

PRIMARY SCHOOL ATTENDANCE IN HONDURAS

Arjun S. Bedi
and
Jeffery H. Marshall

August 2000

Working Paper Series No. 320

ABSTRACT

Honduras has recorded impressive gains in expanding educational access in the 1990s, with the result that primary education is available to almost all children. With improved access the focus has shifted to quality and efficiency issues. Previous research suggests that academic achievement is still quite low while repetition and school desertion rates continue to remain high. An important cause of these outcomes appears to lie in patterns of school attendance. Low levels of school attendance may be responsible for low academic achievement, which in turn is linked to high repetition and desertion rates. Recognizing this probable chain of events, this paper focuses on the school attendance decision. We rely on recently collected data from a national sample of Honduran primary schools to specify and estimate a model of school attendance. We find that increases in the expected benefits of attending school exert a strong impact on the school attendance decision.

CONTENTS

0

1. Introduction	1
2. Costs, Benefits and School Attendance – An Analytical Framework	5
2.1. Empirical Specification	6
2.2. Costs of attending school	7
2.3. Benefits of attending school	8
3. Data Description and Specification	9
4. Results	13
4.1. School Attendance Patterns	13
4.2. Educational Production Functions	16
4.3. School Attendance	18
5. Conclusion	21
References	22
Tables	23

1. INTRODUCTION

Investments in education are widely recognized as a key component of a country's development strategy. Increases in the quantity and quality of educational provision have been associated with a wide range of benefits including enhanced productivity, reduced poverty and income inequality, improved health and economic growth.¹ Spurred by such evidence, governments in developing countries devote a substantial fraction of their total expenditure to the education sector.²

Honduras is no exception, as successive governments have invested substantial resources in education. Between 1993 and 1996 public expenditure on education accounted for 16.5% of total government outlays (UNDP, 1999). These investments have expanded coverage and access at all levels and, as a result, the gross enrollment ratio for Hondurans aged 6-23 (in primary, secondary and tertiary education) has risen from 47.0% in 1980 to 60.0% in 1995 (UNESCO, 1999).³ The expansion in educational opportunities has been especially notable in the pre-primary and primary sectors, and recent initiatives have targeted the most needy populations in Honduras.⁴

However, these impressive gains in coverage have been tempered somewhat by low levels of academic achievement and high rates of grade repetition and desertion. A recent national application of criterion-referenced Spanish and mathematics exams resulted in national averages in grade 3 of 39.7% and 35.9% in Spanish and mathematics, respectively

¹ See Lockheed, Verspoor and associates (1991) for a detailed review.

² According to UNDP (1999), between 1993 and 1996 the average (unweighted) expenditure on education as a percentage of total government expenditure for developing countries was 14.8 %.

³ Gross enrollment ratios at the primary and secondary level were 112% and 32% respectively. For a more detailed overview of Honduran educational coverage, see Edwards (1995), Edwards, Fuller, and Parandekar (1996) and Van Steenwyck and Mejia (1996).

⁴ Examples include the expansion of community-based preschools (*Centros Comunitarios de Iniciación Escolar*—CCIEs) that are designed to reach rural populations (World Bank, 1995; UNICEF, 1992) and the recent creation of 500 community-based primary schools (*Proyecto Hondureño de Educación Comunitaria*—PROHECO) based on the EDUCO (*Educación con Participación de la Comunidad*) program in El Salvador.

(UMCE, 1998).⁵ In 1996, official grade repetition rates were 19.4% in grade 1, 11.6% in grade 2, and 8.6% in grade 3, although real rates are likely to be even higher (Van Steenwyck, 1997; Marshall, 2000). Although difficult to calculate, the World Bank (1995) estimates that about 5.0% of Honduran primary school students desert annually.

While part of the reason for low achievement and consequently high repetition and desertion rates may lie in student ability, it is likely that such outcomes are largely driven by other factors. In a review of the Honduran education system the World Bank identified low attendance rates and poor school inputs as the two main factors responsible for high repetition rates, which in turn was identified as the most important cause of the high dropout rate (World Bank, 1995).

The economic contribution of children to families in developing countries (especially in rural areas) and accordingly the opportunity cost associated with school attendance may be substantial. Attendance will suffer when parents perceive that the return associated with time spent in school does not justify the loss of a child's economic contribution. Parental perceptions of school inputs may also affect the attendance cost-benefit calculus as low quality teachers or limited availability of teaching materials may attenuate the expected benefits from attending school. A reduction in days attended probably exerts a negative influence on academic achievement and increases the probability of repetition and desertion. There may also be a more direct link between school attendance and repetition as some school systems require minimum levels of attendance before allowing a student to appear for exams (see Jacoby, 1994).

Recognition of these indirect and direct links between school attendance and educational outcomes suggests that a focus on the factors underlying the school attendance

⁵ The UMCE exams are designed to measure the implemented curriculum and are not intended for pass/fail decisions, so deciding on a cutoff point for passing and failing is an inherently arbitrary exercise. Nevertheless, the exams do pretend to measure "mastery" for each curriculum component that is covered, which usually include three questions per component. Students who can answer at least two of the three questions correctly are considered to have mastered the component. These aggregated averages show that few Honduran primary students are mastering a significant portion of the curriculum, according to the UMCE standards. Furthermore, since the UMCE exams comprise multiple-choice questions with four options, averages below 40 percent are indicative of low levels of achievement.

decision itself may be quite useful in understanding the dynamics of human capital formation

in a developing country. While there is a substantial amount of literature that has examined the determinants of school enrollment, test scores, grade repetition and desertion there is limited work on the factors that motivate school attendance (for a review see Strauss and Thomas, 1995).⁶ A handful of authors have included a measure of school attendance as an explanatory variable in educational production functions. One example is Tan, Lane and Coustere (1997), who use data from the Philippines and find that the number of days missed has a negative and statistically significant effect on mathematics test scores. Fuller, et al. (1999) also include this variable as a regressor in their analysis of the determinants of academic achievement in Brazil, although the point estimates are not significant. While including attendance as an explanatory variable does highlight the influence of attendance on test scores it ignores the potential endogeneity of school attendance and achievement and does not permit an analysis of the factors that determine school attendance.⁷

This paper builds on our earlier work (Bedi and Marshall, 1999) and focuses primarily on the school attendance decision. We use nationally representative data collected by the External Unit for the Measurement of School Quality (*Unidad Externa de Medición de la Calidad de la Educación* – UMCE) in Honduras to specify and estimate a model of primary school attendance. We assume that parents determine the particular pattern of school attendance for their children on the basis of expected gains and the costs of attending school. We proceed in two steps. First, we estimate the effect of child, family and school characteristics on test scores and obtain predicted test scores. In the second step we estimate a school attendance model which includes predicted test scores as a measure of the expected gains of school attendance. This paper improves on our earlier work in two ways. First, the UMCE database includes a nationally representative sample of primary schools, while our earlier work was based on a limited sample drawn from one province of rural Honduras, which limited the generalizability of our results. Second, the measure of

⁶ Enrolling in school is clearly the first step before one may begin to examine patterns of school attendance. In the Honduran context where there is almost universal enrollment the important issue is not whether a child is enrolled in school but how often does a child attend school.

⁷ Although the main aim of their paper is to compare test scores across school types, Jimenez and Sawada (1999) do provide estimates of school attendance regressions that allow them to examine the determinants of attendance as well as make comparisons across school types.

school attendance that we use in this paper allows a decomposition of school attendance into demand and supply components. This feature of the data is discussed in a subsequent section.

The next section of the paper introduces the analytical framework that we use to motivate our empirical work. Section III describes the data and the variables used in the analysis. Section IV presents results and section V concludes.

2. COSTS, BENEFITS AND SCHOOL ATTENDANCE –AN ANALYTICAL FRAMEWORK

School attendance patterns in developing countries vary substantially across households. Some children may never enter school while others may attend only part-time. The degree of part-time schooling may vary from missing a few weeks to missing several months. The variation in attendance patterns suggests that parents evaluate differently the costs and benefits of attending school and that this evaluation for the same household may also vary according to the particular time of the year. For instance, during the harvest season the opportunity costs of attending school may far outweigh the benefits, resulting in temporary withdrawal, while at other times the benefits may outweigh the costs and result in regular school attendance.⁸ Thus, school attendance over the year may be viewed as the consequence of a daily household decision where a child attends school on a particular day if the expected benefits from attending school on that day are greater than the associated costs.

To formalize these notions and to motivate our empirical work this section presents a framework tailored to our needs.⁹ Consider that the school year consists of n days and it is day i of the school year. We assume that each household has a utility function defined over b_i and c_i , where b_i denotes the benefits associated with attending school on day i , and c_i is

⁸ This is especially true in rural Honduras where children may drop out of school for several months during the harvest season only to return the next year. See World Bank (1995), p. 9. School attendance patterns in our data are discussed later on in the text.

⁹ The framework used here is similar to those in Gertler and Van Der Gaag (1988) and Gertler and Glewwe (1992).

household consumption on day i . While attending school yields benefits it comes at a cost. Direct and opportunity costs associated with school attendance lower resources available for household consumption. Accordingly, household utility on day i conditional on school attendance (denoted by subscript 1) is given as,

$$U_{i1} = U(b_i, c_{i1}). \quad (1)$$

The associated budget constraint is,

$$y_i = c_{i1} + p_i, \quad (2)$$

where y_i is household income, and p_i represents the total cost associated with school attendance.

In a similar fashion the utility associated with not attending school may be defined by,

$$U_{i0} = U(c_{i0}). \quad (3)$$

The budget constraint is $y_i = c_{i0}$. Given the utility associated with both options, households choose the option that yields the highest utility. The solution to the daily unconditional utility maximizing problem is

$$U_i^* = \max(U_{i1}, U_{i0}), \quad (4)$$

where U_i^* is the maximum utility. Alternatively, school attendance may be defined in terms of a dichotomous variable, a_i , where $a_i = 1$ if a child attends school and 0 otherwise. A child attends school i.e., $a_i = 1$ if $U_{i1} > U_{i0}$. Summing up the outcomes of these daily decisions, over the school year, leads to the observed pattern of school attendance.

2.1. Empirical Specification

Since our purpose is to empirically explore the role of expected gains and costs on the school attendance decision we proceed by specifying linear forms of the conditional utility function. For the schooling option,

$$U_{i1} = \beta_1 b_i + \beta_2 c_{i1} + \varepsilon_{i1} \quad (5)$$

where the β 's are coefficients to be estimated and ε_{i1} is assumed to be a mean zero, normally distributed error term with positive variance. Since $c_{i1} = y_i - p_i$, we may rewrite (5) to obtain,

$$U_{i1} = \beta_1 b_i + \beta_2 (y_i - p_i) + \varepsilon_{i1} \quad (6)$$

The utility function for the non-schooling option is,

$$U_{i0} = \beta_2 y_i + \varepsilon_{i0} \quad (7)$$

Thus, an individual attends school, i.e. $a_i = 1$ if $\beta_1 b_i - \beta_2 p_i + \varepsilon_{i1} - \varepsilon_{i0} > 0$.

The chances of attending school on a particular day may be expressed in terms of a linear probability model that may be written as,

$$a_i = \beta_1 b_i - \beta_2 p_i + \varepsilon_{ia}, \quad (8)$$

where ε_{ia} is a normally distributed, mean zero, positive variance composite error term.¹⁰

Equation (8) depicts the probability of attending school on any particular day. Since we are interested in the yearly pattern of school attendance we may sum up the outcome of the daily attendance decision over the course of the school year,

$$\sum_{i=1}^n a_i = \sum_{i=1}^n (\beta_1 b_i - \beta_2 p_i + \varepsilon_{ia}) \quad (9)$$

to yield,

$$A = \beta_1 B + \beta_2 P + \varepsilon_A, \quad (10)$$

where yearly school attendance A , depends on B , the expected benefits associated with school attendance over the school year and P , the yearly costs of attending school.

2.2. Costs of attending school

The total cost (P) of sending a child to school includes monetary (direct) and indirect or opportunity costs. Since education is largely subsidized the main cost incurred by households is likely to be in the form of opportunity costs. Attending school reduces a child's availability for work in and outside the home. If a child makes substantial contributions to family income, or plays an important role in supporting other working members, then the opportunity cost of attending school is likely to be high and this may curtail the

¹⁰ In this linear utility specification, income has been differenced out of the decision rule and does not directly affect the school attendance decision. Household endowments are assumed to influence the school attendance decision through opportunity costs.

attractiveness of the schooling option.¹¹

Both of these cost components are likely to differ across households. For instance, direct costs may vary due to differences in transportation costs. Opportunity costs and the value of a child's time may also differ due to personal characteristics of the child (age, sex) and the value that parents place on a child's time. Since we do not directly observe the costs of attending school we allow P to depend on a vector of child, family and other characteristics that capture the cost of attending school.

2.3. Benefits of attending school

Parents have to ascertain the total benefits (B) associated with school attendance. We consider two types of benefits that may influence parental decision-making. The main benefit associated with attending school is likely to be the expected addition to a child's human capital. To capture this effect we need a measure of the human capital gains associated with school attendance. For this study we incorporate a measure that is widely used to indicate the benefits derived from education: test scores. However, using actual test scores is obviously incorrect due to the potential endogeneity between test scores and attendance. In order to derive an appropriate measure of human capital gains we proceed by estimating educational production functions, one for each subject, for those students for whom test score data are available. These test score equations are specified as,

$$H = \delta Z + \varepsilon_H, \tag{11}$$

where H is a measure of human capital or in this case test scores, Z is a vector of individual, family and school characteristics that influence H , and ε_H is an error term. Estimates from the educational production functions are used to predict test scores for each individual. These predicted values (\hat{H}) are included in equation (10) in order to capture the human capital benefits associated with attending school.

¹¹ For example, Patrinos and Psacharopoulos (1995) show that child earnings account for 27.8% of total income in urban households in Paraguay, while Patrinos and Psacharopoulos (1997) show that child labor contributes 17.7% of household income in rural Peru. A number of studies have also demonstrated that the presence of younger siblings in the household may affect educational outcomes. Using Honduran household survey data, Edwards, Fuller and Parandekar (1996) show that children are likely to delay initial enrollment, and attain fewer years of education, when an infant sibling is present in the home.

In addition to school characteristics that have an impact on test scores, there may be other school characteristics that do not affect academic achievement but do signal the quality (Q) of a school and directly influence the benefits that parents attribute to school attendance. For instance, whether a school has a telephone connection or a sports field may not directly influence academic achievement. However, these are easily observed signals that may be used by parents to judge the quality of a school and in turn may directly influence the benefits that parents associate with school attendance. Thus, some school inputs and facilities may directly influence parental evaluation of the benefits associated with school attendance while others may exert an influence on benefits through their impact on test scores.

To account for the different kinds of benefits that parents associate with school attendance, equation (10) may be adjusted to accommodate both the expected human capital benefits (\hat{H}) and direct benefits (Q) and may be rewritten as,

$$A = \beta_{1\hat{H}}\hat{H} + \beta_{2Q}Q + \beta_3P + \varepsilon_A, \quad (12)$$

where $\beta_{1\hat{H}}$ is a coefficient to be estimated and β_{2Q} and β_3 are conformable coefficient vectors to be estimated. As this equation depicts, school attendance is treated as a function of expected human capital benefits, other benefits and costs.

3. DATA DESCRIPTION AND SPECIFICATION

The data used in this paper are drawn from the second national application of standardized tests administered by the UMCE in October 1998 and March 1999.¹² The sample includes 586 schools from 17 of 18 states¹³ (*departamentos*) and represents approximately seven percent of Honduran primary schools. In this application Spanish and mathematics tests were administered to students in grades 2 and 4. In addition, second and fourth grade teachers were administered questionnaires, as were school directors in each school. Test

¹² The two primary functions of the UMCE are to develop and apply standardized tests covering the basic learning objectives and study the factors associated with academic achievement, especially in student cohorts. See UMCE (1997,1998) and World Bank (1995) for descriptions of the project.

¹³ *Gracias a Dios*, an extremely isolated state in eastern Honduras, was not included in the sample due to the relatively small number of schools and students in this region and the difficulty of school access.

administration personnel filled out observation instruments in each school detailing variables such as school type, enrollment, days worked, school characteristics (including materials and hardware) and special programs. They also copied student data including days missed during the school year, work attitudes, and Spanish and mathematics grades from teacher grade books in second and fourth grades. Finally, parents with children in second or fourth grades were interviewed to collect data on parental characteristics such as education and work experiences, attitudes towards education and specific problems in their school, among other variables. Further details on the data are available in UMCE (1999).

In this paper we restrict our analysis to students in second and fourth grade for whom we have complete information on test scores, attendance, child, family and school characteristics. These restrictions result in sample sizes of 7,026 for grade 2 Spanish and 6,934 for grade 2 mathematics. The grade 2 attendance equation is estimated over a sample of 6,135 observations. For grade 4 the sample sizes are 5,359 for Spanish, 5,024 for mathematics and 4,501 for the attendance equation.¹⁴ Descriptive statistics and variable definitions by grade are provided in Tables 1 and 2.

Turning to the empirical implementation, two educational production functions – one for each subject – are estimated for each grade, yielding a total of four equations. The dependent variables (H) are standardized test scores on Spanish and mathematics examinations. The independent variables are classified into child, family, teacher and classroom/school characteristics. The child-specific variables include age and sex, as well as an indicator of whether Spanish or mathematics is a child's favorite subject among the four main subjects that also include sciences and social studies. This variable is included to control for the effect that inclination towards a particular subject may have on test scores. Family characteristics include variables that reflect parental attitudes towards education, such as whether a child attended preschool, whether parents help children with their

¹⁴ The number of observations available for estimating the attendance equation estimations drops because in some schools we have data on children (including test scores), parents, teachers and schools but the teacher grade books were not available and therefore it was impossible to construct the attendance dependent variable.

homework and parental participation in school activities. Household resources are measured by mean years of education of the parents and an index of household wealth.

Several variables are included to control for the quality of instruction received by the child. An instructor's knowledge of the subject is captured by years of education, years of experience, the number of teaching seminars attended and test scores on Spanish and mathematics examinations. The test score variable is expected to be a more accurate and current measure of an instructor's knowledge. In an attempt to control for unobserved teaching attributes and skills our regressions include a self-reported measure of teacher's self-confidence. Class characteristics include the size of the class, an indicator of whether several grades are taught simultaneously in the same grade (i.e. a multigrade class room), and the availability of textbooks.

The dependent variable (A) in the school attendance equation (12) is the number of days that a child attends school during the school year. This measure is created by subtracting the number of days that a child misses school from the number of days the school was in operation during the 1998 school year. A detailed analysis of this variable is provided in the following section. Attendance is specified as a function of the benefits and costs associated with schooling. Benefits are represented by the expected human capital gains from schooling, \hat{H} , (obtained from the production function estimates) and by school facilities that signal the quality (Q) of the school and from which parents may directly derive benefits. The specific school facility variables include the presence of a telephone connection, an electricity connection, and a composite variable that sums other school facilities that a school may possess. Since we do not have any direct measures of opportunity costs (such as child wages) we specify opportunity costs (P) as a function of child and family characteristics. The child characteristics include the age and sex of the child. Variation in parental evaluation of child time is controlled by the same set of family characteristics that are included in the educational production functions. School related characteristics that may influence the cost of school attendance are the time taken to get to school and whether the school has a lunch program.

Before turning to the results a number of econometric issues must be dealt with. First, the educational production functions are estimated using data only on those students who attended school on the day the tests were administered. Using data only on test takers

may result in inconsistent estimates, since children who are not in attendance on the day of the test may not have the same characteristics as those who were in school on the day of the test.¹⁵ To account for this source of bias we estimate selection corrected educational production functions.¹⁶

Another concern while estimating educational production functions is the potential endogeneity of school inputs. If parents migrate in response to differences in school inputs, or if educational planners distribute school inputs to compensate for low student achievement, then estimates of the effect of school input characteristics may be biased. Endogeneity due to parental migration appears to be quite unlikely in the Honduran context. Analysis of household survey data indicates that 90 percent of individuals migrate in order to find work and the remaining migrate for family reasons (see Bedi, 1997). Endogeneity due to the second source is quite possible. However, examination of our data does not reveal any clear and consistent pattern between the distribution of school resources and the general economic characteristics of a region. For instance, in grade 2, schools in metropolitan areas have slightly larger class sizes (approximately 40 students) as compared to schools in towns (around 38 students), however test scores of teachers is slightly higher in

¹⁵ There are other potential sources of selection bias that may have a bearing on the academic achievement estimations. For instance, children may never enroll on school or many children may drop out of school before reaching grade 2/grade 4 and thus those who do enroll in school and do reach these grades may be regarded as a selective cohort of students. Since almost all children enroll in school any selection bias from this source is likely to be small. However, since a large number of students do leave the school system before reaching grade 2/grade 4 our analysis should be viewed as conditional and restricted to those individuals who have reached grade 2/grade 4.

¹⁶ While we have complete information on 7,206 individuals for grade 2 Spanish and 6,934 for grade 2 mathematics, there are around 4,500 additional students enrolled in grade 2 who did not take the exam. The selection-corrected specifications are estimated over this larger sample. We lose some of these observations since we do not have complete information on all of the explanatory variables. The selection correction estimates are based on sample sizes of 10,050 and 9,981 for grade 2 Spanish and grade 2 mathematics respectively. In grade 4 there are around 3,000 students who did not attend school on the day of the exam. Here we also lose some observations due to missing information on the explanatory variables and the selection corrected estimates for grade 4 Spanish and mathematics are estimated over sample sizes of 7,220 and 7,078 respectively.

metropolitan schools (85% versus 81 percent).¹⁷

Finally, the school specific nature of our data raises the possibility that students attending the same school (i.e. sharing the same observable characteristic) may also share the same unobservable characteristics, which may lead to the presence of intra-school error correlation (see Moulton, 1986). Although least squares is still consistent, the presence of these effects results in biased standard errors and consequently misleading statistical inference. To account for this, an appropriate robust variance-covariance matrix is computed, and the reported *t*-statistics for our various estimates are based on this adjusted matrix.

4. RESULTS

The results are divided into three sections. In the first part we examine patterns of school attendance and present a decomposition of attendance into demand and supply side components. The second section discusses the determinants of academic achievement in mathematics and Spanish, while the third section presents estimates of the school attendance equation.

4.1 School Attendance Patterns

The total number of days that a child attends school is determined by parental or child demand for schooling and the supply of schooling. While a decomposition of the days missed due to demand and supply factors does not affect the basic notion that school attendance is important in determining educational outcomes, identifying whether a child misses school due to demand or supply factors is important from a policy perspective. If the main problem is one of low demand for schooling then the appropriate response may be policies designed to lower costs of schooling or a policy of enhanced investments in

¹⁷ Notwithstanding these patterns it is possible that the distribution of school inputs and student achievement are systematically related. Due to lack of suitable data, we are unable to use an instrumental variables (IV) approach. However, we are able to provide some clues on the direction of the potential bias by referring to our earlier work (Bedi and Marshall, 1999). In our previous paper we found that the coefficients on the school inputs in IV regressions were systematically larger (in absolute terms) than the corresponding OLS estimates.

school inputs to increase the expected returns from schooling. On the other hand, limited supply of schooling would suggest another set of policy responses.

An important feature of the data used in this paper is that we have information not only on the number of days that each student missed during the 1998 school year, but also the number of days that each school in the UMCE sample was open during the school year.¹⁸ This information combined with our knowledge of the number of days that a school is expected to remain open helps us identify whether a child misses school due to lack of demand for schooling or due to lack of school supply. In terms of an equation, D_m , the days missed may be decomposed into

$$D_m = (D_o - D_a) + (172 - D_o), \quad (13)$$

where D_o is the number of days that school is offered and D_a represents days attended. The first term in parentheses on the right hand side represents days missed due to lack of demand for schooling. School is offered on those days but parents do not send their children to school. The second term represents days missed due to lack of school supply. Schools are meant to be operating 172 days during the school year and the gap between the expected days of operation and the days that a school is actually offered represents days missed due to lack of supply.¹⁹

The patterns of school attendance and a demand-supply decomposition of days missed for students in grade 2 and grade 4 are displayed in Tables 3a and 3b. On average a child attends 143-144 days of school which translates into a loss of around five weeks of schooling. The number of days attended varies from 121 at the 10th percentile to 161 at the

¹⁸ This additional information was collected as a consequence of our initial investigation into this issue when we were forced to assume that each school remained open for 160 days and the entire school attendance issue was treated as a demand side problem. As the data used in this study clearly demonstrate, this is an unrealistic assumption, especially for a nationally-representative sample of primary schools.

¹⁹ This decomposition may lead to an exaggeration of the supply side of the problem. If parental (lack of) demand for schooling on a particular day coincides with a day that school is not offered it will lead to an overemphasis of the supply side of the problem. Except for noting this possibility we are not able to offer any additional information on the extent of such an exaggeration.

90th percentile for grade 2 and 124 and 162 for the same percentiles in grade 4, that is, a range of about 1.5-9.5 weeks of missed classes.²⁰

Of particular interest to the present analysis is the decomposition of days missed into demand-supply components. This decomposition on the basis of equation (13) is displayed in Tables 3a and 3b. For grade 2, at the mean, the 29 missed days may be decomposed into 10 days missed due to lack of demand for schooling and 19 days missed due to lack of supply. A similar pattern prevails for grade 4, with 9 days accounted for by lack of demand and 19 days missed due to lack of supply. For both grades, the decomposition at the mean or at different points of the distribution clearly indicates that while both demand and supply factors are responsible for limiting school attendance a larger share of the problem (between 60-71 percent) may be attributed to lack of school supply.

While we can't provide much quantitative evidence on the factors responsible for the supply of schooling, some clues may be gleaned from the data as well as from direct observation during the data gathering exercise. There is substantial district variation in the supply of schooling.²¹ This variation may be related to differences in policy regarding teacher pay collection, local holidays, and teacher meetings. For example, in some rural areas teachers are allowed three days to collect their monthly pay from the district administrative office. Some teachers use their weekends to collect their pay while others collect their pay on weekdays. The frequency of teacher training seminars/meetings also varies substantially. Teachers in some districts attend one seminar/meeting a week while in other districts there are very few seminars/meetings. Another explanation for regional differences in supply lies in different levels of school supervision across districts. In several ru

²⁰ An immediate question that arises is whether these missed days translate into lower levels of human capital or are students able to make up despite missing school? Although educational achievement and school attendance are endogenous, to establish the effect of school attendance on educational outcomes we estimated educational production functions, for both subjects and both grades that included number of days attended as a regressor. For all four regressions we found that the number of days attended had a positive impact on test scores and was statistically significant. For grade 2 the point estimates indicated that an increase in school attendance by one standard deviation (14 days) would increase educational achievement by 3.1 to 4.4 percent. For grade 4 the corresponding calculation yielded increases of 2.2 to 4.4 percent. A detailed examination of the effect of attendance on test scores is available in Marshall and White (2000).

²¹ Within the same state the supply of schooling may vary from 110 to 172 days in a school year.

ral areas there is almost no supervision and school closures may be unknown to district supervisors (Marshall and White, 2000).

As is clear from the decomposition, a complete study of school attendance requires an analysis of the demand for and the supply of schooling. Due to lack of detailed information on the supply side, the focus in our paper is largely on factors that inhibit school attendance from the demand side. Despite the limited supply side analysis, the ability to identify the relative influence of demand and supply factors in determining school attendance is by itself an important step in incorporating supply side issues into an analysis of school attendance.

4.2. Educational Production Functions

Least squares and selection-corrected estimates of the four educational production functions for grade 2 and grade 4 are presented in Tables 4 and 5 respectively. The selection corrected results control for the potential bias that may arise from estimating regressions over samples of students who attended school on the days that the tests were administered.²² For grade 2 the selection patterns for both Spanish and mathematics are positive and statistically significant, indicating that unobserved characteristics that have a bearing on test scores are positively correlated with school attendance. For grade 4 the selection patterns are not as clear. A comparison of the two sets of results (OLS versus selection corrected) shows that there are very few sign changes on the estimated coefficients and while the magnitude of the coefficients for grade 2 are substantially affected, for grade 4 the changes are minimal. Since we have detected some signs of selection bias the discussions in this section are based on the selection corrected estimates.²³

We now examine the effect of various sets of characteristics on test scores. The age of a child does not appear to play an important role in determining test scores, al

²² The selection-corrected estimates are maximum likelihood estimates based on Heckman's (1979) correction procedure. The probability of attending school on the day of the test was specified as a function of individual, school and teacher variables. Detailed estimates are available on request.

²³ It is well known that the selection procedure is quite sensitive to specification of the selection equation as well as to departures from normality. We experimented with some minor changes in specification but the

though there are clear gender differences. Girls score higher than boys by about 1-1.7 points on Spanish examinations, although in mathematics, at least for grade 4, it appears that boys perform better. The favorite subject dummy variable is included in order to control for subject-specific unobservables such as hard work, motivation or interest. This variable is positive and significant for mathematics but negative (and significant) for Spanish, which implies that lower-scoring students may be choosing Spanish as a sort of default option since they may find it to be the easiest subject.

Across both grades, attending preschool does not appear to confer any significant advantages, although it is possible that the inclusion of several other family characteristics makes it difficult to isolate the effect of attending preschool.²⁴ Parental attitudes towards a child's education are measured by two variables: parental help and parental activity in school. Parental help appears to boost performance by about 1-1.8 points in grade 2 (in both subjects) and in grade 4 - Spanish, but the dissipation of this effect for mathematics in grade 4 probably reflects limited parental ability. The point estimates for parental activity in the school are generally positive but not significant. Finally, the parental education and wealth variables serve as proxies for household access to resources, attitudes towards education and genetic ability. Not surprisingly, the effect of both of these variables is positive and statistically significant across all sets of estimates.

The teacher characteristics include conventional measures such as teacher education and experience. In addition, we also have information on the current stock of teacher knowledge as measured by test scores on Spanish and math exams. The number of teaching seminars attended controls for knowledge of teaching skills while unobserved teaching ability may be captured by a self-reported measure of teacher confidence. For grade 2, the education level of the instructor and the number of seminars attended exert a positive ef

limited availability of identifying information precludes a thorough examination. A prudent approach may be to repose greater confidence in those coefficients that are robust to changes in the estimation approach.

²⁴ Although we recognize the possibility that attending preschool may be an endogenous decision we do not pursue this issue as it is not the focus of our work. An informal examination of the effect of this variable was carried out by examining results that excluded this variable. As may be expected there were no substantial differences in our estimates.

fect on test scores. For grade 4, none of the teacher characteristics appears to be systematically related to performance.

Classroom characteristics indicate whether several grades are simultaneously being taught in the same classroom, whether there are enough textbooks available and the number of students in the class. Except for class size, which indicates that a reduction in the number of students in a class by one standard deviation would increase mean test scores by 3.6-4.3 percent for grade 2, and around 1.4 percent for grade 4, none of the other class characteristics are associated with test performance.

Despite the size of the sample and the presence of several variables that are not conventionally available we are unable to identify clear-cut policy relevant variables that influence educational achievement. There could be several data related reasons for the limited effects of the teacher and classroom characteristics on test performance. For instance, there may be limited variation in these characteristics. A look at the means and standard deviations of these variables suggests that while this may be true for some of the variables (e.g. teacher education) there does appear to be substantial variation in the data. Correlation amongst these variables also does not appear to be a serious problem.²⁵ Another possible explanation is that the most important school and teacher variables that determine academic achievement are unobserved (such as teacher ability or school management) in our data set.²⁶

4.3. School Attendance²⁷

As depicted in equation (12) we specify school attendance as a function of opportunity costs (P), the benefits in the form of expected human capital gains (\hat{H}) and the quality of school facilities (Q) which may influence attendance but may not have a direct

²⁵ Correlation among these variables does not appear to be substantial as the condition number for the correlation matrix of the teacher and class variables is less than 5.

²⁶ Regressions that include child and family characteristics and indicator variables for each school increase the explanatory power of the models from a range of 0.113-0.188 to a range of 0.24-0.29. This suggests that while differences across schools do influence test scores we are unable to detect the factors that are responsible for these differences.

bearing on student achievement. The estimates of the school attendance equation for grades 2 and 4 are presented in Tables 6 and 7, respectively. The independent variables in these estimates comprise three sets of variables corresponding to each of the three elements that may have a bearing on school attendance. In Tables 6 and 7 the child, family and the first two school characteristics (time taken to reach school and whether the school has a lunch program) are associated with the opportunity costs of school attendance. The next three variables, the presence of a telephone connection, electricity, and a composite variable of all other school facilities, comprise the Q vector. The final variable is our measure of the expected human capital gains of attending school. We have two measures of \hat{H} , one from the OLS estimates and the other from the selection corrected estimates of the educational production functions. Estimates of the attendance equation with both of these measures are presented in each table.

Identification is always an issue in such models. The dilemma is deciding which variables belong solely in the educational production function and may legitimately be excluded from the school attendance equation. We present two sets of estimates based on different exclusion restrictions. The first two columns of Table 6 and Table 7 present estimates based on the assumption that teacher and classroom characteristics influence the attendance decision only through their effect on test scores. However, this assumption, especially for the classroom characteristics (which may be easily observed) is probably too strong. Acknowledging this possibility, the set of estimates in columns 3 and 4 includes classroom characteristics in the attendance equation. This set of estimates is based on excluding only the set of teacher characteristics.²⁸

We now turn to a discussion of the estimates. For all specifications in both grades age reduces the number of days attended. A 12-year-old child in grade 2 may miss three to

²⁷ Although we recognize that a large part of the variation in school attendance is due to differences in the supply of schooling we would like to reiterate that our attention in this section is restricted to factors that motivate the demand for schooling.

²⁸ Even this exclusion restriction may not be appropriate for all the teacher characteristics. It is possible that the educational qualifications or experience of a teacher may have a direct bearing on school attendance. However, for some of the other teacher characteristics such as seminars attended, self-confidence and test scores, which are difficult for parents to observe, it may be reasonable to argue that they are legitimate exclusions from the attendance equation.

six more days of school as compared to a 7-year-old colleague. This effect translates into a 2-4 percent reduction in attendance. It is clear that as children age they become more valuable and it becomes less likely that they will attend school. Given that many children are forced to leave school when they reach adolescence the process of accumulating years of schooling may be characterized as a race against time.

None of the family characteristics appear to be significantly associated with attendance. While this may seem surprising, it is likely that these variables are exerting their influence on the attendance decision through their effect on expected test scores.²⁹ As expected, the time taken to get to school is negatively associated with school attendance in both grades, although the effect of this variable is not statistically significant, which suggests that most children do not spend a prohibitively long time travelling to school. As the descriptive statistics show the mean travel time is only around 15-16 minutes. The presence of a school lunch program defrays the costs of attending school and is expected to be positively associated with attendance, however the estimates show no evidence of a systematic link.

We hypothesized that parents are more likely to send their children to school if it has better facilities. There is some evidence to support this hypothesis. In grade 2, the presence of electricity in the school is associated with around three more days of school attendance. The effect of this set of variables is slightly higher in grade 4, where the availability of a telephone connection appears to encourage five more days of school attendance. Of course these results should not be interpreted literally but with a view that school facilities that may not have a direct bearing on test scores may nevertheless send out easily observed quality signals that may encourage parents to send their children to school.

The novel empirical result in these estimates is the effect of expected achievement gains on the school attendance decision. Consistent with our hypothesis, the expected achievement variable suggests that the benefits derived from attending school plays an important role in shaping parental decisions. Across both grades and all specifications the

effect of this variable is statistically significant. An increase in the average score by about five points increases school attendance in grade 2 by three to eight days. A similar increase for grade 4 results in increases in school attendance by about five to seven days. These effects translate into an attendance increases of between 2% and 5%, and show that expected achievement gains exert an important influence on the demand for schooling. This finding is similar to Willis and Rosen (1979) who found that expected earnings influence the decision to attend college.

Overall, the school attendance estimates show that opportunity costs play a role in hindering school attendance, but their impact may not be as important as the expected benefits of attending school. The results also suggest that rather than trying to reduce opportunity costs, policies designed to increase school attendance may achieve their objective by encouraging appropriate investments in school inputs. These investments may take the form of school facilities that have a direct bearing on school attendance or on school inputs that influence educational achievement which in turn will encourage greater attendance. While this link appears to be straightforward, identifying the appropriate school inputs that improve student achievement is still a difficult task.

5. CONCLUSION

In recent years the substantial investments in the education sector in Honduras have resulted in a rapid expansion of the primary education system. Access to primary school is nearly universal and almost all children enroll in primary school. Thus, the focal issue now is not whether a child enrolls in school but conditional on enrollment, how often does a child attend school?

The frequency of school attendance is an important determinant of academic achievement, which is subsequently linked to repetition and desertion. Noting this chain of events this paper used recently collected data from a nationally drawn sample of schools to specify and estimate a model of school attendance. Attendance was treated as a function of opportunity costs and the benefits associated with school attendance. The results from

²⁹ Specifications that do not include the predicted human capital variables display a strong link between some

the estimated models suggested that while opportunity costs as measured by the age of a child played a role, perhaps a more important factor in determining school attendance is the expected human capital benefit associated with school attendance. The effect of expected achievement gains is important from a policy perspective, as it suggests that investments in school inputs may be used to achieve the same objective as programs designed to reduce opportunity costs of school attendance.³⁰

While the paper treated school attendance primarily from a demand perspective, an important aspect of the paper was a decomposition of the school days missed into demand and supply components. This decomposition indicated that between 60-71 percent of the days missed during the school year may be attributed to the limited supply of schooling. The magnitude of the supply side effect clearly indicates that a better understanding of school attendance patterns in Honduras requires a detailed analysis of the factors that motivate the supply of schooling. Furthermore, it also suggests that a more fruitful approach towards raising school attendance may lie not in reducing opportunity costs or increasing expenditure on school inputs but in policies that target the supply of schooling.

REFERENCES

- Bedi A.S., 1997. The importance of school quality as a determinant of earnings in a developing country. *International Journal of Educational Development* 17, 427--37.
- Bedi, A.S., J.H. Marshall, 1999. School attendance and student achievement: Evidence from Rural Honduras. *Economic Development and Cultural Change* 47, 657-682.
- Edwards, J.H.Y., 1995. The status of primary education in Honduras. USAID, Tegucigalpa, Honduras.
- Edwards, J.H.Y., Fuller, B., Parandekar, S., 1996. Primary education efficiency in Honduras: What Remains to be Done? USAID, Tegucigalpa, Honduras.

of the family characteristics, especially family wealth and school attendance patterns.

³⁰ Before following this course of action one requires an assessment of the exact school inputs towards which expenditure should be directed, as well as a cost effectiveness analysis of programs designed to reduced opportunity costs versus programs geared towards enhancing school inputs.

- Fuller, B., et al., 1999. How to raise children's early literacy? The influence of family, teacher, and classroom in northeast Brazil. *Comparative Education Review* 43, 1-35.
- Gertler, P., van der Gaag, J., 1988. Measuring the willingness to pay for social services in developing countries. LSMS Working Paper No.45, The World Bank, Washington D.C.
- Gertler, P., Glewwe, P., 1989. The willingness to pay for education in developing countries: Evidence from Rural Peru. LSMS Working Paper No. 54, The World Bank, Washington D.C.
- Gomes-Neto, J.B., Hanushek, E.A., 1994. The causes and consequences of grade repetition: Evidence from Brazil. *Economic Development and Cultural Change* 43, 117-148.
- Heckman, J.J., 1979. Sample selection bias as a specification error. *Econometrica* 47, 153-62.
- Jacoby, H., 1994. Borrowing constraints and progress through school: Evidence from Peru. *Review of Economics and Statistics* 76, 151-60.
- Jimenez, E. Sawada, Y., 1999. Do community-managed schools work? An evaluation of El Salvador's EDUCO program. *World Bank Economic Review* 13, 415-441.
- Lockheed, M., Verspoor A., and Associates, 1991. Improving primary education in developing countries. Oxford University Press, London.
- Marshall, J.H. 2000. Grade failure, repetition and teacher evaluation in Honduran primary schools. Unpublished manuscript.
- Marshall, J.H., White, K.A., 2000. Academic achievement, school attendance and teacher quality in Honduras: An empirical analysis. Unpublished manuscript.
- Moulton, B.R., 1986. Random group effects and the precision of regression estimates. *Journal of Econometrics* 32, 385-97.
- Patrinos, H.A., Psacharopoulos, G., 1995. Educational performance and child labor in Paraguay. *International Journal of Educational Development* 15, 47-60.
- Patrinos, H.A., Psacharopoulos, G., 1997. Family size, schooling and child labor in Peru- An empirical analysis. *Population Economics* 10, 387-405.
- Strauss, J., Thomas, D., 1995. Human resources, empirical modeling of household and family decisions. In: Behraman, J., Srinivasan, T.N. (Eds.), *Handbook of Development Economics* Vol. 3A North-Holland, Amsterdam.

- UMCE, 1998. Factores asociados con el rendimiento escolar en la escuela primaria hondureña” (Factors associated with primary school academic achievement in Honduras.Tegucigalpa, Honduras: UPNFM/MOE.
- UNDP, 1999. Human Development Report. Oxford University Press. Oxford.
- UNESCO, 1999. Statistical yearbook.
- Tan, J.P., Lane, J., Coustere, P.,1997, Putting inputs to work in elementary schools: What can be done in the Philippines? *Economic Development and Cultural Change* 45, 857-79.
- Van Steenwyck, N., Meija, M.T., 1996, *Honduras 2000: Recursos Humanos y Perspectivas Para el Crecimiento Económico* (Honduras 2000: Human Resources and Perspectives on Economic Growth). USAID, Tegucigalpa, Honduras.
- Van Steenwyck, N., 1997. Los Rendimientos Académicos y la Reprobación de Alumnos (Academic Achievement and Repetition). USAID/PEEP, Tegucigalpa, Honduras.
- Willis, R., Rosen, S., 1979. Education and self-selection. *Journal of Political Economy* 87, S1-S36.
- World Bank, 1995. Staff Appraisal Report: Honduras Basic Education Project. The World Bank, Washington, DC.