (3)	3.970	4.257	3.658	27.340		
Baseline hazard (4)	7.064	7.604	3.617	23.751		
Baseline hazard (5)	5.477	5.646	4.017	23.978		
Baseline hazard (6)	7.734	7.526	3.088	13.117		
Baseline hazard (7)	7.231	6.736	2.881	10.736		
Baseline hazard (8)			1.118	2.236		
Girl	-1.039	-4.675	-0.382	-4.638	0.815	0.983
Age	0.023	0.688	0.026	1.830	-0.402	-2.906
Mother's education	0.051	0.680	0.137	7.529	0.192	0.849
Father's education	0.055	1.325	0.086	7.600	0.112	0.825
Household assets missing	-2.322	-2.086	0.512	1.024	6.449	1.165
Log(household assets)	-0.251	-2.397	0.052	1.171	0.497	1.004
Modern latrine	0.008	0.012	0.647	7.597	1.711	1.935
Cultivating household	0.227	1.073				
Owns farm land	0.275	1.193	0.518	5.471	0.059	0.074
Owns non-farm business	0.052	0.248	0.133	1.584	0.498	0.618
Older siblings	-0.056	-0.966	-0.011	-0.603	0.257	1.547
Younger siblings	0.159	1.688	0.016	0.352	0.328	0.935
Village outside Matlab	0.205	0.527	0.494	2.620	-2.404	-1.251
Tubewell in village	-0.132	-0.593	0.183	1.864	-0.450	-0.557
Health facility in village	-0.580	-0.691	1.214	10.569	3.497	1.536
Industry in village	-0.147	-0.548	0.275	1.783	1.707	2.298
Primary school in village	-0.211	-1.092	0.040	0.282	-0.168	-0.205
Secondary school in village	-0.823	-1.188	0.607	5.999	1.837	1.089
Distance to Matlab capital	-0.015	-0.705	-0.035	-2.629	-0.027	-0.332
θ_w	1.000					
θ_s			1.000		1.000	
$Corr(\theta_w, \theta_{s0})$	-0.103					
$Corr(\theta_w, \theta_{s1})$	-0.732					
$Corr(\theta_{s0}, \theta_{s1})$	0.132					
Log-L	-8512.369					

Notes: See notes to table 4.1.

Table D.2. Primary School Level

	Work equation				School equations			
	W_{ip}		$S_{ip}(0,0)$		$S_{ip}(0,1)$		$S_{ip}(1,1)$	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept	-2.149	-1.756	-6.942	-4.236	-14.742	-1.996	-10.914	-1.323
Baseline hazard (1)			-4.152	-22.125	-0.531	-0.893		
Girl	-0.012	-0.057	-0.613	-2.909	0.484	0.489		
Age	-0.027	-0.793	0.200	5.176	0.671	2.697		
School entry age	0.226	3.509	-0.221	-2.298	-1.705	-4.388		
Compulsory school policy	-0.497	-2.375	1.286	5.360	2.483	1.489		
Free tuition policy	-0.059	-0.249	0.294	1.183	5.494	2.349		
Free tuition policy * girl	-0.017	-0.057	0.804	2.857	-4.262	-1.295		
Mother's education	-0.070	-2.278	0.144	3.712	1.502	5.156		
Father's education	-0.040	-2.070	0.121	4.543	0.046	0.371		
Household assets missing	-0.737	-0.926	2.535	2.865	-9.388	-2.206		
Log(household assets)	-0.110	-1.602	0.301	3.822	-0.527	-1.529		
Modern latrine	0.067	0.484	0.585	3.366	7.395	5.546		
Cultivating household	0.350	2.120						
Owns farm land	0.375	2.063	0.445	2.332	3.284	3.464		
Owns non-farm business	0.488	3.850	0.295	2.059	4.333	2.907		
Older siblings	-0.045	-1.414	-0.015	-0.431	0.736	3.726		
Younger siblings	-0.016	-0.253	0.026	0.359	-0.325	-1.046		
Village outside Matlab	0.425	1.139	1.345	3.261	0.283	0.086		
Tubewell in village	0.175	0.985	0.645	3.464	5.059	3.732		
Health facility in village	0.332	2.006	0.955	4.415	0.631	0.557		
Industry in village	-0.104	-0.757	0.294	1.897	-1.308	-2.031		
Primary school in village	0.151	0.908	-0.304	-1.723	1.267	0.620		
Secondary school in village	-0.175	-1.161	-0.197	-1.192	5.619	2.975		
Distance to Matlab capital	0.049	3.372	-0.025	-1.478	-0.214	-2.441		
θ_w	-0.759	-4.548						
θ_s			0.545	1.304	10.192	6.189		

Notes: See notes to table 4.2.

Table D.3. Secondary School Level

	Work equation		School equations					
	W_{ts}		$S_{ts}(0,0,0)$		$S_{ts}(0,0,1)$		$S_{ts}(0,1,1)/S_{ts}(1,1,1)$	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept	6.251	0.555	-34.584	-2.481	-47.513	-3.747	-23.346	-1.020
Baseline hazard (1)			-2.686	-3.033	-0.466	-0.690	-3.658	-5.303
Girl	-7.434	-3.412	-13.077	-3.630	-2.580	-1.611	-0.033	-0.017
Age	-0.308	-0.978	0.747	1.086	0.356	1.351	0.576	1.134
School entry age	2.943	3.552	0.146	0.207	-1.374	-2.680	-0.582	-2.272
Grade repetitions in primary	3.312	2.782	-1.361	-0.898	-10.947	-3.210	-1.847	-0.248
Free tuition policy	-13.039	-3.702	-3.653	-1.665	5.092	3.116	1.382	0.532
Free tuition policy * girl	15.685	4.099	17.978	3.160	-4.585	-2.475	0.082	0.038
Mother's education	-1.614	-4.127	1.507	3.691	1.511	4.451	-0.040	-0.050
Father's education	0.616	3.333	0.368	1.478	0.213	1.182	-0.024	-0.123
Household assets missing	-18.705	-2.712	-4.380	-0.309	12.966	1.468	10.216	2.130
Log(household assets)	-2.188	-3.132	0.278	0.312	1.895	2.407	1.044	2.875
Modern latrine	1.117	1.387	1.729	1.862	-2.476	-1.974	-0.715	-0.142
Cultivating household	3.730	2.008						
Owens farm land	-1.346	-0.677	6.655	1.355	5.560	3.038	-0.089	-0.043
Owens non-farm business	7.635	3.854	3.089	2.999	3.360	2.925	0.213	0.067
Older siblings	-1.853	-3.775	0.576	2.132	0.002	0.005	-0.299	-1.427
Younger siblings	0.281	0.741	0.020	0.067	-1.268	-2.153	0.393	0.938
Village outside Matlab	1.638	0.519	2.872	1.594	8.144	2.982	2.796	0.709
Tubewell in village	4.741	2.484	-0.881	-0.581	1.002	0.724	1.812	0.452
Health facility in village	-2.516	-2.399	2.575	2.089	5.541	3.777	0.308	0.200
Industry in village	-0.821	-0.784	2.757	2.511	-2.914	-2.436	1.220	1.744
Secondary school in village	7.211	3.832	2.503	2.573	-2.380	-1.924	1.322	0.320
Distance to Matlab capital	0.390	2.647	0.363	2.860	-0.025	-0.172	-0.095	-1.515
θ_w	-15.387	-4.260						
θ_s			10.208	1.822	17.157	4.488	-0.081	-0.004

Notes: See notes to table 4.3.