Living Standards
Measurement Study
Working Paper No. 91

LSM- 91
NOV. 1992

Estimating the Determinants of Cognitive Achievement in Low-Income Countries

The Case of Ghana

Paul Glewwe
Hanan Jacoby

Public Disclosure Authorized

LSMS Working Papers

No. 19	*The Conceptual Basis of Measures of Household Welfare and Their Implied Survey Data Requirements*
No. 20	*Statistical Experimentation for Household Surveys: Two Case Studies of Hong Kong*
No. 21	*The Collection of Price Data for the Measurement of Living Standards*
No. 22	*Household Expenditure Surveys: Some Methodological Issues*
No. 23	*Collecting Panel Data in Developing Countries: Does It Make Sense?*
No. 24	*Measuring and Analyzing Levels of Living in Developing Countries: An Annotated Questionnaire*
No. 25	*The Demand for Urban Housing in the Ivory Coast*
No. 26	*The Côte d'Ivoire Living Standards Survey: Design and Implementation*
No. 27	*The Role of Employment and Earnings in Analyzing Levels of Living: A General Methodology with Applications to Malaysia and Thailand*
No. 28	*Analysis of Household Expenditures*
No. 29	*The Distribution of Welfare in Côte d'Ivoire in 1985*
No. 30	*Quality, Quantity, and Spatial Variation of Price: Estimating Price Elasticities from Cross-Sectional Data*
No. 31	*Financing the Health Sector in Peru*
No. 32	*Informal Sector, Labor Markets, and Returns to Education in Peru*
No. 33	*Wage Determinants in Côte d'Ivoire*
No. 34	*Guidelines for Adapting the LSMS Living Standards Questionnaires to Local Conditions*
No. 35	*The Demand for Medical Care in Developing Countries: Quantity Rationing in Rural Côte d'Ivoire*
No. 36	*Labor Market Activity in Côte d'Ivoire and Peru*
No. 37	*Health Care Financing and the Demand for Medical Care*
No. 38	*Wage Determinants and School Attainment among Men in Peru*
No. 39	*The Allocation of Goods within the Household: Adults, Children, and Gender*
No. 40	*The Effects of Household and Community Characteristics on the Nutrition of Preschool Children: Evidence from Rural Côte d'Ivoire*
No. 41	*Public-Private Sector Wage Differentials in Peru, 1985–86*
No. 42	*The Distribution of Welfare in Peru in 1985–86*
No. 43	*Profits from Self-Employment: A Case Study of Côte d'Ivoire*
No. 44	*The Living Standards Survey and Price Policy Reform: A Study of Cocoa and Coffee Production in Côte d'Ivoire*
No. 45	*Measuring the Willingness to Pay for Social Services in Developing Countries*
No. 46	*Nonagricultural Family Enterprises in Côte d'Ivoire: A Descriptive Analysis*
No. 47	*The Poor during Adjustment: A Case Study of Côte d'Ivoire*
No. 48	*Confronting Poverty in Developing Countries: Definitions, Information, and Policies*
No. 49	*Sample Designs for the Living Standards Surveys in Ghana and Mauritania/Plans de sondage pour les enquêtes sur le niveau de vie au Ghana et en Mauritanie*
No. 50	*Food Subsidies: A Case Study of Price Reform in Morocco* (also in French, 50F)
No. 51	*Child Anthropometry in Côte d'Ivoire: Estimates from Two Surveys, 1985 and 1986*
No. 52	*Public-Private Sector Wage Comparisons and Moonlighting in Developing Countries: Evidence from Côte d'Ivoire and Peru*
No. 53	*Socioeconomic Determinants of Fertility in Côte d'Ivoire*
No. 54	*The Willingness to Pay for Education in Developing Countries: Evidence from Rural Peru*

(List continues on the inside back cover)

Estimating the Determinants of Cognitive Achievement in Low-Income Countries

The Case of Ghana

The Living Standards Measurement Study

The Living Standards Measurement Study (LSMS) was established by the World Bank in 1980 to explore ways of improving the type and quality of household data collected by statistical offices in developing countries. Its goal is to foster increased use of household data as a basis for policy decisionmaking. Specifically, the LSMS is working to develop new methods to monitor progress in raising levels of living, to identify the consequences for households of past and proposed government policies, and to improve communications between survey statisticians, analysts, and policymakers.

The LSMS Working Paper series was started to disseminate intermediate products from the LSMS. Publications in the series include critical surveys covering different aspects of the LSMS data collection program and reports on improved methodologies for using Living Standards Survey (LSS) data. More recent publications recommend specific survey, questionnaire, and data processing designs, and demonstrate the breadth of policy analysis that can be carried out using LSS data.

LSMS Working Paper
Number 91

Estimating the Determinants of Cognitive Achievement in Low-Income Countries

The Case of Ghana

Paul Glewwe
Hanan Jacoby

The World Bank
Washington, D.C.

First printing November 1992

The complete backlist of publications from the World Bank is shown in the annual *Index of Publications*, which contains an alphabetical title list (with full ordering information) and indexes of subjects, authors, and countries and regions. The latest edition is available free of charge from the Distribution Unit, Office of the Publisher, Department F, The World Bank, 1818 H Street, N.W., Washington, D.C. 20433, U.S.A., or from Publications, The World Bank, 66, avenue d'Iéna, 75116 Paris, France.

ISSN: 0253-4517

Paul Glewwe is an economist in the Poverty Analysis and Policy Division of the Population and Human Resources Department of the World Bank. Hanan Jacoby is Assistant Professor in the Department of Economics at the University of Rochester.

Library of Congress Cataloging-in-Publication Data

Glewwe, Paul, 1958–
Estimating the determinants of cognitive achievement in low-income countries : the case of Ghana / Paul Glewwe, Hanan Jacoby.
p. cm. — (LSMS working paper, ISSN 0253-4517 ; no. 91)
Includes bibliographical references (p.).
ISBN 0-8213-2271-0
1. Middle schools—Ghana—Case studies. 2. Academic achievement—Economic aspects—Ghana—Case studies. 3. Education—Economic aspects—Ghana—Case studies. 4. School, Choice of—Ghana—Case studies. 5. School attendance—Ghana—case studies. I. Jacoby, Hanan. II. Title. III. Series.
LB1623.53.G4G55 1992
373.2'36'09667—DC20 92-33539
CIP

ABSTRACT

The objective of this paper is to assess the determinants of student achievement in middle schools in Ghana, with special attention given to school characteristics. We present a model of human capital accumulation which includes decisions on how long to attend school, which school to attend, and how much human capital to accumulate. This provides a framework for controlling for selectivity into middle schools, which is often ignored in the human capital production function literature. We also explicitly account for the fact that many children attend school only sporadically, which reduces their cognitive achievements but, according to our model, is a rational response among credit constrained households. We then estimate, for the cohort of children aged 12 to 18, the probability that they are in middle school, their choice of which middle school to attend, and the determinants of achievement in reading and mathematics skills in Ghana's middle schools. In addition to specific findings regarding which school characteristics contribute to such achievement, we find some evidence indicating that households in Ghana suffer from credit constraints and fairly strong evidence that sample selectivity is taking place and hence may distort estimates that do not account for it.

ACKNOWLEDGEMENTS

We wish to thank Tracy Jones and Nathalie Leboucher for excellent research assistance, and we have benefitted from the comments of Jere Behrman, Margaret Grosh, Rick Hanushek, Manny Jimenez, Marlaine Lockheed, Peter Moock, Walter Oi, George Psacharopoulos, and Wim Vijverberg.

FOREWORD

Education is widely regarded as a key to economic development. Because government is the main provider of education in low income countries, policy-makers need to know which educational investments are most effective in boosting student achievement. This paper develops a methodology for estimating the determinants of student achievement that takes into account parents' choices about their children's schooling, and applies it to data on middle school students from Ghana.

This paper is part of a broader program of research in the Population and Human Resources (PHR) Department on the extent of poverty in developing countries and on policies to reduce poverty. This research program is located in the Poverty Analysis and Policy Division. The data used here are from the Ghana Living Standards Survey, which is one of the Living Standards Measurement Study (LSMS) household surveys which the World Bank has implemented in many developing countries. Aside from the findings in this study for Ghana, one of the other objectives of this and similar work using LSMS data is to demonstrate the need for and usefulness of household data collection efforts in other developing countries.

Ann O. Hamilton
Director
Population and Human Resources Department

TABLE OF CONTENTS

LIST OF TABLES

I. Introduction

Education is widely regarded as a key to economic development. Because the government is the main provider of education in low income countries, policy-makers need to know which educational investments are most effective in boosting student achievement; for example, whether at the margin it is more cost effective to provide textbooks, improve teacher training, or repair deteriorated school buildings. Many empirical studies in both developed and developing countries purport to estimate the effect of schooling investments on student achievement (cf. Hanushek, 1986). These studies usually estimate production functions, essentially ad hoc relationships between measures of student achievement and school characteristics, with only tenuous links to economic theory. They overlook the fact that parents make choices about their children's education, deciding which school to send them to, how many years they will attend, and even how regularly they attend. These choices are presumably made to maximize some objective function. Failing to account for optimizing behavior could lead to false inferences about the relative effectiveness of different government educational investments.

In lower income countries, these biases may be large because the economic barriers to schooling are great. In particular, poor parents are unlikely to be able to borrow against their children's future earnings to finance potentially profitable human capital investments. Indeed, children living in poverty throughout the developing world often start primary school well after age six, drop out of school at young ages, and attend classes sporadically. This paper presents a model of educational decisions in poor households that incorporates these various margins of schooling choice, some of which have received little or no previous theoretical attention. The model is then applied to the problem of estimating the determinants of student achievement, as measured by test scores, and some of its implications are tested along the way.

To illustrate the possible biases inherent in the standard approach to this estimation, consider an empirical specification of a typical ad hoc educational production function, F,

$$(1.1)\quad h_i = F(SC_i,\ FB_i,\ A_i,\ Grade_i, h_{0i}) + e_i$$

where h_i is a measure of the student's cognitive skills (e.g. achievement test score), SC_i is a vector of school and teacher characteristics, FB_i is a vector of family background variables, A_i is a measure of "innate ability" (e.g. an IQ test), $Grade_i$ is years of school attendance, h_{0i} is a measure of pre-school cognitive skills, and e_i is a random error representing unobserved determinants of achievement, such as motivation, which are (often implicitly) assumed to be uncorrelated with the other variables.

Suppose (1.1) is estimated by ordinary least squares (OLS) on a sample of, say, middle school students. If parents choose among schools with different characteristics, and this choice is based partly on unobserved child motivation, then SC_i and e_i in (1.1) will be correlated, biasing the OLS parameter estimates. A positive association between school quality and student achievement may simply be due to the fact that more motivated students go to better schools. This problem could be severe in developing countries, where, unlike in the United States, parents often face no legal restrictions on their choice of public schools. The hypothesis that school characteristics are subject to choice (and thus endogenous) has been virtually ignored in the educational production function literature.[1]

If a sample of students in a particular grade or set of grades is unrepresentative of the population of children of the associated age range, an additional bias in the estimation of (1.1) could arise. For example, the OLS estimates of the SC coefficients in (1.1) would not describe the marginal ("treatment") effect on student achievement of increasing school quality because such an increase could draw less motivated students into school. Again, this selection bias could be substantial in developing countries, where early school dropouts and late starters are very common.

A final issue is that parents may not send their children to school full time. If school attendance is sporadic, years of school attendance, $Grade_i$, will overestimate the amount of child time devoted to human capital formation. The "intensity" of schooling is an omitted variable in (1.1), one that could well be correlated with included variables. Since actual (cumulative) school attendance is never observed in practice, this omitted variable problem would appear intractable.

Longitudinal data may improve matters, but it is no panacea. With a panel of students, fixed individual effects can be removed from (1.1), thereby controlling for the sorting of high ability children into better schools. However, as a practical matter, a sufficiently large panel data set is expensive, particularly if collected in conjunction with a survey of schools. Secondly, as has just been observed, early dropping out of school and hence panel attrition is clearly nonrandom. Thirdly, as shown in the next section, omitted schooling intensity is not an individual fixed effect, but varies according to number of years in school.

---

[1] Two exceptions are Jimenez et al, (1988) and (1991), who consider the choice between public and private schools.

The theoretical model in section II provides a way around the unobserved school attendance problem, as well as an overall framework for estimating the determinants of student achievement. This framework is outlined in section III. Our model is estimated, and its implications tested, using household and school level survey data from Ghana. The data are discussed in section IV, along with some details of the estimation. The empirical results appear in Section V, and Section VI concludes the paper and summarizes the findings.

II. A Theoretical Framework

This section presents a unified theoretical treatment of parental schooling decisions in poor households. Our approach is to model these choices as a two-stage optimization problem. In the first stage parents choose for each child, among the schools available, the best school as determined by school quality and attendance costs (school fees and travel time). In the second stage, conditional on this choice of school, parents decide on the timing of human capital investment; that is, when to enroll the child in school, how long to send the child to school, and how intensively (in terms of daily attendance) to educate each child.[2] The bulk of this section is devoted to this second stage problem, emphasizing the importance of assumptions about the nature of credit markets. The first stage problem is examined briefly at the end. An appendix to this paper provides full mathematical details.

The Optimal Timing of Human Capital Investment

Parents are assumed to exercise full control over all family income while they are living and to make all schooling decisions, yet they are assumed altruistic in the sense that child consumption appears as an argument in their

---

[2]If no new information were obtained while the child attended school, the two stages could be worked out simultaneously at the time of school selection. More realistically, parents first choose the school (using "best estimates" of optimal years of schooling for each schooling option) and update their initial estimate of optimal years of schooling over time for the school chosen.

preference function. The parents' instantaneous concave utility, U, is further assumed to depend only on the sum of parental and child consumption, C(t). Parents die at time T, at which point the human capital of their children is assumed to be of no value to them.[3]

Human capital, H(t), is interpreted as a homogeneous stock of embodied knowledge (see, e.g., Ben-Porath 1967), and corresponds to achievement test scores in the empirical work. Children have one unit of time to allocate between schooling and other productive activities. Schooling time consists of classroom time, S(t), and travel time, dS(t), which is assumed proportional to classroom time. Total school time is thus $(1+d)S(t) \in [0,1+d]$. Classroom time is assumed to be the only input into human capital production, which is characterized by a constant returns to scale technology:

$$(2.1) \quad \dot{H}(t) = bH(t)S(t)$$

where b is a productivity parameter, partly reflecting the student's innate ability, but also related to the quality of the school chosen.

The benefit of schooling is the increment to human capital as specified in (2.1). The costs of schooling consist of enrollment and other fees and the opportunity cost of time devoted to schooling, including travel time. Fees can be modeled by subtracting a fixed amount, f, from household income while the child attends school. Such tuition-like fees are more realistic than

---

[3]To relax this perhaps unpalatable assumption, a "salvage" function that attaches utility to the human capital of the child at time T can be appended to the overall parental objective function. However, the cost of this modification is enormous analytical complexity. Financial bequests may be incorporated in a similar fashion without much difficulty, but they are ignored here (i.e., the utility of the terminal stock of assets is set to zero) without changing the basic thrust of the analysis.

daily attendance fees. Children rent their stock of human capital at a constant rental price w; i.e., they work at a wage rate $wH(t)$ for D hours per day and contribute $wH(t)\{D-S(t)(1+d)\}$ to current family income. Parents also contribute an amount $y(t)$ to family income at each point in time. Although $y(t)$ may well rise or fall over time, we assume it is constant $(y(t)=y)$ to avoid needless complication.[4]

In the initial phase, from time 0 to t_0, which starts when the child is at the minimum age of primary school entry, parents keep the child at home. Household income is just $y + wDH_0$, where H_0 is the child's initial stock of human capital; the assumption being that young children (say, age six and up) contribute to family income in a meaningful way, perhaps by doing chores or tending livestock. We explicitly model this phase because late entry into primary school is such a pervasive phenomenon in many low income countries (including Ghana). If the household can borrow and lend freely at the market interest rate, r, then during this initial period its net asset stock, $A(t)$, will evolve according to

$$(2.2) \quad \dot{A}(t) = rA(t) + y + wDH_0 - C(t) \qquad 0 \le t \le t_0$$

In the second phase, $[t_0,t_1]$ the child is enrolled in school and parents pay the school fee f, so that

$$(2.3) \quad \dot{A}(t) = rA(t) + y - f + wH(t)[D - S(t)(1 + d)] - C(t)$$

---

[4] Jacoby (1991) considers nonconstant parental income. Endogenous parental labor supply is a trivial extension of the model as long as utility is assumed additively separable in consumption and parents' leisure.

In the final phase, $[t_1,T]$, the child has quit school and is working full-time at wage rate $wDH(t_1)$, i.e.,

$$(2.4) \quad \dot{A}(t) = rA(t) + y + wDH(t_1) - C(t).$$

We ignore on-the-job training.

Case of unconstrained borrowing

As a benchmark, consider the case where there are no constraints on parents' ability to borrow against the future earnings of the child. Parents choose t_0, t_1, $C(t)$ and $S(t)$ to maximize discounted lifetime utility,

$$(2.5) \quad \int_0^T U(C(t))\, e^{-\delta t} dt$$

subject to (2.1)-(2.4), the time constraint $0 \leq S(t) \leq 1$, and the boundary conditions $H(0) = H_0$, $A(0) = A_0 = 0$, and $A(T) = 0$. The last condition is implied by nonsatiation and the assumption that bequests have no value. The rate of time preference is denoted by δ, and it is assumed that $b > \max(r,\delta)$; otherwise, human capital investment will never be undertaken.[5]

The model is easy to solve when borrowing is allowed, since there is a separation between consumption and human capital investment decisions. We show in the Appendix that delays are never optimal; if schooling is a worthwhile investment, it pays to begin it immediately. As it stands, the model also implies that children always attend school full-time, which is

[5]This condition is necessary. It is also sufficient if $f = 0$.

clearly at odds with the reality of most low income countries.

Case of Constrained Borrowing

To obtain schooling patterns more in line with those observed in the data, we now impose a borrowing constraint of the form $A(t) \geq 0$.[6] Not all households are necessarily affected by such a borrowing constraint, since some parents may be willing to save even while sending their children to school. To focus the exposition, we examine the case of an "impatient" household, i.e., one with $\delta > r$, that starts off with zero assets ($A_0 = 0$).[7]

Figure 1 illustrates the generic solution to the problem with borrowing constraints (see Appendix 1 for mathematical details). Consider first what happens during the three phases represented by constraints (2.2)-(2.4):

$[0,t_0]$: The household has a constant income stream, $y + wDH_0$, but anticipates a fall in income at t_0 to $y - f + wH_0[D - S(t_0)(1 + d)]$. Therefore, at least during the latter part of this interval, parents save, allowing consumption to fall smoothly, rather than abruptly at t_0.

$[t_0,t_1]$: The child is enrolled in school. Initially, the household can dip into its savings to finance consumption, so the child is sent to school full-time. Even after savings from the previous phase have been exhausted, children may still attend full-time for a while, from t_0 to t_s, if parental income net of the school fee is sufficiently high. But eventually parents begin gradually withdrawing their child from school and putting him to work.

---

[6]A similar analysis applies if assets are constrained to lie above a negative lower bound.

[7]The case of the patient household ($r < \delta$) is more complicated, but it can also lead to part-time school attendance, especially if parental income is low (see Jacoby 1991).

The absence of credit, in linking household consumption with the child earnings profile, raises the cost of consuming today versus tomorrow, making parents appear more patient.

$[t_1, T]$: Income is constant at $y + wDH(t_1)$ and consumption simply follows income because the household is assumed impatient.

Turn now to the choice of school starting and quitting times. Unlike the case of unconstrained borrowing, here delays may be optimal. It may be worthwhile delaying entry into the labor force so as to postpone paying the school fee. However, the larger parental income is relative to the school fee, the less advantageous a delay would seem to be, and the lower t_0^* (see appendix). Similarly, the choice of t_1 depends on parental income. Notice from Figure 1 that both S(t) and C(t) have a discontinuity at t_1. The reason is the fixed school fee. Although parents want to avoid jumps in consumption, low school attendance does not make sense in the face of a substantial enrollment fee. In selecting t_1, parents balance the marginal savings of school fees against the higher future consumption that schooling yields.[8]

Based on our discussion thus far, we can write the optimal switching time functions as follows

$$(2.6) \quad t_i^* = \tau_i(y, f, wH_0, D, b, d, r, \delta, T) \qquad \text{for } i = 0, 1.$$

---

[8]So far we have assumed that credit constrained households start out with no assets ($A_0 = 0$). Of course, the type of assets (e.g., land, livestock or cash) will determine whether, or how fast, they can be liquidated over time, or whether they can be used as collateral for limited borrowing. In either case, greater asset holdings will tend to mitigate the delay in school enrollment and allow parents to send their child to school on a full-time basis for a longer period; however, the precise effect is difficult to determine analytically.

Our model does not yield closed form solutions for t_0^* and t_1^*, even in very restrictive special cases. An exception, however, is when $f = 0$, in which case $t_0^* = 0$ and $t_1^* = T + \frac{1}{\delta}\ln(1 - \delta/b')$. The implication is that even if borrowing is restricted, as long as school fees are negligible, neither the school starting time nor quitting time depends on parental income. In fact, there is no incentive at all to postpone entry into school.

We now characterize the optimal stock of human capital for children enrolled in school as a function of potentially observable variables. In general, this function has the form

$$(2.7)\qquad \begin{aligned} \tilde{H}(t) &= H_0 e^{b'(t-t_0)} && \text{for } t_0 \le t \le t_s \\ &= g(y,f,w,H_0,D_0,b,d,r,\delta,T,t) && \text{for } t_s \le t \le t_1. \end{aligned}$$

where t_s is the point at which the optimal rate of school attendance, $S^*(t)$, falls below unity (see Figure 1). The second line of (2.8) reflects the fact that $S^*(t)$ is a function of the same variables that determine t_1^* and t_0^*. Note that the former variable has been optimized out of the "demand" function, $\tilde{H}(t)$, which depends only on exogenous variables.

As with the τ_i functions, it is not possible to sign the partial derivatives of g in full generality.[9] The main point is that, unlike the case of unconstrained borrowing where $t_s = t_1$, with a borrowing constraint the optimal stock of human capital depends on parental income, school fees, and the rental price of human capital.

---

[9]However, for the case of logarithmic utility and small f, the intuitively plausible signs are obtained; $\tilde{H}(t)$ is positively related to y, b and H_0 and negatively related to w and d.

Multi-child Households

So far we have focused on a single child household, saying nothing about how the timing of human capital investment of a given child is affected by the presence of siblings. Analyzing multi-child households is complicated (see Jacoby, 1991), but one point is clear. If parents can freely borrow against the future earnings of their children, then the timing of children's human capital investment would be independent of the presence of siblings and their age structure. In general, however, the desire to smooth consumption coupled with borrowing constraints implies that at any point time, the school attendance of a given child depends upon the attendance of his or her siblings, and vice versa. Fixed school fees, for example, would lead to a "staggering" effect; i.e., children close in age would not be enrolled in school together. At the very least, then, family demographics--e.g., the number of siblings and their age structure--should appear in the τ_i and g functions discussed above.

The Optimal Choice of School

Turn now to the first stage of the parental decision problem, choosing which school to attend. Given (2.7) and (2.8), lifetime indirect utility conditional on schooling choice can be written as

$$(2.8) \quad V = V(y,f,w,H_0,D,b,d,r,\delta,T).$$

Schools differ according to the quality of instruction (which affects learning efficiency, b), travel time or distance (d), fees (f), and unobserved characteristics. Each school i under consideration offers lifetime utility $V_i = V(y,w,H_0,D,r,\delta,T,b_i,d_i,f_i) + \varepsilon_i$, where ε_i is assumed to be an additive random component. Thus, first stage optimization amounts to choosing the school among the N available that maximizes utility (V).

III. Empirical Implementation

In the introduction we pointed out how estimates of the effects of schooling investments on student achievement based on (1.1) may be misleading in low income countries because: 1) school quality is subject to choice; 2) children in school are not a random sample of the total child population; and 3) the intensity of school attendance is unobserved and varies in the population. Our model of Section II provides a framework for estimating a human capital demand function that is consistent with optimizing behavior, as well as allowing a test of the effects of borrowing constraints on the timing of human capital investments. In this section we specify the empirical counterparts to the human capital demand function, (2.8), to the school starting and quitting time functions, (2.7), and to the school selection model just delineated.

The Human Capital Demand Function

The human capital demand function (2.8) bears some resemblance to the ad hoc education production function given by (1.1). Both contain school characteristics, such as teacher competence and textbook availability; in (2.8) these enter through the learning efficiency parameter, which is school specific. Students' innate ability also enters (2.8) through b, as does parents' education and innate ability, to the extent that parents help and encourage child learning. Years in school also appears in both specifications, but in (2.8) it is interpreted as elapsed time.[10] As such, grades completed is an exogenous variable for children still attending school. Finally, the initial stock of human capital, H_0, appears in (2.8), as it has in some education production function studies in the form of pre-school test scores.

A key difference between (1.1) and (2.8) is in the treatment and interpretation of economic variables. In the standard approach, indicators of the household's economic status are often included in the specification to proxy the unobserved component of family background. In our model, parental income, assets, school fees, and even child wages play no role in the human capital stock function of children attending school, as long as the household is able to borrow freely against future child earnings.[11] On the other hand, these variables do enter (2.8) when the household cannot borrow and has reached the part-time schooling phase. At the very least, our model places a different interpretation on the economic variables; but, given adequate controls for parental ability and education, we argue that <u>most</u> of the residual correlation between economic variables and student achievement, if any exists, can be attributed to economic constraints on human capital investment, not to unobserved components of the home environment.

Empirical implementation of the model requires data on parental income, a

---

[10]For school systems in which barriers to grade advancement exist and grade repetition is prominent, grades completed would no longer represent elapsed time. As will be seen in the empirical section of this paper, grade repetition is minimal in the primary and middle schools of Ghana.

[11]This is because parental utility in (2.5) depends only on consumption. If parents valued the education of their children for its own sake, higher income could lead to greater demand for education even in the absence of credit constraints. We return to this point in the empirical estimates of this paper.

variable that can be extremely difficult to measure in developing countries (such as in Ghana) because self-employment is pervasive. Our approach is to replace parental income with total household consumption expenditures, which is relatively easy to measure. According to the model, total consumption follows total household income throughout the period in which the child is enrolled in school. Consumption may thus be expressed as follows

$$(3.1)\quad C(t) = \begin{cases} y - f & \text{for } t_0 \leq t \leq t_s \\ \phi(y, f, w, H_0, D, b, d, r, \delta, T, t) & \text{for } t_s \leq t \leq t_1 \end{cases}$$

Solving (3.1) for y and substituting into (2.8) yields

$$(3.2)\quad \tilde{H}(t) = G(C(t), f, w, H_0, D, b, d, r, \delta, T, t) \quad \text{for } t_s \leq t \leq t_1$$

The same strategy is used for the switching time functions τ, so that C(t) replaces y in (2.7). The joint determination of household consumption and human capital investment implies that C(t) will be statistically endogenous in the regressions based on these three relations, a problem that will be dealt with in the next section.

Sample Selectivity

The human capital demand function derived above holds only for children still attending school.[12] But, as was argued in the introduction, a sample of

---

[12]In particular, years in school is no longer exogenous for children who have quit school. Moreover, there is likely to be a looser connection between currently observed economic variables (e.g., household consumption) and school achievement for these children.

students is certainly not a random sample of all children in a particular age range; and, as seen below, this is especially true of middle school students in Ghana. Suppose middle school consists of grades seven through ten. There are four reasons a child in a given age range would not be attending middle school: (1) they never started school in the first place, (2) they entered school late, and have not yet reached grade seven, (3) they have quit school, and (4) they have gone beyond grade ten already. In terms of the theoretical model, (1) implies $t > t_1^* = t_0^* = 0$, (2) means $t < t_1^*$ and $(t - t_0^*) < 7$, (3) means $t > t_1^*$, and (4) indicates that $(t_0^* + 10) < t < t_1^*$.

By estimating the functions τ_0 and τ_1, it is possible to calculate the probabilities of being in each of these categories, thus completely characterizing the sample selection mechanism. We utilize a pair of ordered probit models adapted from Lillard and King (1984), one for school attainment $(t_1^* - t_0^*)$ and the other for the number of years late in starting school (t_0^*). These estimates are used to construct selectivity correction terms analogous to those in Heckman (1979), which are then inserted into the human capital demand function. Notice, that the selection criteria, based on (2.7), contain no variable that is not also in (3.2), which suggests an identification problem. However, the next section explains how identification might be achieved through exclusion restrictions.

School Selection

Another kind of selectivity arises because schools are chosen by parents based on a comparison of conditional indirect utilities. This choice is determined by differences in schools, both in costs and in quality. Standard discrete choice models can be used here as well, though now multinomial models must be considered because the number of schooling options may be greater than 2. Conditional means, constructed from multinomial logit and probit models, can be inserted into the human capital regressions (2.8) to control for the sorting of children into schools (see Maddala, 1983). However, if there is correlation between the school selection rule and the sample selectivity rules described above, matters become more complicated. Tests for such correlation must be performed, and if the null hypothesis of no correlation is rejected, appropriate measures must be employed (see Maddala, 1983, p.282).

Our theoretical model in section II says that all three of these selectivity rules emanate from the same underlying utility maximization problem. Unfortunately, the full set of restrictions imposed by the theory across the selection rules and the human capital demand equation are difficult, if not impossible, to test or impose. Joint estimation of all the relations is also infeasible. But a two-step procedure, where sample selectivity regressions are estimated first, followed by estimates of the human capital demand function, is relatively straightforward.

IV. Context, Data, and Estimation

Education in Ghana

Since the mid 1980's Ghana has begun to institute widespread educational reforms, many aimed at raising the quality of education in primary and middle schools.[13] Yet school quality remains appallingly low in rural areas. A recent World Bank Report (1989, p.18) points out that "in many of the remote areas, especially in the northern half of the country, the large majority (often more than 80%) of children completing grade 6...were completely illiterate." During Ghana's economic decline, which reached bottom in 1983-84, funds for basic materials such as textbooks dried up, and only recently have they reappeared in many Ghanaian schools. Teacher quality also declined substantially: "Due to the economic decline of the 1970's, large numbers of trained and highly qualified teachers left the country and were replaced by untrained teachers in primary and middle schools. The result has been over a decade of ineffective basic education..." (p.27).

Almost every child (93%) attends government (public) schools at the primary level (grades 1-6). The next level of schooling is middle schools, which encompass grades 7-10.[14] Most middle school students (94%) also attend government schools. After middle school children enter the secondary stream,

---

[13] The World Bank lent Ghana $35 million in the form of an education sector adjustment credit in 1987 and $50 million more in 1990.

[14] Ghana has now converted all its middle schools to junior secondary schools. At the time the data used here were collected, the first two grades of middle school (7 and 8) had been transformed into junior secondary grades (the main practical effect is that the curriculum was changed) while the last two grades of middle school (9 and 10) were still in place.

which consists of five years up to the time when the GCE O Level exams are taken and another two years after which the GCE A level exams are taken. Entry into the secondary stream is determined by performance on the Common Entrance Examination (CEE), which most students do not pass. Thus children proceed through the first 10 years of education without encountering any formal obstacle preventing them from advancing to the next grade.[15] The model described in the Section II is best suited for the first 10 years, since it does not incorporate screening tests to advance to higher levels of education.

Data and Sample Selection

The data used in this paper are from the Ghana Living Standards Survey (GLSS), a comprehensive household survey implemented by the Ghana Statistical Service (GSS) on random samples of households throughout Ghana. Details of the first year of the survey (1987-88) can be found in Scott and Amenuvegbe (1989), and Glewwe and Twum-Baah (1990). One half (1586) of the households interviewed in the second year (1988-89) participated in an exercise in which persons from age 9 to 55 were given written tests covering reading (in English), mathematics and abstract thinking.[16] In addition, information was collected on local schools, whether attended by household members or not.

Table 1 provides mean scores on both very simple (8 questions each) mathematics and reading tests as well as more difficult tests (29 questions

---

[15] Any child who has completed primary school could take the CEE; passing the exam allows the child to go immediately to the secondary stream (i.e., bypass middle school). This rarely happens, but even so, our sample selectivity techniques account for this by treating such children as "early starters".

[16] All individuals took the last test, but only those with three or more years of schooling took the mathematics and reading tests.

Table 1: Mean Test Scores - All Ghana

GRADE	Simple Math		Simple Reading		Long Math		Long Reading	
Primary								
3	2.57	(162)	0.16	(162)	0.87	(162)	0.18	(162)
4	3.61	(160)	1.02	(160)	1.73	(155)	1.00	(158)
5	4.12	(121)	1.12	(121)	2.18	(118)	1.07	(121)
6	4.61	(109)	2.00	(109)	4.06	(105)	2.72	(108)
Middle/JSS								
7	4.80	(109)	3.09	(109)	3.89	(104)	3.98	(105)
8	4.84	(114)	3.56	(114)	4.39	(110)	4.21	(112)
9	5.61	(83)	4.65	(83)	6.55	(80)	6.54	(82)
10	5.95	(152)	5.66	(152)	8.29	(144)	9.86	(146)
Secondary								
1	6.52	(23)	6.17	(23)	9.71	(21)	12.19	(21)
2	6.95	(19)	7.32	(19)	12.53	(17)	17.31	(16)
3	6.67	(15)	7.40	(15)	12.07	(15)	19.27	(15)
4	6.85	(13)	7.46	(13)	13.58	(12)	20.00	(12)
5	7.20	(30)	7.53	(30)	15.96	(23)	21.48	(23)
A-Level								
1	7.75	(4)	7.50	(4)	22.25	(4)	24.50	(4)
2	7.56	(9)	7.89	(9)	20.63	(8)	23.38	(8)

Notes: 1. Sample sizes of means are in parentheses.

2. Persons who scored 4 or less on a simple test were given a score of zero for the respective long test.

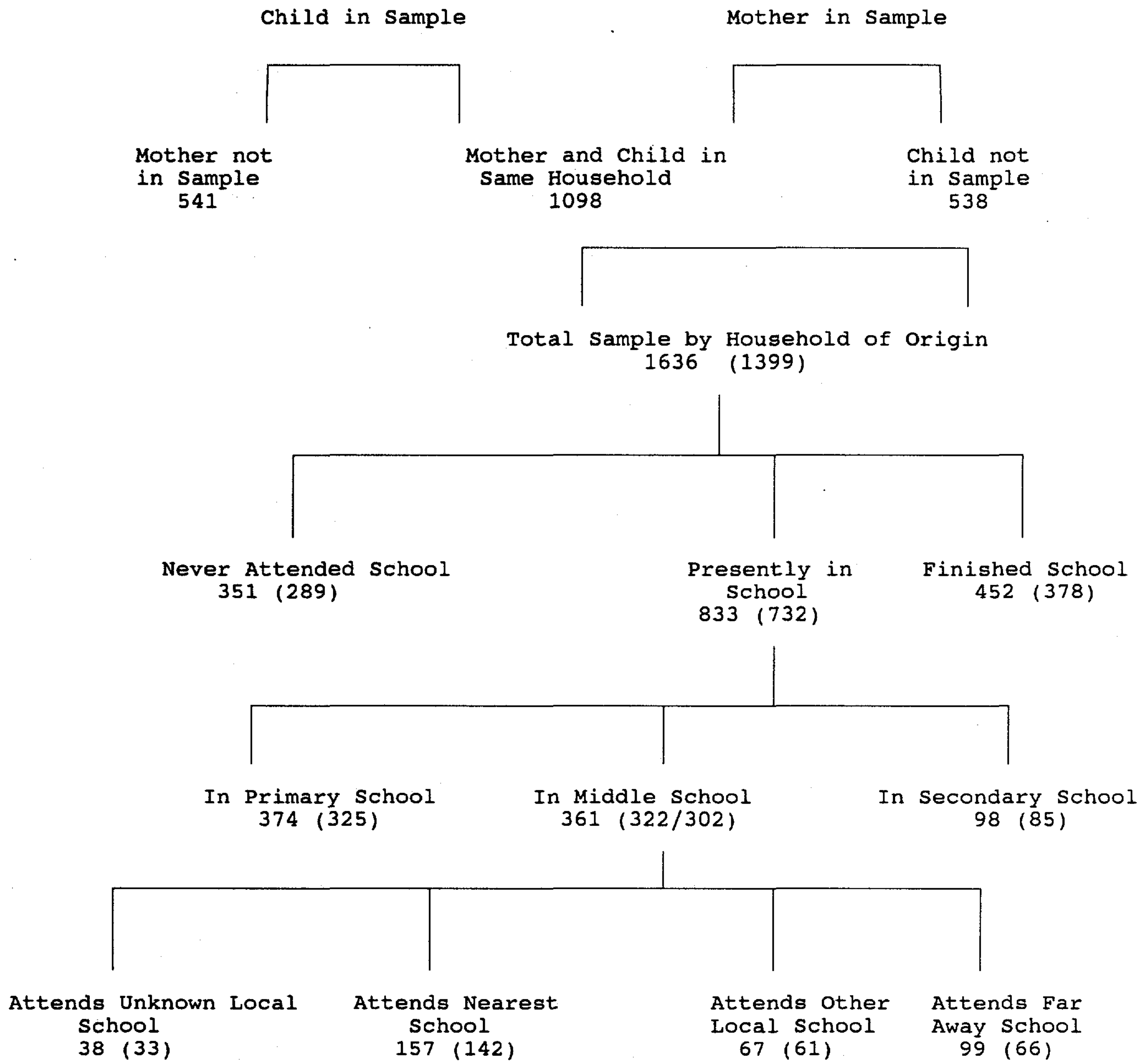
for reading and 36 questions for mathematics). The mean scores on the simple reading test are particularly shocking. Even after 6 years of school the mean score was 2 out of 8, which is exactly what one would expect if the students guessed randomly on the test (the simple reading test was multiple choice, with four possible responses for each item).[17] Obviously, many students simply gave up and did not complete the test. Middle school scores are still low, but there is more variation to analyze.

This paper focuses on those children who have completed primary school (grades 1-6) and have thereby made decisions regarding middle school (grades 7-10). We take as our initial sample all children aged 11 to 20, which includes virtually all children presently attending middle school. As seen in Table 2, there are 1639 such students in the 1586 households, yet it is worth noting that 541 of these are living in households away from their mothers; in Ghana, as in many other West African countries, it is common for older children to live away from their parents (cf. Ainsworth, 1990). Many of these children have left their parents in order to attend school, and are living with relatives or family friends. We assume that it is the parents who make schooling decisions for their children, so we require data on children's households of *origin* in the empirical analysis (in cases where the mother and father do not live in the same household we choose the mother's household as the household of origin). Hence, the 541 children living away from their mothers are dropped from the sample, since the GLSS data tell us very little about their households of origin. Of course, the remaining 1098 children who do live in their household of origin are not a random sample of the population

---

[17] Although in the earlier primary grades native African languages are promoted, English language instruction is introduced by the 4th grade and receives a great deal of emphasis thereafter.

Table 2: Sample of Children 11-20, Ghana Living Standards Survey



Notes: 1. Numbers in parentheses indicate those children with complete data. For example, of the sample of 1636 children by household of origin, we lose 41 from 2 clusters for which school data are missing and one cluster that has no middle school, 147 due to father's education missing and 47 due to other missing variables.

2. Of the 361 children presently in middle school, 322 have adequate data for estimation in the 2 ordered probit models (which require only cluster means of school characteristics) and 302 have adequate data for estimates of the schooling choice model.

of 11 to 20 year olds in Ghana. To them we add 538 children aged 11-20 who have moved away from their parents' households; the GLSS contains information on these children, including whether or not they are presently enrolled in school. Thus we have a random sample based on the household of origin.

Of the 1639 children in our sample, 833 are currently enrolled in school and 803 are not. Of those not in school, 452 have finished schooling and 351 have had no schooling at all (i.e., $t_1^* = t_0^* = 0$). Of those in school, 374 are still in primary school, 361 are in middle school and 98 are in secondary school or beyond. Among the middle school students are 157 who attend the nearest school, 67 who attend other local schools and 99 attending non-local schools (living away from home). Finally, 38 go to a local school but we do not know which one. The numbers in parentheses in Table 2 indicate those children for whom the data are complete.

Specification of the School Attainment and Late Starting Regressions

To control for selectivity into our sample of children aged 11-20 currently attending middle school we estimate the determinants of years of schooling completed ($= t_1^* - t_0^*$), as well as the determinants of the number of years after the normal age of primary school entry the child began school (t_0^*). Years late are thus calculated by subtracting years completed plus six from the child's age. Grade attainment can be estimated for all children in the sample, while years late can only be estimated for those who have attended school at some point (since late starting is undefined for persons who never attend school). Assuming that the unobserved components in both decision rules are normally distributed, the models can be estimated using ordered probits. We also estimate the two ordered probits jointly to test for correlation across decision rules.

A key point for identifying selectivity correction terms in the cognitive skills regressions below is the way school characteristics are specified in the discrete choice models. In the ordered probits, we include the mean characteristics of all local schools, as opposed to the characteristics of the schools actually attended because: 1) specific school characteristics are not known for children who have never attended school and for many of the children who have finished their schooling; and 2) we have argued that these specific school characteristics are themselves objects of choice, whereas the means of characteristics are not subject to this criticism.[18] Mean characteristics of local primary schools are also used (and thus also serve as identifying variables) in both probits, since school quality and attendance costs over the whole life-cycle influence schooling decisions.

Our choice of explanatory variables is guided by (2.7). Recall that the learning efficiency parameter b depends in part on the characteristics of the relevant schools. We have detailed middle and primary school questionnaires that provide a potentially vast array of school characteristics. There are so many that we must reduce the list of regressors considerably. Appealing to a World Bank review of evidence on the effect of school characteristics on student achievement (Fuller, 1986), we focus on the availability of textbooks and other instructional materials as well as teacher education and training.[19] The GLSS also records data on the travel time to each school and the school

[18]The fact that children move away from home to attend non-local schools is not a problem since we examine local school variables of the household of origin. The GLSS migration data indicate that it is rare for entire households to move for the purpose of securing a better school for the children.

[19]We divide the number of textbooks in the school by the number of rooms rather than, as is commonly done, by the number of students, because we essentially treat the number of students in a school as endogenous in this paper.

fee (d and f, respectively, in (2.7)).

Now turn to the individual and household level variables (all variables used in the regressions are defined in Table 3). Age and sex of the child are included along with sets of regional, religious and ethnic dummy variables. Guided by the theoretical model, we use the logarithm of per-capita consumption expenditures (as calculated by Glewwe and Twum-Baah, 1990) to control for family income. But first we regress log expenditures on a set of instruments alluded to earlier, and form the residuals.[20] Using the Two-Stage Conditional Maximum Likelihood argument of Rivers and Vuong (1988), these estimated residuals are included along with actual expenditures as regressors in all our discrete choice models to control for the endogeneity of consumption.

Also, in keeping with the theoretical discussion, the village level male agricultural wage is included to capture variation in the rental price of human capital.[21] Landholdings, which are predominantly inherited, are used to proxy initial asset holdings. Number of siblings and the birth order of the child are also included, for reasons indicated in the previous section.

Finally, although some researchers (e.g., Boissiere, Knight and Sabot, 1985) have used scores on the Raven's (abstract thinking) test as indicators of innate ability, a child's score on the test was never intended to be interpreted as being independent of his or her educational background. However, one can estimate an indicator of innate ability using the Raven's

---

[20] Instruments include mean village per capita expenditures, the schooling, sex and age of the head of household, household size and other household composition variables, and the value of livestock and of capital of household enterprises.

[21] The wage of unskilled agricultural labor can be thought of as the intercept term in a standard wage equation.

Table 3: Variable Definitions

COAST, FOREST, SAVANNAH - Dummy variables indicating residence of mother's household. Accra, the capital, is the omitted variable.

SEMI-URBAN, RURAL - Dummy variables indicating degree of urbanization of mother's household. Rural areas are communities with under 1500 residents and semi-urban communities have 1500-4999 residents. The omitted variables are urban communities (more than 5000 residents).

SEX - Dummy variable indicating female individual.

AGE - Age of individual in completed years.

ABILITY - Fixed effect estimate of innate ability based on parent's and child's scores on Raven's Progressive Coloured Matrices test.

MOTHER'S/FATHER'S YEARS SCHOOLING - Number of years of schooling attained by each of the child's parents.

ABILITY MISSING - Dummy variable indicating no fixed effect ability estimate available for child. In such cases ABILITY is set equal to the sample mean.

MUSLIM, CATHOLIC, PROTESTANT, OTHER CHRISTIAN, TRADITIONAL RELIGION - Dummy variables indicating religion of head of household.

AKAN, EWE, GA-ADANGBE - Dummy variables indicating ethnic group of household.

MALE AGRICULTURAL WAGE - Daily real wage rate for male agricultural labor, in Cedis.

WAGE DATA MISSING - Dummy variable indicating that data on male agricultural wage are not available.

CHILD ORDER - Rank of child in terms of siblings on mother's side.

SIBLINGS - Number of children the child's mother has had, excluding those that died.

LOG PER CAPITA EXPENDITURE - Natural logarithm of consumption expenditures per capita of the mother's household.

EXPENDITURE RESIDUAL - Difference between actual and estimated log per capita expenditures.

LAND OWNERSHIP - Acres of agricultural land owned by the household.

TRAVEL TIME - Travel time from the mother's household to the nearest school of a given type, in minutes.

AVG. TEACHER EXPERIENCE/SCHOOLING/TRAINING - School-wide averages of teachers' years of teaching experience, general schooling and teacher training, respectively.

BLACKBOARDS - Fraction of school's classrooms that have blackboards.

LIBRARY - Dummy variable indicating that the school has a library.

Table 3 Continued:

BOOKS/ROOM - Number of textbooks divided by number of classrooms. In middle schools this is broken down into new JSS curriculum and old Middle School curriculum.

LACK OF DESKS - Indicates that the school lacks desks or tables for some children.

ENROLLMENT FEE - Annual school fee, which for middle schools is separated into JSS grades (7 and 8) and Middle School grades (9 and 10).

JSS ONLY - Indicates the school was newly organized and only has JSS grades (7 and 8).

DENY ADMISSION - Dummy variable taking the value of one if school does not admit all children who want to enroll, zero otherwise.

SHED/UNUSABLE/LEAKY CLASSROOMS - Indicates proportion of total classroom which are of very simple construction, are completely unusable, or cannot be used when it rains, respectively.

NO WATER/NO ELECTRICITY - Dummy variables indicating that school has no water or no electricity, respectively.

PRIVATE - Dummy variable indicate that the school is private.

COMPLETE SCHOOL - Dummy variable indicating that the primary school has all 6 grades.

PMATH/PREAD - Village averages of mathematics and reading scores among children now attending local primary schools.

scores by assuming that innate ability is one of several factors that jointly produce the Raven's scores and that innate ability is inherited by children from their parents. Assume that scores of parents and children depend on age, years in school and innate ability:

$$(3.1)\quad R_i = I_i + \beta_1 YS_i + \beta_2 AGE_i + \beta_3 AGE_i^2 + u_i \qquad \text{for } i=f\ m$$

$$R_c = (I_m + I_f)/2 + \beta_1 YS_c + \beta_2 AGE_c + \beta_3 AGE_c^2 + u_c + \eta_c$$

where the subscripts f, m and c indicate father, mother and child, respectively, I_i is innate ability, YS_i is years of schooling and the inheritance of innate ability is a simple average of the parents innate ability plus a random term η. The innate ability terms in (3.1) can be estimated as a family fixed effect in any household for which Raven's scores are available for both parents and at least one child.[22] These estimates of innate ability can then be used in the subsequent estimation steps.

Specification of the Schooling Choice Decision

Because schools are chosen, their characteristics cannot be treated as exogenous when estimating their impact on the attainment of cognitive skills. We adopt a switching regression model for accumulation of cognitive skills analogous to those used in estimating the determinants of wages when the type of job is an individual choice. This requires estimates of the probabilities of choosing different schools, i.e., a discrete choice model. A practical concern is how to aggregate a large number of options to get a tractable

---

[22] If only one parent has a Raven score that score alone is used.

empirical model. We aggregate the household's myriad options into three schooling choices: the school nearest to the household in terms of travel time; all other schools within walking distance of the household; and all other (non-local) schools in Ghana.[23] Means of local school quality variables (excluding the nearest school) are used in the case of the other local school option. The utility from the faraway option is normalized to zero, as its school characteristics do not vary over households. The coefficients on the individual and household level characteristics are allowed to vary across options,[24] but those on the school characteristics cannot be allowed to vary across schooling options.

In order to assess the robustness of our switching regression to distributional assumptions, we estimate two models: the multinomial logit and multinomial probit. The former assumes that the error term in equation (2.11) follows a type I extreme value distribution while the latter assumes that it follows a normal distribution. Both models can be estimated so as to allow for correlation of the residuals across schooling choices.

Estimation of Human Capital Accumulation

Our ultimate objective is to estimate the determinants of cognitive skills as measured by test scores. Denote the test score of individual i in school j by h_{ij}, the vector of individual and household characteristics by X_i,

---

[23]For about 10% of the sample no other local school option exists. Also, for 11% of the sample we know that a local schooling option has been chosen but we do not know if it is the nearest or another local school. In both cases it is straightforward to accommodate the observations in the likelihood function.

[24]Actually, one can only estimate the difference in the coefficients on the individual and household characteristics; we identify these differences by setting the coefficients on the faraway option equal to zero.

and the vector of characteristics of the school attended by Z_{ij}. Since we do not observe test scores for children who have left home to attend a far away school, we have the following switching regression model

$$(3.2)\quad h_{in} = \beta_0 + X_i'\beta_1 + Z_{in}'\beta_2 + u_{in} \qquad \text{if } j = \text{nearest school}$$

$$h_{il} = \beta_0 + X_i'\beta_1 + Z_{il}'\beta_2 + u_{il} \qquad \text{if } j = \text{other local school}$$

Notice that the coefficients on all of the regressors are constrained to be equal across choices, but that the error terms are different, each with a different (nonzero) conditional mean. We replace each conditional mean term by the appropriate inverse Mills' ratios (denote these as λ_n and λ_l, respectively), constructed from the previously estimated discrete choice models. Now, dropping the school subscript, we can write the model as

$$(3.3)\quad h_i = \beta_0 + X_i'\beta_1 + Z_i'\beta_2 + \delta_n D_i\lambda_n + \delta(1 - D_i)\lambda_l + e_i$$

where D_i is a dummy variable indicating whether the child i attends the nearest school (and equals zero if he or she attends another local school). In addition to λ_n and λ_l, two other selectivity terms are included in (3.3) to control for selection into the sample, as described in the previous section. We assume zero correlation across decision rules, although we later (partially) test this assumption.

Besides excluding from (3.3) the characteristics of middle schools that the child chose not to attend, local primary school characteristics can also be used to identify the selectivity terms. Our model suggests that primary school characteristics enter (2.8) only through H_0, the pre-middle school stock of human capital. We control for H_0 in (3.3) by using the mean test scores of local primary school students. Any residual effect of primary school characteristics on test scores is thus assumed to be an artifact of selectivity; namely, children who went to better primary schools are more likely to go on to middle school.

V. Estimation Results

Determinants of Middle School Enrollment

We first estimate ordered probits for school attainment and late starting. The separate estimates are reported in Tables 4 and 5, respectively. Both tables show two estimates, a "full model" that includes a large number of individual and school characteristics, and a "parsimonious model" that drops those having low t-statistics in the "full model".[25] Joint estimation (not reported here) revealed no significant correlation across the random components in these two decision rules. Specification tests of the distributional assumption of these two ordered probit models were mixed. The hypothesis of a normal distribution of the error term could not be rejected for the late starting estimates, while for the school attainment estimate the hypothesis could be rejected at the 95% significance level but not the 99% significance level.[26]

Focusing on the statistically significant coefficients in Table 4, children in the Northern Savannah area of Ghana, and to a lesser extent children in semi-urban areas display lower school attainment. Children from Muslim, "other Christian," and traditional religion households have lower attainment than those from Catholic and Protestant Christian households (and

---

[25] Specifically, individual level variables with t-statistics less than 1 was dropped and school variables with a t-statistics below 0.5 were dropped (school variables were treated less parsimoniously because they are of particular interest). In some cases this rule was modified for variables that are naturally paired or grouped.

[26] The χ^2 statistics (2 degrees of freedom) were 1.77 for (parsimonious) late starting and 6.84 for (parsimonious) school attainment. See Glewwe (1991a) for details of this test.

Table 4: Ordered Probit, School Attainment

Individual Variables	Full Model Coefficient	Full Model t-stat	Parsimonious Model Coefficient	Parsimonious Model t-stat	Mean	Standard Deviation
Constant	-1.3331	-0.65	-1.5518	-1.26	1.0000	--
Coast	-0.2827	-0.82	--	--	0.1951	--
Forest	-0.3195	-1.00	-0.0716	-0.60	0.5139	--
Savannah	-1.0999	-2.68	-0.8190	-4.95	0.2187	--
Semi-Urban	-0.2708	-1.59	-0.2285	-1.83	0.2244	--
Rural	-0.0796	-0.44	--	--	0.4482	--
Sex	-0.7271	-9.15	-0.7196	-9.26	0.4882	--
Age	-0.0176	-1.08	-0.0176	-1.13	15.2030	2.77
Mother's Years Schooling	0.0436	3.16	0.0455	3.44	2.5540	4.32
Father's Years Schooling	0.0560	5.47	0.0581	6.02	5.2538	5.88
Ability	0.0586	3.77	0.0582	3.89	12.5142	3.04
Ability Missing	-0.0747	-0.81	--	--	0.3059	--
Muslim	-0.5180	-1.98	-0.4912	-3.44	0.1523	--
Catholic	-0.0309	-0.12	--	--	0.1515	--
Protestant	-0.2502	-1.01	-0.2096	-1.53	0.2287	--
Other Christian	-0.3788	-1.56	-0.3430	-2.37	0.1773	--
Traditional Religion	-0.4624	-1.91	-0.4244	-3.43	0.2466	--
Akan	0.0368	0.26	--	--	0.4546	--
Ewe	0.3841	2.21	0.3273	2.54	0.1744	--
Ga-Adangbe	0.5852	1.83	0.5602	2.14	0.0415	--
Male Agricultural Wage	-0.0017	-2.00	-0.0016	-2.08	131.3324	168.80
Wage Data Missing	-1.1033	-3.29	-1.0338	-3.52	0.5904	--
Child Order	0.1088	3.27	0.1080	3.48	2.9192	1.63
Siblings	-0.0478	-1.88	-0.0440	-1.97	5.5061	2.08
Log Per Capita Expenditures	0.2815	2.69	0.2821	3.45	10.7970	0.57
Expenditure Residual	-0.0031	-0.03	--	-	0.0020	0.53
Land Ownership	0.0004	1.17	0.3596	1.20	7.3661	14.47
Cluster Means of Local Middle Schools						
Travel Time to Nearest Middle School	-0.0038	-1.63	-0.0045	-2.23	24.6160	33.20
Average Teacher Experience (years)	0.0615	1.93	0.0645	2.96	6.3681	3.14
Average Teacher Schooling (years)	0.0144	0.20	--	--	13.5110	1.19
Average Teacher Training (years)	0.0616	0.50	-0.0634	-0.74	2.3990	0.61
Blackboard	0.6069	1.41	0.6350	1.74	0.8887	0.20
Library	0.2849	0.98	0.3123	1.20	0.0838	0.20
Books per Room (Middle)	-0.0109	-1.30	-0.0011	-1.73	9.4860	7.79
Books per Room (JSS)	0.0011	0.43	--	-	49.2190	26.75
Lack of Desks	-0.2012	-0.56	-0.1338	-0.67	0.1077	0.26
Enrollment Fee	0.2488	1.21	0.1588	0.95	223.9900	411.49
JSS Only	-0.0910	-0.32	--	--	0.1695	0.30
Deny Admission	0.3594	1.05	0.3873	1.40	0.1051	0.17
"Shed" Classrooms	-0.0929	-0.25	--	--	0.1160	0.18
Unusable Classrooms	-0.0225	-0.04	--	--	0.0857	0.16
Leaky Classrooms	-0.4603	-1.63	-0.4894	-2.18	0.1530	0.21
No Water	0.1155	0.35	--	--	0.8601	0.28
No Electricity	-0.1726	-0.62	-0.1442	-0.70	0.8555	0.28
Private	0.1682	0.25	--	--	0.0220	0.10
Cluster Means of Local Primary Schools						
Travel Time to Nearest Primary School	-0.0039	-0.97	-0.0036	-0.97	12.6760	14.02
Average Teacher Experience (years)	-0.0264	-1.02	-0.0243	-1.32	9.1339	3.23
Average Teacher Schooling (years)	-0.0317	-0.79	-0.0221	-0.71	11.6835	1.76
Average Teacher Training (years)	0.0233	0.22	--	--	2.5874	0.98
Blackboards	0.3914	0.70	0.4878	0.99	0.9039	0.12
Books/per Room	0.0000	0.00	--	--	51.7810	31.06
Lack of Desks	0.1026	0.36	--	--	0.2449	0.36
Enrollment Fee	-0.3811	-0.68	-0.2575	-0.52	122.8060	128.23
Deny Admission	-0.2958	-1.51	-0.2560	-1.51	0.2732	0.30
Complete School (All 6 Grades)	-0.5068	-0.66	-0.4712	-0.73	0.9571	0.11
"Shed" Classrooms	-0.4280	-1.13	-0.3850	-1.32	0.1815	0.21
Unusable Classrooms	0.6929	1.18	0.5737	1.20	0.0705	0.11
Leaking Classrooms	-0.7350	-1.87	-0.6663	-2.38	0.2772	0.22
No Water	0.6243	1.97	0.4836	1.99	0.8782	0.23
No Electricity	0.0087	0.02	--	--	0.9136	0.18
Private	-0.3501	-0.62	-0.1795	-0.43	0.0554	0.14
Log Likelihood	-1,352.741		-1,354.907			
Sample Size	1,399		1,399			

Note: The dependent variable is years of schooling completed. The tuition variables were divided by 1000 in the estimation but the means are the actual means.

those from "other religion" households, the omitted category), possibly reflecting the fact that most schools in Ghana were organized by Catholic and Protestant churches and many retain at least nominal affiliation to those churches. Finally, Ewe and Ga-Adange children go further in school than other ethnic groups. Girls have significantly lower educational attainment than boys. Both mother's and father's educational levels enhance school attainment. Also, the innate ability measure, calculated as a fixed effect from the child's Raven score (see Table 1 of Appendix 2 for this regression), is positively associated with attainment.

Both log per capita expenditures (which represents family income) and agricultural wages have the expected signs.[27] Thus, poorer households remove their children from school earlier, supporting the proposition that they cannot borrow to finance human capital investment.[28] Also, children in high wage villages appear to face greater opportunity costs of schooling. Household demographic structure also affects child school attainment. Later born children go to school longer than their older siblings; in terms of the model, parents appear to remove their older children from school earlier to finance the subsequent schooling of younger siblings. Also, children from larger families have lower school attainment. Both these results are consistent with the credit constraints hypothesis.

The remaining variables in Table 4 are indicators of school quality and

---

[27]Note that the insignificance of the expenditure residual indicates that conventional standard errors are appropriate.

[28]However, other interpretations are possible. For example, schooling may be a consumption good as well as an investment good. Also, income may simply be serving as a proxy for unobserved family background factors. Here we do not attempt to choose between the two interpretations.

school costs. On the cost side, travel time to the nearest middle school has a significant negative effect on school attainment, but middle school tuition has no significant effect. The primary school travel time and enrollment fee variables have no significant effect, though they have the expected signs.[29]

Although many of the middle school quality variables have statistically insignificant coefficients, the highest t-statistics greater than 2 show the expected signs. Specifically, teacher experience has a positive effect on school attainment while leaking classrooms (those that cannot be used when it rains) have a negative effect. Of the two variables significant only at the 10% level, one has the expected sign (blackboards) but the other does not (books per room-middle). Primary school quality variables also have generally weak effects on school attainment (only two have coefficients significantly different from zero at the 5% level and no additional ones at the 10% level), which may reflect the fact that primary school quality changed since schooling decisions where made for the children in our sample. Leaking classrooms and private schools all have a negative effect on school attainment, as one would expect, but lack of water has an unexpected positive effect.

Table 5 displays the results of the ordered probit for late starting for 1110 individuals who formerly attended or are currently enrolled in school. If children frequently repeated grades, then our measure of years late (age - grades completed - 6) would confound late starting and grade repetition, and a child's age would be positively related to years late. But grade repetition is uncommon in Ghana, and indeed the age of the child has a negative effect on late starting in Table 5.

---

[29]The mean primary school enrollment fee is only 123 Cedis (about forty U.S. cents) and the mean travel time to the nearest primary school is only 13 minutes.

Table 5: Ordered Probit, Late Starting

Individual Variables	Full Model Coefficient	Full Model t-stat	Parsimonious Model Coefficient	Parsimonious Model t-stat	Mean	Standard Deviation
Constant	3.5886	1.86	3.9705	2.56	1.0000	-
Coast	0.4465	1.64	0.1994	1.87	0.2081	-
Forest	0.2418	0.93	--	--	0.5838	-
Savannah	-0.0157	-0.05	--	--	0.1216	-
Semi-Urban	-0.0879	-0.53	--	--	0.2279	-
Rural	-0.2371	-1.34	-0.1493	-1.18	0.4054	-
Sex	-0.3130	-4.02	-0.2922	-3.89	0.4468	-
Age (years)	-0.0807	-5.41	-0.0835	-5.94	15.1342	2.82
Mother's Schooling (years)	-0.0304	-2.87	-0.0332	-3.65	3.0892	4.61
Father's Schooling (years)	0.0011	0.14	--	--	6.3009	5.97
Ability Estimate	-0.0273	-1.95	-0.0323	-2.46	12.6009	3.23
Ability Missing	-0.0370	-0.42	--	--	0.2928	-
Muslim	0.0423	0.18	--	--	0.1324	-
Catholic	-0.0362	-0.17	--	--	0.1631	-
Protestant	-0.1522	-0.73	--	--	0.2757	-
Other Christian	-0.2904	-1.39	-0.2278	-2.23	0.1973	-
Traditional Religion	0.2533	1.20	0.2556	2.46	0.1802	-
Akan	-0.1159	-0.83	--	--	0.5234	-
Ewe	-0.2413	-1.57	-0.1562	-1.45	0.2054	-
Ga-Adangbe	0.1151	0.51	--	--	0.0486	-
Male Agricultural Wage	0.0031	3.55	0.0031	4.28	136.9634	170.81
Wage Data Missing	1.1516	3.21	1.1855	4.19	0.5784	-
Child Order	0.0074	0.24	--	--	2.9829	1.66
Siblings	-0.0288	-1.07	-0.0253	-1.31	5.5252	2.06
Log Per Capita Expenditures	-0.2674	-2.28	-0.3054	-2.86	10.8382	0.56
Expenditure Residual	0.0692	0.58	--	--	-0.0043	0.50
Land Ownership	0.0001	0.35	--	--	63.1432	139.97
Cluster Means of Local Middle Schools						
Travel Time to Nearest Middle School	0.0038	1.65	0.0044	2.44	19.0405	25.41
Average Teacher Experience (years)	0.0091	0.31	--	--	6.9382	3.02
Average Teacher Schooling (years)	0.1619	2.63	0.1618	3.27	13.5881	1.10
Average Teacher Training (years)	0.0989	0.95	0.1072	1.51	2.4149	0.57
Blackboard	0.3579	0.86	0.2953	0.82	0.9119	0.14
Library	-0.0730	-0.29	--	--	0.0961	0.22
Books per Room (Middle)	-0.0069	-1.00	-0.0046	-0.74	9.5497	7.55
Lack of Desks	-0.0362	-0.10	--	--	0.0997	0.24
Enrollment Fee	-0.2625	-1.48	-0.2248	-2.20	241.8324	444.74
JSS Only	0.2556	1.02	0.1884	0.96	0.1215	0.24
Deny Admission	-0.3212	-0.95	-0.2512	-0.90	0.1083	0.18
"Shed" Classrooms	0.0538	0.17	--	--	0.1225	0.18
Unusable Classrooms	-0.0666	-0.14	--	--	0.0699	0.14
Leaky Classrooms	-0.4404	-1.68	-0.4272	-1.98	0.1493	0.20
No Water	-0.5285	-1.93	-0.4629	-2.27	0.8486	0.30
No Electricity	0.2743	1.24	0.2455	1.33	0.8324	0.29
Private	-0.1865	-0.33	--	--	0.0222	0.09
Cluster Means of Local Primary Schools						
Travel Time to Nearest Primary School	0.0001	0.01	--	--	11.9937	13.35
Average Teacher Experience (years)	-0.0271	-1.13	-0.0204	-1.12	9.5620	3.13
Average Teacher Schooling (years)	-0.0562	-1.70	-0.0532	-1.89	11.8293	1.89
Average Teacher Training (years)	-0.1273	-1.39	-0.1125	-1.92	2.6597	0.99
Blackboard	-0.1168	-0.24	--	--	0.9136	0.11
Books per Room	0.0014	0.64	0.0019	1.11	52.2338	30.56
Lack of Desks	0.1861	0.66	0.1285	0.94	0.1984	0.32
Enrollment Fee	0.1351	0.25	--	--	134.5566	131.43
Deny Admission	0.2274	1.30	0.2827	1.95	0.2860	0.30
Complete School (All 6 Grades)	-0.9042	-1.18	-0.9577	-1.59	0.9732	0.08
"Shed" Classrooms	0.5306	1.61	0.5162	2.32	0.1778	0.20
Unusable Classrooms	-0.1203	-0.21	--	--	0.0716	0.11
Leaky Classrooms	-0.1601	-0.45	--	--	0.2578	0.22
No Water	0.6495	2.10	0.5940	2.71	0.8682	0.24
No Electricity	0.3331	1.03	0.2670	0.97	0.9009	0.19
Private	0.1532	0.33	--	--	0.0572	0.15
Log Likelihood	-1,468.956		-1,474.498			
Sample Size	1,110		1,110			

Note: The dependent variable is years delayed enrollment after age six. The enrollment fee variables were divided by 1000 in the estimation, but the means are the actual means.

Although Table 4 shows that girls do not go as far as boys in school, girls have a significant and substantial tendency to start at a younger age. Mother's schooling also has a strong negative effect on late starting and greater child innate ability leads to earlier enrollment in school. Children from "other Christian" households are less likely to start late, while children from households practicing traditional African religions are more likely to be late starters. There is no significant impact by ethnic group.

Table 5 also indicates an important role for economic variables in explaining late entry of children into school. Children in high wage villages are much more likely to start late, as are children from poorer households. Both of these findings are consistent with the theoretical model when school fees are present.[30] In contrast to the school attainment estimates, household demographic structure has no significant effect on the timing of human capital investment.

In looking at the estimated effects of school characteristics it should be kept in mind that, since late starting is a decision made prior to primary school, primary school characteristics are likely to make more of a difference than middle school characteristics. Five of the primary school quality coefficients have t-statistics of about 1.9 or higher, and all have the expected signs, in that higher quality leads to earlier school entry. More experienced and more highly trained teachers reduce late starting, while selective admissions, "shed" classrooms and lack of electricity make children start later.

Although 5 middle school characteristic variables have t-statistics of

---

[30] Although viewing schooling as a consumption good may explain the positive correlation between schooling and household expenditure levels in Table 4, it is not clear that this would explain the negative correlation of expenditure with late starting.

2.0 or higher, these effects are for the most part perverse. Greater travel time to the nearest middle school leads to more late starting, as one would expect. Better educated middle school teachers, on the other hand, promote late starting, while middle school enrollment fees, lack of water and leaking classrooms have negative effects. If middle enrollment fees are capturing some unmeasured aspects of school quality, then the negative effect is not surprising. The other results are more difficult to explain.

To summarize this subsection, the two decision rules governing the selection into the sample of children aged 11 to 20 who presently attend middle school are by and large consistent with expectations, although a few exceptions are found. Since some of the school quality characteristics affect how far children go in school and at what age they start, care should be taken when estimating the impact of school quality on cognitive skills. Before proceeding with this estimation, the next subsection examines how families who have decided to send their children to middle school choose among the various schools available.

Choice of Middle School

Next we analyze the school choice of the 302 children currently enrolled in middle school. Multinomial logit and probit estimates are given in Tables 6 and 7, respectively. The parameter estimates are very robust to these distributional assumptions, so they will be discussed together. Variables are dropped from the logit and probit specifications using the same rule as in Tables 4 and 5. Note also that the parsimonious version of the logit model is nested, i.e. the utilities of the nearest and other local options are allowed to be correlated.[31] This correlation coefficient (sigma) is significantly different from zero and one, implying that nesting is necessary.

In reading Tables 6 and 7, note that the underlying choice model implies that school attributes that are attractive to parents should have positive coefficients while those that are unattractive should have negative ones. Almost all of the 12 school variables in the tables have the expected signs, although not always significantly so. At roughly the 5% level, the presence of a library and (middle school) textbooks make schools more attractive, while greater travel time, better trained teachers, and a lack of water make them less attractive. The negative effect of teacher training may reflect the poor state of teacher training in Ghana (see World Bank, 1989), although a negative coefficient seems implausible.

In both the nearest and other local columns, the parameters on the individual variables show the relative utility of the option with respect to the faraway option. For example, the (weakly significant) negative coefficients on fathers' schooling indicate that better educated parents derive higher utility by sending their children to far-away schools. Interestingly, having more siblings increases the chances a child will attend school away from the household of origin, even after controlling for the income of the household (using the expenditure variable). Finally, higher income also encourages parents to educate their children outside their village.

---

[31] They are viewed as more similar to each other than to the faraway school option because the latter implies that the child moves away from home.

Table 6: Multinominal Logit Model of Choice of Middle School

School Variables	Full Model Coeff.	Full Model t-stat	Parsimonious Model Coeff.	Parsimonious Model t-stat
Travel Time to School (in minutes)	-0.0124	-1.17	-0.0220	-1.96
Average Teacher Experience (years)	0.0212	0.41	--	--
Average Teacher Schooling (years)	-0.0532	-0.26	--	--
Average Teacher Training (years)	-0.6041	-1.55	-0.7419	-2.09
Blackboard	1.0871	0.57	0.4535	0.34
Library	1.5378	1.68	2.1125	2.67
Books per Room (Middle)	0.0304	1.21	0.0309	1.66
Books per Room (JSS)	-0.0005	-0.07	--	--
Lack of Desks	-0.2508	-0.29	--	--
Enrollment Fee	-0.0001	-0.09	--	--
JSS Only	-0.8569	-0.87	-0.8724	-0.94
Deny Admission	-0.8237	-0.94	-0.5184	-0.79
"Shed" Classrooms	-0.7857	-0.84	-0.4225	-0.60
Unusable Classrooms	-0.4225	-0.29	--	--
Leaky Classrooms	-0.3257	-0.35	--	--
No Electric	-1.0436	-1.31	-0.9174	-1.47
No Water	-2.4505	-1.72	-3.0456	-2.88
Private	2.1607	0.97	2.1774	1.05

Other Variable	Full Model Nearest Coeff.	Full Model Nearest t-stat	Full Model Other Local Coeff.	Full Model Other Local t-stat	Parsimonious Nearest Coeff.	Parsimonious Nearest t-stat	Parsimonious Other Local Coeff.	Parsimonious Other Local t-stat
Constant	27.2001	2.78	26.3863	2.38	32.6017	3.24	30.5537	2.52
Coast	-0.4296	-0.25	-1.3797	-0.94	--	--	--	--
Forest	-0.3390	-0.20	0.0239	0.02	--	--	--	--
Savannah	1.6002	0.72	0.3507	0.14	--	--	--	--
Semi-Urban	1.3040	1.16	-1.9927	-1.36	1.0653	1.11	-2.3943	-1.97
Rural	1.7507	1.66	-1.5623	-1.21	1.6995	1.92	-1.1913	-1.15
Sex	-0.0494	-0.10	-0.6102	-0.91	--	--	--	--
Age (years)	-0.1129	-0.84	-0.0684	-0.45	--	--	--	--
Mother's Schooling (years)	-0.0891	-1.32	0.0373	0.41	-0.1046	-1.30	0.0233	0.27
Father's Schooling (years)	-0.0544	-1.14	-0.0957	-1.37	-0.0578	-0.98	-0.1444	-1.84
Ability	-0.0908	-1.33	-0.1223	-1.52	-0.1026	-1.15	-0.1151	-1.23
Ability Data Missing	-1.2803	-2.26	-0.2239	-0.25	-1.5596	-1.78	-1.2127	-1.24
Child Order	0.1034	0.59	-0.1742	-0.70	--	--	--	--
Siblings	-0.3776	-2.33	-0.3294	-1.62	-0.3437	-1.99	-0.4057	-2.06
Muslim	0.7416	0.44	0.9731	0.40	--	--	--	--
Catholic	1.1924	0.73	2.3430	1.03	-0.9371	-0.80	-0.2828	-0.21
Protestant	1.0493	0.71	2.1625	1.00	0.4232	0.57	-0.0083	-0.01
Other Christian	1.1120	0.75	1.9226	0.94	--	--	--	--
Traditional Religion	0.1278	0.08	1.9169	0.84	--	--	--	--
Akan	-0.5844	-0.58	0.1990	0.13	--	--	--	--
Ewe	-0.1650	-0.16	-1.8433	-1.36	-1.0445	-1.27	-0.3011	-0.30
Ga-Adangbe	-0.3802	-0.24	-2.9712	-1.28	-0.8402	-1.35	1.1243	1.43
Male Agricultural Wage	0.0036	0.56	0.0060	0.57	--	--	--	--
Wage Data Missing	1.4702	0.66	2.9245	0.82	--	--	--	--
Log Per Capita Expenditures	-1.7413	-2.43	-1.7347	-2.06	-2.0161	-2.19	-1.7437	-1.94
Expenditure Residual	2.3825	2.75	2.4011	2.18	3.0567	2.41	2.6539	1.99
Land Owned	0.0069	0.22	0.0447	0.94	--	-	-	--
Sigma	--				0.6084	3.24		
Log Likelihood	-186.554				-200.153			
Sample Size	302				302			

Note: The value of sigma was set equal to 1 in the full model. Allowing sigma to vary led to a failure to converge properly. This problem did not arise in the parsimonious model.

Table 7: Multinominal Probit Model of Choice of Middle School

School Variables	Full Model Coeff.	Full Model t-stat	Parsimonious Model Coeff.	Parsimonious Model t-stat
Travel Time to School (in minutes)	-0.0065	-1.12	-0.0074	-1.71
Average Teacher Experience (years)	0.0125	0.43	--	--
Average Teacher Schooling (years)	-0.0286	-0.26	--	--
Average Teacher Training (years)	-0.3421	-1.59	-0.2748	-1.94
Blackboard	0.6069	0.58	0.1468	0.25
Library	0.8121	1.61	0.8092	2.45
Books per Room (Middle)	0.0174	1.25	0.0171	2.03
Books per Room (JSS)	-0.0014	-0.29	--	--
Lack of Desks	-0.1647	-0.49	--	--
Enrollment Fee	-0.0000	-0.06	--	--
JSS Only	-0.3580	-0.56	-0.3467	-0.85
Deny Admission	-0.4403	-0.89	-0.3470	-1.22
"Shed" Classrooms	-0.4526	-0.88	-0.1671	-0.48
Unusable Classrooms	-0.2233	-0.28	--	--
Leaky Classrooms	-0.0344	-0.07	--	--
No Electric	-0.5376	-1.23	-0.4779	-1.56
No Water	-1.3359	-1.74	-1.2680	-3.16
Private	1.0926	0.93	0.9786	0.98

Other Variable	Full Model Nearest Coeff.	Full Model Nearest t-stat	Full Model Other Local Coeff.	Full Model Other Local t-stat	Parsimonious Nearest Coeff.	Parsimonious Nearest t-stat	Parsimonious Other Local Coeff.	Parsimonious Other Local t-stat
Constant	14.3461	2.65	13.3048	2.18	10.3286	2.81	11.4225	3.22
Coast	-0.0784	-0.09	-0.5624	-0.71	--	--	--	--
Forest	-0.0926	-0.10	0.0754	0.10	--	--	--	--
Savannah	1.0266	0.86	0.3252	0.24	--	--	--	--
Semi-Urban	0.7996	1.28	-1.0661	-1.36	0.7643	2.15	-0.9289	-2.19
Rural	1.0276	1.78	-0.7991	-1.15	0.9590	2.95	-0.5356	-1.40
Sex	-0.0231	-0.09	-0.3289	-0.90	--	--	--	--
Age (years)	-0.0658	-0.87	-0.0393	-0.48	--	--	--	--
Mother's Schooling (years)	-0.0497	-1.33	0.0235	0.48	-0.0354	-1.26	0.0248	0.72
Father's Schooling (years)	-0.0274	-1.01	-0.0529	-1.40	-0.0185	-0.96	-0.0477	-1.69
Ability	-0.0463	-1.26	-0.0640	-1.49	-0.0272	-0.95	-0.0450	-1.52
Ability Data Missing	-0.6354	-2.00	-0.0463	-0.09	-0.5491	-2.27	-0.1629	-0.49
Child Order	0.0713	0.72	-0.0774	-0.57	--	--	--	--
Siblings	-0.1944	-2.17	-0.1638	-1.45	-0.1292	-2.18	-0.1639	-2.29
Muslim	0.3479	0.37	0.6864	0.52	--	--	--	--
Catholic	0.5951	0.89	1.2579	1.01	--	--	--	--
Protestant	0.4997	0.62	1.0864	0.94	--	--	--	--
Other Christian	0.5007	0.62	0.8954	0.81	--	--	--	--
Traditional Religion	-0.0005	-0.00	0.9885	0.80	--	--	--	--
Akan	-0.3198	-0.58	0.2399	0.28	--	--	--	--
Ewe	-0.0367	-0.06	-0.8742	-1.17	0.2817	1.02	-0.6480	-1.90
Ga-Adangbe	-0.1053	-0.12	-1.4126	-1.16	0.0318	0.46	-1.3673	-1.69
Male Agricultural Wage	0.0014	0.40	0.0028	0.49	--	--	--	--
Wage Data Missing	0.5803	0.48	1.3875	0.68	--	--	--	--
Log Per Capita Expenditures	-0.9088	-2.31	-0.8654	-1.90	-0.6403	-2.12	-0.6250	-2.24
Expenditure Residual	1.1924	2.52	1.1324	1.92	0.9858	2.74	0.8482	2.21
Land Owned	0.0000	0.02	0.0021	0.86	--	--	--	-
Log Likelihood		-187.527				-203.328		
Sample Size		302				302		

Determinants of Mathematics and Reading Test Scores

Our sample of students enrolled in middle school totals 302, but 66 of these are living away from home and 73 others did not take the achievement tests. We are thus left with only 163 children for our analysis of test scores. Tables 8 and 9 report several versions of the mathematics and reading test score regressions, respectively. The first column of each table presents "full model" results. A second specification dropped all variables in the first with t-statistics less than 0.5 (see Table A2 of Appendix 2). Finally, all the variables with t-statistics less than 1.0 were dropped - this is repeated in the second column of each Table. Columns 1 and 2 control for sample selectivity using the estimates of the previous subsections; the two selectivity terms for middle school choice (i.e. for nearest school and other local school) are formed from the multinomial logit model (using the multinomial probit gives essentially the same results, see Tables A3 and A4, Appendix 2). Finally, the third column presents results without controlling for sample selectivity, keeping only the variables with t-statistics greater than 1.0.

Looking at the mathematics scores in Table 8, children in the forest region have significantly lower scores relative to other children in Ghana. Girls do significantly worse than boys, a common finding in many countries. More years of schooling leads to higher math achievement, and the schooling of the mother has a strong positive effect. Family composition does appear to play a role in determining mathematics skills, with children from larger families doing worse on the test (t value of 1.59), and later born children doing better than older ones. These findings can be interpreted in terms of the theoretical model as saying that borrowing constrained parents allocate

Table 8: Determinants of Mathematics Achievement

Household Variables	Full Model Coeff.	Full Model t-stat	Parsimonious Model Coeff.	Parsimonious Model t-stat	Ignoring Selectivity Coeff.	Ignoring Selectivity t-stat
Constant	-0.8749	-0.04	-0.6662	-0.05	-2.3183	-0.18
Coast	-1.9734	-0.50	--	--	--	--
Forest	-5.4555	-1.44	-3.1138	-3.21	-3.2101	-3.32
Savannah	-2.7436	-0.66	--	--	--	--
Semi-Urban	0.9963	0.41	--	--	--	--
Rural	0.6467	0.33	--	--	--	--
Sex	-3.8966	-2.49	-3.9558	-4.25	-3.8157	-4.32
Schooling of Child (years)	1.5587	2.82	1.5844	4.13	1.5772	4.11
Mother's Schooling (years)	0.3995	2.66	0.4298	4.26	0.4294	4.26
Father's Schooling (years)	-0.1531	-0.97	--	--	--	--
Ability Estimate	0.0932	0.47	--	--	--	--
Ability Data Missing	-0.6777	-0.50	--	--	--	--
Primary Math Score	0.1079	0.26	--	--	--	--
Muslim	0.3767	0.08	--	--	--	--
Catholic	1.1674	0.25	--	--	--	--
Protestant	1.2578	0.28	--	--	--	--
Other Christian	-0.1288	-0.03	--	--	--	--
Traditional Religion	-1.7143	-0.41	--	--	--	--
Akan	0.8284	0.47	--	--	--	--
Ewe	0.0120	0.01	--	--	--	--
Ga-Adangbe	-0.5506	-0.15	--	--	--	--
Male Agricultural Wage	-0.0028	-0.26	--	--	--	--
Wage Data Missing	-1.3946	-0.36	--	--	--	--
Child Order	0.5112	1.26	0.6706	2.21	0.6873	2.26
Siblings	-0.3588	-0.94	-0.4281	-1.59	-0.4372	-1.63
Log Per Capita Expenditures	-1.5186	-0.87	-1.3777	-1.39	-1.3325	-1.35
Land Ownership	0.0020	0.34	--	--	--	--
School Variables						
Travel Time	0.0286	0.61	--	--	--	--
Average Teacher Experience (years)	-0.0788	-0.52	--	--	--	--
Average Teacher Schooling (years)	0.6251	1.20	0.5844	1.58	0.6121	1.66
Average Teacher Training (years)	0.4906	0.48	--	--	--	--
Blackboard	11.9740	2.19	9.4305	2.83	9.4321	2.83
Library	4.6536	1.90	3.6111	2.09	3.6089	2.09
Books Per Room (Middle)	-0.0563	-0.87	-0.0581	-1.39	-0.0565	-1.36
Books Per Room (JSS)	-0.0089	-0.35	--	--	--	--
Lack of Desks	-0.4898	-0.16	--	--	--	--
JSS Only	-2.7306	-0.90	-4.0153	-2.07	-4.0460	-2.09
Shed Classrooms	-4.2733	-1.33	-3.4278	-1.61	-3.4327	-1.61
Unusable Classrooms	3.6550	0.93	3.4563	1.39	3.7530	1.51
Leaky Classrooms	-8.2353	-2.99	-5.4491	-2.78	-5.1934	-2.66
Enrollment Fees	0.0098	1.32	0.0083	1.82	0.0081	1.77
Deny Admission	-0.8575	-0.39	--	--	--	--
No Water	0.6982	0.22	--	--	--	--
No Electricity	0.2672	0.12	--	--	--	--
Private School	-2.1063	-0.30	--	--	--	--
Selectivity Correction Terms						
Lambda (Nearest School)	0.8969	0.60	--	--	--	--
Lambda (Other Local)	0.4140	0.48	--	--	--	--
Lambda (Optimal Years)	-0.1546	-0.03	--	--	--	--
Lambda (Late Start)	0.6367	1.38	0.4363	1.14	--	--
Sample Size	163		163		163	
R^2	0.4356		0.3825		0.3770	

Note: The dependent variable, score on the mathematics test, had a mean of 9.57 and a standard deviation of 5.30.

fewer resources per child the more children they have, and that the school attendance of earlier born children suffers because they are put to work to finance the schooling of younger siblings.[32]

Turning to the school variables, of the teacher variables only teacher schooling has weakly significant effect (t-value of 1.58) on the attainment of mathematics skills,[33] but the presence of blackboards and libraries both have significant positive effects. Of the remaining school variables, only two are statistically significant at the 5% level - leaking classrooms and the fact that a school is a newly built junior secondary school, both have very strong negative effects. The interpretation of the leaking classrooms variable is quite simple. Schools with such classrooms are closed whenever it rains, so that children spend fewer days in school per year. The junior secondary school variable suggests some problem with these recently built schools.

Although perhaps anticlimactic, none of the selectivity correction terms were statistically significant. Comparison of columns 2 and 3 in Table 8 reveals that controlling for selectivity has very little effect on the different coefficient estimates. Although this is good news in the sense that previous estimates of this kind may not be seriously biased by failing to control for sample selectivity, it must be kept in mind that this result may not be robust to other countries.

Finally, turn to the results for reading test scores in Table 9.

---

[32] Hanushek (1992) argues, however, that the important constraint is on parental time, not financial resources, and that later born siblings get less parental time.

[33] Most teachers in the sampled schools also took the same mathematics, reading and Raven's tests, but in preliminary regressions these variables never had significant effects on either the mathematics or reading tests of students. These variables were dropped because they reduce our sample size.

Table 9: Determinants of Reading Achievement

Household Variables	Full Model Coeff.	Full Model t-stat	Parsimonious Model Coeff.	Parsimonious Model t-stat	Ignoring Selectivity Coeff.	Ignoring Selectivity t-stat
Constant	-40.5686	-1.33	-33.9858	-4.49	-33.1698	-4.38
Coast	-1.0237	-0.21	--	--	--	--
Forest	-5.2917	-1.14	-4.9883	-3.30	-4.5567	-3.08
Savannah	0.1145	0.02	--	--	--	--
Semi-Urban	6.6673	2.16	6.7603	2.75	6.3248	2.59
Rural	2.6837	1.07	3.0336	1.43	2.8596	1.35
Sex	-2.5606	-1.25	-2.6043	-2.16	-2.6487	-2.19
Schooling of Child (years)	2.3956	3.33	2.4133	4.66	2.4540	4.73
Mother's Schooling (years)	0.3176	1.65	0.2396	1.60	0.2438	1.63
Father's Schooling (years)	-0.1273	-0.62	--	--	--	--
Ability Estimate	0.5347	2.02	0.4378	2.44	0.4531	2.52
Ability Data Missing	0.0578	0.03	--	--	--	--
Primary Reading Score	0.4968	0.76	0.4808	1.23	0.5794	1.51
Muslim	-5.2614	-0.89	-3.2756	-1.66	-3.5025	-1.77
Catholic	-4.1495	-0.69	-2.4893	-1.43	-2.6495	-1.52
Protestant	-2.4841	-0.42	--	--	--	--
Other Christian	-1.6639	-0.29	--	--	--	--
Traditional Religion	-5.0406	-0.90	-3.7515	-2.16	-3.3238	-1.94
Akan	1.3010	0.57	--	--	--	--
Ewe	5.1544	1.72	3.6349	2.38	3.2577	2.17
Ga-Adangbe	3.8374	0.78	5.0129	1.35	4.2977	1.16
Male Agricultural Wage	0.0123	0.88	0.0048	0.44	0.0043	0.39
Wage Data Missing	3.8058	0.78	2.3469	0.59	2.3189	0.58
Child Order	0.5616	1.08	0.5438	1.32	0.4724	1.15
Siblings	0.6568	1.34	0.4971	1.40	0.5285	1.48
Log Per Capita Expenditures	1.1763	0.49	--	--	--	--
Land Ownership	-0.0006	-0.08	--	--	--	--
School Variables						
Travel Time	0.0221	0.32	--	--	--	--
Average Teacher Experience (years)	-0.1170	-0.60	--	--	--	--
Average Teacher Schooling (years)	-0.3602	-0.53	--	--	--	--
Average Teacher Training (years)	1.1016	0.92	0.9810	1.08	0.9323	1.03
Blackboard	11.4017	1.56	12.9494	2.44	11.3820	2.20
Library	-2.3441	-0.74	--	--	--	--
Books Per Room (Middle)	-0.0591	-0.70	--	--	--	--
Books Per Room (JSS)	0.0368	1.10	0.0383	1.71	0.0416	1.86
Lack of Desks	3.4286	0.84	--	--	--	--
JSS Only	0.5938	0.16	3.1830	1.17	2.4345	0.91
Shed Classrooms	-0.0395	-0.01	--	--	--	--
Unusable Classrooms	-8.3485	-1.61	-5.1403	-1.46	-5.4275	-1.54
Leaky Classrooms	-8.9487	-2.58	-7.7800	-3.04	-6.9268	-2.79
Enrollment Fees	-0.0092	-0.94	-0.0059	-0.91	-0.0049	-0.77
Deny Admission	0.6735	0.24	--	--	--	--
No Water	0.9294	0.21	--	--	--	--
No Electricity	-2.9313	-0.90	-2.9292	-1.50	-2.5192	-1.31
Private School	12.1776	1.36	11.9132	1.67	12.1928	1.70
Selectivity Correction Terms						
Lambda (Nearest School)	1.1989	0.63	--	--	--	--
Lambda (Other Local)	1.0063	0.84	1.0163	1.32	--	--
Lambda (Optimal Years)	-0.4237	-0.07	--	--	--	--
Lambda (Late Start)	0.3745	0.63	--	--	--	--
Sample Size	163		163		163	
R^2	0.5418		0.5176		0.5114	

Note: The dependent variable, score on reading test, had a mean of 7.09 (out of 29) and a standard deviation of 6.68.

Children in forest areas do significantly worse, while those in semi-urban areas perform relatively well. As with the math test, girls show lower reading skills than boys, which may indicate that girls attend school more sporadically than boys. Years in school has a positive impact on reading skills, but the positive impact of mother's schooling is only marginally significant. The fixed effect estimator of innate ability has a significantly positive impact, unlike the case with mathematics scores. Muslims and children from households practicing traditional African religions appear to do worse on the reading test, perhaps reflecting less exposure to the English language in the home. Finally, Ewe children in Ghana seem to read better than those from other ethnic groups.

Turning to the school variables, none of the teacher characteristics has any significant impact on test scores. As with mathematics scores, the coefficient on blackboards was significantly positive. The number of (JSS) books per room has a positive effect but is statistically significant only at the 10% level. Leaking classrooms again has a strong negative effect on reading. Finally, private schools appear to be more effective than public schools.

The lack of significance of the selectivity correction terms in the reading score regression again indicates that nonrandom selection does not seriously bias the results shown here. A comparison of columns 2 and 3 shows almost no change in the coefficient estimates. However, one must be cautious about dismissing the problem of sample selectivity in these types of estimates because low sample size in both test score regressions means that more subtle biases probably could not be detected.

Finally, as a general point for both Tables 8 and 9, economic variables, such as the agricultural wage and income and land ownership, have little or no direct effect on cognitive skills. The main effect of these variables operates indirectly, through school attainment and the school starting time; i.e., through the sample selection mechanism.

Direct and Indirect Effects of Improving School Quality

Improvements in school quality can raise the skills of students in two distinct ways. First, for a given number of years of school attendance, such improvements can raise cognitive achievement by making each year of schooling more effective. This direct effect is measured in Tables 8 and 9. Second, school quality raises skills because parents decide to send their children to school for a longer period of time. This indirect effect is evident in the ordered probit results in Table 4. Table 10 presents estimates of both these effects for six hypothetical schooling improvements. The indirect effect is calculated using the change in the expected years of schooling when all other variables are evaluated at their sample means.

Reducing travel time and raising average teacher experience have only indirect effects (the parameters in Tables 8 and 9 were not statistically different from zero). Their overall impact is relatively modest. Providing a school library has indirect effects and a direct effect only for the mathematics score. Its impact on test scores is also modest. Repairing schools in which all classrooms cannot be used when it rains and provision of blackboards to schools where none now exist have both direct and indirect effects, and the former are much stronger than the latter. Finally, providing more textbooks would have only a direct effect on the reading score, and a relatively modest one at that. Among all the possible avenues for raising cognitive achievement given in Table 10, repairing schools with leaking classrooms and providing schools with blackboards where none now exist will

Table 10: Impact of Raising Middle School Quality on Cognitive Achievement

	Mathematics Score			Reading Score		
Policy Option	Indirect	Direct	Total	Indirect	Direct	Total
1. Reducing Travel Time from 2 hours to 0 hours	+3.44	-	+3.44	+5.24	-	+5.24
2. Raising Average Teacher Experience from 2 years to 10 years	+3.19	-	+3.19	+4.87	-	+4.87
3. Providing a School Library	+1.88	+3.61	+5.49	+2.87	-	+2.87
4. Repairing Classrooms in Schools where all classrooms cannot be used when it rains	+3.10	+5.45	+8.55	+4.73	+7.78	+12.51
5. Providing Blackboards in Schools where none presently exist	+4.05	+9.43	+13.48	+6.16	+12.95	+19.11
6. Providing 50 more (JSS) Textbooks per room in Schools which now only have 25 per room	-	-	-	-	+1.92	+1.92

Notes: 1. Indirect effect is calculated as the product of the change in expected years of schooling (using the parameters from Table 4) due to improved school quality and the parameters on years schooling from Tables 8 and 9.

2. For purposes of comparison the mean mathematics and reading scores for our sample of 163 students presently in middle school are 9.57 (out of 36) and 7.09 (out of 29), respectively.

clearly have the largest impact. However, most schools in Ghana already have sufficient blackboards (cf. variable means in Table 4) so this will only apply to a limited number of schools, and a similar situation exists regarding repairing leaking classrooms.

In viewing these policy simulations, one should keep in mind that quality attributes such as blackboards may partly be picking up the effects of school administrative skills or other unobservables that are positively correlated with student achievement and progress through school. If so, the achievement gains reported in Table 10 will be overly optimistic. Finally, it should be noted that for policy analysis one also requires information on the costs of these options.[34]

---

[34] These costs also include the number of years that specific items (e.g. blackboards and textbooks) can be expected to remain usable.

VI. Summary and Conclusions

Investments in schooling in developing countries are one of the most important ways of raising productivity and, ultimately, the standard of living in those countries. Choosing which investments to make requires knowledge of their impact on the educational outcomes of students. This paper has presented a model for examining the relationship between schooling investments and schooling outcomes which, unlike almost all previous work, explicitly accounts for the optimizing behavior of families.

The following findings are the most important. First, we find support for the hypothesis that, in Ghana, many families are credit constrained in that their income levels and, to a lesser extent, their composition, affect the timing of human capital investment. Second, we find clear evidence that school quality affects both school attainment and the tendency to start school at a late age. These indirect effects are often ignored in attempts to measure the impact of schooling investments on human capital outcomes. Third, we find no strong evidence that sample selectivity is biasing our estimates of the determinants' cognitive outcomes. Finally, turning to specific schooling investments, we find that blackboards have a positive impact and classrooms that are unusable when it rains have a strong negative impact on the acquisition of both reading and mathematics skills. Reading skills also suffer in the absence of textbooks, and both skills benefit from reduced travel time and higher teacher experience.

References

Ainsworth, Martha. 1990. "Economic Aspects of Child Fostering in Côte d'Ivoire." Mimeo. Africa Technical Department. The World Bank. Washington, D.C.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." Journal of Political Economy. vol. 75. pp. 352-365.

Boissiere, Maurice, John Knight and Richard Sabot. 1985. "Earnings, Schooling, Ability, and Cognitive Skills." American Economic Review, vol. 75, no. 5, pp. 1016-1030.

Fuller, Bruce. 1986. Raising School Quality in Developing Countries: What Investments Boost Learning? The World Bank, Washington, D.C.

Glewwe, Paul, and Kwaku A. Twum-Baah. 1990. "The Distribution of Welfare in Ghana in 1987-88." Living Standards Measurement Study Working Paper No. 75. The World Bank. Washington, D.C.

Glewwe, Paul. 1991. "A Test of the Normality Assumption of the Ordered Probit Discrete Choice Model." Population and Human Resources Department. The World Bank. Washington, D.C.

Glewwe, Paul. 1992. "The Three-Choice Multinomial Probit Model with Selectivity Corrections." Econometric Theory (Problems and Solutions Series), vol. 8, no. 2.

Hanushek, Eric. 1986. "The Economics of Schooling." Journal of Economic Literature. vol. 24. pp. 1141-77.

Hanushek, Eric. 1992. "The Trade-off Between Child Quantity and Quality," Journal of Political Economy vol. 100, pp. 84-117.

Heckman, James. 1979. "Sample Selection and Specification as a Specification Error." Econometrica, vol. 47, pp. 153-61.

Jacoby, Hanan. 1991. "Borrowing Constraints and School Attendance: Theory and Evidence from a Low Income Country." Mimeo. Department of Economics. University of Rochester. Rochester, NY.

Jimenez, Emmanuel, Marlaine Lockheed and Nougnuch Wattanawah. 1988. "The Relative Efficiency of Private and Public Schools: The Case of Thailand." World Bank Economic Review, vol. 2, pp. 139-64.

Jimenez, Emmanuel, Marlaine Lockheed, and Victor Pacqueo. 1991. "The Relative Efficiency of Private and Public Schools in Developing Countries." World Bank Research Observer, vol. 6, no. 2.

Kamien, Morton, and Nancy Schwarz. 1981. Dynamic Optimization. North Holland, Amsterdam.

Lillard, Lee, and Elizabeth King. 1984. "Methods of Analyzing Schooling Choice with Household Survey Data." The Rand Corporation, Santa Monica, CA.

Maddala, G.S. 1983. Limited - Dependent and Qualitative Models in Econometrics. Cambridge University Press.

Raven, J C., J.H. Court and J. Raven. 1984 "Manual for Raven's Progressive Matrices and Vocabulary Scales." H.K. Lewis & Co., London.

Rivers, Douglas, and Quang Vuong. 1988. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." Journal of Econometrics, vol. 39, pp. 347-66.

Scott, Christopher, and Ben Amenuvegbe. 1989. "Sample Designs for the Living Standards Survey in Ghana and Mauritania." Living Standards Measurement Study Working Paper No. 49. The World Bank. Washington, D.C.

Tomiyama, Ken. 1985. "Two-Stage Optimal Control Problems and Optimality Conditions." Journal of Economic Dynamics and Control, vol. 9, pp. 317-37.

World Bank. 1989. "Basic Education for Self-Employment and Rural Development." African Country Department IV. Washington, D.C. World Bank. 1990. World Development Report, Washington, D.C.

Appendix 1: A Dynamic Model of School Attainment

Maximization of discounted lifetime utility, as given in equation (2.5) in the text, is best done using the mathematics of optimal control (cf. Kamien and Schwartz, 1991). Formally, we want to maximize discounted life-cycle utility:

$$\int_0^T U(C(t))\,e^{-\delta t}dt \tag{A.1}$$

where U is assumed to be strictly concave ($U' > 0$ and $U'' < 0$), subject to two differential equation contraints:

$$\dot{H}(t) = bH(t)S(t) \tag{A.2}$$

$$\begin{aligned} \dot{A}(t) &= rA(t) + y + wH(t)D - C(t) & \text{if } S(t) = 0 \\ &= rA(t) + y - f + wH(t)[D-(1+d)S(t)] - C(t) & \text{if } 0 < S(t) \le 1 \end{aligned} \tag{A.3}$$

H(t) and A(t) are state variables while S(t) and C(t) are control variables. The following state and control variable constraints and initial conditions hold:

(A.4) Control Constraints: $0 \le S(t) \le 1$
$C(t) \ge 0$

(A.5) State Constraint: $A(t) \ge 0$ if the household is credit contrained

(A.6) Boundary Conditions: $H(0) = H_0$
$A(0) = 0$

For simplicity, assume that consumption is strictly greater than zero ($C(t) > 0$); this amounts to assuming that the utility function U has a shape such that it is never optimal for consumption to equal zero, which is in fact what one observes in the real world. The assumption that the household begins with no assets ($A(0) = 0$) can be modified; the impact of such a change will be (non-rigorously) discussed at various points. The assumption that assets at the time of the parents' death (T) have no value implies that parents will consume all assets before they die, so that A(T) = 0. This is intuitively obvious.

In this appendix both the credit contrained and unconstrained cases will be examined. It is instructive to begin with the unconstrained case.

Case 1: No Credit Contraints

When parents face no credit constraints they can separate their consumption and human capital investment decisions. The latter is done to maximize the present discounted value of life-cycle wealth, and then the former is done conditional on life-cycle wealth. Thus to understand schooling decisions one need only examine the maximization of life cycle wealth. If school fees do not exist (f=0), the maximand is:[1]

$$\int_0^T (y + wH(t)\,[D - S(t)\,(1+d)])\,e^{-rt}dt \tag{A.7}$$

subject to (A.2), (A.4) and $H(0) = H_0$. The Hamiltonian equation is:

$$H_u = (y + wH(t)[D - S(t)(1+d)])e^{-rt} + \mu(t)bH(t)S(t) \tag{A.8}$$

The following necessary conditions hold:

$$\begin{array}{ll} \mu(t)b \geq (1+d)we^{-rt} & \text{if } S(t) = 1 \\ \mu(t)b = (1+d)we^{-rt} & \text{if } 0 < S(t) < 1 \\ \mu(t)b \leq (1+d)we^{-rt} & \text{if } S(t) = 0 \end{array} \tag{A.9}$$

$$\mu'(t) = -\mu(t)bS(t) - w[D - (1+d)S(t)]e^{-rt} \tag{A.10}$$

[1]Positive school fees can also be handled, as will be discussed below.

Suppose $0 < S(t) < 1$. Solving for $S(t)$ in (A.10) gives:

(A.11) $$S(t) = \frac{-\mu'(t) - Dwe^{-rt}}{\mu(t)b - (1+d)we^{-rt}}$$

Substituting the second line of (A.9) into (A.11) results in division by zero. Thus it is never the case that $0 < S(t) < 1$.[2] Either the child attends school full-time or not at all. Intuitively, this will also hold for school fees greater than zero because part-time schooling becomes even more "wasteful" since the same school fees are paid for full-time and part-time schooling.

Because schooling is always full-time, the expression to optimize becomes:

$$\begin{aligned} LCW &= \int_0^{t_0} wH_0De^{-rt}dt + \int_{t_0}^{t_1}\left([D-(1+d)]\,wH_0e^{b(t-t_0)} - f\right)e^{-rt}dt + \int_{t_1}^{T} wH_0De^{b(t_1-t_0)}e^{-rt}dt \\ &= -\frac{1}{r}wH_0D(e^{-rt_0}-1) + \frac{1}{b-r}[D-(1+d)]\,wH_0e^{-bt_0}[e^{t_1(b-r)}-e^{t_0(b-r)}] \\ &\quad + \frac{f}{r}(e^{-rt_1}-e^{-rt_0}) - \frac{1}{r}wH_0De^{b(t_1-t_0)}(e^{-rT}-e^{-rt_1}) \end{aligned}$$

(A.12)

Differentiating this expression by t_0 and t_1, respectively, yields:

(A.13) $$\frac{\partial LCI}{\partial t_0} = fe^{-rt_0} + wH_0\left(De^{-rt_0} + \frac{[D-(1+d)]}{b-r}(re^{-rt_0}-be^{t_1(b-r)-bt_0}) + \frac{b}{r}(e^{-rT}-e^{-rt_1})De^{b(t_1-t_0)}\right)$$

(A.14) $$\frac{\partial LCI}{\partial t_1} = -fe^{-rt_1} - (1+d)\,wH_0e^{(b-r)t_1-bt_0} - \frac{b}{r}wH_0D(e^{-rT}-e^{-rt_1})\,e^{b(t_1-t_0)}$$

Setting the derivative with respect to t_1 equal to zero (the solution must be an interior one if schooling is profitable), substitute (A.14) into (A.13):

---

[2]Alternatively, one can show that (A.9) and (A.10) imply that $\mu(t) = (Dw/r)e^{-rt} + k$, where k is a constant less than zero. Inserting this into (A.9) implies that $0 < S(t) < 1$ can only occur for more than an instant of time if $b/r = (1+d)/D$. Later it will be seen that this implies that schooling is not a profitable investment.

(A.15) $$\frac{\partial LCI}{\partial t_0} = f(e^{-rt_0}-e^{-rt_1}) + wH_0\left(De^{-rt_0} - (1+d)\,e^{b(t_1-t_0)}e^{-rt_1} + \frac{[D-(1+d)]}{b-r}[re^{-rt_0}-be^{(b-r)t_1-bt_0}]\right)$$

For any human capital investments to be profitable, the following must hold:

(A.16) $$\int_0^{t_0} wH_0De^{-rt}dt + \int_{t_0}^{t_1}\left([D-(1+d)]\,wH_0e^{b(t-t_0)} - f\right)e^{-rt}dt + \int_{t_1}^{T} wH_0De^{b(t_1-t_0)}e^{-rt}dt > \int_0^{T} wH_0De^{-rt}dt$$

which holds if and only if

(A.17) $$f(e^{-rt_0}-e^{-rt_1}) + wH_0\left(\frac{-r[D-(1+d)]}{b-r}(e^{(b-r)t_1-bt_0}-e^{-rt_0}) + D[e^{-rt_0} + e^{b(t_1-t_0)}(e^{-rT}-e^{-rt_1}) - e^{-rT}]\right) < 0$$

Inserting (A.17) into (A.15) yields:

(A.18) $$\frac{\partial LCI}{\partial t_0} + wH_0De^{-rT}(e^{b(t_1-t_0)} - 1) < 0$$

Equation (A.18) implies that the derivative of life cycle income with respect to t_0 is always negative, thus t_0^* must equal 0. Households without credit constraints always begin their child's schooling immediately and send them to school full-time. By Pontryagin's maximum principle all schooling must be completed during one long period of attendance; suppose a child attends full time and then leaves at some time t_1, if further schooling is optimal it must begin immediately at t_1, leaving no gaps in attendance.

The only decision for the unconstrained household is to choose the time to remove the child from school (t_1). If school fees equal zero, then:

$$t_1^* = T + \log[1 - r(1+d)/bD]/r \tag{A.19}$$

But if school fees are positive, one cannot solve for t_1^* explicitly from (A.14). However, after setting $t_0=0$, the right hand side of (A.14) can be set equal to zero and the resulting equation totally differentiated to yield:

$$\begin{aligned}
0 = {} & [-(1+d)\,e^{(b-r)\,t_1^*} - \frac{b}{r}De^{bt_1^*}(e^{-rT}-e^{-rt_1^*})]\,(wdH_0 + H_0dw) + [\frac{b}{r}wH_0De^{bt_1^*}re^{-rT}]\,dT \\
& + [-e^{-rt_1^*}]\,df + [-wH_0e^{(b-r)\,t_1^*}]\,dd + [\frac{-b}{r}wH_0e^{bt_1^*}(e^{-rT}-e^{-rt_1^*}]\,dD \\
& + [fre^{-rt_1^*} - (1+d)\,wH_0\,(b-r)\,e^{(b-r)\,t_1^*} - \frac{b}{r}wH_0Dbe^{bt_1^*}(e^{-rT}-e^{-rt_1^*}) - \frac{b}{r}wH_0De^{bt_1^*}re^{-rt_1^*}]\,dt_1^* \\
& + [ft_1^*e^{-rt_1^*} + (1+d)\,wH_0t_1^*e^{(b-r)\,t_1^*} + \frac{b}{r^2}wH_0De^{bt_1^*}(e^{-rT}-e^{-rt_1^*}) + \frac{b}{r}wH_0De^{bt_1^*}(Te^{-rT}-t_1^*e^{-rt_1^*})]\,dr \\
& + [-(1+d)\,wH_0t_1^*e^{(b-r)\,t_1^*} - \frac{1}{r}wH_0De^{bt_1^*}(e^{-rT}-e^{-rt_1^*}) - \frac{b}{r}wH_0D(e^{-rT}-e^{-rt_1^*})\,t_1^*e^{bt_1^*}]\,db
\end{aligned}$$

(A.20)

All the terms in brackets in (A.20) can be signed. The condition that the second derivative of (A.12) with respect to t_1^* be negative implies that the expression associated with dt_1^* is negative. Clearly, those associated with df and dd are negative and those associated with dD and dT are positive. Setting (A.14) equal to zero (using $t_0^* = 0$) implies that the terms associated with $wdH_0 + H_0dw$ and db are positive. Finally, using (A.14) and the fact that $\frac{1}{r}(1-e^{r(T-t_1^*)}) + T-t_1^*$ is negative (it equals 0 if $T=t_1^*$ and its derivative with respect to T is negative when $T > t_1^*$) shows that the terms associated with dr is negative. Thus the amount of time spent in school, t_1^*, increases when D, T, w, H_0 or b increase but decreases when f, d or r increase.

A further useful result emerges when t_0 is set to zero in (A.17):

$$\text{(A.21)} \qquad \frac{f}{wH_0}(e^{-rt_1^*}-1) + D(e^{-rT}-e^{bt_1^*-rT}) + (e^{(b-r)t_1^*}-1)\,[D\frac{d}{b-r} - (1+d)\frac{r}{b-r}] > 0$$

The first two terms are negative, so schooling is only profitable if the third term is positive. If b > r, the third term is positive. If b < r, then it is positive only if Db > (1+d)r (which also holds if b > r). Note that this condition is necessary but not sufficient.

Case 2: Credit Constrained Household

<u>Case of No School Fees</u>

Now turn to the case where (A.5) is imposed, i.e. the household has no access to credit. Begin with the case where school fees equal zero (f=0). The Hamiltonian is simply:

$$\text{(A.22)} \qquad H = U(C(t))e^{-\delta t} + \mu(t)bH(t)S(t) + \eta(t)A(t) + \lambda(t)[rA(t) + y + wH(t)[D - S(t)(1+d)] - C(t)]$$

The following necessary conditions hold:

$$\text{(A.23)} \qquad \lambda(t) = U'(C(t))e^{-\delta t}$$

$$\text{(A.24)} \qquad \begin{array}{ll} \mu(t)b - \lambda(t)(1+d)w = 0 & \text{if } 0 < S(t) < 1 \\ \mu(t)b - \lambda(t)(1+d)w > 0 & \text{if } S(t) = 1 \\ \mu(t)b - \lambda(t)(1+d)w < 0 & \text{if } S(t) = 0 \end{array}$$

$$\text{(A.25)} \qquad \mu'(t) = -\mu(t)bS(t) - \lambda(t)w[D - S(t)(1+d)]$$

(A.26) $\lambda'(t) = -\eta(t) - r\lambda(t)$

(A.27) $\eta(t)A(t) = 0$ and $\eta(t) \geq 0$

(A.28) $\lambda(T)A(T) = 0$ and $\lambda(T) \geq 0$

(A.29) $\mu(T) = 0$

Differentiate (A.23) with respect to t and combine with (A.26):

(A.30) $$C'(t) = \frac{(\delta - r) - \eta(t)/\lambda(t)}{U''(C(t))/U'(C(t))}$$

By (A.27) $A(t) > 0$ implies that $\eta(t)=0$. For an impatient household ($\delta-r > 0$), consumption will decline if assets are held ($A(t) > 0$).

Suppose $0 < S(t) < 1$. Differentiate the first line of (A.24) with respect to t, then use (A.25) and (A.26) to get:

(A.31) $$-\mu(t)bS(t) - \lambda(t)w[D - S(t)(1+d)] = -(1+d)[\eta(t) - r\lambda(t)](w/b)$$

which can be rearranged, using (A.24) as :

(A.32) $$\mu(t)[Db/(1+d) - r] = \eta(t)(1+d)(w/b)$$

For any schooling to be done at all, it was shown above that (A.21) implies $Db > (1+d)r$. Since $\mu(t) > 0$ (since $\lambda(t) > 0$ by (A.23)), it follows that $\eta(t) > 0$, which implies that $A(t) = 0$, no assets are held.

Suppose $S(t) = 1$ at time zero. Assuming initial assets equal zero ($A(0) = 0$), the only way assets can be accumulated is if consumption falls below income. However, if this occurs consumption must fall further due to (A.30).

Assets will continue to accumulate, ad infinitum, because income can only increase (even more so if $S(t)$ falls below 1) and consumption declines further. This contradicts the result that assets must equal zero at time T, thus no assets are accumulated. This finding holds whenever $S(t) = 1$ and no assets exist.

Suppose $S(t) = 0$ at time zero. If assets are accumulated while $S(t)$ still equals zero, a shift to part-time schooling ($0 < S(t) < 1$) is not possible since it can only occur if assets equal zero. Thus the only way to accumulate assets is to accumulate them when $S(t)=0$ and then consumed when $S(t) = 1$. Note that the Hamiltonian must be continuous at any point where $S(t)$ "jumps" from 0 to 1. Equating the Hamiltonians at this point yields (recall $\eta(t) = 0$ when $A(t) > 0$):

$$\text{(A.33)} \quad U(C(t))e^{-\delta t} + \lambda(t)[rA(t) + y + wH(t)D - C(t)] = U(C(t)e^{-\delta t} + \mu(t)bH(t) + \lambda(t)[rA(t) + y + wH(t)[D - (1+d)] - C(t)]$$

$C(t)$ cannot "jump" (since $\lambda(t)$ would thereby not be continuous) and neither can $A(t)$ nor $H(t)$. Thus one gets:

$$\text{(A.34)} \qquad \mu(t)b = \lambda(t)w(1+d)$$

Both $\mu(t)$ and $\lambda(t)$ are differentiable, so as in the case where $0 < S(t) < 1$ we can derive (A.32), which implies $A(t) = 0$, contradicting the assumption that assets have been accumulated. Thus no assets can be accumulated when $S(t)=0$. This finding holds any whenever $S(t)$ becomes zero and no assets exist.

To summarize, no assets will be accumulated by impatient households, so the Hamiltonian in (A.22) can be simplified to:

$$H = U(y + wH(t)[D - S(t)(1+d)])e^{-\delta t} + \mu(t)bH(t)S(t) \tag{A.35}$$

where S(t) is the (only) control variable and H(t) the (only) state variable. The necessary conditions for a maximum are:

$$\begin{aligned} &-U'(y + wH(t)[D - S(t)(1+d)])e^{-\delta t}w(1+d) + \mu(t)b = 0 \quad \text{if} \quad 0 < S(t) < 1 \\ &-U'(y + wH(t)[D - S(t)(1+d)])e^{-\delta t}w(1+d) + \mu(t)b > 0 \quad \text{if} \quad S(t) = 1 \\ &-U'(y + wH(t)[D - S(t)(1+d)])e^{-\delta t}w(1+d) + \mu(t)b < 0 \quad \text{if} \quad S(t) = 0 \end{aligned} \tag{A.36}$$

$$\mu'(t) = -U'(y + wH(t)[D - S(t)(1+d)])e^{-\delta t}w[D - S(t)(1+d)] - \mu(t)bS(t) \tag{A.37}$$

$$\mu(T) = 0 \tag{A.38}$$

Examining (A.36) when t = T, with reference to (A.38), shows that S(t) must equal zero at time T. Note as well that a "jump" from S(t) = 0 to S(t) = 1 (or vice versa) at some point t implies, by the continuity of $\mu(t)$ and using (A.36):

$$U'(y + wH(t)D)e^{-\delta t}w(1+d) \geq \mu(t)b \geq U'(y + wH(t)[D - (1+d)])e^{-\delta t}w(1+d). \tag{A.39}$$

This relation cannot hold since U" < 0, thus such jumps cannot occur. More generally, the continuity of $\mu(t)$ implies that when schooling moves from full-time to part-time or from part-time to no schooling at all, S(t) must be continuous. No "jumps" in S(t) will occur for impatient households.

When S(t) = 0 it will remain at zero until time T. To see this, set S(t) = 0 in (A.37) and solve for $\mu(t)$, noting that consumption is constant: $\mu(t)=U'(y+wH(t)D)e^{-\delta t}wD/\delta + k$, where k is a constant term. At time T, the strict inequality must hold in the last line of (A.36), because of (A.38). Thus the point at which a child leaves school permanently must be at some time before T, call it t_1. From t_1 to T, U' is simply a constant (because no more human capital is being accumulated and income is being consumed as it is earned), thus the solution to $\mu(t)$ above implies that $k < 0$. By the continuity of $\mu(t)$, the equality holds for (A.36) at time t_1. Inserting the solution for $\mu(t)$ into (A.36) at time t_1 gives:

$$\text{(A.40)} \qquad [bD/\delta - (1+d)]U'(y+wH(t_1)D)e^{-\delta t} = -kb$$

which implies that $bD/\delta - (1+d)$ is positive.[3] Using this result, and inserting $\mu(t)$ into the last line of (A.36), $\mu(t)b$ falls faster over time than the first term in (A.36), thus $S(t)$ must remain equal to zero once it becomes zero. Finally, this solution for $\mu(t)$, along with (A.23) and the first term of of (A.24) yields (the first term can be used by the continuity of $\mu(t)$):

$$\text{(A.41)} \qquad t_1^* = T + \ln[1 - \delta(1+d)/bD]/\delta$$

This expression is similar to that of the unconstrained case where school fees equal zero, except r is replaced by δ. Thus the total time in school declines when d and δ increase and rises as T, b and D increase.

Schooling can begin either as part time ($0 < S(t) < 1$) or full time ($S(t) = 1$). If full time, a part time phase must be entered before schooling completely ends. If part time, schooling may later become full time but will eventually become part time again and then finish before time T. Intuitively, if schooling is ever full time one would expect an initial period of full time schooling, then by part time schooling, followed by leaving school forever. To see when this is possible, examine the point at which $S(t)$ switches from full

---

[3]If it were not positive then no schooling would ever take place.

time to part time. At this point the second line of (A.36) becomes an equality. Differentiate with respect to t yields, using (A.37) (when $S(t) = 1$):

$$\text{(A.42)}\quad U''(y+wH(t)[D-(1+d)])e^{-\delta t}w^2(1+d)/b\{H(t)S(t)b[D-(1+d)]-H(t)S'(t)(1+d)\}$$
$$= \mu(t)[\delta-bD/(1+d)]$$

Upon rearranging terms one obtains:

$$\text{(A.43)}\qquad S'(t) = b(\frac{D}{1+d} - 1) - \frac{\mu(t)b(\delta - \frac{bD}{1+d})}{U''(y+wH(t)[D-(1+d)])e^{-\delta t}w^2(1+d)^2H(t)}$$

The first term is positive if $D > (1+d)$, and the second is negative since $U'' < 0$ and $bD/\delta > 1+d$). If the working day for the child (D) equals the school day plus travel time (1+d), then $S'(t) < 0$ in the neighborhood of $S(t)=1$; once the child moves from full-time schooling to part-time schooling, full-time schooling will never reccur. Further, if U" is a constant (implying that the third derivative of U equals 0), the second term falls in absolute value over time because $H'(t) \geq 0$ and $\mu(t) < 0$. Thus, once schooling moves from full-time to part-time, it can never be full-time again; if it did $S'(t) \geq 0$ at that point and would thereafter remain positive, which would force schooling to be full-time forever, which is not optimal.

Positive School Fees

When school fees exist a technical wrinkle emerges in the mathematics - the integrand when the child is in school is different then when he or she is not in school. Tomiyama (1985) shows how to handle such situations. Below it will be shown that once a child leaves school he or she will not return. This implies that there are three phases to consider, from time zero until the child enters school ($0 \le t \le t_0$), from the time the child enters school until leaving school permanently ($t_0 \le t \le t_1$), and the time the child leaves school until the parents die ($t_1 \le t \le T$). Tomiyama shows that, within these 3 periods, the standard necessary conditions still hold, as one would expect. More importantly, he showed which conditions must hold at the boundary points (t_0 and t_1) to ensure that those points are optimally chosen.

Begin with the third time period ($t_1 \le t \le T$) for the impatient household. It will be seen below that it is optimal for assets to equal zero at time t_1. Since $S(t) = 0$, no school fee is paid, so the mathematics of the "no fees" case applies. Briefly, no assets will be accumulated and consumption will equal income up to time T, and once S(t) becomes zero it will remain zero.[4] Note for later reference that the optimal path of utility in phase three is simply $H_3^* = U_3^*(t) = U(y+H(t_1))e^{-\delta t}$.

Now turn to the second time period ($t_0 \le t \le t_1$). This corresponds to the cases of $S(t) = 1$ and $0 < S(t) < 1$ from the no fees case. However, as will be seen below, some assets will have been accumulated prior to t_0. It is instructive to view this period in detail. The Hamiltonian is:

$$\text{(A.44)} \qquad H_2 = U(C(t))e^{-\delta t} + \mu(t)(bH(t)S(t)) + \eta(t)A(t) + \lambda(t)[rA(t) + y - f + wH(t)[D-S(t)(1+d)] - C(t)]$$

The following necessary conditions hold for $t_0 < t < t_1$:

[4]This proves that all schooling must take place at one interval of time, as asserted in the previous paragraph. Note that the essential "trick" is that $bD/\delta-(1+d) > 0$.

(A.45) $\lambda(t) = U'(C(t))e^{-\delta t}$

(A.46) $\mu(t)b - \lambda(t)(1+d)w = 0$ if $0 < S(t) < 1$
$\mu(t)b - \lambda(t)(1+d)w > 0$ if $S(t) = 1$

(A.47) $\mu'(t) = -\mu(t)bS(t) - \lambda(t)w[D-S(t)(1+d)]$

(A.48) $\lambda'(t) = -\eta(t) - r\lambda(t)$

(A.49) $\eta(t)A(t) = 0$ and $\eta(t) \geq 0$

Since assets at time t_0 are greater than zero, by (A.49) it must be that $\eta(t)=0$ beginning at time t_0. From (A.48) it is clear that $\lambda(t)=\lambda_0 e^{-rt}$ + a constant. Suppose that $0 < S(t) < 1$. Inserting the first line of (A.46) into (A.47) ultimately yields $\mu(t) = \mu_0 e^{-qt}$ + a constant, where $q = bD/(1+d)$. But then (A.46) cannot hold in general.[5] Thus $S(t) = 1$, at least until assets are exhausted. Schooling must initially be full-time. Suppose that assets are exhausted before time t_1. At that time, call it t", consumption will equal current income, unless the household starts saving again. If saving does occur, consumption falls below current income, the child continues to attend school full-time, and assets will accumulate indefinitely (because $\eta(t)=0$ when $A(t)>0$, (A.29) will hold, causing consumption to fall indefinitely, while in fact income will rise) until time t_1. But this holding of positive assets at time t_1 is not optimal, because consumption immediately before time t_1 is below that immediately after t_1, and at the margin assets scheduled to be saved until time t_1 (and consumed afterwards) contribute less to life-cycle utility than they would if they were consumed before t_1, since $U'(C(t))$ is higher at a lower value of $C(t)$ and the rate of time preference (δ) is greater than the rate of growth in assets

[5]It will only hold for more than one instant of time if $bD/(1+d) = r$. But it was shown in the uncontrained case that schooling is only profitable if $bD/(1+r) > r$.

(r).[6] Thus after time t" the household consumes current income and does not save.

If household assets are exhausted before time t_1, no savings should be undertaken but another choice arises - the household may either send the child to school full-time or part-time. More specifically, at time t_1 the child may be attending full-time or part-time. Without further information on U, either could be possible. Suppose it is part-time, that is $0 < S(t_1) < 1$. According to Tomiyama (1985), at the switching point t_1 the Hamiltonians for periods 2 and 3 must be equal in the limit as t approaches t_1 from both sides.[7] Setting the Hamiltonians equal to each other at time t_1 yields (using A.46):

$$\text{(A.50)} \quad \frac{U(y + wDH(t_1^+)) - U(y - f + wH(t_1^-)[D - S(t_1^-)(1+d)])}{U'(y - f + wH(t_1^-)[D - S(t_1^-)(1+d)])} = w(1+d)H(t_1^-)S(t_1^-)$$

where "+" indicates approaching t_1 from above (period 3) and "-" indicates from above (period 2). Clearly, if $f > 0$ then $S(t_1^-)$ must approach a number greater than zero (in the limit) as t approaches t_1 from below. This implies that there will be a "jump" from part-time schooling to no schooling at all at t_1 (S(t) will not be continuous). The existence of a school fee means that S(t) cannot decline gradually until it equals zero, since for very small levels of S(t) the cost in terms of the school fee outweighs any benefit of additional human capital accumulation.

It is possible that assets will not run out before time t_1. In this case the child must attend school full-time all the way up to time t_1, but assets must

---

[6]Formally, it is easy to show that $\delta[U(ye^{-r\epsilon} + wH(t_1))e^{\delta(t+\epsilon)}]/\delta y$ is always less than $\delta[U(y - f - A'(t) + wH(t)S(t)[D - (1+d)])e^{-\delta t}]/\delta y$, where ϵ is a fixed amount of time.

[7]$\mu(t)$ must also be continuous in the same way, but $\lambda(t)$ is not continuous and in fact is irrelevent because all assets are exhausted and it is never optimal to build them up.

run out again exactly at time t_1. The reason they must run out is that, if they did not, consumption of those assets in period 3 would not be optimal relative to consuming them in period 1, when consumption is lower both because non-asset income is lower and because saving is taking place.

Finally, examine the first time period ($t_0 \leq t \leq t_1$). The Hamiltonian here is simply:

$$\text{(A.51)} \qquad H_1 = U(C(t))e^{-\delta t} + \lambda(t)[rA(t) + y + DwH_0 - C(t)] + \eta(t)A(t)$$

The necessary conditions that hold for time 0 to t_0 are equations (A.45), (A.48) and (A.49). Assuming that it is optimal for a period 1 to exist (this may not be the case for some households), some savings must take place before t_0, which in turn implies that consumption must decline during the first time period. To show that this is the case, examine what happens if no assets are accumulated before time t_0. Tomiyama's condition that the Hamiltonians must be equal at time t_0 yields:

$$\text{(A.52)} \qquad U(C(t_0^-))e^{-\delta t_0^-} = U(C(t_0^+))e^{-\delta t_0^+} + \mu(t_0^+)bH_0S(t_0^+))$$

$$+ \lambda(t_0^+)\,[y - f + wH_0[D-S(t_0^+)(1+d) - C(t)]$$

If assets are accumulated beginning at point t_0, by (A.29) consumption will drop further and assets will then be accumulated indefinitely. This will not be optimal because assets will remain at time t_1. Thus no assets will be accumulated, which implies that the term in brackets in (A.52) equals zero. It is clear from (A.45) and (A.46) that μ is positive when approaching t_0 from above, which implies that consumption levels on both sides of equation (A.52) are not equal. But this violates the other condition of Tomiyama that $\lambda(t_0^-) = \lambda(t_0^+)$

(since (A.45) holds in both periods 1 and 2). Thus it must be the case that some savings takes place before time t_0. When savings begins, call it time t_s, consumption declines and continues to decline because (A.29) will hold and $\eta(t)=0$. At t_0 there will be no jump in consumption because $\lambda(t_0^-) = \lambda(t_0^+)$.

To conclude for the positive school fees case, we have shown that it is plausible for late starting to be optimal, which was never the case when school fees equal zero. Whether or not it actually occurs depends on the parameter values.

Appendix 2: Estimations Not Reported in the Text

Table A1: Determinants of Score on Raven Test

Independent Variable	Coefficient	t-statistic
Age	0.3654	4.88
Age^2	-0.0050	-4.49
School	0.8820	4.92
$School^2$	0.0179	2.28
Age x School	-0.0060	-2.46
Father's Schooling	-0.0225	-0.50
Mother's Schooling	-0.1272	-1.39
Sex	-0.3159	-0.47
Sex x Age	-0.0146	-0.55
Sex x School	-0.2375	-3.02
Sample size = 1736		
R^2 = 0.9696		

Note: The dependent variable has a mean of 18.66 and a standard deviation of 4.54.

Table A2: Estimates of Semi-Parsimonious Achievement Test Scores

	Mathematics		Reading	
Household Variables	Coefficient	t-statistic	Coefficient	t-statistic
Constant	2.2968	0.17	-27.9533	-2.41
Forest	-3.5051	-2.98	-4.8680	-2.90
Savannah	-1.2942	-0.70	.--	.--
Semi-Urban	.--	.--	6.6673	2.58
Rural	.--	.--	2.6219	1.19
Sex	-3.9427	-4.32	-2.6959	-2.12
Schooling of Child (years)	1.5197	3.87	2.3443	4.29
Mother's Schooling (years)	0.4609	4.06	0.2979	1.86
Father's Schooling (years)	-0.0724	-0.86	-0.1215	-0.94
Ability Estimate	.--	.--	0.4614	2.42
Primary Reading Score	.--	.--	0.5434	1.33
Muslim	.--	.--	-2.6875	-1.26
Catholic	.--	.--	-2.1052	-1.15
Traditional Religion	.--	.--	-3.5156	-1.84
Akan	.--	.--	1.4578	0.81
Ewe	.--	.--	5.3548	2.27
Ga-Adangbe	.--	.--	4.4358	1.05
Male Agricultural Wage	.--	.--	0.0141	1.14
Wage Data Missing	.--	.--	4.6870	1.07
Child Order	0.7397	2.33	0.6090	1.42
Siblings	-0.4266	-1.54	0.5393	1.44
Log Per Capita Expenditure	-1.4758	-1.44	.--	.--
School Variables				
Travel Time	0.0255	0.83	.--	.--
Avg. Teacher Experience (yrs)	-0.0902	-0.78	.--	.--
Avg. Teacher Schooling (years)	0.6438	1.65	-0.4274	-0.73
Avg. Teacher Training (years)	.--	.--	1.1682	1.16
Blackboard	8.4334	2.43	11.6512	1.97
Library	3.7461	2.01	-2.3233	-0.84
Books per Room (middle)	-0.0651	-1.49	-0.0489	-0.70
Books per Room (JSS)	.--	.--	0.0405	1.65
Lack of Desks	.--	.--	3.1594	1.04
JSS Only	-4.0497	-2.07	.--	.--
Shed Classrooms	-3.9187	-1.77	.--	.--
Unusable Classrooms	3.3819	1.29	-6.5465	-1.75
Leaking Classrooms	-6.4200	-3.04	-9.7815	-3.28
Enrollment Fees	0.0085	1.79	-0.0080	-1.08
No Electricity	.--	.--	-3.4668	-1.60
Private	.--	.--	11.7328	1.47
Selectivity Correction Terms				
Lambda (Nearest School)	0.7000	0.68	1.3625	0.83
Lambda (Other Local)	.--	.--	1.2852	1.54
Lambda (School Attainment)	.--	.--	.--	.--
Lambda (Late Starting)	0.4991	1.27	0.4512	0.84
Sample Size	163		163	
R^2	0.3932		0.5368	

Table A3: Determinants of Math Achievement (Probit)

	Full Model		Parsimonious Model		Streamline Model	
Household Variables	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Constant	8.3883	0.34	7.1088	0.50	5.0535	0.39
Coast	-3.1267	-0.81	-1.9609	-0.81	.--	.--
Forest	-6.8840	-1.85	-5.2084	-2.44	-3.5265	-3.09
Savannah	-3.5532	-0.87	-2.7093	-1.04	-1.0522	-0.58
Semi-Urban	1.4269	0.59	1.0308	0.79	.--	.--
Rural	0.6459	0.31	.--	.--	.--	.--
Sex	-3.4959	-2.22	-3.8338	-4.15	-3.7446	-4.18
Schooling of Child	1.6505	2.96	1.5162	3.82	1.5237	3.91
Mother's Schooling (years)	0.4006	2.63	0.4561	4.00	0.4242	4.11
Father's Schooling (years)	-0.2010	-1.26	-0.0803	-0.95	.--	.--
Ability Estimate	0.0669	0.33	.--	.--	.--	.--
Ability Data Missing	-0.6527	-0.49	.--	.--	.--	.--
Primary Math Score	0.0719	0.17	.--	.--	.--	.--
Muslim	0.7155	0.16	.--	.--	.--	.--
Catholic	1.5544	0.33	.--	.--	.--	.--
Protestant	1.7357	0.38	.--	.--	.--	.--
Other Christian	0.4269	0.10	.--	.--	.--	.--
Traditional Religion	-1.2727	-0.30	.--	.--	.--	.--
Akan	0.6130	0.35	.--	.--	.--	.--
Ewe	-0.1567	-0.07	.--	.--	.--	.--
Ga-Adangbe	-0.4800	-0.13	.--	.--	.--	.--
Male Agricultural Wage	-0.0019	-0.18	.--	.--	.--	.--
Wage Data Missing	-0.8547	-0.22	.--	.--	.--	.--
Child Order	0.5470	1.34	0.6935	2.21	0.7371	2.41
Siblings	-0.4769	-1.16	-0.4433	-1.61	-0.4813	-1.78
Log Per Capita Expenditures	-2.1977	-1.24	-1.8843	-1.79	-1.6692	-1.70
Land Ownership	0.0014	0.24	.--	.--	.--	.--
School Variables						
Travel Time	0.0227	0.48	.--	.--	.--	.--
Average Teacher Exp. (years)	-0.0790	-0.52	-0.0795	-0.67	.--	.--
Average Teacher Sch. (years)	0.6017	1.13	0.6673	1.64	0.5077	1.34
Average Teacher Trg. (years)	0.3491	0.34	.--	.--	.--	.--
Blackboard	12.6661	2.30	9.7825	2.51	8.9489	2.62
Library	4.4424	1.80	3.4558	1.85	3.2678	1.83
Books Per Room (Middle)	-0.0686	-1.06	-0.0662	-1.51	-0.0586	-1.38
Books Per Room (JSS)	-0.0124	-0.48	.--	.--	.--	.--
Lack of Desks	0.5850	0.19	.--	.--	.--	.--
JSS Only	-3.0084	-0.99	-4.7703	-2.27	-4.1378	-2.13
Shed Classrooms	-4.5796	-1.41	-4.0100	-1.78	-3.4002	-1.58
Unusable Classrooms	4.2331	0.98	3.0405	1.14	3.7288	1.46
Leaking Classrooms	-8.2166	-3.00	-6.6882	-3.13	-6.2007	-2.98
Enrollment Fees	0.0085	1.12	0.0085	1.70	0.0081	1.77
Deny Admission	-0.5374	-0.23	.--	.--	.--	.--
No Water	0.7441	0.24	.--	.--	.--	.--
No Electricity	0.2167	0.09	.--	.--	.--	.--
Private School	-2.3748	-0.34	.--	.--	.--	.--
Sigma 21	0.7077	0.62	0.7605	1.06	0.7238	1.02
Sigma 31	-0.6048	-0.27	.--	.--	.--	.--
Lambda (Optimal Years)	-1.4197	-0.29	.--	.--	.--	.--
Lambda (Late Start)	0.6115	1.33	0.5055	1.28	0.4537	1.17
R^3	0.4364		0.4006		0.3901	
Sample Size	163		163		163	

Table A4: Determinants of Reading Achievement (Probit)

	Full Model		Parsimonious Model		Streamline Model	
Household Variables	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Constant	-33.3266	-1.08	-29.2886	-2.49	-33.5762	-4.61
Coast	-2.2279	-0.46	.--	.--	.--	.--
Forest	-6.4348	-1.39	-4.3904	-2.58	-4.2360	-2.93
Savannah	-0.6127	-0.12	.--	.--	.--	.--
Semi-Urban	7.3225	2.39	6.9916	2.67	6.2935	2.77
Rural	3.1040	1.21	2.8464	1.28	2.3328	1.15
Sex	-2.0051	-0.98	-2.5094	-1.94	-2.6683	-2.19
Schooling of Child	2.5049	3.47	2.3025	4.21	2.3483	4.53
Mother's Schooling (years)	0.2967	1.52	0.2981	1.85	0.2918	1.86
Father's Schooling (years)	-0.1753	-0.85	-0.1459	-1.12	-0.0930	-0.75
Ability Estimate	0.5356	1.98	0.5123	2.72	0.4863	2.68
Ability Data Missing	0.1038	0.06	.--	.--	.--	.--
Primary Reading Score	0.6128	0.94	0.5123	1.81	0.6508	1.68
Muslim	-4.3070	-0.72	-2.3774	-1.15	-2.4289	-1.21
Catholic	-2.9393	-0.48	.--	.--	.--	.--
Protestant	-0.7721	-0.13	.--	.--	.--	.--
Other Christian	-0.1881	-0.03	.--	.--	.--	.--
Traditional Religion	-3.5994	-0.65	-2.5349	-1.40	-2.3120	-1.32
Akan	1.3179	0.58	2.1936	1.22	1.8459	1.09
Ewe	4.9405	1.66	5.6352	2.40	4.5571	2.23
Ga-Adangbe	4.2808	0.88	5.6169	1.36	5.9955	1.49
Male Agricultural Wage	0.0115	0.81	0.0181	1.44	0.0101	0.91
Wage Data Missing	4.3743	0.89	5.7591	1.30	3.3911	0.84
Child Order	0.5420	1.04	0.5929	1.36	0.4886	1.21
Siblings	0.6350	1.19	0.5029	1.33	0.5847	1.68
Log Per Capita Expenditures	0.4836	0.20	.--	.--	.--	.--
Land Ownership	-0.0011	-0.15	.--	.--	.--	.--
School Variables						
Travel Time	-0.0274	-0.40	.--	.--	.--	.--
Average Teacher Exp. (years)	-0.1111	-0.57	-0.1344	-0.77	.--	.--
Average Teacher Sch. (years)	-0.4002	-0.57	-0.4547	-0.76	.--	.--
Average Teacher Trg. (years)	1.1139	0.92	1.3772	1.37	1.3446	1.43
Blackboard	11.5771	1.57	10.8373	1.84	9.2053	1.77
Library	-3.0889	-0.98	-3.5471	-1.30	-3.9961	-1.51
Books Per Room (Middle)	-0.0652	-0.78	-0.0412	-0.59	.--	.--
Books Per Room (JSS)	0.0345	1.04	0.0383	1.35	0.0383	1.71
Lack of Desks	4.6594	1.16	2.7775	0.93	.--	.--
JSS Only	0.4771	0.13	.--	.--	.--	.--
Shed Classrooms	-0.8854	-0.21	.--	.--	.--	.--
Unusable Classrooms	-8.3231	-1.47	-7.0380	-1.87	-6.4239	-1.77
Leaking Classrooms	-8.4940	-2.47	-9.0372	-3.06	-7.3547	-2.75
Enrollment Fees	-0.0100	-0.99	-0.0080	-1.09	-0.0059	-0.91
Deny Admission	1.2117	0.41	.--	.--	.--	.--
No Water	1.8159	0.41	.--	.--	.--	.--
No Electricity	-2.9721	-0.89	-3.4458	-1.52	-3.4436	-1.70
Private School	12.0310	1.34	13.1730	1.66	14.8195	1.94
Sigma 21	1.0061	0.63	1.0534	1.01	0.6288	0.66
Sigma 31	0.9617	0.35	.--	.--	.--	.--
Lambda (Optimal Years)	-1.9153	-0.30	.--	.--	.--	.--
Lambda (Late Start)	0.3798	0.64	0.3870	0.72	.--	.--
R^3	0.5400		0.5265		0.5147	
Sample Size	163		163		163	

Distributors of World Bank Publications

ARGENTINA
Carlos Hirsch, SRL
Galeria Guemes
Florida 165, 4th Floor-Ofc. 453/465
1333 Buenos Aires

AUSTRALIA, PAPUA NEW GUINEA, FIJI, SOLOMON ISLANDS, VANUATU, AND WESTERN SAMOA
D.A. Books & Journals
648 Whitehorse Road
Mitcham 3132
Victoria

AUSTRIA
Gerold and Co.
Graben 31
A-1011 Wien

BANGLADESH
Micro Industries Development Assistance Society (MIDAS)
House 5, Road 16
Dhanmondi R/Area
Dhaka 1209

Branch offices:
156, Nur Ahmed Sarak
Chittagong 4000

76, K.D.A. Avenue
Kulna 9100

BELGIUM
Jean De Lannoy
Av. du Roi 202
1060 Brussels

CANADA
Le Diffuseur
C.P. 85, 1501B rue Ampère
Boucherville, Québec
J4B 5E6

CHINA
China Financial & Economic Publishing House
8, Da Fo Si Dong Jie
Beijing

COLOMBIA
Infoenlace Ltda.
Apartado Aereo 34270
Bogota D.E.

COTE D'IVOIRE
Centre d'Edition et de Diffusion Africaines (CEDA)
04 B.P. 541
Abidjan 04 Plateau

CYPRUS
Cyprus College Bookstore
6, Diogenes Street, Engomi
P.O. Box 2006
Nicosia

DENMARK
SamfundsLitteratur
Rosenoerns Allé 11
DK-1970 Frederiksberg C

DOMINICAN REPUBLIC
Editora Taller, C. por A.
Restauración e Isabel la Católica 309
Apartado de Correos 2190 Z-1
Santo Domingo

EGYPT, ARAB REPUBLIC OF
Al Ahram
Al Galaa Street
Cairo

The Middle East Observer
41, Sherif Street
Cairo

EL SALVADOR
Fusades
Alam Dr. Manuel Enrique Araujo #3530
Edificio SISA, ler. Piso
San Salvador 011

FINLAND
Akateeminen Kirjakauppa
P.O. Box 128
SF-00101 Helsinki 10

FRANCE
World Bank Publications
66, avenue d'Iéna
75116 Paris

GERMANY
UNO-Verlag
Poppelsdorfer Allee 55
D-5300 Bonn 1

GUATEMALA
Librerias Piedra Santa
5a. Calle 7-55
Zona 1
Guatemala City

HONG KONG, MACAO
Asia 2000 Ltd.
46-48 Wyndham Street
Winning Centre
2nd Floor
Central Hong Kong

INDIA
Allied Publishers Private Ltd.
751 Mount Road
Madras - 600 002

Branch offices:
15 J.N. Heredia Marg
Ballard Estate
Bombay - 400 038

13/14 Asaf Ali Road
New Delhi - 110 002

17 Chittaranjan Avenue
Calcutta - 700 072

Jayadeva Hostel Building
5th Main Road Gandhinagar
Bangalore - 560 009

3-5-1129 Kachiguda Cross Road
Hyderabad - 500 027

Prarthana Flats, 2nd Floor
Near Thakore Baug, Navrangpura
Ahmedabad - 380 009

Patiala House
16-A Ashok Marg
Lucknow - 226 001

Central Bazaar Road
60 Bajaj Nagar
Nagpur 440010

INDONESIA
Pt. Indira Limited
Jl. Sam Ratulangi 37
P.O. Box 181
Jakarta Pusat

ISRAEL
Yozmot Literature Ltd.
P.O. Box 56055
Tel Aviv 61560
Israel

ITALY
Licosa Commissionaria Sansoni SPA
Via Duca Di Calabria, 1/1
Casella Postale 552
50125 Firenze

JAPAN
Eastern Book Service
Hongo 3-Chome, Bunkyo-ku 113
Tokyo

KENYA
Africa Book Service (E.A.) Ltd.
Quaran House, Mfangano Street
P.O. Box 45245
Nairobi

KOREA, REPUBLIC OF
Pan Korea Book Corporation
P.O. Box 101, Kwangwhamun
Seoul

MALAYSIA
University of Malaya Cooperative Bookshop, Limited
P.O. Box 1127, Jalan Pantai Baru
59700 Kuala Lumpur

MEXICO
INFOTEC
Apartado Postal 22-860
14060 Tlalpan, Mexico D.F.

NETHERLANDS
De Lindeboom/InOr-Publikaties
P.O. Box 202
7480 AE Haaksbergen

NEW ZEALAND
EBSCO NZ Ltd.
Private Mail Bag 99914
New Market
Auckland

NIGERIA
University Press Limited
Three Crowns Building Jericho
Private Mail Bag 5095
Ibadan

NORWAY
Narvesen Information Center
Book Department
P.O. Box 6125 Etterstad
N-0602 Oslo 6

PAKISTAN
Mirza Book Agency
65, Shahrah-e-Quaid-e-Azam
P.O. Box No. 729
Lahore 54000

PERU
Editorial Desarrollo SA
Apartado 3824
Lima 1

PHILIPPINES
International Book Center
Fifth Floor, Filipinas Life Building
Ayala Avenue, Makati
Metro Manila

POLAND
ORPAN
Palac Kultury i Nauki
00-901 Warzawa

PORTUGAL
Livraria Portugal
Rua Do Carmo 70-74
1200 Lisbon

SAUDI ARABIA, QATAR
Jarir Book Store
P.O. Box 3196
Riyadh 11471

SINGAPORE, TAIWAN, MYANMAR, BRUNEI
Information Publications Private, Ltd.
02-06 1st Fl., Pei-Fu Industrial Bldg.
24 New Industrial Road
Singapore 1953

SOUTH AFRICA, BOTSWANA
For single titles:
Oxford University Press Southern Africa
P.O. Box 1141
Cape Town 8000

For subscription orders:
International Subscription Service
P.O. Box 41095
Craighall
Johannesburg 2024

SPAIN
Mundi-Prensa Libros, S.A.
Castello 37
28001 Madrid

Librería Internacional AEDOS
Consell de Cent, 391
08009 Barcelona

SRI LANKA AND THE MALDIVES
Lake House Bookshop
P.O. Box 244
100, Sir Chittampalam A. Gardiner Mawatha
Colombo 2

SWEDEN
For single titles:
Fritzes Fackboksforetaget
Regeringsgatan 12, Box 16356
S-103 27 Stockholm

For subscription orders:
Wennergren-Williams AB
Box 30004
S-104 25 Stockholm

SWITZERLAND
For single titles:
Librairie Payot
1, rue de Bourg
CH 1002 Lausanne

For subscription orders:
Librairie Payot
Service des Abonnements
Case postale 3312
CH 1002 Lausanne

TANZANIA
Oxford University Press
P.O. Box 5299
Maktaba Road
Dar es Salaam

THAILAND
Central Department Store
306 Silom Road
Bangkok

TRINIDAD & TOBAGO, ANTIGUA BARBUDA, BARBADOS, DOMINICA, GRENADA, GUYANA, JAMAICA, MONTSERRAT, ST. KITTS & NEVIS, ST. LUCIA, ST. VINCENT & GRENADINES
Systematics Studies Unit
#9 Watts Street
Curepe
Trinidad, West Indies

UNITED KINGDOM
Microinfo Ltd.
P.O. Box 3
Alton, Hampshire GU34 2PG
England

VENEZUELA
Libreria del Este
Aptdo. 60.337
Caracas 1060-A

The World Bank

Headquarters
1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.

Telephone: (202) 477-1234
Facsimile: (202) 477-6391
Telex: WUI 64145 WORLDBANK
RCA 248423 WORLDBK
Cable Address: INTBAFRAD
WASHINGTONDC

European Office
66, avenue d'Iéna
75116 Paris, France

Telephone: (1) 40.69.30.00
Facsimile: (1) 40.69.30.66
Telex: 640651

Tokyo Office
Kokusai Building
1-1, Marunouchi 3-chome
Chiyoda-ku, Tokyo 100, Japan

Telephone: (3) 3214-5001
Facsimile: (3) 3214-3657
Telex: 26838