Rochester Center for
Economic Research

Borrowing Constraints and School Attendance: Theory and Evidence from a Low Income Country

Jacoby, Hanan

Working Paper No. 284
June 1991

University of
Rochester

Borrowing Constraints and School Attendance: Theory and Evidence from a Low Income Country

Hanan Jacoby

Rochester Center for Economic Research
Working Paper No. 284

June 1991

Borrowing Constraints and School Attendance: Theory and Evidence from a Low Income Country

Hanan Jacoby

University of Rochester

and

The Rochester Center for Economic Research

Working Paper No. 284

June 1991

ABSTRACT

This paper adapts a simple human capital investment model to the circumstances of low income countries, allowing for the possibility that households face borrowing constraints. Because of parents' desire to smooth consumption over time, children are gradually withdrawn from school and put to work, leading poor children to fall behind richer children in school. This paper also makes the first attempt to study the timing of child school attendance in a multi-child household. The main implications of the model are tested on household survey data from Peru, and support for them is found.

Acknowledgments: I wish to thank John Boyd, Prajit Dutta, Paul Glewwe, Changyong Rhee and Emmanuel Skoufias for helpful comments. I am also grateful to seminar participants at Penn. State, Pittsburgh, Rochester, Yale and The World Bank for useful suggestions on an earlier version of this paper. All remaining errors are mine.

I. Introduction

Low educational attainment in the developing world along with high estimated rates of return to schooling are often cited as justification for public investment in more and better quality schools. This same evidence, however, indicates barriers to private investment in human capital, and suggests that at the margin resources may be better spent on encouraging individuals to fully utilize existing schooling facilities. Lending support to this latter view is the fact that children in low income countries tend to progress through school more slowly than their counterparts in industrialized nations. Why do children fall behind in school? This paper argues that such behavior arises out of the desire of poor families to smooth consumption over time. Unable to borrow against their children's future earnings, parents are compelled to gradually withdraw them from school and put them to work.

Most schooling demand studies for low income countries are cast in a static time allocation framework in which child schooling directly enters the family utility function (see, e.g., Rosenzweig and Evenson 1977). Previous empirical work focuses almost exclusively on either grade attainment or current school enrollment, ignoring the phenomena of grade repetition and late entry into primary school, so pervasive in low income countries.[1] Because the static framework predicts only how total schooling will respond to wage and income changes, not how school attendance responds at each point in time, it cannot explain why children fall behind in school rather than simply quit school at an earlier age. Moreover, with cross-sectional data one faces the

[1] Lockheed and Verspoor (1989) report that in 1985 the median rate of grade repetition among primary school students in low income countries was seventeen percent, about eleven percent in lower middle income countries, and only two percent in high income countries.

dilemma that wage and income variables relevant to the original schooling decision are not observed contemporaneously with educational attainment.

This paper adapts a simple human capital investment model to the circumstances of low income countries. Unlike Ben-Porath (1967), Heckman (1976) and Blinder and Weiss (1976), the model considers the possibility that households are constrained in their ability to borrow, and, as a consequence, do not send their children to school full-time. In a second departure from previous models, children are treated as members of families with other sources of income besides child earnings, allowing an examination of the effects of household poverty and borrowing constraints on school attendance patterns over time. This paper also makes the first attempt to study the timing of each child's human capital investment in a multi-child household.[2]

Peru provides an interesting venue for the empirical test of the model. Under education reforms adopted there in the late 1960s, students could only advance a grade by completing a given series of lessons (King and Bellew, 1989). Thus, the substantial grade repetition rates found in the Peruvian Living Standards Survey of 1985-86 are due primarily to poor school attendance. Delayed entry of children into primary school is also highly evident in Peru.

The next section develops a model of the allocation of human capital investment over time <u>and</u> across several heterogeneous children. To get a handle on the allocation over time, the theoretical discussion focuses first on a single child household, and then extends the model to a multi-child

---

[2]Becker and Lewis (1973) and Behrman, et al. (1982) both discuss the interaction between number of children and education per child, but neither model sheds any light on how the <u>timing</u> of human capital investment interacts with number of siblings.

household. Section III sets out an estimation strategy utilizing cross-sectional data to test the main empirical implications of the model. The results of these tests, reported in section IV, lend qualified support to the basic proposition that borrowing constraints affect child schooling patterns in Peru. Section V sums up the paper.

II. The Model

A. Single-Child Households

The household consists of two overlapping generations, parents (i.e., two parents with a common objective function) and a child. Parents are assumed to exercise full control over all family income while they are living and to make all schooling decisions; in particular, they decide on the fraction of time, S(t), the child spends in school each year. Parents are assumed altruistic in the sense that they value child consumption, and their instantaneous utility is further assumed to be a concave function of the sum of parental and child consumption, C(t), i.e., u = U(C(t)).[3]

Take time zero to be the date the child is first eligible to enroll in school. Since parents die at time T, their problem is to choose C(t) and S(t) to maximize discounted lifetime utility (δ is the rate of time preference)

$$(2.1)\quad \int_0^T U(C(t))\, e^{-\delta t} dt \qquad \text{subject to}$$

$$(2.2)\quad \dot{A}(t) = rA(t) + y(t) + wH(t)[1 - S(t)] - C(t)$$

[3]This formulation is convenient and will be useful in the empirical work below. However, other specifications yield similar theoretical results; e.g., $u = U(C^p(t) + \alpha C^c(t))$, where the superscripts refer to parents and children, respectively, and α is a positive constant; or the additively separable case, $u = U(C^p(t)) + V(C^c(t))$.

$$(2.3)\quad \dot{H}(t) = bH(t)S(t) \qquad 0 \le S(t) \le 1,$$

$$\text{and}\quad H(0) = H_0\ ,\quad A(0) = A_0,\quad A(T) = 0,\quad b > \max(r,\delta)$$

The human capital of the child, H(t), is given its common interpretation as a homogeneous stock of embodied knowledge, and parents are assumed to place no value on human capital after their death.[4] Financial bequests are also ignored here (i.e., the utility of the terminal stock of assets is set to zero) without changing the basic thrust of the analysis.

Constraint (2.2) is the law of motion for net assets, A(t), assuming that the household is able to freely borrow and save at interest rate, r. Household income has two components, child earnings and parents' income. When the child is not enrolled full-time in school (i.e., S(t) < 1), he rents his stock of human capital at the constant rental price w; in other words, he works at wage wH(t), contributing wH(t)[1-S(t)] to family income. Note that the forgone earnings of the child, wH(t)S(t), is assumed to be the only cost of going to school.[5] Meanwhile, parents' income, y(t), is assumed exogenously given and may rise or fall over time.[6]

---

[4]To relax this perhaps unpalatable assumption, a "salvage" function representing the utility of the terminal stock of human capital can be appended to (2.1). However, the cost of this modification is enormous analytical complexity. Essentially, the shadow price of human capital would no longer be independent of the initial stock (see the Appendix). Time spent in school would also now depend on the parameters of the salvage value function.

[5]School fees do not appear to be quantitatively important in Peru (see Gertler and Glewwe 1989).

[6]Endogenous parental labor supply is a trivial extension of the model as long as utility is assumed additively separable in consumption and parents' leisure. Also, since y(t) does not directly affect the necessary conditions for the maximization of the parental objective function, it can be allowed to evolve over time in an arbitrary manner. However, in order to explicitly solve for the optimal policy, the form of y(t) must be specified. Since time zero is taken to be the normal age at which children enter primary school,

Human capital is accumulated according to equation (2.3). The human capital stock is assumed to be self-productive, and increments to the stock are assumed to be produced with a constant returns to scale technology, where b is a Hicks-neutral learning "efficiency" parameter. While b may be interpreted as a student's ability, it could also reflect school quality. The assumption that $b > \max(r, \delta)$ insures that human capital investment will be undertaken. Human capital depreciation is ignored here for simplicity, as are other purchased inputs into human capital production.

Case of unconstrained borrowing

The solution to the optimal control problem is illustrated in figure 1 (mathematical details can be found in part A of the Appendix). The constant returns to scale assumption leads to a "bang bang" control; since the opportunity cost of time is always rising, human capital investment is immediately pushed to the limit, as shown in panel (a). Thus, the student begins full-time schooling at time zero and continues until the quitting time

$$(2.4) \quad t_1 = t_1^* = T + \frac{1}{r}\ln(1-r/b)$$

where t_1^* falls with the market interest rate and rises with the efficiency of human capital production b.

Meanwhile, as illustrated in panel (b), unrestricted borrowing allows consumption to rise or fall over time depending on whether or not r exceeds δ. The degree of borrowing or saving at a given point in time depends on $r - \delta$, y(t) and the three boundary conditions. For example, if parents start out with no assets ($A_0 = 0$) and have a flat income profile ($y(t) = y$), then the

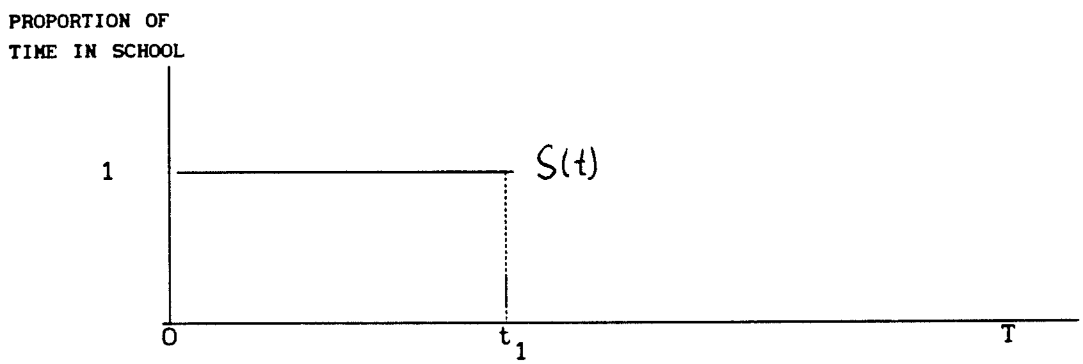
---

y(t) may well fall over [0,T], as the peak of parental productivity may already be behind them.

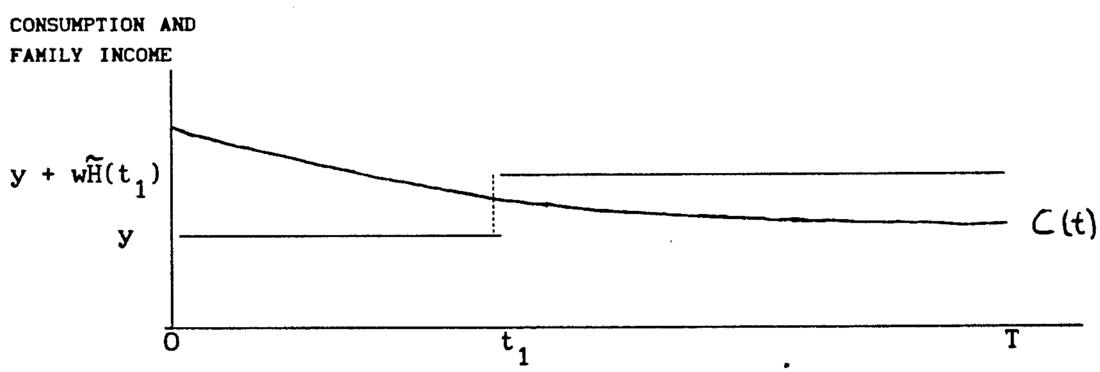
FIGURE 1

TIMING OF HUMAN CAPITAL INVESTMENT WHEN BORROWING IS UNCONSTRAINED

Panel (a)



Panel (b)



following results emerge. If $\delta > r$, they will borrow while the child attends school to finance high consumption, and then repay the debt when the child leaves school and begins working (this is the case depicted in figure 1). If $r > \delta$, the household may also borrow early on, but saving while the child attends school is possible as well. The latter is more likely to be the case when parents desire consumption to rise quickly, and expect only a small jump in income when their child quits school.

The S(t) path in panel (a) is a reasonable characterization of school attendance patterns in developed countries; children enter school immediately and rarely fall behind for economic reasons. However, in most developing countries, including Peru, this model would do great violence to the data. One way to obtain part-time and declining school attendance would be to allow for diminishing returns to scale in human capital production, as in the original Ben-Porath model.[7] But it is hard to believe that diminishing returns set in as early as primary school and are the main cause of low attendance. Moreover, diminishing returns cannot explain differences in schooling patterns between children of rich and poor families, unless the latter have faster diminishing returns than the former. The theory below generates part-time school attendance by adding an economic constraint rather than from a parametric restriction on the human capital production function.

Case of constrained borrowing

Reconsider the utility maximization problem with the additional constraint that the family's net asset position can never fall below zero, so

---

[7]That is, modify (2.3) to $\dot{H}(t) = b[H(t)S(t)]^{\sigma}$ where $0 < \sigma < 1$.

that $A(t) \geq 0$.[8] A similar analysis applies if assets are constrained to lie above a negative lower bound. Not all households are affected by this new constraint. As was just noted, some households will be willing to save even when sending their children to school. Take the case of the "impatient" household (i.e., $\delta > r$) first.

To simplify the exposition, assume as before that $A_0 = 0$ and $y(t) = y$ (the effect of modifying these assumptions is discussed below). A logarithmic utility function $U(C) = \log(C)$ is used to obtain an explicit solution to the model, though most results hold more generally (see Appendix). Households with these preferences are interested in maintaining a constant rate of growth (decline) in consumption. When $\delta > r$, the borrowing constraint is binding throughout the child schooling period--and, indeed, throughout the whole life cycle--regardless of parental income.

Binding borrowing constraints do not rule out a period of full-time schooling, which may occur at the beginning of the student's career. Once again the first order conditions and other mathematics are relegated to the appendix and the solution to the model is illustrated graphically in figure 2. Panel (a) shows an initial period $[0, t_0]$ of full-time school attendance. Eventually the opportunity cost of child time rises to the point where full-time schooling is no longer optimal and the child is gradually withdrawn from school over the period $[t_0, t_1]$.[9] The reason for the child's gradual

---

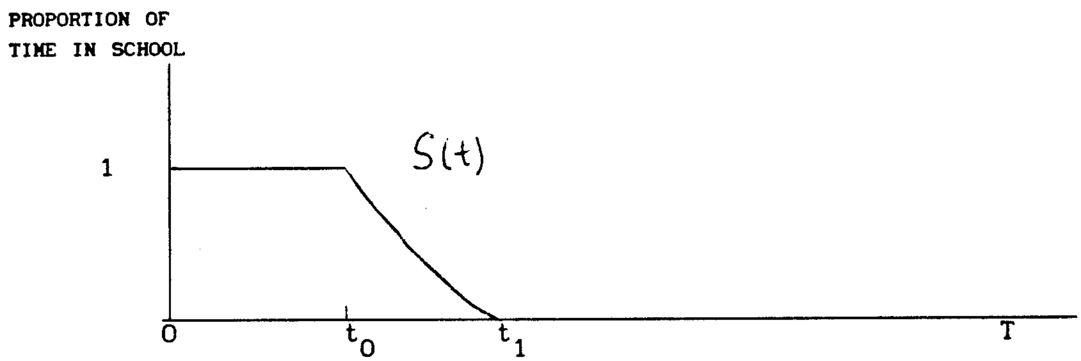
[8] Weiss discusses this constraint in his 1986 survey. Wallace and Ihnen (1975) and Johnson (1978) impose borrowing constraints in a human capital investment model, but try to maintain credit market separation by allowing students to obtain consumption loans but not investment loans, a distinction even they admit is somewhat artificial. In these models, borrowing constraints become effective only if there are direct costs of schooling, such as tuition.

[9] A simple comparative dynamics analysis (available upon request) shows that $S(t)$ must fall monotonically for all $y(t)$ and any concave utility function.

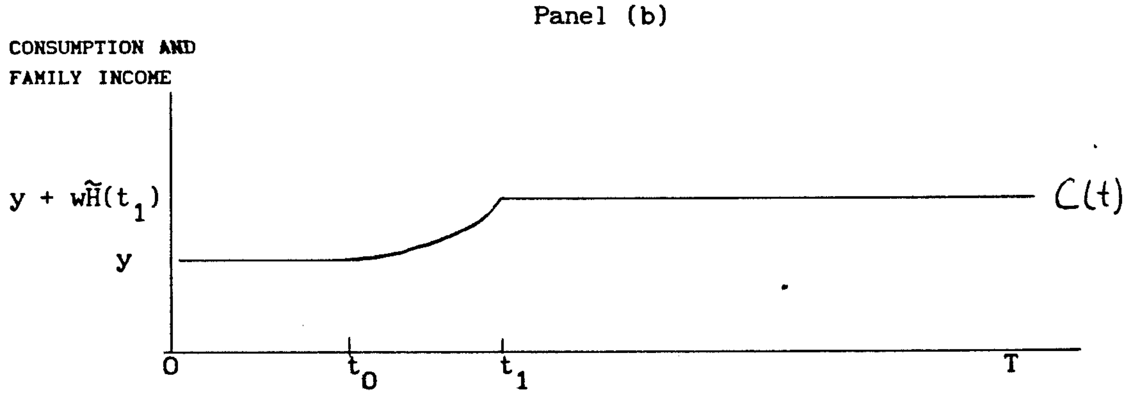
FIGURE 2

TIMING OF HUMAN CAPITAL INVESTMENT WHEN BORROWING IS CONSTRAINED: CASE OF THE IMPATIENT, SINGLE-CHILD HOUSEHOLD

Panel (a)



Panel (b)



entry into the labor force is that it is the only means by which the family can smooth consumption in the absence of credit markets. As shown in panel (b), prior to t_0, consumption simply follows parental income ($C(t) = y$). After t_1, however, $C(t) = y + w\tilde{H}(t_1)$, where $\tilde{H}$ is the child's optimal human capital stock function. Without a period of part-time schooling, consumption would jump when the child quits school.[10,11]

The optimal number of years spent in school, t_1^*, is again given by (2.4), except that in the case of no borrowing the interest rate r is replaced by the rate of time preference δ. Meanwhile, an expression for the full-time/part-time switching point, t_0^*, can only be obtained by solving the model explicitly, using the budget constraint, first order conditions and the human capital production function. With logarithmic utility, t_0^* solves[12]

$$(2.5) \quad \frac{b}{\delta}\,[1 - e^{-\delta(T-to)}] - 1 = \frac{wH_0}{y}\, e^{bto}.$$

---

[10] Perfect consumption smoothing, in this case, rests on the supposition that school attendance can be continuously adjusted toward zero, and should be considered only an approximation to reality. If there exists an institutionally mandated lower bound on school attendance or fixed costs of starting a grade, for example, then some jump in family consumption is inevitable.

[11] Given y, wH_0, b and δ, a jump in consumption at t_1 is optimal for at most one r. In this "knife-edge" case, the household saves prior to t_1, sending their child to school full time, and dissaves after t_1.

[12] After much tedious calculation, the optimal human capital stock function is

$$\tilde{H}(t) = H_0\, e^{bt} \qquad 0 \le t \le t_0$$

$$\tilde{H}(t) = \frac{y}{w}\,\{\, \frac{b}{\delta}\, e^{(b-\delta)(t-to)}[1 - e^{-\delta(T-t)}] - 1\} \qquad t_0 \le t \le t_1$$

The fact that both branches of this function are equal at t_0 gives (2.5).

Notice that t_0^* is decreasing in wH_0/y and that as the latter approaches zero (e.g., as parental income gets very large), t_0^* approaches $t_1^* = T + \frac{1}{\delta}\ln(1-\delta/b)$, the school quitting time. The intuition is simple: As wH_0/y increases the proportional jump in consumption that would occur if the child left school all at once would be greater, so the household desires a longer period $[t_0, t_1]$ of part-time child school attendance to smooth consumption. At one extreme, very high income households behave just like unconstrained households in that $t_0^* \cong t_1^*$. At the other extreme of parental income, $t_0^* = 0$ and the child begins school with only part time attendance ($S^*(0) < 1$).

The optimal attendance profile for the part-time schooling phase is

$$(2.6) \qquad S^*(t) = \frac{1 - \delta/b - e^{-\delta(T-t)}}{1 - e^{-\delta(T-t)} - \frac{\delta}{b} e^{-(b-\delta)(t-t_0^*)}} \qquad \text{for } t_0 \le t \le t_1$$

Comparative statics with respect to the economic parameters of the model are simple because y and wH_0 enter $S^*(t)$ only through t_0^*. Thus $S^*(t)$ is decreasing in wH_0/y.[13] When $t_0^* = 0$, $S^*(t)$ takes a slightly different form than (2.6), but it is still decreasing in wH_0/y.

Now turn to the more complicated case of the patient household. When $r > \delta$, the borrowing constraint is not necessarily binding throughout the schooling phase and it is not binding at all in the post-schooling phase (under the assumptions made in this subsection). Depending on the parameters of the model, two possible scenarios can occur. In both scenarios, the optimal school quitting time t_1^* is given by (2.4), exactly as in the unconstrained case. In the first scenario, the borrowing constraint is

---

[13] Neither $\partial t_0^*/\partial b$ nor $\partial S^*/\partial b$ can be signed, and post-school earnings, $w\tilde{H}(t_1^*)$, is increasing in both y and b.

binding at only one point, namely t_1. Here there is a discontinuity in the consumption profile, while before and after t_1 consumption grows at a constant rate (i.e., $r - \delta$ for logarithmic utility). In other words, the family is willing to save (consume little) while the child is in school and experience a jump in consumption thereafter. This scenario is more likely to occur when child earnings are small relative to parental income, since then the consumption jump would be relatively small (see below).

In the second scenario, the household again resorts to child school attendance as a consumption smoothing device. As illustrated in figure 3, the household may start out by saving and subsequently dissaving while sending their child to school full time. At t_0 savings run out, the borrowing constraint begins to bind, and the level of consumption is maintained by a discrete fall in school attendance. As parents continue to gradually withdraw their child from school, consumption grows with full income as it did in figure 2. At the quitting time, t_1^*, the household will again face a constant income stream, $y + w\tilde{H}(t_1)$. To achieve rising consumption beyond t_1, $S(t)$ jumps downward to zero at t_1 and parents save the child's extra earnings. Savings subsequently decline and turn negative as parents approach T.

Assuming logarithmic utility, some strenuous algebra shows that for this second scenario t_0^* solves

$$(2.7)\quad \frac{b}{r}\,\frac{1 - e^{rto}}{1 - e^{\delta to}}\left\{e^{-\delta(t_1-t_0)}[1 - e^{-\delta(T-t_1)}] + e^{(b-\delta)t_0}[1 - e^{-\delta(t_1-t_0)}]\right\} - 1 = \frac{wH_0}{y}\,e^{bto},$$

where $t_1 = t_1^* = T + \frac{1}{r}\ln(1-r/b)$.[14] Total differentiation of (2.7) yields a

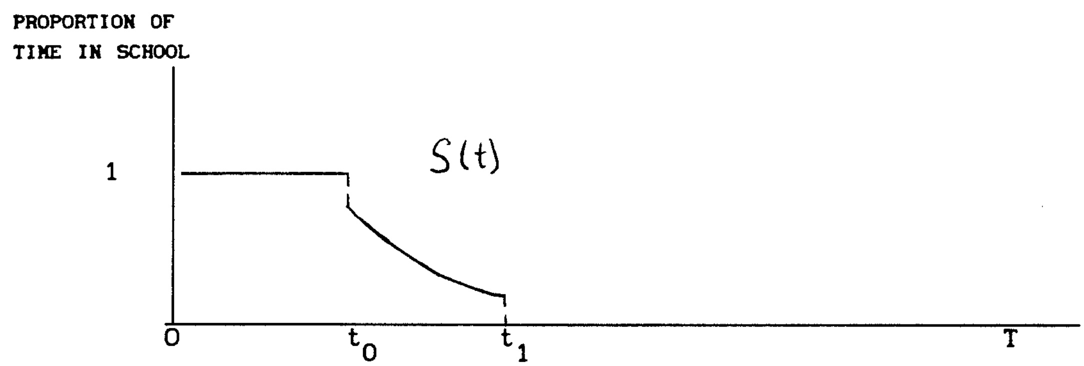
---

[14]Note that by setting $t_0 = t_1$ in (2.7), one can solve for the critical value

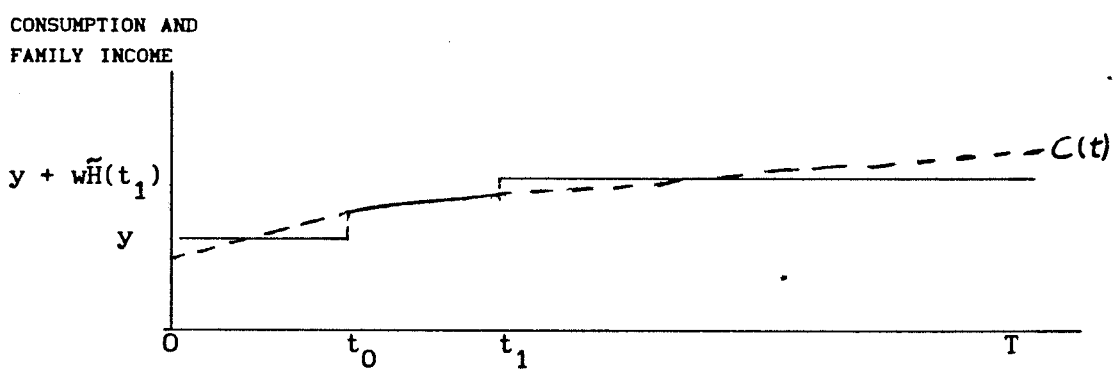
FIGURE 3

TIMING OF HUMAN CAPITAL INVESTMENT WHEN BORROWING IS CONSTRAINED: CASE OF THE PATIENT, SINGLE-CHILD HOUSEHOLD

Panel (a)



Panel (b)



messy expression for $\partial t_0^*/\partial(wH_0/y)$ that cannot be signed analytically, except when evaluated at $t_0 = 0$, in which case it is negative. More generally, given values for the parameters, (2.7) can be solved numerically and the derivative can be evaluated. For plausible parameter values,[15] the sign of the derivative is found to be consistently negative, as it is when $\delta > r$.

Briefly, consider the consequences of relaxing the zero initial assets and constant parental income assumptions. Because the household may either consume out of initial assets or perhaps use them as collateral for loans (i.e., the borrowing constraint becomes $A(t) \geq -A_0$), greater initial assets allow the household to come closer to achieving desired (unconstrained) consumption, and will thus lead to a postponement, and ultimately an elimination, of the part-time schooling phase.

To examine the effects of nonconstant parental income, return to the logarithmic utility example and assume parental income grows (declines) at a constant rate γ (i.e., $y(t)=y(0)\exp(\gamma t)$). For the impatient household, the borrowing constraint no longer binds initially if $\gamma \leq r - \delta$; i.e., if income falls as fast or faster than parents would like consumption to decline in the absence of borrowing constraints. In this case, even impatient households will save while sending their child to school full-time, though eventually the constraint may bind again, resulting in part-time schooling. Rising parental income ($\gamma > 0$) only aggravates the borrowing constraint for impatient

---

of wH_0/y, below which the first scenario (i.e., a jump in consumption) results.

[15]For example, when $wH_0/y = .05$, $r=.042$, $\delta=.04$, $b=.055$, and $T = 40$, then $t_0 = 3.8$ and $t_1 = 5.7$, the latter being a reasonable number for time spent in school for the average Peruvian.

households, and makes it more likely to bind for patient households as well.

To summarize, part-time child school attendance is more likely to occur the more impatient the household is (the larger δ-r), the lower initial assets, the lower (initial) parental income y, the faster parental income grows, and the larger are the initial potential earnings of the child (wH_0). The main empirically testable implication of the theory is that for households that cannot borrow, those with low income begin withdrawing their children from school sooner than those with high income, while schooling patterns are independent of income for unconstrained households. Note finally, that, regardless of the parameter configuration, the model is incapable of explaining late entry into primary school. If schooling is worthwhile at all, it pays to begin immediately and with maximum attendance.[16]

B. Multi-Child Households

So far the analysis has ignored the effects of siblings on child schooling patterns. Assume that the number and the spacing of children is given, so that parents jointly solve for the optimal school attendance paths of their children, conditional on the demographic structure of the household.[17]

---

[16]The age of school entry may differ across children because of variation in the rate of "maturity". But, to fit the data (see below) one would need to explain why the rate of maturity is correlated with income. Fixed attendance costs, such as school fees, may also lead to delays, but they are not significant in Peru (see above). Glewwe and Jacoby (1991) solve a model similar to the one presented here with school fees. They show that delays can occur with or without borrowing constraints, and that the length of the delay is independent of parental income in the constrained case, but not in the unconstrained case.

[17]One can argue that, while the spacing of children may be subject to choice, the timing of fertility is <u>less</u> controllable than the timing of child schooling. Also note that, unlike Behrman, et al. (1982), parents are assumed not to care directly about equality of education across their children, although the possibility that they equalize consumption across children is not

The basic intuition is that if parents can freely borrow against the future earnings of their children--i.e., with credit market separation--the timing of each child's human capital investment would be independent of the demographic structure of the household. With a borrowing constraint, on the other hand, school attendance paths of siblings would generally be interdependent.

Suppose a household has K already born children. Let **a** = $\{a_0, a_1, \ldots, a_{k-1}\}$ be the vector of dates that each of the K children can first enter primary school; $a_{i+1}-a_i$ is simply the age gap between adjacent children. Besides age, children may differ in their learning efficiency, **b** = $\{b_1, \ldots, b_k\}$, in the rental price of their human capital, **w** = $\{w_1, \ldots, w_k\}$, and in their initial human capital stocks, $\mathbf{H_0}=\{H_{01}, \ldots, H_{0k}\}$. Normalizing $a_0=0$ (for the moment), the parents' problem is to choose C(t) and a vector of school attendance paths, **S**(t), to maximize (2.1) subject to

$$(2.8) \quad \dot{A}(t) = rA(t) + y(t) + \sum_{i=1}^{K} I(a_{i-1}, t) w_i H_i(t)[1 - S_i(t)] - C(t)$$

$$\text{where } I(a_{i-1}, t) = \begin{cases} 1 & \text{if } t \geq a_{i-1} \\ 0 & \text{otherwise} \end{cases} \qquad \text{and } a_0 = 0,$$

$$(2.9) \quad \dot{H}_i(t) = b_i H_i(t) S_i(t) \qquad 0 \leq S_i(t) \leq 1$$

$$\text{and } H_i(0) = H_{0i} \text{ for } i = 1, \ldots, K, \quad A(t) \geq 0, \; A(0) = A_0, \; A(T) = 0.$$

Again, for the purposes of exposition, assume that $\delta > r$, $y(t) = y$, and $A_0 = 0$. Also, assume that there are just two children, differing in age by a_1 years (the general case of K children follows as an extension). The problem can be split into two separate problems on the intervals $[0, a_1]$ and $[a_1, T]$, respectively. The second stage problem, in which two school-age children are

---

ruled out (see, e.g., Becker and Tomes 1976).

present, can be solved first, conditional on the stocks of human capital as of time a_1.[18] Then, the first stage problem can be solved, treating the maximized value of lifetime utility in the second stage as a bequest function. This technique is known as two-stage optimal control, and Tomiyama (1985) gives the necessary conditions for an optimum (see part B of the appendix).

It is useful to think of the second stage problem itself as a two-stage problem in which parents first decide on total resources devoted to human capital investment (i.e, choose $R(t)=w_1H_1(t)S_1(t) + w_2H_2(t)S_2(t)$), and then decide on how to allocate the investment across children. As in the single child case, $R(a_1)$ depends crucially on parental income, and while high income parents may send both children to school full-time initially, low income parents may not even be willing to send one of them to school full-time.

Given $R(a_1)$, the intrahousehold allocation of human capital investment over time is determined by comparative advantage. With constant returns to scale in (2.9), parents never send two children to school part-time simultaneously. Instead, the child with the highest w/b (the one with comparative advantage at work) begins withdrawing from school first, and only after he is finished does the child with next highest w/b begin withdrawing.

Now consider the first stage problem, which is simply a variation on the single child model. The key difference is that the shadow price of the first child's human capital at the endpoint a_1 equals the marginal value of the human capital he "bequeaths" to the next stage (see Tomiyama, 1985). Also, the marginal utility of wealth must be continuous at a_1, which implies that the consumption path must be continuous. A number of scenarios consistent

---

[18]The assumption that $\delta > r$ means that no assets are accumulated in the initial phase, so that $A(a_1)=0$.

with these optimality conditions are possible, depending on the value of the parameters, initial human capital stocks and parental income. Four of these cases are illustrated in figure 4. In all the scenarios, child 1 refers to the older sibling and child 2 to the younger.

In panel (a) of figure 4 the older child quits school completely before his sibling ever starts. The older child quits sooner than he would have had he had no younger sibling. This case is more likely to occur for large a_1, and as a_1 approaches T it reverts to the single child model. Panels (b) and (c) illustrate two intermediate cases where the older child starts withdrawing from school before the younger sibling even starts in order to smooth consumption. In panel (b) the assumption is that $w_1/b_1 > w_2/b_2$, so that the younger child has a comparative advantage in school and the older child finances his schooling.[19] By contrast, panel (c) assumes $w_1/b_1 < w_2/b_2$, and the older child helps smooth consumption only until the younger child comes of age. At that point the younger child takes over the earning responsibility, paving the way for his older, more able, sibling to pursue his studies. Finally, panel (d) of figure 4 shows what happens when a_1 is small and $R(a_1)$ is sufficiently large. Since both children are sent to school full time at a_1, the older child is not required to work in the first stage to smooth consumption. In the second stage, the child with the highest w/b (in this example, the oldest one) is withdrawn from school first, as described above.

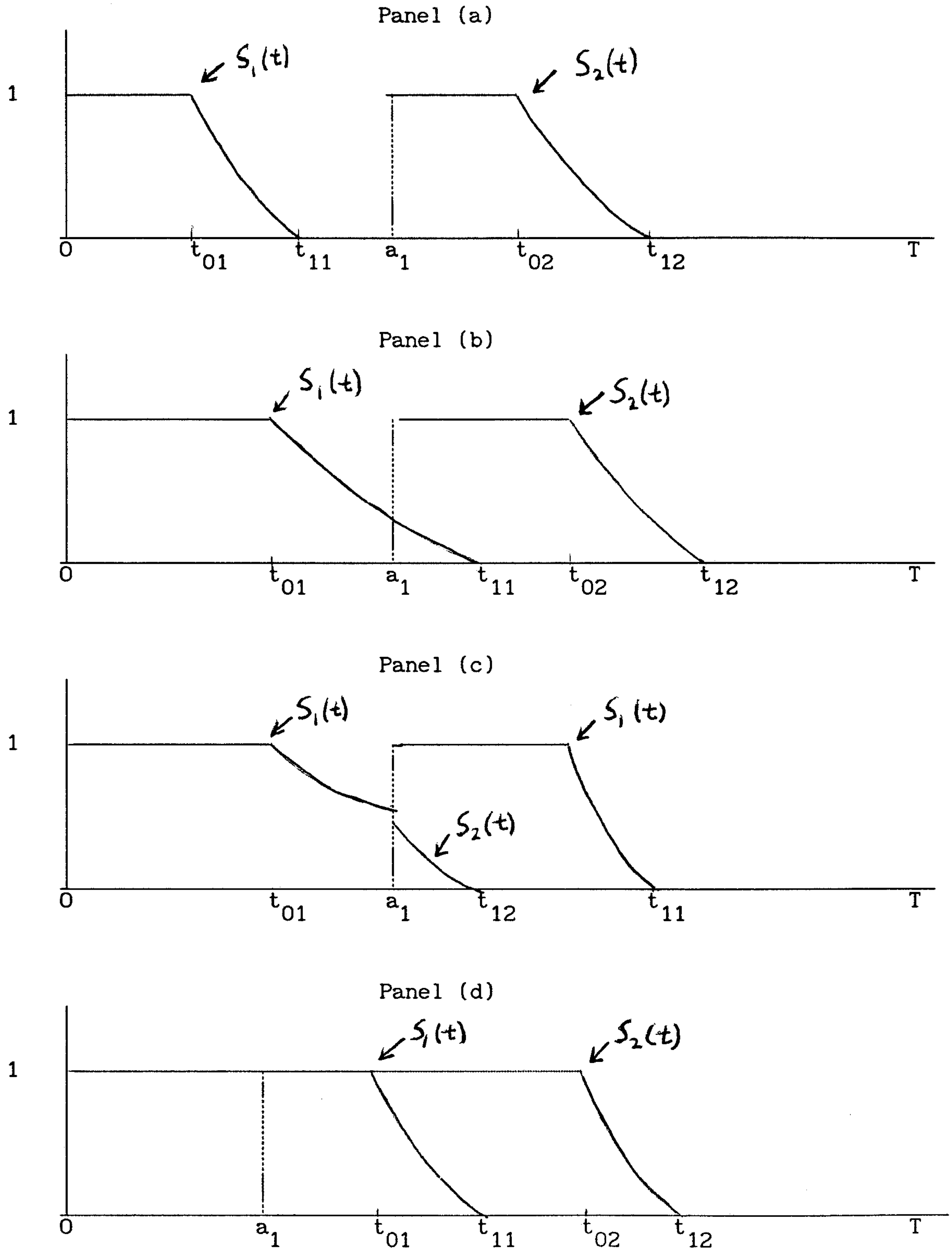
In general, the point in time at which child i first starts withdrawing from school, t^*_{0i}, depends on the vector of age gaps; i.e.,

---

[19]This scenario roughly corresponds to Greenhalgh's (1985) argument that in Taiwanese society girls with younger male siblings often leave school early to work, increasing the resources available for the education of their brothers.

FIGURE 4

TIMING OF HUMAN CAPITAL INVESTMENT WHEN BORROWING IS CONSTRAINED:
CASE OF THE IMPATIENT, TWO-CHILD HOUSEHOLD



$$(2.10)\quad t_{0i}^{*} = t_{0i}^{*}(\mathbf{a}-a_i, \mathbf{b}, \mathbf{w}, H_0, y, \delta, T). \qquad i = 1, \ldots, K.$$

However, since t_{0i}^{*} cannot be solved for explicitly, comparative statics results for changes in the age gaps are enormously difficult to obtain. Instead, assuming a utility function (here the logarithmic is used) and choosing the parameter values, for each scenario in figure 4 numerical methods can be used to solve the appropriate system of nonlinear equations for the switching points. Notice that smaller age gaps between siblings arise either because the older child is born later in the parents' life-cycle, or because the younger child is born earlier. The former can be viewed as an increase the left endpoint in the first stage problem, a_0, holding a_1 fixed, while the latter involves a decrease in a_1, holding a_0 fixed.

The effect of narrowing the age gap by the first route can be seen by calculating t_{01}^{*} and t_{02}^{*} for different values of a_0. For reasonable parameter configurations, an increase in a_0 (and hence a fall in a_1-a_0) in scenario (a) of figure 4 leads to a relatively small fall in $t_{02}^{*} - a_1$, but a large fall in $t_{01}^{*} - a_0$. Intuitively, the larger a_0 is the less human capital the older child accumulates before the younger child starts school. As a result, when the older child enters the work force, he contributes less to family income, leading parents to begin withdrawing the younger child from school earlier.

When the age gap is narrowed by decreasing a_1, holding a_0 fixed, the effect on attendance patterns is more complicated. The numerical solution of scenario (a) reveals that $t_{01}^{*}-a_0$ is a U-shaped function of a_1, while $t_{02}^{*}-a_1$ declines monotonically as a_1 increases. As the age gap between siblings shrinks, parents invest more in the younger child, because the returns will be realized over a longer horizon. Initially, the older child's education suffers and he begins withdrawing sooner. But when the siblings are sufficiently close together in age (but still far enough apart to be in scenario (a)), the older child starts withdrawing from school later (t_{01} increases) as a_1 falls, eventually leading to either scenario (b) or (c).

With all the theoretical possibilities, the impact of variations in family structure on the timing of child school attendance cannot be stated unequivocally, though it seems reasonable to expect that the smaller the age gap between a given child and adjacent siblings, the sooner that child will start leaving school. It is important to keep in mind that in the absence of borrowing constraints the presence of older or younger siblings, regardless of age, has no impact whatsoever on school attendance.

While the multi-child household model can generate various attendance profiles, including rising attendance over time (see panel (c) of figure 4), the goal of explaining late entry into primary school remains elusive. One would think that parents might delay enrolling younger children to allow an older sibling to finish school. But delays do not occur in this model because, as in the single child case, opportunity costs are lowest at the date of initial eligibility. Still, it seems worthwhile to confront the data with the question of whether the close spacing of siblings contributes to delays in school enrollment, in addition to a more rapid withdrawal from school on the part of those already enrolled.

III. An Empirical Strategy for Cross-Sectional Data

Estimation of a dynamic schooling model would seem to require longitudinal data on time spent in school. Yet, for developing countries, sufficiently long panels with reliable information on child time allocation are scarce, if not nonexistent. The strategy proposed in this section involves looking at how a child is progressing through school at a given point in time, allowing the model to be estimated on a single cross-section of data.

Focusing on children still attending or who have not yet started school obviates the empirical dilemma mentioned in the introduction; namely, that economic variables observed after schooling has been completed may be only weakly related to the original schooling decisions.

To test for behavior implied by borrowing constraints, it would also appear necessary to restrict the analysis to a sample of children from constrained households. But, because these tests amount to looking for deviations from behavior implied by credit market separation, it is not essential that the sample be composed entirely of constrained households. In fact, equation (2.5) shows that as income rises, constrained households behave more and more like unconstrained ones anyway. Nevertheless, section IV attempts to differentiate households by their access to credit markets and to see whether the schooling patterns of their children differ.

Grade repetition

Although the switching point, t_0^*, between full and part-time schooling cannot be observed directly in a cross section, it can be estimated. Suppose that for each child attending school the number of years, t, spent in school is known, along with the actual number of grades completed, G. If $t > G$--i.e., if a grade has been repeated--then under the assumption that students are held back because of abnormally low attendance (see the introduction) it must be the case that $t > t_0^*$. Suppose further that a child who has not yet repeated a grade ($t = G$) has $t \le t_0^*$; i.e., assume the child is a full-time student.[20] Repetition of first grade implies that $t_0^* = 0$ or $S^*(0) \le 1$. Since

[20]Clearly, there is the possibility of misclassification: a child reporting zero grades repeated may have already begun falling behind in school. This measurement problem is not inherent in cross-sectional data, but is due to the fact that schooling is typically measured in whole numbered grades. The threshold point could be better determined with accurate data on yearly hours

$Pr(t > G) = Pr(t > t_0^*)$, specifying the conditional distribution for t_0^* produces a standard discrete choice model.[21]

No attempt is made here to impose the complicated functional form restrictions on t_0^* implied by the theory.[22] Rather, a linear version of the general t_0^* function given by (2.10) forms the basis of the empirical specification. Thus, t_0^* is assumed to vary across individuals in the sample according to differences in (1) school quality and student ability reflected in an individual specific learning efficiency parameter, b_i; (2) household income; (3) initial child wages, wH_0; (4) the age gaps between adjacent siblings, and, (5) tastes for and "noneconomic" barriers to schooling (e.g., sex discrimination, class barriers, etc.).

Learning efficiency is specified as follows

(3.1) $b_i = \gamma' X_i + u_i$,

---

in school. In lieu of such data, a partial solution to the problem is offered in the next section.

[21]The identification of t_0^* in a cross-section relies on the (untestable) assumption that grade repetition is not the result of a temporary withdrawal from school, but represents the movement down the life-cycle school attendance profile, at least for most children in the sample. If one is not willing to accept this assumption, the following argument may be more appealing. Suppose parental income is subject to stochastic fluctuations (e.g., bad harvests). If parents cannot borrow ex post, they would also tend to withdraw their children from school and put them to work in response to unanticipated income shocks, exactly as they do to smooth consumption over the life-cycle. The extent to which constrained parents withdraw their children from school in lean times will be negatively related to the level of income and positively related to the value of child time.

[22]Jacoby (1990) shows that a t_0^* function similar to the one estimated in this paper can be considered a linear approximation to the solution of (2.5). Thus, by maintaining the assumptions underlying (2.5), the structural parameters of the model (e.g., b and δ) can be identified. The identifying restrictions are much more complicated, however, for the multi-child case.

where X_i is a vector of observed characteristics, such as school quality and parents' education, and u_i is a mean zero random error, representing the child's innate ability and unobserved household and school characteristics.

Measurement of household income is problematic in countries like Peru because self-employment is so pervasive. However, the assumption of borrowing constraints provides a possible way around the lack of income data. Consider the logarithmic utility, constant parental income, example discussed in section III. In that case, household consumption, while the child is attending school, is given by

$$(3.2)\quad \begin{aligned} C(t) &= y && \text{for } 0 \le t \le t_0. \\ &= y\, e^{(b-\delta)(t-to)} && \text{for } t_0 \le t \le t_1, \end{aligned}$$

A similar relationship holds in the multi-child case. Equation (3.2) gives a justification for using household expenditures, data on which are available in many LDC household surveys, as a proxy for income in the t_0^* function, (2.10).

Of course, the use of expenditure data creates its own problem. Since household consumption is jointly determined with t_0^* in the model, the former will be endogenous in an equation describing the probability that $t > t_0^*$. Households facing greater noneconomic barriers to schooling, having less taste for schooling, or less able children will send their children to school less, put them to work more, and thus be able to consume more. Fortunately, it is not difficult to find instruments for family consumption expenditures, with parental occupation and wage rate variables being the obvious candidates. Mean expenditures of other households in the same village is also a useful instrument in the LDC context.

The switching point t_0^* also depends (negatively) upon the own initial wage, wH_0. While the value of a primary school age child's time may be influenced by local labor market conditions, and thus captured by regional or village level dummies, in Peru, and in most other developing countries, the labor market for young children is very thin. Children's main role is in home production or on the family farm. Proxies for wH_0 available from the survey data are discussed in the next section.

Late starting and self-selection

In addition to the repetition of grades by those children attending primary school, children may also start school late and drop out of primary school before completion. For the age range considered in this paper, school dropouts are very few in number, but there are a substantial number of late starters. Late starting is important to model for two reasons. First, it is interesting to see whether parental income and household demographic structure affect when a given child starts school, as might be expected if parents face borrowing constraints. Second, because it excludes late starters, the sample of children chosen for the grade repetition analysis will generally not be a random sample of children in that age range. If the late starting decision is correlated with the intensity of subsequent school attendance, as seems plausible, then estimates of the grade repetition model will be biased.

For the school starting decision, let t_s^* be the optimal starting time and assume it depends upon the same observable variables as t_0^*, as well as a random component representing unobserved child ability, measurement error and cultural factors. The overall empirical model, incorporating (3.1), is

$$(3.3) \quad t_{ji}^* = Z_i\beta_j + \varepsilon_{ij} \qquad j = 0,\ s$$

From the point of view of estimating the grade repetition model, the focus is on the probability of having repeated a grade conditional on being in school, or $\Pr(t > t_0^* \mid \text{age} > t_s^*)$. Assuming ε_0 and ε_s are jointly normal, mean zero errors, the joint unconditional probability is

$$(3.4)\quad \Pr(\ t > t_0^* \text{ and age} > t_s^*) = \Phi_2(t - Z\beta_0,\ \text{age} - Z\beta_s,\ \rho)$$

where Φ_2 is the bivariate normal cdf, and ρ is the correlation coefficient. Since grade repetition is not observed for children who have not yet started school, this probability is estimated as a bivariate probit with selection (see van de Ven and van Praag, 1981, for the appropriate likelihood function). Also, to deal with the endogenous regressors in the probit model, a Two-Stage Conditional Maximum Likelihood estimator is used (Rivers and Vuong 1988).

IV. Estimation

A. Data and variables

The data set used in the empirical investigation is the Peruvian Living Standards Survey (PLSS), a nationwide cross-sectional survey of about 5,000 households conducted by Peru's Institut National de Estatistica and The World Bank in 1985-86. The initial sample consists of the 4,257 children between the ages of seven and twelve, inclusive, who live with their parents,[23] and who have usable household expenditure data. Seven through twelve are the most common ages of enrollment in Peru's five grade primary school sequence.

A crucial conditioning variable in the grade repetition equation is time spent in school, denoted by t. In Peru, neither grade attainment nor age minus six adequately measure t. Number of grades completed does not take into account grade repetition, which implies a higher t, and age - 6 overstates t for the many children who started primary school after age six. Fortunately,

---

[23]Only limited information is available on children living away from home. Children of rural families may migrate to urban areas to pursue secondary and post-secondary education. Sample selection bias due to nonrandom migration is minimized here by restricting attention to primary school students.

the PLSS asks for the number of grades each child has repeated; so t can be calculated as grades completed plus grades repeated, which is independent of when the child started school. However, since the PLSS only asks for the number of grades repeated in the most recent level (primary, secondary, or post-secondary) in which a grade was completed, t cannot be calculated for those children who last completed a grade in secondary school or beyond. The exclusion of secondary school students, of which there are only 87 in the initial sample, creates the problem that the remaining students observed to have spent more than five years in school will not be representative of that population; i.e., they will all be laggards. Thus, these additional 70 children are eliminated, leaving 4,100 observations.

The PLSS asks whether each child is attending school at present, and, if not, whether he or she has attended during the past twelve months. If both questions are answered in the negative, the child is assumed to have quit school permanently. If the child is not currently attending, but has attended during the past year, then the child is considered still in school. However, unless the family was interviewed during Peru's summer vacation, children in this latter category are apparently not attending full-time.[24] These children, in addition to those who have actually repeated at least one grade, are assumed to be falling behind in school (i.e., $t > t_0^*$). An advantage of expanding the definition of being behind in school in this way is that children who have not yet completed first grade may be picked up as falling behind (these are children for whom $S^*(0) < 1$).

---

[24]A problem with this interpretation is that the twelve month reference period may span two academic years, so that it is possible that the child has already dropped out of school completely by the interview date. However, there are very few dropouts in the seven to twelve age range, so this problem would not appear to be serious.

Table 1 summarizes the status of the 4,100 children in the sample, illustrating large rural-urban differences in schooling patterns. Note that just 25 of these children have dropped out of school, and, of the remaining 4,075, a total of 3,654 or about 90 percent are enrolled. Among the 421 children who are not enrolled in school, there are clearly some who will never attend--the incidence of nonattendance at each age never falls below 3.4 percent (the figure for eleven year-olds). Still, the majority of the seven, eight and nine year-olds who are not enrolled will enter school eventually. For estimation purposes, all children who have not started first grade by age seven are assumed to be late starters. Of the 3,654 children who have started school, a total of 896 or 24.5 percent, have fallen behind in school by the above definition--160 because they are not presently in school though they have attended during the past year.

A final point about the use of t as a regressor in the empirical model has to do with measurement error. Because the dependent variable in the analysis is based on whether the child has repeated a grade or not, and t is constructed by summing grades repeated and completed, errors in reporting grade repetition may lead to large biases in the coefficient estimates. Thus, rather than inserting t directly into the probit model, t is regressed on the set of instruments, including the age of the child, in the first stage. As in Rivers and Vuong (1988), the estimated residual from this regression is then included along with t in the bivariate probit specifications.

The means of the other regressors used in the analysis are reported in column one of table 3. Real monthly household expenditures is created from disaggregated data on several expenditure items.[25] To construct the age gap

---

[25]This variable is due to Glewwe (1987). Expenditure items include (1) regularly purchased non-food items and food consumed outside the household

variables, children are matched to their older and younger, immediately adjacent, siblings six to eighteen years old (note the larger age range as compared to that of the sample).[26] Similarly, variables are constructed to record the sex (i.e., 1 if female) of the next oldest and youngest sibling.[27] The regressions also control for number of older and younger siblings, and the number of pre-school age children (zero to five years old).

To address the issue of heterogeneity in the initial value of child time, wH_0, one must understand what children do in Peru. According to studies of the Peruvian Highlands by Collins (1983) and Deere (1983), two of the main activities of primary school age children are the caring for younger siblings and livestock production activities (herding, gathering forage, etc.).[28] The inclusion of the number of pre-school age children (see above) captures this first activity, while a measure of the size of the household's herd of

---

within the last two weeks; (2) clothing, household goods and maintenance, medicines and other irregular expenditures within the last three months; (3) food expenditures within the last two weeks; (4) estimated rental value of durable goods (based on depreciation of reported present value); (5) value of food produced and consumed by the household in the last three months; (6) value of payments in kind (food and non-food) received by household members; and (7) actual housing rental payments or imputed rents for owner occupied housing based on hedonic rent equations. To obtain real monthly expenditures, nominal expenditures for each household are deflated using regional monthly price indices in order to correct for Peru's rampant inflation over the survey period (the country's consumer price index rose by over seventy percent), and then normalized by a regional price level.

[26] If the child in question had no older or younger siblings, then the age gaps are set to the maximum gap plus one year (i.e., 19 - age or age - 4).

[27] When either no older or younger siblings are present, the sex variable is given the value of .5.

[28] Jacoby (forthcoming) uses a production function approach and the PLSS data to show that, in the Peruvian Highlands, young children contribute more to livestock output than they do to crop output, and more to livestock output than even teenage children do.

livestock captures the second. Using the PLSS data, a herd size variable is constructed by adding the number of cows, goats, sheep and llamas the household has. Families, however, may decide to sell off livestock to finance a child's education; in other words, herd size cannot be treated as exogenous with respect to child schooling decisions. Thus, the mean of herd size within each cluster[29] is used in the regressions (though this turns out not to make much difference in the results), with the idea that cluster fixed effects (e.g., climate and altitude) are uncorrelated with the unobservables in (3.3).

The learning efficiency characteristics, X_i, consist of the years of completed schooling of the child's mother and father along with a set of school quality variables. Three indicators of the quality of the last primary school attended by each individual are available: (1) the number of teacher per grade offered;[30] (2) a dummy for whether the school had writing facilities (i.e., desks) for each student; and (3) a dummy for whether the individual had textbooks for their personal use, which may include books purchased by the child's family as well as those provided by the school. Rather than use these individual level quality variables directly, the means across children within each cluster are taken. Using cluster means reduces the possibility of finding a spurious positive relationship between school quality and progress through school, arising from the fact that more able children, who are less likely to fall behind, go to schools with better facilities and are more

---

[29] PLSS households are organized into about 350 small and dispersed geographic units called "clusters", which in rural areas correspond roughly to villages.

[30] The survey asks for the highest grade offered and the number of teachers in the last primary school attended. Since there are a few primary schools that apparently do not offer five grades, the number of teachers per grade is used. Ultimately, the variable of interest is the number of teachers per student, but this is simply not available in the survey.

likely to be furnished books by their parents.

Finally, a dummy variable indicating whether or not the individual resides in a rural area (defined as a village with fewer than 2,000 inhabitants), along with a set of twelve regional dummies, is included in all specifications. It will be interesting to see whether the rural dummy has a significant coefficient once household consumption and school quality (both lower on average in rural areas), as well as the value of child time (probably higher in the countryside), are taken into account.

B. Estimation results

In the first stage of the estimation procedure, t and log(C(t)) are regressed on all of the exogenous variables mentioned above and a large set of variables excluded from the grade repetition and late starting equations (see table 2). Among these excluded variables are the age and sex of the household head, dummies for the occupational sector of the head, dummies for the mother's occupational category, total household size and the number of elderly people in the household, the size of the family's land holdings (both irrigated and non-irrigated), mean wages of men and women in the cluster of residence and mean household expenditures in that cluster. The whole sample of 4,100 children is used in the first stage to avoid sample selection bias.

The next step is to estimate probability (3.4) on the full sample of 4,075 children who have not yet dropped out of school.[31] The first column of table 3 displays the results for grade repetition. The log expenditure

---

[31]There are substantially fewer households than children in the sample, and the multi-child theoretical model takes this into account. Still, unobservables are likely to be correlated across children within the same household. Ignoring this correlation, however, does not affect consistency of the estimates, leading only to a loss in efficiency.

variable has a negative and highly significant coefficient, meaning that children in higher income households are less likely to have repeated a grade for each t, a clear rejection of unrestricted borrowing and support of the alternative hypothesis suggested here. Note that t itself has the obvious positive effect on the repetition probability, and the hypothesis that both log(C(t)) and t are statistically exogenous is rejected.

The variables associated with adjacent siblings show some interesting results. As the age gap between the child and his or her next oldest sibling shrinks, the probability of grade repetition increases (the coefficient's p-value = .053), just as the theoretical model of borrowing constraints suggests. If this next oldest sibling is female, the probability of repeating also falls. This finding is consistent with the story told in figure 4, when girls are treated as though their w/b is higher relative to that of boys. Surprisingly, though, the selectivity corrected estimates show that girls are not significantly more likely to repeat grades than boys, conditional on being enrolled in school. Instead, what seems to be driving the result is the fact that girls older than twelve are more likely to drop out of school early than boys older than twelve.[32] The age and gender effects for adjacent younger siblings are not significantly different from zero.[33]

Of the other demographic variables, none has a significant impact on grade repetition except for the number of children under five years old, with

---

[32]To be precise, girls age thirteen to eighteen, conditional on attending school ever, have a higher dropout rate at each age than boys of the same age. Overall, 23.7 percent of the girls dropout compared to 16.8 percent of the boys.

[33]An identical model with these variables replaced by the averages of age gaps between the child and all his older and younger siblings, and the proportions of females among them, gives virtually the same results (with marginally lower significance), and is not reported.

a strong negative effect on school progress. However, it is unlikely that this represents a spacing effect, because of the insignificance of the younger adjacent sibling variables already mentioned. A more plausible story is that the presence of infants increases the initial value of child time, promoting grade repetition. Meanwhile, the cluster average herd size, the other child productivity variable, has no significant effect in column two.

Parental education has the expected negative impact on grade repetition, and both the coefficients on mother's and father's schooling are significant at conventional levels. The effect of mother's schooling is larger in magnitude than father's schooling. None of the school quality variables are significant in the grade repetition equation. Finally, the rural dummy is also insignificant; apparently, income differences have captured much of the urban-rural differences in repetition rates evidenced in table 1.

Column three of table 3 contains the estimates of the probability that a child has started school. Again, income, as proxied by household expenditures, has a strong positive effect on school progress: children from higher income families are more likely to be in school at any given age. Gender now has a significant effect, with girls more likely to start late than boys. As seen in table 1, though, in rural areas the gender difference in school enrollment is concentrated at the later ages; more girls apparently never attend school at all. None of the household demographic variables has any impact on the school starting time, except the number of pre-school age children once again. Thus, although parental income certainly plays a role in the decision, the close spacing of siblings does not seem to be an explanation for delays in school enrollment.

Along with the significant effect of children under five years old, the size of herd variable shows up very significantly in the enrollment probability. Children from villages with more livestock per household are less likely to have started school. Again, these findings point to an economic explanation of late starting, rather than to the argument that children simply mature at different rates.

Interestingly, school quality variables have more of an impact on school starting than they do on grade repetition, and the books variable becomes significant in column 2, with greater access to books encouraging enrollment. Better educated parents start their children in school sooner, and mother's schooling again has a larger effect than the father's.

Note finally, the extremely large magnitude and significance of the estimated correlation coefficient in the bivariate probit. The negative correlation between ε_0 and ε_s makes sense if the error terms contain unobserved ability. Abler children are more likely to start school on time and to progress through school faster. Comparing the bivariate probit, which controls for selectivity, with univariate probit estimates of the grade repetition equation shows that sample selectivity bias leads to a very significant overestimate of the coefficient on time in school. The other coefficients, however, are not greatly affected.

Splitting the sample

In the estimates in table 3, all households are assumed to have the same access, or lack of access, to credit markets. But if households have different borrowing opportunities, then pooling the data would tend to obscure the effect of income, for example, on schooling patterns. Two approaches to splitting the sample into constrained and unconstrained groups are possible using the PLSS data. The indirect approach, used by Hayashi (1985), Zeldes (1989) and others to investigate the effect of liquidity constraints on consumption behavior, involves identifying households by their asset holdings or savings. The direct approach splits the sample on the basis of a household's response to a survey question about its access to credit markets. The indirect approach runs into difficulties in a poor country with undeveloped financial markets and high inflation. In such an environment, households, particularly in rural areas, may accumulate wealth primarily in the form of physical assets (e.g., land and livestock), rather than financial assets (e.g., savings accounts). These physical assets may be used either as collateral for loans, or traded to smooth consumption.

Fortunately, the PLSS provides information on a broad array of household assets, including those used on family farms and business enterprises. An aggregate measure of current assets is constructed by adding the following: (1) cash savings, (2) ninety percent of the value of durable goods, (3) agricultural assets (value of land + value of farm animals + value of farm equipment),[34] and (4) value of physical capital used in family enterprises. Various ways of splitting the sample based on aggregate assets lead to similar conclusions. In one attempt, 1,360 children whose households' assets exceed median yearly consumption expenditures (9,660 Intis) are put in the "unconstrained" group, and the 2,699 remaining children are assigned to the "constrained" group.[35]

The estimates of the model on these two subsamples appear in table 4. The correlation coefficient in the bivariate probit for the unconstrained

---

[34] Because of missing data, in the case of land, cluster medians of land values per hectare of dry and irrigated land, respectively, are used to calculate total value of land for each household.

[35] The mean total asset value is 19,588 Intis, but the median is only 4,355. There are sixteen observations in the sample with missing asset data.

group failed to converge to a value greater than -1,[36] so only the univariate probits are reported in this case (as was seen, the parameters of interest are not significantly affected by selectivity bias anyway). Overall, there is no evidence that the high asset group behaves differently than the low asset group.[37] In fact, the coefficient on log expenditures is actually larger in magnitude for the former group in both decision rules. If one wishes to stand by the model, then the conclusion from this exercise must be that either assets are poorly measured, or that using current asset holdings does not adequately separate constrained from unconstrained households, or that all households are actually constrained.

Before concluding that all households in Peru are constrained, however, the direct approach should be tried. The PLSS asks each household whether any consumption credit is available to it.[38] Of the 4,998 households answering the survey, 831 or about seventeen percent responded in the affirmative (the percentage is about the same in the current sample of children). Before splitting the sample on the basis of this criterion, one might ask which household characteristics determine access to credit. Consider the probit estimates reported in the third column of table 2, in which the same set of instruments used in the first stage expenditures regression are used to

---

[36]This finding may be due to the lack of exclusion restrictions in the model; ρ is identified solely from the nonlinearity. Although there is no clear theoretical justification for any exclusion restrictions in this model, when some variables are arbitrarily excluded from one equation or another ρ still moves outside the parameter space.

[37]Essentially, the correlation between assets and the school progress indicators (grade repetition and enrollment) is very weak.

[38]The exact question is: "Do you have any consumption credit available to your household (or enterprises owned by your household)? For example, credit cards, cooperatives, etc.?" (Grootaert and Arriagada, 1986).

predict the availability of credit to the household (credit=1). The most significant coefficients reveal that households more likely to have access to consumption credit live in clusters with higher average expenditures. Education seems to improve access to credit, as does the fact that the household head works in the service sector (the omitted category), which includes government, and not in the agricultural, manufacturing, construction, commerce or transportation sectors.

Given these reasonable results, the model is estimated separately on the two groups of households. The estimates for the so-called constrained group (3,380 children) are recorded in the first and second columns, and the corresponding estimates for the unconstrained group (695 children) appear alongside in the next two columns.[39] While the coefficient on log expenditures for the unconstrained group in the grade repetition decision is negative, the t-ratio is only .64. The corresponding estimate for the constrained group has a t-ratio of nearly four. Likewise, for the late starting decision the point estimates of the expenditure coefficients and the effects of herd size are similar across the two groups, but only for the constrained children are the coefficients significant. These findings seem to support the contention that the school attendance patterns of children in unconstrained households are independent of current family income and the value of child time. Results for the sibling variables, however, are inconclusive.

---

[39] Once again, the bivariate probit for the smaller, unconstrained, group had difficulties converging ($\hat{\rho}$ is very close to the boundary). In this case, though, eventually one of convergence criteria in the LIMDEP package was met. In spite of these difficulties, the bivariate probit estimates and standard errors reported in table 5 are in line with the univariate probit results.

V. Summary and Conclusions

This paper has argued that poverty, in conjunction with constraints on the ability to borrow, can explain school attendance patterns in less developed countries. Cross-sectional variation in these patterns across primary school age children in Peru seems generally consistent with the model. Lower income families begin withdrawing their children from school earlier, and there is some evidence that school progress suffers in households that place a greater value on child time. The model of multi-child households delivers readily testable implications, and some empirical support is found for the hypothesis that close spacing of siblings leads parents to hasten the withdrawal of children from school. Splitting the sample into ostensibly constrained and unconstrained groups leads to ambiguous results. When the split is based on asset holdings, it appears as though all households are constrained. But when the split is based on self-reported credit constraints, it is found that economic variables do not affect schooling decisions of unconstrained households, as would be expected.

Clearly, much more theoretical and empirical work needs to be done to explain the complicated schooling patterns found in poor countries. This paper takes a first step in viewing phenomena such as grade repetition and late starting as the outcomes of economic choices. Future research must address the impact of unanticipated income fluctuations. If households cannot borrow ex-post, shocks to parental income may lead to temporary withdrawals of children from school. This phenomenon may even help explain late starting, but an empirical investigation must await adequate longitudinal data on child time allocation and family income.

References

Becker, G. S. and H. G. Lewis (1973): "On the interaction between the quantity and quality of children," *Journal of Political Economy*, 81: S279-S288.

———————— and N. Tomes (1976): "Child endowments and quantity and quality of children," *Journal of Political Economy*, 84:S143-S162.

Behrman, J., R. A. Pollack and P. Taubman (1982): "Parental preferences and provision for progeny," *Journal of Political Economy*, 90:52-73.

Ben-Porath, Y. (1967): "The production of human capital and the life cycle of earnings," *Journal of Political Economy*, 75: 352-365.

Blinder, A. and Y. Weiss (1976): "Human capital and labor supply: a synthesis," *Journal of Political Economy*, 84: 449-472.

Collins, J. (1983): "Fertility determinants in a High Andes community," *Population and Development Review*, 9: 61-75.

Deere, C. (1983): "The allocation of familial labor and the formation of peasant household income in the Peruvian Sierra," in *Women and Poverty in the Third World*, ed. by M. Bovinic, M. Lycette, and W. McGreevey. Baltimore: Johns Hopkins University Press.

Gertler, P. and P. Glewwe (1989): "The willingness to pay for education in developing countries: Evidence from rural Peru," LSMS working paper #54, The World Bank.

Glewwe, P. (1987): "The distribution of welfare in Peru in 1985-86," LSMS Working Paper #42, The World Bank, December.

Glewwe, P. and H. Jacoby (1991): "Student achievement and schooling choice in low income countries: Evidence from Ghana," unpublished manuscript.

Greenhalgh, S. (1985): "Sexual stratification: The other side of 'Growth with Equity' in East Asia," *Population and Development Review*, 11: 265-314.

Grootaert, C. and A-M. Arriagada (1986): "The Peruvian Living Standards Survey: An annotated questionnaire," The World Bank, Washington D.C.

Hayashi, F. (1985): "The effect of liquidity constraints on consumption: A cross-sectional analysis," *Quarterly Journal of Economics*, 100: 183-206.

Heckman, J. (1976): "A life cycle model of earnings learning and consumption," *Journal of Political Economy*, 84(Supplement): 511-544.

Jacoby, H. (1990): "Borrowing constraints and school attendance: Theory and evidence from a low income country, University of Rochester, June.

———————— : "Productivity of men and women and the sexual division of labor in peasant agriculture of the Peruvian Sierra," *Journal of Development Economics*, forthcoming.

Johnson, T. (1978): "Time in school: the case of the prudent patron," *American Economic Review*, 68: 862-872.

King, E. and R. Bellew (1989): "The effects of Peru's push to improve education," P.P.R. Working Paper #172, The World Bank.

Lockheed, M. and A. Verspoor (1989): *Improving Primary Education in Developing Countries: A Review of Policy Options*, The World Bank.

Rivers D. and Q. H. Vuong (1988): "Limited information estimators and exogeneity tests for simultaneous probit models," *Journal of Econometrics*

Rosenzweig, M. and R. Evenson (1977): "Fertility, schooling and the economic contribution of children in rural India: an econometric analysis," *Econometrica*, 45: 1065-1080.

Tomiyama, K. (1985): "Two-stage optimal control problems and optimality conditions," *Journal of Economic Dynamics and Control*, 9: 317-337.

van de Ven, W. and B. van Praag (1981): "The demand for deductibles in private health insurance," *Journal of Econometrics*, 17: 229-252.

Wallace, T. and L. Ihnen (1975): "The theory of human capital accumulations with alternative loan markets," *Journal of Political Economy*, 83: 157-184.

Weiss, Y. (1986): "The determination of life cycle earnings: a survey," *Handbook of Labor Economics*. Ed. by O. Ashenfelter and R. Layard. Amsterdam: Elsevier Science Publishers.

Zeldes, S. (1989): "Consumption and liquidity constraints: An empirical investigation," *Journal of Political Economy*, 97: 305-346.

Appendix

A. *Single-Child Households*

(i) Unconstrained borrowing

Define multiplier functions $\mu(t)$ and $\lambda(t)$ associated with the human capital production constraint and the budget constraint, respectively. In addition to these laws of motion for the state variables, necessary conditions for the maximization of the corresponding Hamiltonian are

$$\text{(A.1)} \quad U'(C(t))\, e^{-\delta t} = \lambda(t)$$

$$\text{(A.2)} \quad \mu(t)b \begin{cases} < \lambda(t)w & \text{if } S(t) = 0 \\ = & \text{if } 0 < S(t) < 1 \\ > & \text{if } S(t) = 1 \end{cases}$$

$$\text{(A.3)} \quad \dot{\lambda}(t) = -r\lambda(t)$$

$$\text{(A.4)} \quad \dot{\mu}(t) = -\lambda(t)w - S(t)(\mu(t)b - \lambda(t)w)$$

The boundary condition on the shadow price of human capital, $\mu(T) = 0$, derives from the assumption that human capital has no value at the end of the lifetime. As, for example, in Heckman (1976), it is useful to write $g(t)=\mu(t)/\lambda(t)$ and interpret $g(t)$ as the shadow price of human capital relative to nonhuman capital.

Using (A.2)-(A.4) it easily shown that

$$\text{(A.5)} \quad \dot{g}(t) = -(b - r)g(t) \qquad \text{for } S(t) = 1$$

$$\text{(A.6)} \quad \dot{g}(t) = -w + rg(t) \qquad \text{for } S(t) = 0$$

$$\text{(A.7)} \quad g(t) \text{ satisfies (A.5), (A.6) and } g(t) = w/b \qquad \text{for } 0 < S(t) < 1$$

Clearly, $S(t)$ cannot take on a value between zero and one over an interval of time, since (A.7) implies $\dot{g}(t) = 0$ which is inconsistent with both (A.5) and (A.6) unless $b = r$. The optimal $S(t)$ is thus a "bang bang" control. If $S(t)$ is set to unity from $t = 0$ to t_1 and falls to zero thereafter, the solutions to (A.5) and (A.6) are as follows

(A.8) $g(t) = w/b\ e^{(b-r)(t_1-t)}$ for $S(t) = 1$

(A.9) $g(t) = w/r\ [1 - e^{r(t-T)}]$ for $S(t) = 0$

The switching point, $t_1 = t_1^* = T + \ln(1 - r/b)/r$, is determined by the fact that $g(t_1)=w/b$.[40]

(ii) Constrained borrowing

Associate another multiplier function $\eta(t)$ with the constraint $A(t) \geq 0$, where $\eta(t)A(t) = 0$. The only change in the first order conditions appears in equation (A.3), which becomes

(A.10) $\dot{\lambda}(t) = -r\lambda(t) - \eta(t)$

Consequently, the time path of $g(t)$ is determined by

(A.11) $\dot{g}(t) = -(b - r - \eta(t)/\lambda(t))g(t)$ for $S(t) = 1$

(A.12) $\dot{g}(t) = -w + (r + \eta(t)/\lambda(t))g(t)$ for $S(t) = 0$

(A.13) $g(t)$ satisfies (A.11), (A.12) and $g(t) = w/b$ for $0 < S(t) < 1$

Differentiation of condition (A.1) together with (A.10) implies

(A.14) $\dot{C}(t) = -(r - \delta + \eta(t)/\lambda(t))\ U'/U''$.

Now if constrained parents choose $S(t) = 1$ over the interval $[0, t_0]$, then consumption must equal family income $y(t)$, and, in particular, $\dot{C}(t) = \dot{y}(t)$. Therefore, from (A.14), $\eta(t)/\lambda(t) = \delta - r - \dot{y}(t)U''/U'$. For the special case used in the text, $\dot{y}(t)=0$, $\eta(t)/\lambda(t) = \delta - r$, and,, by (A.11), $\dot{g}(t) = -(b - \delta)g(t)$. Solving the latter, using the boundary condition $g(t_0) = w/b$, produces

---

[1]If the individual instead began life with zero investment and entered school full-time only as the age of retirement approached, then (A.5) and the boundary condition $g(T)=0$ would imply that g is zero throughout the investment phase, which requires an impermissible jump (a drop actually) in g at the switching point. It is easily proven that the Hamiltonian is continuous at the switching point in the former case, so that (A.8) and (A.9) characterize the optimal path.

$$(A.15)\quad g(t) = (w/b)\, e^{-(b-\delta)(t-t_0)} \qquad \text{for } 0 \le t \le t_0$$

Proceeding through the life-cycle, suppose a period of part-time schooling is undertaken over the interval $[t_0, t_1]$. According to (A.13), during this phase $\dot{g}(t)=0$, and therefore by (A.11), $\eta(t)/\lambda(t) = b - r$. The growth of constrained consumption is then given by $\dot{C}(t) = -(b - \delta)U'/U'' > 0$. Because the relative shadow price of human capital is constant over $[t_0, t_1]$, to satisfy the terminal condition $g(T)=0$ there must be a period of zero investment where $g(t)$ falls. Analyzing this last phase in the life-cycle will determine how long the total schooling period (i.e., $[0, t_1]$) lasts.

After t_1, total family income (inclusive of child earnings) is constant over time. Therefore, as long as $\delta > r$, the household wishes its consumption to remain constant, which again implies that $\eta(t)/\lambda(t) = \delta - r$. Inserting this equality into (A.12), and using $g(T)=0$ gives

$$(A.16)\quad g(t) = w/\delta\ [1 - e^{\delta(t-T)}] \qquad \text{for } t > t_1$$

Again, the fact that $g(t_1)=w/b$ uniquely determines $t_1^* = T + \frac{1}{\delta}\ln(1-\delta/b)$.

B. Two-Child Households

(i) Stage two

The analysis proceeds exactly as in the last section, except that there are two types of human capital, $H_1(t)$ and $H_2(t)$, with associated shadow prices, $g_i(t) = \mu_i(t)/\lambda(t)$ $i=1,2$, and that the left endpoint is a_1. Consider the solution depicted in panel (d) of figure 4 (all other stage two scenarios are special cases). It is assumed that $w_1/b_1 > w_2/b_2$.

If $S_1(t)=S_2(t) = 1$ over the interval $[a_1,t_{01}]$, then $\eta(t)/\lambda(t) = \delta - r$ and $\dot{g}_1(t) = -(b_1 - \delta)g_1(t)$. Since child 1 begins withdrawing first at t_{01}, $g_1(t)$ is the relevant shadow price, initially.

Between t_{01} and t_{11} child 1 is withdrawn from school and $g_1(t)=w_1/b_1$. Meanwhile consumption grows at rate $\dot{C}(t) = -(b_1 - \delta)U'/U''$. After child 1 quits school completely, it will be seen momentarily that there must be a period of full-time schooling for child 2, from t_{11} to t_{02} (unless $w_1/b_1 = w_2/b_2$). First note that on this interval $g_2(t)$ is the relevant shadow price, and $\dot{g}_2(t) = -(b_2 - \delta)g_2(t)$.

Finally, from t_{02} to t_{12}, the second child withdraws, $g_2(t)=w_2/b_2$, and consumption grows at rate $\dot{C}(t) = -(b_2 - \delta)U'/U''$. Once both children have left school, family income and consumption stay constant until T.

Working backwards again, it is possible to derive all the switching points. Using the fact that $g_2(t) = w_2/\delta\ [1 - e^{\delta(t-T)}]$ on $[t_{12},T]$ and $g(t_{12})=w_2/b_2$ delivers the last switching point, $t^*_{12} = T + \frac{1}{\delta}\ln(1-\delta/b_2)$.

Recall that when child 1 leaves school, the relevant shadow price of human capital switches from $g_1(t)$ to $g_2(t)$. Therefore, it must be true that $g_1(t_{11})=g_2(t_{11})$, which means

$$\text{(B.1)}\quad g_1(t_{11})= w_1/b_1 = (w_2/b_2)\ e^{-(b_2-\delta)(t_{11}-t_{02})} = g_2(t_{11})$$

or $t_{11}-t_{02} = \frac{1}{(b_2-\delta)} \log(b_1w_2/b_2w_1)$. In other words, the period of full-time schooling of child 2, following the withdrawal of child 1, is longer the greater the difference in their w/b.

The remaining switching points can be derived by assuming a form for U (e.g., logarithmic), and solving the differential equations for the stocks of human capital over the relevant intervals.

(ii) Stage one

Applying Tomiyama's (1985) results, the endpoint condition on $\mu_1(t)$ becomes

$$\text{(B.2)}\quad \mu_1(a_1) = \frac{\partial}{\partial H_1(a_1)} \int_{a_1}^{T} U(C^*(t))dt,$$

where $C^*(t)$ is the optimized value of consumption. Furthermore, it is necessary that $\mu_1(t)$ and $\lambda(t)$ be continuous at a_1.

Even by solving the second stage explicitly and obtaining $C^*(t)$, it is still impossible to solve the first stage explicitly because of the complicated endpoint condition. The best that can be done is to obtain a set of nonlinear equations that determine the first stage switching times as a function of the parameters and a_1.

TABLE 1

SCHOOL ATTENDANCE STATUS OF PLSS SAMPLE

Age:	7	8	9	10	11	12	7-12
Urban Males							
Late	15	4	1	2	4	2	28
Behind	4	25	43	48	48	41	209
Total	174	192	191	180	167	150	1054
Urban Females							
Late	27	10	3	4	2	1	47
Behind	7	26	31	48	44	38	194
Total	187	184	156	188	180	131	1026
Rural Males							
Late	58	51	27	13	5	8	162
Behind	10	27	51	46	67	64	265
Total	185	184	201	142	163	163	1038
Rural Females							
Late	61	48	25	20	11	19	184
Behind	11	26	43	45	41	62	228
Total	160	191	162	170	138	161	982
All Children							
Late	161	113	56	39	22	30	421
Behind	32	104	168	187	200	205	896
Dropout	1	1	3	4	6	10	25
Total	706	751	710	680	648	605	4100

Notes: Late refers to number of children who have not started school. Behind means child has repeated a grade or is enrolled in school but is not currently attending. Dropout means child is no longer enrolled. Totals are for respective category. Rural is defined as a village with fewer than 2,000 inhabitants.

TABLE 2

SELECTED COEFFICIENT ESTIMATES FROM FIRST STAGE REGRESSIONS

Regressor	Log Expend.	Yrs. in school	Pr(credit = 1)
log expenditures (cluster mean)	.360 (.022)	.0230 (.042)	.309 (.060)
Adult male wage (cluster mean)	.0008 (.0005)	-.00048 (.0009)	-.00057 (.0016)
Adult female wage (cluster mean)	.0134 (.004)	-.0031 (.007)	.0208 (.010)
Age of head	-.0005 (.0010)	.0016 (.002)	-.0020 (.003)
Sex of head (1 if female)	-.207 (.029)	-.129 (.057)	-.097 (.084)
Head's occupation in agriculture	.0586 (.035)	-.198 (.066)	-.333 (.094)
Head's occupation in mining	.230 (.070)	-.237 (.134)	.019 (.187)
Head's occupation in manufacturing	-.0460 (.036)	-.0034 (.069)	-.406 (.094)
Head's occupation in construction	-.0912 (.045)	-.037 (.086)	-.516 (.127)
Head's occupation in commerce	.0149 (.034)	-.0012 (.065)	-.514 (.089)
Head's occupation in transportation	.0204 (.046)	.0750 (.088)	-.622 (.127)
Head's occupation in financial	-.064 (.070)	-.079 (.134)	.0051 (.161)
Father's schooling	.0335 (.025)	.0412 (.007)	.0210 (.009)
Mother's schooling	.0251 (.004)	.0235 (.007)	.0254 (.010)
Household size	.113 (.008)	-.0042 (.014)	.0290 (.021)
Number of adults 65 ≤ age ≤ 74	-.0514 (.031)	.102 (.059)	-.196 (.087)
Number of adults 75 ≤ age ≤ 99	-.102 (.043)	-.139 (.082)	-.433 (.143)
Dry land (Hectares)	-.00002 (.00016)	-.00007 (.0003)	-.0038 (.0022)
Irrigated land (Hectares)	.0185 (.0030)	.0134 (.006)	.00041 (.0085)
Age of child		.684 (.013)	
R^2	.46	.59	

Notes: Standard errors in parentheses. All regressions include occupational dummies for the mother and twelve regional dummies, as well as the instruments listed in table 3. Columns 1 and 3 include household level demographic variables rather than individual level variables in column 2.

TABLE 3

DETERMINANTS OF SCHOOL PROGRESS: FULL SAMPLE RESULTS

Regressor	Means	$Pr(t>t_0^* \mid age>t_s^*)$	$Pr(age>t_s^*)$
Log expenditures	7.38 (.77)	-.373 (.094)	.411 (.125)
Expenditures residual	----	.271 (.100)	-.350 (.137)
Years in school	2.07 (1.71)	.143 (.038)	----
Years in school residual	----	.217 (.038)	----
Age of child	9.39 (1.68)	----	.265 (.046)
Sex of child	.488 (.500)	.0186 (.0465)	-.135 (.062)
Age gap to next oldest sibling	4.55 (3.45)	-.0200 (.0103)	.0085 (.0129)
Sex of next oldest sibling	.494 (.422)	-.1047 (.0551)	.100 (.072)
Age gap to next youngest sibling	3.02 (1.56)	-.0023 (.0024)	-.0024 (.0492)
Sex of next youngest sibling	.497 (.339)	-.0745 (.0640)	.0264 (.1050)
Number of older siblings (6-18)	1.54 (1.38)	.0043 (.0276)	-.0349 (.0329)
Number of younger siblings (6-18)	.662 (.844)	.0048 (.0498)	.0775 (.0910)
Number of children age < 6	1.13 (1.08)	.108 (.023)	-.0739 (.0280)
Herd size (cluster mean)	7.10 (13.0)	.00132 (.00205)	-.00683 (.00262)
Teachers per grade (cluster mean)	1.49 (.73)	.0274 (.0625)	.0894 (.0797)
Writing facilities? (cluster mean)	.896 (.149)	.108 (.187)	.305 (.196)
Textbooks? (cluster mean)	.770 (.174)	-.062 (.188)	.586 (.223)
Father's schooling	5.63 (4.24)	-.253 (.009)	.0494 (.0137)
Mother's schooling	3.84 (4.04)	-.0562 (.0106)	.0643 (.0164)
Rural?	.490 (.500)	.0285 (.0768)	-.087 (.107)
Correlation coeff.	----	-.951 (.032)	
Sample size	4,075	3,654	4,075

Notes: All equations include the twelve regional dummies and a constant. Standard errors are in parentheses, and are not adjusted for the fact that the first stage residuals are generated regressors.

TABLE 4

DETERMINANTS OF SCHOOL PROGRESS: SAMPLE SPLIT BASED ON ASSET HOLDINGS

Regressor	Constrained group $Pr(t>t_0^*\|age>t_s^*)$	Constrained group $Pr(age>t_s^*)$	Unconstrained group $Pr(t>t_0^*\|age>t_s^*)$	Unconstrained group $Pr(age>t_s^*)$
Log expenditures	-.243 (.119)	.333 (.164)	-.585 (.188)	.836 (.244)
Expenditures residual	.204 (.125)	-.292 (.179)	.403 (.203)	-.775 (.265)
Years in school	.175 (.045)	----	.298 (.063)	----
Years in school residual	.174 (.048)	----	.188 (.073)	----
Age of child	----	.277 (.062)	----	.266 (.077)
Sex of child	.060 (.060)	-.196 (.081)	-.081 (.089)	-.040 (.111)
Age gap to next oldest sibling	-.020 (.013)	.013 (.017)	-.014 (.021)	-.004 (.024)
Sex of next oldest sibling	-.079 (.070)	.058 (.092)	-.190 (.104)	.091 (.125)
Age gap to next youngest sibling	.0023 (.0294)	-.074 (.064)	-.004 (.043)	.115 (.084)
Sex of next youngest sibling	-.104 (.081)	-.002 (.141)	-.035 (.116)	.060 (.173)
Number of older siblings (6-18)	.011 (.036)	-.050 (.045)	.006 (.053)	-.047 (.060)
Number of younger siblings (6-18)	-.025 (.063)	-.002 (.120)	.045 (.087)	.173 (.155)
Number of children age < 6	.100 (.029)	-.085 (.037)	.159 (.045)	-.033 (.051)
Herd size (cluster mean)	.0005 (.0027)	-.0066 (.0036)	.0005 (.0038)	-.0093 (.0045)
Teachers per grade (cluster mean)	.0052 (.077)	.184 (.095)	.001 (.128)	-.258 (.179)
Writing facilities? (cluster mean)	-.315 (.255)	.158 (.273)	.695 (.338)	.177 (.324)
Textbooks? (cluster mean)	.285 (.271)	.342 (.307)	-.133 (.298)	1.20 (.33)
Father's schooling	-.032 (.011)	.047 (.016)	.008 (.020)	.063 (.030)
Mother's schooling	-.064 (.013)	.048 (.020)	-.051 (.021)	.099 (.036)
Rural?	.021 (.094)	-.065 (.127)	.065 (.162)	-.529 (.285)
Correlation coeff.		-.914 (.059)		----
Sample size	2445	2699	1196	1360

Notes: See notes to table 3.

TABLE 5

DETERMINANTS OF SCHOOL PROGRESS: SAMPLE SPLIT BASED ON SELF-REPORTING

	Constrained group		Unconstrained group	
Regressor	$Pr(t>t_0^* \mid age>t_s^*)$	$Pr(age>t_s^*)$	$Pr(t>t_0^* \mid age>t_s^*)$	$Pr(age>t_s^*)$
Log expenditures	-.402 (.103)	.424 (.145)	-.190 (.297)	.421 (.493)
Expenditures residual	.280 (.111)	-.357 (.158)	.204 (.292)	-.406 (.459)
Years in school	.157 (.038)	----	.091 (.095)	----
Years in school residual	.199 (.041)	----	.332 (.119)	----
Age of child	----	.256 (.048)	----	.441 (.327)
Sex of child	.033 (.051)	-.157 (.067)	-.060 (.134)	.174 (.263)
Age gap to next oldest sibling	-.014 (.011)	.010 (.014)	-.049 (.030)	-.038 (.049)
Sex of next oldest sibling	-.101 (.060)	.074 (.077)	-.148 (.169)	.578 (.318)
Age gap to next youngest sibling	.009 (.026)	-.074 (.064)	.020 (.065)	-.260 (.324)
Sex of next youngest sibling	-.096 (.070)	.014 (.051)	.0002 (.183)	.007 (.402)
Number of older siblings (6-18)	.015 (.030)	-.036 (.037)	-.030 (.084)	.146 (.142)
Number of younger siblings (6-18)	-.008 (.054)	.106 (.094)	.074 (.142)	-.322 (.581)
Number of children age < 6	.112 (.025)	-.062 (.031)	.060 (.066)	-.176 (.103)
Herd size (cluster mean)	.0013 (.0022)	-.0061 (.0027)	.0022 (.0080)	-.0084 (.0228)
Teachers per grade (cluster mean)	.063 (.066)	.023 (.086)	-.199 (.240)	.681 (.463)
Writing facilities? (cluster mean)	.070 (.200)	.347 (.211)	.343 (.697)	-.250 (.875)
Textbooks? (cluster mean)	-.075 (.200)	.637 (.237)	.099 (.723)	-.26 (1.3)
Father's schooling	-.030 (.010)	.060 (.016)	-.014 (.026)	-.003 (.046)
Mother's schooling	-.056 (.012)	.049 (.019)	-.064 (.028)	.119 (.063)
Rural?	.030 (.083)	-.203 (.119)	.095 (.275)	.745 (.506)
Correlation coeff.		-.958 (.030)		-.993 (.198)
Sample size	3002	3380	652	695

Notes: See notes to table 3. Constrained group has credit=0.