Children's Health and Achievement in School

Jere R. Behrman
Victor Lavy
<table>
<thead>
<tr>
<th>No.</th>
<th>Title and Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>Deaton, Quality, Quantity, and Spatial Variation of Price: Estimating Price Elasticities from Cross-Sectional Data</td>
</tr>
<tr>
<td>31</td>
<td>Suarez-Berenguela, Financing the Health Sector in Peru</td>
</tr>
<tr>
<td>32</td>
<td>Suarez-Berenguela, Informal Sector, Labor Markets, and Returns to Education in Peru</td>
</tr>
<tr>
<td>33</td>
<td>van der Gaag and Vijverberg, Wage Determinants in Côte d'Ivoire</td>
</tr>
<tr>
<td>34</td>
<td>Ainsworth and van der Gaag, Guidelines for Adapting the LSMS Living Standards Questionnaires to Local Conditions</td>
</tr>
<tr>
<td>35</td>
<td>Dor and van der Gaag, The Demand for Medical Care in Developing Countries: Quantity Rationing in Rural Côte d'Ivoire</td>
</tr>
<tr>
<td>36</td>
<td>Newman, Labor Market Activity in Côte d'Ivoire and Peru</td>
</tr>
<tr>
<td>37</td>
<td>Gertler, Locay, Sanderson, Dor, and van der Gaag, Health Care Financing and the Demand for Medical Care</td>
</tr>
<tr>
<td>38</td>
<td>Stelcner, Arriagada, and Moock, Wage Determinants and School Attainment among Men in Peru</td>
</tr>
<tr>
<td>39</td>
<td>Deaton, The Allocation of Goods within the Household: Adults, Children, and Gender</td>
</tr>
<tr>
<td>40</td>
<td>Strauss, The Effects of Household and Community Characteristics on the Nutrition of Preschool Children: Evidence from Rural Côte d'Ivoire</td>
</tr>
<tr>
<td>41</td>
<td>Stelcner, van der Gaag, and Vijverberg, Public-Private Sector Wage Differentials in Peru, 1985–86</td>
</tr>
<tr>
<td>42</td>
<td>Glewwe, The Distribution of Welfare in Peru in 1985–86</td>
</tr>
<tr>
<td>43</td>
<td>Vijverberg, Profits from Self-Employment: A Case Study of Côte d'Ivoire</td>
</tr>
<tr>
<td>44</td>
<td>Deaton and Benjamin, The Living Standards Survey and Price Policy Reform: A Study of Cocoa and Coffee Production in Côte d'Ivoire</td>
</tr>
<tr>
<td>45</td>
<td>Gertler and van der Gaag, Measuring the Willingness to Pay for Social Services in Developing Countries</td>
</tr>
<tr>
<td>46</td>
<td>Vijverberg, Nonfarm Household Enterprises in Côte d'Ivoire: A Descriptive Analysis</td>
</tr>
<tr>
<td>47</td>
<td>Glewwe and de Tray, The Poor during Adjustment: A Case Study of Côte d'Ivoire</td>
</tr>
<tr>
<td>48</td>
<td>Glewwe and van der Gaag, Confronting Poverty in Developing Countries: Definitions, Information, and Policies</td>
</tr>
<tr>
<td>49</td>
<td>Scott and Amenuvegbe, 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</td>
</tr>
<tr>
<td>50</td>
<td>Laraki, Food Subsidies: A Case Study of Price Reform in Morocco (also in French, 50F)</td>
</tr>
<tr>
<td>51</td>
<td>Strauss and Mehra, Child Anthropometry in Côte d'Ivoire: Estimates from Two Surveys, 1985 and 1986</td>
</tr>
<tr>
<td>52</td>
<td>van der Gaag, Stelcner, and Vijverberg, Public-Private Sector Wage Comparisons and Moonlighting in Developing Countries: Evidence from Côte d'Ivoire and Peru</td>
</tr>
<tr>
<td>53</td>
<td>Ainsworth, Socioeconomic Determinants of Fertility in Côte d'Ivoire</td>
</tr>
<tr>
<td>54</td>
<td>Gertler and Glewwe, The Willingness to Pay for Education in Developing Countries: Evidence from Rural Peru</td>
</tr>
<tr>
<td>55</td>
<td>Levy and Newman, Rigidité des salaires: Données microéconomiques et macroéconomiques sur l'ajustement du marché du travail dans le secteur moderne (in French only)</td>
</tr>
<tr>
<td>56</td>
<td>Glewwe and de Tray, The Poor in Latin America during Adjustment: A Case Study of Peru</td>
</tr>
<tr>
<td>57</td>
<td>Alderman and Gertler, The Substitutability of Public and Private Health Care for the Treatment of Children in Pakistan</td>
</tr>
<tr>
<td>58</td>
<td>Rosenhouse, Identifying the Poor: Is “Headship” a Useful Concept?</td>
</tr>
<tr>
<td>59</td>
<td>Vijverberg, Labor Market Performance as a Determinant of Migration</td>
</tr>
<tr>
<td>60</td>
<td>Jimenez and Cox, The Relative Effectiveness of Private and Public Schools: Evidence from Two Developing Countries</td>
</tr>
<tr>
<td>61</td>
<td>Kakwani, Large Sample Distribution of Several Inequality Measures: With Application to Côte d'Ivoire</td>
</tr>
<tr>
<td>62</td>
<td>Kakwani, Testing for Significance of Poverty Differences: With Application to Côte d'Ivoire</td>
</tr>
<tr>
<td>63</td>
<td>Kakwani, Poverty and Economic Growth: With Application to Côte d'Ivoire</td>
</tr>
<tr>
<td>64</td>
<td>Moock, Musgrove, and Stelcner, Education and Earnings in Peru's Informal Nonfarm Family Enterprises</td>
</tr>
<tr>
<td>65</td>
<td>Alderman and Kozel, Formal and Informal Sector Wage Determination in Urban Low-Income Neighborhoods in Pakistan</td>
</tr>
</tbody>
</table>

(List continues on the inside back cover)
Children’s Health and Achievement in School
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.
Children's Health and Achievement in School

Jere R. Behrman
Victor Lavy

The World Bank
Washington, D.C.
## Contents

**Foreword** .................................................. vii

**Abstract** .................................................. ix

**Acknowledgments** ............................................ xi

**Introduction** ............................................... 1

**Modeling the Relation Between Child Health and Schooling Success** ............ 4

**Data** .......................................................... 10

- Cognitive Achievement Test Scores ........................................ 10
- Child Health .................................................. 12
- Grades of Schooling at Time of Cognitive Achievement Test .................. 13
- Pre-School Ability ............................................. 13
- Parental Schooling ............................................. 14
- Child Age .................................................... 14
- Child sex ..................................................... 14
- Variables for Reduced-Form Estimates for Child Health and Child Schooling Attainment at Time the Cognitive Achievement Test was Taken ................. 14

**Estimates of Cognitive Achievement Production Function with Instrumental Variable Control for Simultaneity** ........................................... 15

- OLS versus Instrumental Variable Estimates with Basic Sample .................. 15
- OLS versus Instrumental Variable Estimates with Alternative Samples or Dependent Variables ................................................................. 19
- Variations on the Basic Estimates ............................................. 19

**Estimates of Cognitive Achievement Production Function with Control for Unobserved Family and Community Effects** .................................. 22

- Family and Community Fixed Effects Estimates for Basic Sample .................. 23
- Family and Community Fixed Effects Estimates for Other Samples and Definitions of Dependent Variables .................................................. 26
- Other Aspects of These Estimates ............................................. 27

**Might Child Health Affect Cognitive Achievement through Schooling Attainment?** ... 30
Conclusions .......................................................... 33

References .......................................................... 35

Tables

Table 1: Descriptive Statistics for Ghanaian Children Aged 10-17 in Living Standards Measurement Survey (LSMS) Sample .................. 11

Table 2: Cognitive Achievement Production Function Estimates, Ghanaian LSMS: CLS and Simultaneous Estimates .................. 16

Table 3: Percentages of Total Variance that is within Family for Key Child Characteristics Variables, Ghanaian LSMS .................. 23

Table 4: Cognitive Achievement Production Function Estimates for Sibling Data and Family and Community Fixed Estimates, Ghanaian LSMS .......... 24
Foreword

The extent to which there are synergies between investment in education and health of children is very important for prioritizing the allocation of resources for human capital investments. The literature that has investigated the impact of child health and nutrition on child schooling success has concluded that such effects are positive and important over the range of child health observed in children who attend school, in addition to effects on who attends school. However, this literature does not control for the probable endogeneity of the determination of child health. This paper analyses the implication of the failure to control for such endogeneity and provide new evidence on the interrelationship between child health and schooling.

This paper is part of broader research effort in the Policy Research Department (PRD) that examines the effect of the quality of social services on human capital investment outcomes. This work is located in the Poverty and Human resources division. The data used 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.

Lyn Squire
Director
Policy Research Department
Abstract

Casual observations suggest that extremely poor child health is detrimental to educational achievement. There also is a widespread perception that available systematic evidence supports a strong role of child health on child schooling success for variations in child health above extremely poor levels, which underlies in part strong advocacy for improving child health since such improvements are claimed to have strong fairly immediate effects on child education and, through this channel, important long-run effects on labor productivity.

However, in fact the evidence is quite limited about the impact on education of child health within the range of health usually observed among school children. Previous studies based on socioeconomic survey data that purport to support the important role of child health on child schooling success fail to incorporate into their analysis the probable endogenous nature of child health. Most such studies also are limited because of fairly limited measures of schooling achievement, such as schooling attendance, though some do use better indicators such as school grades or test performances.

On a priori grounds it would seem that child health and child schooling are determined simultaneously by households given their observed and unobserved characteristics and those of the community in which they are. If so, failure to control for such household allocations in estimates of the impact of child health on child schooling is likely to lead to biased estimates of that effect in the standard estimates that do not control for such allocations. The direction of this bias, however, may be positive or negative depending on which of a number of household allocation behaviors dominate.

This paper explores the a priori nature of the possible biases and then presents some illustrative empirical analysis of these effects using some rich data for this purpose from the Ghanaian Living Standard Measurement Study (LSMS). These explorations lead to four major conclusions for this data set. First, the failure to control for estimation problems as in previous studies leads to a considerable bias in the estimated impact of child health on child schooling success. Second, instrumental variable estimates based on observed family and community characteristics similar to those often used in other studies suggest that the direction of this bias in standard estimates without control for simultaneity is downward. Third, estimates with family and community fixed effects (to control for factors such as parental time and the general learning environment), however, suggest that the direction of the bias in standard estimates is upward and that the true effects of the range of observed child health on school success is nil despite the strong association that leads to the appearance of an effect in standard OLS estimates or with instrumented level estimates using family and community variables. Fourth, exploration of the possibility that child health may affect child cognitive achievement through schooling attainment also does not reveal a significant positive impact of child health on child schooling.
Consideration of the relations that usually have been used to investigate such a possibility, moreover, suggests that the coefficients that are estimated are not, in contrast to the usual claim, coefficients that represent the impact of child health on child schooling.

Thus, despite the OLS and instrumented level estimates, this paper concludes that for this sample there is not evidence of an impact of the observed range of child health on child cognitive achievement. It also concludes that the striking difference between the instrumental variable instruments, using a set of instruments that are fairly typical for this type of study, and the family and community fixed effects estimates raises the question of whether other studies that have depended on similarly instrumented estimates may not be subject to similar problems.
Acknowledgments

We thank for useful comments Harold Alderman, Nancy Birdsall, Anil Deolalikar, Paul Glewwe, Estelle James, Dean Jolliffe, Mark Pitt, Steve Stein, John Strauss, Duncan Thomas, members of the Population Studies Workshop at the University of Pennsylvania (particularly Mark Rosenzweig and Andrew Foster), members of the World Bank World Development Report 1993 team (particularly Dean Jamison, Kenneth Hill, and Philip Musgrove), and an anonymous referee. The authors alone, and not the World Bank, are responsible for the contents of this paper.
Introduction

Common sense and casual observations suggest that extremely poor child health is detrimental to educational achievement. There also is a widespread perception that available systematic evidence supports a strong role of child health on child schooling success for variations in child health above extremely poor levels, including the levels observed among children actually attending schools (see the surveys in Pollitt 1990, Miller and Korenman 1993, and World Bank 1993). In part based on such evidence there is strong advocacy for improving child health since such improvements are claimed to have strong fairly immediate effects on child education and, through this channel, important long-run effects on labor productivity in both developed and developing countries.¹

However, in fact the evidence is quite limited about the impact of child health within the range usually observed on education. True, there are a number of studies based on socioeconomic survey data that purport to support the important role of child health, usually as represented by anthropometric measures, on child schooling success.² But these studies are not as persuasive as usually is claimed because of their failure to incorporate into their analysis the probable endogenous nature of child health. Most such studies also are limited because of fairly limited measures of schooling achievement, such as schooling attendance, though some do use better indicators such as school grades or test performances.

On a priori grounds it would seem that child health and child schooling are determined simultaneously by households given their observed and unobserved characteristics and those of the community in which they are. If so, failure to control for such simultaneity in estimates of the impact of child health on child schooling is likely to lead to biased estimates of that effect in the standard estimates in the literature that do not control for such simultaneity. The direction of this bias, however, is not obvious a priori. On one hand, this bias may arise primarily from unobserved (to researchers) parental or community characteristics such as pro-child quality tastes or productivities that contribute to better child schooling performance and also contribute to better child health. In such cases the standard estimates give upward-biased estimates of the impact of child health on child schooling

---

¹ There is widespread emphasis on large productivity effects of schooling success in the developing countries, though some debate about the implications of estimation problems. For example, see Barro (1991), Behrman (1990a,b,c), Birdsall and Sabot (1993), Colclough (1982), Eisemon (1988), Haddad, Carnoy, Rinaldi and Regel (1990), King and Hill (1993), Mensch, Lentzner, and Preston (1985), Psacharopoulos (1985, 1988), Schultz (1988, 1993), and World Bank (1980, 1981, 1990, 1991, 1993). For developed economies the evidence generally suggests somewhat lower rates of return to schooling than for the developing countries, though some studies argue that recent rates of return to schooling in the United States are from 14 to 30 percent if measurement error is controlled in the estimates (e.g., Ashenfelter and Krueger 1993, Butcher and Case 1992).

² There also are a number of experimental studies on the relation between nutrition or health and school achievement (e.g., Soemantri et al 1985, Soemantri 1989, Pollitt et al 1989, Seshadri and Gopaladas 1989, Nokes et al. 1992a,b). However these studies often are on small, selected samples and tend to focus on micro nutrients such as iron or on specific parasitic infections rather than on more general indicators of health. Behrman (1993) surveys some of these studies.
success. On the other hand, the bias may arise largely from unobserved parental preferences that are heterogeneous regarding the value placed on child health versus child schooling success, unobserved parental preferences for equity across the human resources of their children and associated efforts to compensate for children's endowment differentials, unobserved prices of unobserved health versus schooling inputs that are heterogenous across households, or unobserved child endowments favoring health versus schooling investments that are heterogenous across children. In these cases, the standard estimates give downward-biased estimates of the impact of child health on child schooling success. In all of these cases in which child health reflects household behavior, the basic estimation problem is to isolate the effect of child health on child schooling success through assuring that the representation of child health that is used in the estimates is independent of the disturbance term in the relation being estimated. This disturbance term, in turn, may include unobserved individual, family, and community factors that may operate directly on child cognitive achievement in the cognitive achievement production function or may operate indirectly through unobserved inputs into the cognitive achievement production that are allocated within the household such as parental time and the general learning environment.

For these reasons, the correlations between anthropometric indicators of child health and child schooling achievement in studies of both developing and developed countries such as Freeman, Klein, Townsend and Lechtig (1980), Wilson (1981), Chutikul (1982, 1986), Wolfe (1985), Moock and Leslie (1986), Jamison (1986), Florencio (1988), Glewwe and Jacoby (1992), Gomes-Neto, Hanushek, Leite, and Frota-Bezzera (1992), and Miller and Korenman (1993) are not compelling evidence of the extent of the impact of child health on schooling success despite the fact that this interpretation is quite widespread. Associations between indicators of child health and indicators of child schooling achievement do not demonstrate that child health causes child schooling achievement to the same degree. The true effects may be much smaller or much larger.

We address the issue of the impact of the endogenous determination of child health on child schooling success in this paper. In Section 2 we explore the a priori nature of the possible biases that are summarized above. In Section 3 we describe the data that we use for our illustrative empirical analysis of these effects, the Ghanaian Living Standard Measurement Study (LSMS) data. This data set includes child anthropometric measures to represent child health, cognitive achievement test scores to represent schooling success, a test of pre-school ability to control for that dimension of child endowments, a fairly rich range of household and community characteristics to use for simultaneous estimates, and sibling data to explore unobserved family and community fixed effects. The remaining sections explore

3. Glewwe and Jacoby (1992) are sensitive to possible problems of simultaneity in their investigation of the impact of anthropometric indicators on delayed schooling enrollment in Ghana. However, their estimates do not control for the possibility considered below that there are unobserved variables allocated by the household that affect the dependent variable of interest. Nor do they adequately justify their specification from the perspective of the discussion in Section below of relation (5C).
the empirical relevance of the endogenous determination of child health on child schooling success for this data set.

These explorations lead to four major conclusions for this data set. First, the failure to control for estimation problems as in previous studies leads to a considerable bias in the estimated impact of child health on child schooling success. Second, instrumental variable estimates based on observed family and community characteristics similar to those often used in other studies (under the maintained hypothesis that such characteristics are independent of the disturbance term in the cognitive achievement production function) suggest that the direction of this bias in standard estimates without control for simultaneity is downward, so the second set of factors pertaining to simultaneity described above apparently prevail. Third, estimates with sibling data, however, suggest that the direction of the bias in standard estimates is upward and that the true effects of the range of observed child health on Ghanaian school success is nil despite the strong association that leads to the appearance of an effect in standard OLS estimates or with instrumented level estimates using family and community variables. That is, there are unobserved family and community effects — such as parental time and the general learning environment — that cause upward biases in the standard estimates and in instrumented level estimates through influencing unobserved variables allocated by the household that affect child cognitive achievement. The estimates with sibling data control for these underlying unobserved family and community factors and lead to an estimated effect of child health on child cognitive achievement that is not different from zero that persists whether or not there is control for unobserved child endowments using instrumented within estimates. Fourth, exploration of the possibility that child health may affect child cognitive achievement through schooling attainment also does not reveal a significant positive impact of child health on child schooling. Consideration of the relations that usually have been used to investigate such a possibility, moreover, suggests that the coefficients that are estimated are not, in contrast to the usual claim, coefficients that represent the impact of child health on child schooling.

Thus, despite the OLS and instrumental variable level estimates, we conclude that for this sample there is not evidence of an impact of the observed range of child health on child cognitive achievement. We also conclude that the striking difference between the instrumental variable estimates, using a set of instruments that are fairly typical for this type of study, and the within-family and within-community estimates raises the question of whether other studies that have depended on household production function estimates with instrumental variable estimates — such as Pitt, Rosenzweig and Hassan (1990) and Rosenzweig and Schultz (1983, 1987, 1988) — may not be subject to similar problems.
Modeling the Relation Between Child Health and Schooling Success

We are interested in obtaining the estimated impact of child health on child schooling success. We measure child schooling success in our empirical estimates below by performance on cognitive achievement tests, that have been shown for the data set that we use to have a significant positive relation with adult wages.\(^4\) Thus we are basically interested in obtaining an unbiased estimate of the impact of child health in a production function for cognitive achievement:

\[
(1) \quad CA_c = CA(H_c, S_c, I_c, F_c, C_c, X_c, e_c),
\]

where

- \(CA_c\) is cognitive achievement,
- \(H_c\) is a vector of indicators of child health (e.g., anthropometric indicators such as height standardized for age and sex);
- \(S_c\) is schooling attainment at the time that the cognitive achievement test is taken (years of schooling);
- \(I_c\) is a vector of observed (in the data that we use) predetermined individual characteristics (e.g., age, gender, and ability);
- \(F_c\) is a vector of observed family characteristics that contribute to the learning environment that the child experiences (e.g., parental schooling);
- \(C_c\) is a vector of observed community characteristics that affect child cognitive achievement;
- \(I_c^*\) is a vector of unobserved predetermined individual characteristics that affect cognitive achievement (e.g., motivation, capacity for concentration);
- \(F_c^*\) is a vector of unobserved predetermined household characteristics that affect cognitive achievement (e.g., intellectual atmosphere, nature of conversations);
- \(C_c^*\) is a vector of unobserved community characteristics that affect child cognitive achievement (e.g., general intellectual atmosphere, expected returns to investing in cognitive achievement given structure of local economy);
- \(X_c^*\) is a vector of unobserved resources allocated by the household that affect the child's cognitive achievement (e.g., parental time, reading material);
- \(e_c\) is a stochastic disturbance term;
- and \(c\) refers to the \(c\)th child and his/her family.

We use "observed" to mean observed in the data and "unobserved" to mean not observed in the data, though observed by decision makers in the processes being investigated.

---

\(^4\) See Glewwe (1992). This result also is reported in the other three data sets from developing countries of which we are aware that have such data: for urban East Africa (Kenya and Tanzania) in Knight and Sabot (1990), for rural Pakistan in Alderman, Behrman, Ross and Sabot (1993), for Morocco in Lavy, Spratt and Leboucher (1992). Related evidence for the U.S. is provided in Bishop (1989, 1991), Blackburn and Neumark (1993) and references therein.
The estimation problem is that in ordinary least squares (OLS) estimates of the linear approximation of relation (1), the disturbance term is \( v = I_a + F_a + C_a + X_a + e \). If \( H_x \) depends on \( I_a, F_a \) or \( C_a \) or on the determinants of \( X_a \) as in the reduced-form relation (5) below, the OLS coefficient estimate of \( H_x \) is biased. To understand the possible nature of this bias, it is necessary to consider how this cognitive achievement production function and the determinants of its inputs fit into the household decision making process.

Assume that the household objective function includes among its arguments child cognitive achievement and health, given heterogeneity across households in preferences for child quality versus parental consumption and in preferences for child cognitive achievement versus child health:

\[
U = U(C, H, Z, \ldots | \delta_q, \delta_{ca/h}),
\]

where \( Z \) is parental consumption; \( \delta_q \) reflects heterogeneity in parental preferences concerning all elements of child quality (including child cognitive achievement and child health) versus parental consumption; and \( \delta_{ca/h} \) reflects heterogeneity in parental preferences concerning child cognitive achievement versus child health in the composition of child quality. We have written this objective function to be consistent with there being an unified household preference function that is maximized by the household. If there is household bargaining and bargainers differ in their relative emphasis on child quality versus parental consumption and in their relative emphasis on different components of child quality, then in the reduced forms in relation (5) below \( \delta_q \) and \( \delta_{ca/h} \) could be interpreted as weighted averages of the preferences of the bargainers with the weights reflecting the bargaining power of the bargainers. We are not able to identify whether or not household bargaining occurs with the data that we use in our empirical estimates below, so for simplicity we refer to \( \delta_q \) and \( \delta_{ca/h} \) as parental preference parameters in what follows.

This objective function is maximized subject to production function constraints and a full-income constraint, given household assets, market prices and community characteristics. One critical production function for the present study is that for cognitive achievement in relation (1). Another critical production function is that for child health:

\[
H_c = H(C, I_c, F_c, C_h, I_h, F_h, C_h, X_h, e_c),
\]

---

5. These may be of interest in themselves, or because they affect future productivity and earnings capacities.

6. Perhaps the most visible efforts to data to explore whether who receives income matters are Schultz (1990) and Thomas (1990). These studies use "unearned" income (i.e., income from other than current labor market earnings) in their analyses in order to attempt to represent command over resources of various individuals without confounding price and taste effects through proxies for the price of time or productivity or for materialistic tastes that plague earlier studies in this genre that use schooling or wages. Such a motive for the use of unearned income is commendable, though unearned income may be correlated with time prices and leisure/goods tastes since it comes substantially from asset earnings from past savings (see Behrman 1994 for further discussion). In any case, the data that we use for this study do not identify unearned income by recipient, so we can not follow the approach used in these studies.
where the superscript h means that the variables are defined parallel to those in relation (1) for cognitive achievement, but refer to child health instead of cognitive achievement. The full-income constraint states that full-income is equal to time and other resource expenditures on the items that enter into the relevant production relations and the objective function:

\[ Y = P_y S + P_y C_e + P_y C_u + P_y X_e + P_y Z + P_y N_e + P_y C_e^h + P_y C_e^{hu} + \ldots \]

The constrained maximization of the objective function leads to reduced forms for all outcomes determined by the household (W) — including child cognitive achievement, child health, child schooling attainment, and other variables allocated by the household that enter into the production of cognitive achievement (X) — as dependent on household taste parameters, all predetermined household assets, and prices (including their monetary and their time components):

\[ W = W(I_{e0}, F_{e0}, C_{e0}, I_{e}, F_{e}, C_{e}, I_{u}, F_{u}, C_{u}, X_{e}, F_{x}, C_{x}, \ldots) \]

We now consider, within this simple one-period framework, why ordinary least squares estimates of relation (1) are likely to lead to biased estimates of the effects of child health on child cognitive achievement. The following factors in isolation each lead to an upward bias:

1. **Heterogeneity in tastes regarding child quality versus parental consumption:** Some parents care more about child quality relative to their own consumption than do others. In such a case a higher value of \( \delta_q \) is likely to lead to higher child health, child schooling, and child cognitive achievement in relations (5). This is likely to be reflected in relation (1) in a higher value of \( X_e \), which is unobserved and therefore in the OLS disturbance. As a result, \( H_e \) and that disturbance are positively correlated, so the coefficient estimate for \( H_e \) is upward biased since \( H_e \) represents the correlated part of the unobserved \( X_e \) in the estimates in addition to representing the effect of child health per se.

2. **Heterogeneity in unobserved predetermined family endowments that affect the production of child quality:** Some parents are more productive than others in producing all forms of child quality in ways that are not captured in observed representations of parents’ characteristics such as their schooling. Such parents have more \( F_{e}^u \) and \( F_{e}^{hu} \), which is likely to lead to higher child health, child schooling and child cognitive achievement through relations (1), (3) and (5). If relation (1) is estimated by OLS, therefore, \( H_e \) is positively correlated with the disturbance term.

---

7. Critical inputs into child health status include nutrient intakes, perhaps in interaction with the child's past health status, particularly regarding intestinal disorders. Unfortunately individual nutrients are not observed in the data set that we use, nor in any other data sets of which we are aware that has the other data necessary for our approach. Therefore we do not directly estimate this production function as part of the present study.
since the latter includes $F_e^u$. Thus the coefficient of $H_e$ is upward biased since $H_e$ again proxies for the correlated part of the omitted variable (in this case, $F_e^u$).

3. **Heterogeneity in unobserved predetermined community endowments that affect the production of child quality**: Some communities have characteristics that are more conducive than others in producing all forms of child quality in ways that are not captured in observed representations of communities' characteristics. Such communities have more $C_e^u$ and $C_e^{hu}$, which is likely to lead to higher child health, child schooling and child cognitive achievement through relations (1), (3) and (5). If relation (1) is estimated by OLS, therefore, $H_e$ is positively correlated with the disturbance term since it includes $C_e^u$. Thus the coefficient of $H_e$ is upward biased since $H_e$ again proxies for the correlated part of the omitted variable (in this case, $H_e^u$).

4. **Unobserved predetermined child characteristics affect in the same direction the production of both child health and child cognitive achievement**: Suppose that a child who is more robust has better measured child health and, through greater energy, learns better beyond the effects controlled by observed child health so that $I_C^u$ and $I_e^{hu}$ are positively correlated. In this case, once again, $H_e$ is correlated positively with the residual in relation (1) (this time through $I_C^u$), so its coefficient estimate is biased upwards.

5. **Unobserved heterogeneity across households in access to capital markets**: Poorer households appear to have less capital market access, or access on less favorable terms, than do better-off households. If capital market access limits investment in $H_C$ and in $X_C^u$, $H_e$ is positively correlated with the disturbance term once again, so its OLS coefficient estimate is biased upwards since it proxies for other unobserved investment inputs that affect cognitive achievement.

If there is an upward bias, then the residual in the cognitive achievement production function represents factors such as those just discussed that have effects in the same direction on child health and child schooling. Therefore, introducing this residual into the reduced-form estimates in relation (5) should result in positive coefficient estimates for this residual for both child health and child cognitive achievement.

There also are factors that, in isolation, each lead to a **downward bias**: 

8. For example, see Behrman, Foster and Rosenzweig (1994) and Foster (1993), though Paxson (1993) presents some evidence to the contrary for a somewhat higher-income economy (Thailand) than the one that we consider in our empirical work below.

9. The residual of relevance is not the residual in the production function estimates that control for simultaneity by using, say instrumented health. Instead it is the residual that is calculated by using actual health and consistent point estimates of the parameters in relation (1).
1. **Unobserved heterogeneity in parental tastes regarding the cognitive achievement versus health composition of child quality:** Suppose that because of inherent preferences some parents value highly a child who succeeds intellectually, while others value highly a child who is physically robust. Within the framework presented above, this effect is captured by the parameter $\delta_{ah}$. If cognitive achievement is valued relatively highly due to pure taste considerations, more unobserved resources are likely to be allocated to child cognitive achievement and less to health than in an otherwise identical family with a lower value of this parameter. In this case in OLS estimates of relation (1), $H_c$ is negatively correlated with $X_{c''}$, so its coefficient estimate is biased downwards.

2. **Unobserved heterogeneity across households in the expected returns to cognitive achievement versus child health:** Suppose that some parents want a child who succeeds intellectually, while others want a child who is physically robust because of differential expected returns to such characteristics in different labor markets. For example, cognitive achievement may be rewarded more in more modern parts of an economy and physical robustness more in more traditional primary sectors in which strength and stamina have relatively high payoffs. Within the framework presented above, this effect is parallel to that captured by the parameter $\delta_{ah}$. If cognitive achievement is valued relatively highly due to higher expected returns considerations, more unobserved resources are likely to be allocated to child cognitive achievement and less to health than in an otherwise identical family with a lower value of this parameter. In this case, again, in OLS estimates of relation (1), $H_c$ is negatively correlated with $X_{c''}$, so its coefficient estimate is biased downwards.

3. **Heterogeneity in parental unobserved endowments regarding their relative efficiency in producing child cognitive achievement and child health:** Suppose that some parents are relatively better at creating a home environment that is conducive to cognitive achievement (e.g., through articulating their curiosity about how things work) and others are relatively better at creating a home environment that is conducive to physical health (e.g., through being role models with good health habits). This means that $F_{c''}$ and $F_{c'bu}$ are negatively correlated. As a result, in OLS estimates of relation (1), $H_c$ is negatively correlated with the residual (in this case, the $F_{c''}$ part), so its coefficient estimate is biased downward.

4. **Heterogeneity in community unobserved endowments regarding their relative conduciveness in producing child cognitive achievement and child health:** Suppose that some communities are more conducive to improving cognitive achievement (e.g., through the greater emphasis on intellectual activities and greater links with the rest of the world) and others are more conducive to better physical health (e.g., through less congestion and pollution and more scope for more physical

---

10. Under the plausible assumption that for affecting cognitive achievement alone the most effective use of resources at the margin is not likely to be through improving child health.
activity for children). This means that $C_{Cu}$ and $C_{Bu}$ are negatively correlated. As a result, in OLS estimates of relation (1), $H_e$ is negatively correlated with the residual (in this case, the $C_{Cu}$ part), so its coefficient estimate is biased downward.

5. **Heterogeneity in unobserved child characteristics that affects the production of child cognitive achievement versus child health:** Suppose that some children have more inherent intellectual curiosity that increases their cognitive achievement (beyond the control for ability) and others have greater physical exuberance that improves their physical health (through their relatively active physical life). In this case $I_{Cu}$ and $I_{Bu}$ are negatively correlated. As a result, in OLS estimates of relation (1), $H_e$ is negatively correlated with the residual (in this case, the $I_{Cu}$ part), so its coefficient estimate is biased downward.

6. **Heterogeneity in unobserved prices for unobserved inputs used to produce child cognitive achievement versus child health:** Suppose that $P_{Xc}/P_{Xh}$ varies across households. Then otherwise identical households facing relatively low prices for unobserved inputs used for child cognitive achievement relative to those used for child health tend to purchase more $X_c$ relative to $X_h$. If so, $X_c$ and $X_h$ are negatively correlated and $H_c$ and $X_c$ are negatively correlated. Therefore the OLS estimate of the coefficient of $H_c$ in relation (1) is biased downwards.

If there is an downward bias, then the residual in the cognitive achievement production function represents factors such as those just discussed that have effects in the opposite directions on child health and child cognitive achievement. Therefore, introducing this residual into the reduced-form estimates in relation (5) should result in coefficient estimates for this residual of opposite signs in the child health and child schooling reduced forms.

Thus, this framework suggests that biases may occur in OLS estimates of the coefficient of $H_e$ in relation (1) in either direction. Of course, as usually is the case if considerable complexities are allowed in economic models, it is not possible with existing data sets to sort out all of these possibilities. However we can make some major progress with the data that we use by exploring the first order question of what direction of bias dominates.
Data

The data that we use for our empirical analysis of the possible effects of the simultaneous determination of child health on child schooling success are from the Ghanaian Living Standard Measurement Study (LSMS) data. This data set includes child anthropometric measures to represent child health, cognitive achievement test scores to represent schooling success, a test of pre-school ability to control for that dimension of child endowments, a fairly rich range of household and community characteristics to use for estimation of relation (5) as part of the simultaneous estimates, and sibling data to explore unobserved fixed family and community effects. We limit our sample to children who are in the age range 9-17, as we discuss below with respect to the child age variable.

Table 1 gives the sample means and standard deviations for the variables that we use in our analyses for the approximately (depending on the variable) 1200 observations in our sample. We now discuss briefly the major variables that enter directly into our cognitive achievement production function estimates. For further details on the overall sample and on various schooling-related variables, see Glewwe (1992) and Glewwe and Jacoby (1994).

Cognitive Achievement Test Scores

The cognitive achievement tests\(^{11}\) cover reading and mathematics, with a simple and an advanced subtest for each. The cognitive achievement tests that were used were adapted from tests designed for use in East Africa by the Educational Testing Service in Princeton, NJ as part of the project that is summarized in Knight and Sabot (1990) in which also are published sample questions from these tests. Originally the intention was to administer these tests (and the Raven’s test of pre-school ability described below) only to individuals age nine or older with three or more years of schooling. In the field work, 1467 individuals\(^{12}\) age 9-17 took the Raven’s test and 910 individuals took at least the simple subtests. The age limitation was strictly enforced, but not the schooling limitation: 123 children with less than three years of schooling took cognitive achievement tests and 85 children with more than three years of schooling did not take the cognitive achievement tests because they were not

---

11. The details of the procedures followed in the data collection are based on documentation prepared by (e.g., Glewwe 1991, 1994) and conversations with Paul Glewwe (World Bank project head for the data collection) and Dean Jolliffe (World Bank research assistant for the data preparation for the analysis that Glewwe undertook with these data). The test scores that give us a sample of 1205 are those used (for a different subsample defined by middle-school attendance and without limitation to those for whom local health service characteristics are available) by Glewwe and Jacoby (1994).

12. About a sixth of the eligible children did not take any tests primarily because they were absent. Such children may not be a random draw of the population. But we note that: (i) absence also may be a problem with other studies of cognitive achievement in the literature noted above, (ii) to the extent that such absence reflects household characteristics our family fixed effects estimates controls for the factors determining such sample selectivity, (iii) explicit control for such selectivity through a two-stage procedure does not change materially the estimates presented below, and (iv) we explore how robust our basic results are to other changes in the sample because of the imputation issues discussed below.
Table 1: Descriptive Statistics for Ghanaian Children Aged 10-17 in Living Standards Measurement Survey (LSMS) Sample

<table>
<thead>
<tr>
<th>Endogenous Variables</th>
<th>Entire Sample</th>
<th>Sibling Subsample</th>
<th>Entire Sample</th>
<th>Sibling Subsample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard Deviation</td>
<td>Mean</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Endogenous Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cognitive achievement test overall</td>
<td>7.5</td>
<td>11.7</td>
<td>10.0</td>
<td>12.6</td>
</tr>
<tr>
<td>reading</td>
<td>2.9</td>
<td>6.8</td>
<td>3.9</td>
<td>7.6</td>
</tr>
<tr>
<td>mathematics</td>
<td>4.6</td>
<td>5.9</td>
<td>6.0</td>
<td>6.1</td>
</tr>
<tr>
<td>height (z score)</td>
<td>-1.5</td>
<td>1.2</td>
<td>-1.5</td>
<td>1.2</td>
</tr>
<tr>
<td>years of schooling</td>
<td>4.3</td>
<td>2.5</td>
<td>4.7</td>
<td>2.4</td>
</tr>
<tr>
<td><strong>Predetermined Child Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>0.53</td>
<td>0.50</td>
<td>0.57</td>
<td>0.49</td>
</tr>
<tr>
<td>age (years)</td>
<td>12.8</td>
<td>2.5</td>
<td>12.8</td>
<td>2.4</td>
</tr>
<tr>
<td>pre-school ability</td>
<td>17.5</td>
<td>5.7</td>
<td>17.9</td>
<td>6.1</td>
</tr>
<tr>
<td><strong>Family Background</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log (per capita expenditures)</td>
<td>10.8</td>
<td>0.32</td>
<td>10.8</td>
<td>0.32</td>
</tr>
<tr>
<td>father's years of schooling</td>
<td>5.8</td>
<td>5.9</td>
<td>6.9</td>
<td>5.9</td>
</tr>
<tr>
<td>mother's years of schooling</td>
<td>3.0</td>
<td>4.5</td>
<td>3.5</td>
<td>4.8</td>
</tr>
<tr>
<td>mother's age (years)</td>
<td>27.4</td>
<td>21.0</td>
<td>28.4</td>
<td>20.3</td>
</tr>
<tr>
<td>mother's height</td>
<td>61.4</td>
<td>46.4</td>
<td>63.3</td>
<td>45.4</td>
</tr>
<tr>
<td>father's height</td>
<td>42.5</td>
<td>47.3</td>
<td>44.3</td>
<td>47.7</td>
</tr>
<tr>
<td>mother's height missing</td>
<td>0.36</td>
<td>0.48</td>
<td>0.32</td>
<td>0.47</td>
</tr>
<tr>
<td>father's height missing</td>
<td>0.55</td>
<td>0.49</td>
<td>0.54</td>
<td>0.50</td>
</tr>
<tr>
<td>head of household age (years)</td>
<td>49.8</td>
<td>14.0</td>
<td>49.6</td>
<td>13.1</td>
</tr>
<tr>
<td>head of household sex male</td>
<td>0.66</td>
<td>0.48</td>
<td>0.64</td>
<td>0.48</td>
</tr>
<tr>
<td><strong>Community Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>minutes to nearest middle school</td>
<td>27.5</td>
<td>47.1</td>
<td>20.6</td>
<td>30.2</td>
</tr>
<tr>
<td>minutes to nearest primary school</td>
<td>17.0</td>
<td>23.7</td>
<td>14.3</td>
<td>20.5</td>
</tr>
<tr>
<td>% of households in cluster with:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>private water</td>
<td>0.20</td>
<td>0.34</td>
<td>0.24</td>
<td>0.36</td>
</tr>
<tr>
<td>standpipe</td>
<td>0.06</td>
<td>0.20</td>
<td>0.03</td>
<td>0.14</td>
</tr>
<tr>
<td>water from vendor</td>
<td>0.03</td>
<td>0.11</td>
<td>0.04</td>
<td>0.10</td>
</tr>
<tr>
<td>well with pump</td>
<td>0.10</td>
<td>0.26</td>
<td>0.06</td>
<td>0.18</td>
</tr>
<tr>
<td>flush toilet</td>
<td>0.18</td>
<td>0.30</td>
<td>0.15</td>
<td>0.27</td>
</tr>
<tr>
<td>latrines</td>
<td>0.05</td>
<td>0.14</td>
<td>0.05</td>
<td>0.15</td>
</tr>
<tr>
<td>pan/bucket</td>
<td>0.54</td>
<td>0.39</td>
<td>0.60</td>
<td>0.38</td>
</tr>
<tr>
<td>health facilities:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cluster 0-6 miles from facility</td>
<td>0.78</td>
<td>0.42</td>
<td>0.88</td>
<td>0.32</td>
</tr>
<tr>
<td>ampicillin in stock</td>
<td>0.41</td>
<td>0.49</td>
<td>0.40</td>
<td>0.49</td>
</tr>
<tr>
<td>number of working medical doctors</td>
<td>1.1</td>
<td>2.3</td>
<td>1.0</td>
<td>2.2</td>
</tr>
<tr>
<td>number of beds in facility</td>
<td>20.6</td>
<td>44.3</td>
<td>16.3</td>
<td>36.0</td>
</tr>
<tr>
<td>price per consultation</td>
<td>71.1</td>
<td>104.9</td>
<td>74.8</td>
<td>111.2</td>
</tr>
<tr>
<td>hours open per week</td>
<td>63.9</td>
<td>48.1</td>
<td>69.6</td>
<td>52.1</td>
</tr>
<tr>
<td>postnatal services</td>
<td>0.72</td>
<td>0.44</td>
<td>0.78</td>
<td>0.41</td>
</tr>
<tr>
<td>public facility</td>
<td>0.67</td>
<td>0.47</td>
<td>0.73</td>
<td>0.44</td>
</tr>
<tr>
<td>laboratory in facility</td>
<td>0.37</td>
<td>0.48</td>
<td>0.38</td>
<td>0.48</td>
</tr>
<tr>
<td>offer immunizations</td>
<td>0.61</td>
<td>0.49</td>
<td>0.66</td>
<td>0.47</td>
</tr>
<tr>
<td>geographical area:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>forest zone</td>
<td>0.49</td>
<td>0.50</td>
<td>0.63</td>
<td>0.48</td>
</tr>
<tr>
<td>Savannah zone</td>
<td>0.20</td>
<td>0.40</td>
<td>0.11</td>
<td>0.31</td>
</tr>
<tr>
<td>rural</td>
<td>0.48</td>
<td>0.50</td>
<td>0.45</td>
<td>0.50</td>
</tr>
</tbody>
</table>

a. With missing treated as zero. For those without missing observations for mothers the mean is 95.9 percent and for fathers the mean is 94.4 percent of age-specific international standards.
able to answer any of the simple subtest questions (these individuals were imputed a score of zero in the data files). The advanced subtests, as is standard practice in this type of tests, were supposed to be given only to those who answered correctly at least four questions on the simple tests — about half of those who took the simple reading test and two-thirds of those who took the simple mathematics test. The limitation on giving the advanced subtests only to those who took the simple test generally was followed, though there were some exceptions. Those who did not take the advanced tests because of their poor performance on the simple tests were imputed a score of zero on the advanced tests. The maximum sum of the beginning and advanced test scores for the reading component is 37 and for the mathematics component is 44. The mean performance of the sample of 1205 who are recorded in the data files as having cognitive achievement scores (including imputations) and the other data that we use in our analysis was low on both test scores, with fairly large variances. This includes 438 individuals (36 percent) with zero scores. In our estimates below we present a series of estimates for alternative samples as well as for this sample of 1205 because of the large mass point at zero and because of the issues raised by the sample imputations: (1) using the total cognitive achievement scores (including imputations in cases in which they were made for the advanced tests) for the sample of 910 individuals who took the simple subtests, (2) using the simple subtest results only for the sample of 910 individuals who took the simple subtests, and (3) using a subsample of 881 individuals who had positive cognitive achievement test scores.

**Child Health**

We use the Z score for child height as our basic measure of longer-run child health, as is standard in the nutrition literature (and increasingly common in the economics literature). The mean value for our sample of the Z score for height is -1.5 and the standard deviation is 1.2, indicating that the distribution of heights is concentrated below the age- and sex-specific standards. In some alternative estimates we also have explored the impact of children's body mass index (BMI), a shorter-run index of health. Since some of the

---

13. This subsample includes 120 individuals who were not included in the sample of 1205 in our original estimates because they had missing data on the price of health consultations, one of the instruments used for the estimates. Since that instrument is not statistically important in the first-stage estimates and imposing this restriction reduces the sample size a lot, this instrument was dropped.

14. Among these 910 individuals, 412 were not able to answer any of the simple test questions. These individuals were supposed to be given a test score of zero, but in the field "na" (not available) was indicated instead. We therefore also have made estimates for the 498 individuals who were a subset of the 910 individuals who took the simple subtests and for whom in the field numerical scores were indicated. In this group 2 percent had zero scores.

15. The Z score indicates how many standard deviations a particular child's height is from age- and sex-specific standards. We use the NCCHS (U.S. National Center for Health Statistics) standards.

16. The BMI is defined as the ratio of weight in kilograms to the square of height in centimeters. Cole (1991) and Fogel (1991a,b) survey the use of this indicator.
children in the sample already completed school, but the anthropometric indicators are available only at the time of the survey, the shorter-run BMI probably is a less satisfactory indicator than is the Z score for child height. Also a priori the accumulative effect of longer-run health is likely to be more important in the determination of the stock cognitive achievement variable than the shorter-run health represented by the BMI even for children still in school.

**Grades of Schooling at Time of Cognitive Achievement Test**

At the time of the survey the averaged completed schooling was 4.3 grades, with a standard deviation of 2.5 grades. The use of the 9-17 age group means that the observed completed grades of schooling variable is truncated as a representation of completed or optimal schooling for a number of members of the sample. Of course, in our estimate of the cognitive achievement production function in relation (1) we do not want completed schooling, only schooling completed at the time that the cognitive achievement test was taken. Therefore, in a sense, the truncation of the observed schooling variable is not a problem. Yet, in another sense, it may be. We want to treat grades of schooling at the time of taking the cognitive achievement test as endogenous since that is an implication of the model of Section 3. But relation (5) indicates how to estimate not schooling completed at the time that the cognitive achievement test was taken, but completed schooling. We therefore obtain estimates of schooling completed at the time that the cognitive achievement test was taken in two ways. **First,** we estimate relation (5) for completed grades of schooling with control for truncation and then use the minimum of that estimate and the years elapsed since an individual started school as the years of school at the time that the test was taken. **Second,** we estimate a relation with the same right-side variables for years of school completed at the time that the cognitive achievement test was taken and use this estimate directly in our simultaneous estimates of the cognitive achievement production function. We report here that the estimates of the impact of child health on child cognitive achievement do not differ significantly between these two alternatives.

**Pre-School Ability**

To obtain a measure of pre-school reasoning ability, Raven's (1956) Coloured Progressive Matrices (CPM), a test of reasoning ability that involves the matching of patterns, was administered to everybody in the sample nine years of age or older. The test is designed so that formal schooling does not influence performance, though performance may reflect early childhood environment as well as innate capacity. The maximum score possible on the test is 36. The mean score obtained for our sample is 17.5, with a standard deviation of 5.7. This test has been used to control for pre-school ability in estimates of the

---

17. Knight and Sabot (1990) provide some illustrations from this test. Since the tests were administered to our sample members when they were nine years old and older, rather than pre-school ability, it might be more accurate to refer to these tests as measuring ability that is independent of schooling. But such phraseology is cumbersome, so, with this caveat, we continue to refer to these tests as measurements of pre-school ability.
determination of cognitive achievement in some other studies in the economics literature (e.g., Alderman et al. 1992) and in other literatures (e.g., Nokes et al. 1992b). Since there may be a possibility that this test result is determined endogenously, we also have estimated a set of relations in which the coefficient of this test score is constrained to zero. We report here that this restriction does not change the pattern of estimates of child health of interest in this study.

**Parental Schooling**

One special positive feature of the Ghanaian LSMS is that it asks each individual for his or her parents' schooling. That means that information is available about parental schooling even if one or both parents do not reside in the same household due to death, fostering, marital dissolution, migratory work or whatever. In fact for our sample co-residence with a biological parent(s) does not occur for about a quarter of the children, primarily because fostering is common in West Africa. Therefore having information on parents' schooling whether or not they are co-resident is useful. The mean years of schooling for fathers is 5.8 years, almost twice the mean of 3.0 years for mothers. For both fathers and mothers there is a fair amount of variance in years of schooling.

**Child Age**

Since the critical data on cognitive achievement test scores are available only for children nine years of older, we limit our sample to children who are at least nine years old. In order to focus on a relatively young cohort and to lessen any possible problem of older children having selectively left sample households, we limit the sample to children under 18.

*Child sex:* Slightly more than half (53 percent) of our sample is male.

**Variables for Reduced-Form Estimates for Child Health and Child Schooling Attainment at Time the Cognitive Achievement Test was Taken**

The other variables that are given in Table 1 do not enter directly into our estimated cognitive achievement production functions, but only indirectly through the variables that may be determined simultaneously with cognitive achievement (i.e., years of schooling and child health). Therefore we do not discuss them in any detail here. But we do wish to emphasize that the data set includes a rich array of household and community variables with substantial sample variance that, under the assumption that there is no unobserved household allocated input ($X_c^n$), constitute a plausible set of first-stage instruments.  

18. There is information on the qualities of the local school alternatives that we do not include among these first-stage instruments. Glewwe and Jacoby (1994) posit that the school that a child attends is not necessarily the closest school, but a matter of choice. If so, to include school characteristics in the cognitive achievement production function would require treating as endogenous a number of school characteristics which would add (continued...)
Estimates of Cognitive Achievement Production Function with Instrumental Variable Control for Simultaneity

Table 2 presents alternative estimates of the overall cognitive achievement production function in relation (1). To focus on the point of interest in this paper, we keep the specification simple, with linear terms in the Z score for child height, child schooling at the time the cognitive achievement test was taken, child pre-school ability, child age, child sex, and parental schooling. For all of these alternatives except the first one we treat years of schooling at the time that the cognitive achievement test was given as endogenous.

OLS versus Instrumental Variable Estimates with Basic Sample:

The first three columns in Table 2 give estimates using the largest sample that we consider in Section 3 with 1205 individuals. Column 1 gives OLS estimates with child height and child schooling attainment at the time of the test treated as independent of the disturbance term as in the previous literature. Column 2 gives the estimates in which child schooling attainment, but not child health, is instrumented. Column 3 gives the estimates in which child schooling and child height both are instrumented. These instrumented

18. (...continued)
function would require treating as endogenous a number of school characteristics which would add substantial complexity to our analysis without adding to our investigation of the impact of child health on child schooling success. Therefore we do not include school characteristics in most of our cognitive achievement production function estimates, but also do not include them in our basic set of instruments since they a priori would seem to be correlated with the disturbance term in relation (1). If we do include observed indicators of school quality and treat them as predetermined the patterns of results with respect to the coefficient estimates for child health is not changed significantly.

19. We have explored some nonlinearities. For example, we have included the product of parental schooling in addition to the linear terms for parental schooling. In this case the coefficient estimates for father's and mother's schooling, respectively, are 0.18 and 0.01 (with absolute t values of 2.8 and 0.1) and that for the interaction is 0.022 (with a t value of 2.2). It is of general interest that mother's schooling appears significant only in interaction with father's schooling, but not vice versa. But the relevant point for the present paper is that the introduction of this interaction does not change significantly any of the coefficient estimates for the child's characteristics (including, in particular, child height). Therefore, for simplicity, we focus on the simple linear specification in the text.

20. First-stage estimates of relation (5) for the Z score for child height and for years of completed schooling at the time the cognitive achievement test was taken have as right-side variables 42 variables that represent individual child characteristics (e.g., age, sex, pre-school ability), parental and household characteristics (e.g., years of school of father and mother, household assets), and prices and community characteristics primarily related to details of the travel time prices of primary and middle schools, water and sanitation quality, and travel time to and quality of health clinics. There has been concern expressed in the recent literature (e.g., Bound, Jaeger and Baker 1993, Deaton 1994, Nelson and Startz 1990a,b) about the use of instrumental variable estimates when the first-stage estimates have very little predictive power (e.g., $R^2$ close to 0.00, F tests of magnitude around 1.0 or less). For our first-stage estimates the $R^2$ are 0.65 and 0.78 for child height and child schooling respectively, and the F tests are 49.9 and 93.2 respectively, far above the levels about which this literature expresses concern.
Table 2: Cognitive Achievement Production Function Estimates, Ghanaian LSMS: CLS and Simultaneous Estimates

<table>
<thead>
<tr>
<th>Child Characteristics</th>
<th>(1) Largest Sample with Data Available Including Imputed Test Scores</th>
<th>(2) Total Cognitive Achievement Scores, Only Individuals who Took Easy Tests</th>
<th>(3) Easy Tests Cognitive Achievement Scores, Only Individuals who Took Easy Tests</th>
<th>(4) Only Individuals with Positive Cognitive Achievement Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS Estimates (1)</td>
<td>Height Predetermined (2)</td>
<td>Height Simult. (3)</td>
<td>OLS Estimates (4)</td>
</tr>
<tr>
<td>Height (2 score)</td>
<td>0.6(2.8)</td>
<td>0.6(3.6)</td>
<td>1.9(3.2)*</td>
<td>0.53(1.9)</td>
</tr>
<tr>
<td>Years of schooling</td>
<td>1.6(16.6)</td>
<td>1.4(6.0)`</td>
<td>1.5(6.3)`</td>
<td>2.1(11.6)</td>
</tr>
<tr>
<td>Age</td>
<td>0.018(1.9)</td>
<td>0.032(2.3)</td>
<td>0.033(2.3)</td>
<td>-0.02(1.2)</td>
</tr>
<tr>
<td>Male</td>
<td>1.2(2.6)</td>
<td>1.3(2.7)</td>
<td>1.9(3.4)</td>
<td>1.5(2.3)</td>
</tr>
<tr>
<td>Pre-School Ability</td>
<td>0.68(15.0)</td>
<td>0.71(13.6)</td>
<td>0.64(10.4)</td>
<td>0.77(13.8)</td>
</tr>
<tr>
<td>Parental Schooling</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father</td>
<td>0.17(3.6)</td>
<td>0.21(3.8)</td>
<td>0.19(3.4)</td>
<td>0.18(3.1)</td>
</tr>
<tr>
<td>Mother</td>
<td>0.23(3.7)</td>
<td>0.25(3.9)</td>
<td>0.22(3.2)</td>
<td>0.21(2.9)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-14.8(10.3)</td>
<td>-16.7(8.5)</td>
<td>-14.6(6.6)</td>
<td>-13.2(6.2)</td>
</tr>
<tr>
<td></td>
<td>0.56</td>
<td>0.52</td>
<td>0.51</td>
<td>0.51</td>
</tr>
<tr>
<td>Root MSE</td>
<td>7.8</td>
<td>7.8</td>
<td>8.0</td>
<td>9.0</td>
</tr>
<tr>
<td>F</td>
<td>219.7</td>
<td>188.3</td>
<td>177.4</td>
<td>135.7</td>
</tr>
<tr>
<td>N</td>
<td>1205</td>
<td>1205</td>
<td>1205</td>
<td>910</td>
</tr>
</tbody>
</table>

a. See Table 1 for basic statistics for the underlying data. Absolute values of t statistics are given in parenthesis to the right of the point estimate.
b. Treated as simultaneously determined, with first-stage estimates using (nonendogenous) variables in Table 1 plus school quality indicators for columns 5-7.
c. Treated as simultaneously determined, with first-stage estimates using (nonendogenous) variables in Table 1.
estimates are consistent under the assumption that the instruments are not correlated with the disturbance term in relation (1), i.e. that the estimates are consistent under the assumption that the instruments are not correlated with the $L^u$, $F^u$, $C^u$, and $X^u$ that enter into this cognitive achievement production function relation. In this section we treat as a maintained assumption that these instruments satisfy this condition (though we return to this assumption in the next section). A parallel maintained assumption has been made in a number of other recent studies (e.g., Pitt, Rosenzweig and Hassan 1990, Rosenzweig and Schultz 1983, 1987, 1988, Schultz and Tansel 1993).

First we discuss the coefficient estimates for the variables other than child height in order to see what they imply. We note that for none of these variables do the coefficient estimates differ significantly among the three columns, and for all of them the coefficient estimates seem to be plausible a priori. Every additional year of child schooling increases the overall cognitive achievement test score by about 1.4 to 1.6 points, and every additional point on the pre-school ability test increases the overall cognitive achievement test score by about two-thirds of a point. Child age increases overall cognitive achievement scores significantly, though the point estimates vary a lot (though not significantly) depending upon whether schooling is instrumented. Males have about 1.3 to 1.9 points on the average significantly higher scores than do females. This gender difference may reflect gender differences in male versus female conditioning to succeed on tests or in the allocation of unobserved resources to boys versus girls in the home or in school. An additional year of parental schooling is estimated to increase child cognitive achievement by about 0.2 to 0.3 points, independently of for which parent the higher schooling is observed. This result is of interest in light of the frequent tendency in the literature to assert that mother’s schooling is

---

21. Since most data sets that have been used to investigate the impact of schooling on test scores and other outcomes do not have information on pre-school ability, it is of interest to ask how the coefficient estimates change if pre-school ability is excluded from the specification a priori. For the specification in column two with this restriction in addition, the coefficient estimates for child height, schooling and sex are approximately double those in column two and those for parental schooling are about three-fifths as large (while that for child age remains insignificant). Assuming that the true model includes pre-school ability, therefore, the exclusion of pre-school ability makes other child characteristics appear much more important but — once there is control for the choices regarding those child characteristics — parental schooling less important. But, as we note in Section 3, whether or not the pre-school ability is included does not change importantly the results concerning the bias in the estimated coefficient of child health in the standard estimates.

22. The sign of this estimate differs from that reported for rural Pakistan in Alderman, Behrman, Ross and Sabot (1992). In the latter study, once there is control for pre-school ability and years of school (for both of which the data indicate significantly higher means for males than for females) females have significantly higher cognitive achievement scores. In that case, since schools tend to be single-sex, the possibility of differential allocation of school resources by sex seems sharper and, indeed, teacher/student ratios are much higher for girls’ schools than for boys’ schools.

23. Our residuals are purged of any such gender effect in unobserved resource allocation by the inclusion of the child sex variable in the estimates.
more important than is father's schooling in child development. But for some subsamples for which estimates are reported below (i.e., only children co-resident with biological parents, only males, only females, only families with siblings in the sample) there are significant differences between the estimated effects of paternal versus maternal schooling.

Next, we come to the question of central interest regarding these estimates. Does treatment of child height as simultaneously determined affect the estimated impact of child height on child overall cognitive achievement? These estimates suggest that the answer is yes. The coefficient estimates for child health differ significantly between the first two columns in which child health is treated as predetermined and column three in which the instrumented value is used (e.g., t = 2.2 for the difference between the estimates in columns two and three). Moreover, they vary substantially, with that with control for simultaneity 3.1 times that without such control. Therefore, conditional on the assumption discussed two paragraphs above, these estimates suggest that the second set of possible biases discussed in Section 3 dominate in standard OLS estimates, so that child health has a much more powerful positive effect on child schooling success than indicated by standard OLS estimates. A further illustration of the fact that these estimates are consistent with the second set of possible simultaneity biases is provided by exploring what happens if the residuals from the cognitive achievement production function are introduced into the reduced-form relations for child height and cognitive achievement. As predicted with regard to the second set of possible simultaneity biases in Section 3, the coefficient estimates are opposite in sign in the child health reduced form relation from that in the reduced form relation for the child overall cognitive achievement test score.

We also note that the dependence of the coefficient estimates for child health on whether or not it is treated as simultaneously determined differ from those for child schooling. The comparison between the estimates in which both child health and child schooling are treated as predetermined with the estimates in which both are treated as simultaneously determined (i.e., column 1 versus column 3) reveals no change in the schooling coefficient estimate (though there is somewhat less precision in the point estimate in column 3) but a substantial change in the health coefficient estimate (with somewhat greater precision in column 3). Therefore the change in the estimated health effect does not just reflect that instrumenting any variable with the instruments in Table 1 causes an increase in the estimated coefficient because such a procedure picks up the effects of unobserved choice variables (Xc). This does not happen for the coefficient estimates of the schooling variable.

24. For example, see Colclough (1982), King (1990), King and Hill (1993), Mensch, Lentzner and Preston (1985), Schultz (1989, 1993) and World Bank (1980, 1981). Also see Behrman (1990c) for references to studies other than the present one that also find that mother's schooling does not have a significantly larger effect on child outcomes than does father's schooling.

25. Because of space constraints the full estimates are not presented here. The respective coefficient estimates (with absolute t values in parentheses) are -0.041 (7.5) and 1.3 (53.9).
OLS versus Instrumental Variable Estimates with Alternative Samples or Dependent Variables

The imputations for cognitive achievement scores that are discussed in Section 3 raise the question of whether our estimates with the basic sample reflect some artifact of those imputations or problems because of the mass point at zero that violates the underlying distributional assumptions for the stochastic terms. Therefore, in columns 4-9 in Table 2, we present three additional sets of estimates for three alternative samples/dependent variables: (2) using the total cognitive achievement scores (including imputations in cases in which they were made for the advanced tests) for the sample of 910 individuals who took the simple subtests, (3) using the simple subtest results only for the sample of 910 individuals who took the simple subtests, and (4) using a subsample of 881 individuals who had positive cognitive achievement test scores. For each of these three set of estimates we present both the OLS estimates and estimates with both child schooling and child health instrumented (parallel to columns 1 and 3 for the sample of 1205 individuals).

There are some differences among the four sets of estimates in Table 2. Most striking, but hardly surprising, the estimated impact of most of the variables is substantially less if the test results are limited to those on the simple subtests rather than including both the simple and the advanced test results. But the basic patterns that we discuss above remain the same. That is, if child health is treated as simultaneous with instrumental variables instead of using OLS the estimated coefficient increases substantially (by factors of two to four) and significantly, while that for schooling does not change significantly depending on which estimator is used.

Variations on the Basic Estimates

For all four of the approaches that are summarized in Table 2 we also have undertaken other estimates for the cognitive achievement production function in relation (1) without and with the instrumental variable control for simultaneity to explore the robustness of the results in Table 2. We summarize verbally these estimates here. (1) Children not completing first grade: In our basic sample 16 percent of the children had not attended school enough even to finish the first grade. Perhaps the production technology for cognitive achievement differs for those who do not even finish the first grade. Therefore we estimated probits for completing at least the first grade, based on the same set of predetermined variables that are used for the first-stage estimates of child height and child schooling, to control for this possible selectivity, but find no significant differences in the estimated

26. For the further alternative subsample in which in the field "na" was not used to indicate that the individuals were unable to answer any questions (see note 14 above) the pattern is the same. The coefficient estimates for health and schooling, respectively (with t tests in parentheses) are 0.23 (1.6) and 0.61 (6.4) in the OLS estimates and 0.74 (2.4) and 0.63 (2.5) in the instrumental variable estimates.
coefficients.\textsuperscript{27} (2) School quality indicators in instrument set: The instrument set for the first-stage estimates for the cognitive achievement production function estimates in Table 2 do not include indicators of local school quality because of the possibility that school quality reflects choice, particularly at the middle school level (see Glewwe and Jacoby 1994). If we use an instrument set expanded to include the local school quality indicators, we find no significant differences in the estimated coefficients. (3) School quality in the production function: Since school quality may reflect choice we do not include school quality in our basic estimates of the cognitive achievement production function. But it is of interest to ask whether the same pattern of increases in the child health coefficient occurs if it is treated as simultaneously determined if we include in the production function average community school characteristics: school library available, blackboard available, teacher schooling, teacher experience, governmental expenditure, leaky classrooms.\textsuperscript{28} None of the coefficient estimates are affected significantly by including these schooling quality indicators in the relation. Therefore child health is not just proxying for observed schooling quality indicators that were excluded from our basic specification. (4) Only children co-resident with at least one biological parent: Child fostering for work and school is a fairly common practice in Ghana. A little over a quarter of our sample are not co-resident with even one biological parent. Perhaps the production technology for cognitive achievement differs for those who are co-resident with their biological parents versus those who are not. Estimates for the subsample of children who lived with at least one biological parent at the time of the survey differ notably from those in Table 2 only in that they suggest that mother’s schooling may have more impact on child cognitive achievement for co-resident children, with the point estimate for mother’s schooling almost twice as large as that for father’s schooling and significantly larger at the 10 percent level (t = 1.8). (5) BMI instead of height: As noted in Section 2, a shorter-run anthropometric indicator of child health that can be constructed from the data is the body-mass index (BMI). While there are some who suggest that BMI is a better index for some purposes (e.g. Cole 1991, Fogel 1991a,b), we prefer the longer-run measure of standardized height for our basic exploration since some of the children in our sample have been out of school for several years and since the cognitive achievement dependent variable is a longer-run stock variable. But the basic thrust of the results is not changed if BMI is used instead of child height. The coefficient estimates of the other included variables do not change significantly with the exception of age and the coefficient estimates of cognitive achievement determinants.

\textsuperscript{27} Glewwe and Jacoby (1994) also report that their control for middle school choice in the same sample does not affect their estimates of cognitive achievement determinants.

\textsuperscript{28} These characteristics were selected on the basis of the results in Glewwe and Jacoby (1994). In our estimates only teachers’ schooling is significantly nonzero at standard levels (with a coefficient estimate of 0.5 and a t value of 4.1. (In previous studies of such production functions for developing countries observed variables often do not appear significant although often there are teacher and school fixed effects that are significant, see Harbison and Hanushek 1992.) If average middle-school characteristics are used instead of primary school characteristics the coefficient estimates of child height follow the same pattern (e.g., changing from 0.5 with a t value of 2.4 to 1.6 with a t value of 2.8). If both primary and middle school average characteristics are included, the pattern is similar, though there is greater imprecision for the coefficient estimates of the school quality indicators because of greater multicollinearity.
estimate of child health (BMI) still increases substantially (by a factor of 6) and significantly \( (t = 3.1) \) in the instrumental variable estimates in comparison with OLS estimates.

(6) **Mathematics versus reading test scores:** The overall cognitive achievement test score, as noted in Section 3, is the sum of scores on mathematics and on reading. If these components are considered separately, school per se is relatively more important for increasing the mathematics scores, while pre-school ability, home environment, and maturity are relatively more important for increasing the reading scores. The estimated effect of child height is significantly greater for reading than for mathematics \( (t = 2.6 \text{ for the simultaneous estimates}) \) and control for simultaneity increases the coefficient estimate substantially for the reading score (the multiplicative factor is over three) and significantly \( (t = 2.6) \), but less for the mathematics score (the multiplicative factor is 2.3) and insignificantly even at the 10 percent level \( (t = 1.5) \). (7) **Females versus males:** There are gender roles in Ghanaian society that may affect the relations being estimated. An F test indicates that there are significant differences for the overall estimates by sex beyond differences in the intercepts \( (F = 6.8 \text{ with a critical value at the 1 percent level of 2.6}) \). There are significant gender differences in the relations beyond differences in the intercepts that apparently are concentrated in the parental and child schooling coefficients, but the standard errors are large enough that the differences in individual coefficient estimates are not highly significant. If child health is treated as predetermined rather than simultaneously determined, the effect is basically for males, not females. For males the point estimate increases by a multiplicative factor of 2.9 with control for simultaneity, but for females there is no significant change even at the 25 percent level.

---

29. For example, Thomas (1992) claims that there are stronger own than cross gender intergenerational links between parental schooling and child anthropometrics in Ghana as well as in some other societies. Lavy (1991), however, does not find such patterns for Ghanaian child schooling.
Estimates of Cognitive Achievement Production Function with Control for Unobserved Family and Community Effects

In the previous section we present simultaneous estimates under the maintained assumption that the instruments used are not correlated with the elements of $I^a$, $F^a$, $C^a$, and $X^a$ that enter into the cognitive achievement production function relation. However that maintained assumption would appear to be a strong one since the instruments are family and community characteristics that a priori would appear quite possibly to be associated with $F^a$ and $C^a$, respectively, and also be among the determinants of $X^a$. Observed family characteristics such as family assets, parental age, and parental height, for example, plausibly are associated with unobserved family characteristics that may enter directly into the cognitive achievement production function such as average family genetic endowments and the general household learning environment. Such observed family variables also are likely to be associated with the determinants of unobserved choice inputs into the production of cognitive achievement, such as the time parents spend reading with their children. Observed community characteristics relating to the quality of health services and water also are likely to be associated with the unobserved community learning environment that enters directly into the cognitive achievement production function, as well as be among the reduced-form determinants of unobserved family choice variables that enter in such as parental time allocation, once again. Thus, while the simultaneous estimates that we present in Section 4 are striking regarding their implications for the strong positive impact of child health on child schooling success conditional on the maintained assumption about the instruments used for the simultaneous estimates, this maintained assumption is not obviously valid.

Therefore in this section we present estimates that control for unobserved family and community characteristics through using sibling data and family dummy variables. These estimates eliminate from the disturbance term in relation (1) all unobserved family and community characteristics. Thus, they eliminate any biases in the estimated coefficients of child health due to unobserved family and community factors that are discussed in Section 2 such as heterogeneity in tastes regarding child quality versus parental consumption, heterogeneity in unobserved predetermined family endowments that affect the production of child quality, heterogeneity in family access to credit markets, unobserved heterogeneity in parental tastes regarding the cognitive achievement versus health composition of child quality, unobserved heterogeneity across households in the expected returns to cognitive achievement versus child health, heterogeneity in parental unobserved endowments regarding their relative efficiency in producing child cognitive achievement and child health, and heterogeneity in unobserved prices for inputs used to produce child cognitive achievement versus child health. However, these estimates do not eliminate biases due to unobserved

---

30. A number of other studies have claimed that unobserved family characteristics affect importantly child schooling success (e.g., Behrman, Hrubec, Taubman and Wales 1980, Behrman and Wolfe 1987, Olneck 1977).

31. The importance of unobserved community characteristics in understanding various dimensions of human resource investments in developing countries is emphasized in a number of recent studies (e.g., Rosenzweig and Wolpin 1986, Behrman and Deolalikar 1993, Foster and Roy 1993, Pitt, Rosenzweig and Gibbons 1993).
individual child characteristics such as those that also are discussed in Section 2. Also these estimates are more subject to biases towards zero due to random measurement error since the noise-to-signal ratio is increased by focusing on deviations from family means.\(^{32}\) We below address these problems to the extent that we are able.

**Family and Community Fixed Effects Estimates for Basic Sample**

To undertake estimates that control for family and community we must limit ourselves to the subsample of 727 individuals in 293 families in our data for which there are at least two siblings in the same family — and to smaller subsamples for exploration of the sensitivity of our results to the issues concerning imputations parallel to those in the previous section.\(^{33}\) Of course to undertake such estimates there must be some within-family variance in the relevant variables. If most or all of the variance in the relevant variables reflects differences across families, there would be little or no within-family variance with which to estimate the within-family relations. Therefore we present in Table 3 the share of the total variance that is within-family rather than between-family. Half of the variance in cognitive

### Table 3: Percentages of Total Variance that is within Family for Key Child Characteristics Variables, Ghanaian LSMS\(^{a}\)

<table>
<thead>
<tr>
<th>Child Characteristics</th>
<th>Percent of Total Variance that is Within Family</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Achievement Test</td>
<td>50</td>
</tr>
<tr>
<td>Height (Z score)</td>
<td>37</td>
</tr>
<tr>
<td>Years of Schooling</td>
<td>35</td>
</tr>
<tr>
<td>Pre-School Ability</td>
<td>59</td>
</tr>
</tbody>
</table>

\(^{a}\) See Table 1 for basic statistics for underlying data.

\(^{32}\) The increased bias towards zero in such estimates due to random measurement error has been emphasized at least since Bishop (1976) and Griliches (1977). In a recent study Ashenfelter and Krueger (1993) suggest that the increased random noise-to-signal ratio error, not the control for unobserved endowments, accounts for the substantial reduction of estimated schooling effects in within-identical-twin estimates, though Behrman, Rosenzweig and Taubman (1993) report less strong results and both of these studies are conditional on measurement errors in the reports about the twins' schooling from others (i.e., the other twin in the former, the twins' children in the latter) being uncorrelated with the measurement errors in the twins' self reports. Also, as Behrman (1984) shows, if the measurement error is not random but is systematic (e.g., ignoring the quality dimension of schooling which is associated with the grades of schooling), the within estimates are not necessarily more biased towards zero than the individual estimates.

\(^{33}\) We limit ourselves to the households with at least two biological siblings. If we also include fostered children the sample increases to 825 children, but none of the results reported in this section change significantly.
Table 4: Cognitive Achievement Production Function Estimates for Sibling Data and Family
and Community Fixed Estimates, Ghanaian LSMS

<table>
<thead>
<tr>
<th>(1) Largest Sample with Data Available</th>
<th>(2) Total Cognitive Achievement Scores, Only Individuals who Took Easy Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OLS Estimates:</strong></td>
<td><strong>Family &amp; Community Fixed Effects:</strong></td>
</tr>
<tr>
<td><strong>Height Predetermined (1)</strong></td>
<td><strong>Height Simultaneous (2)</strong></td>
</tr>
<tr>
<td><strong>OLS Estimates:</strong></td>
<td><strong>Family &amp; Community Fixed Effects:</strong></td>
</tr>
<tr>
<td><strong>Height Predetermined (3)</strong></td>
<td><strong>Height Simultaneous (4)</strong></td>
</tr>
<tr>
<td><strong>Height Simultaneous (5)</strong></td>
<td><strong>Height Predetermined (6)</strong></td>
</tr>
</tbody>
</table>

| Child Characteristics                  |                                              |
|----------------------------------------|                                              |
| Height (2 score)                       | 0.7(2.8)                                     |
|                                       | 2.0(3.3)^b                                    |
|                                       | -0.21(0.8)                                    |
|                                       | 0.7(2.0)                                     |
|                                       | 2.3(2.8)^b                                    |
|                                       | -0.10(0.2)                                    |
| Years of schooling                     | 1.7(13.7)^a                                   |
|                                       | 1.5(5.6)^a                                    |
|                                       | 1.6(11.4)                                     |
|                                       | 2.2(9.3)                                     |
|                                       | 1.5(2.1)^a                                    |
|                                       | 1.8(5.5)                                     |
| Age                                    | 0.021(1.8)                                    |
|                                       | 0.038(2.4)                                    |
|                                       | 0.007(0.7)                                    |
|                                       | -0.01(0.4)                                    |
|                                       | 0.05(1.0)                                     |
|                                       | 0.07(0.2)                                     |
| Male                                   | 1.6(2.7)                                     |
|                                       | 2.2(3.4)                                     |
|                                       | 0.19(0.8)                                     |
|                                       | 2.3(3.0)                                     |
|                                       | 3.2(3.4)                                     |
|                                       | 1.77(1.9)                                     |
| Pre-School Ability                     | 0.67(11.9)                                    |
|                                       | 0.64(9.0)                                     |
|                                       | 0.53(10.1)                                    |
|                                       | 0.75(11.1)                                    |
|                                       | 0.75(8.3)                                     |
|                                       | 0.66(7.6)                                     |

| Parental Schooling                     |                                              |
|----------------------------------------|                                              |
| Father                                 | 0.12(1.9)                                    |
|                                       | 0.15(2.0)                                    |
|                                       | 0.11(2.2)                                    |
|                                       | 0.16(2.1)                                    |
| Mother                                 | 0.34(4.0)                                    |
|                                       | 0.32(3.4)                                    |
|                                       | 0.28(3.1)                                    |
|                                       | 0.29(3.0)                                    |
| Intercept                              | -15.1(8.4)                                    |
|                                       | -15.2(6.2)                                    |
|                                       | -15.1(3.8)                                    |
|                                       | 18.8(3.9)                                    |

| R^2                                    | 0.59                                          |
|                                       | 0.54                                          |
|                                       | 0.42                                          |
|                                       | 0.55                                          |
|                                       | 0.50                                          |
| Root MSE                               | 7.6                                           |
|                                       | 7.7                                           |
|                                       | 5.1                                           |
|                                       | 8.8                                           |
|                                       | 8.9                                           |
| F                                      | 153.1                                         |
|                                       | 124.0                                         |
|                                       | 107.4                                         |
|                                       | 100.9                                         |
|                                       | 83.5                                          |
| N                                      | 727                                           |
|                                       | 727                                           |
|                                       | 727                                           |
|                                       | 587                                           |
|                                       | 587                                           |
|                                       | 587                                           |

a. See Table 1 for basic statistics for the underlying data. Absolute values of t statistics are given in parenthesis to the right of the point estimate. These are corrected in the fixed effects estimates for the degrees of freedom used for the fixed effects.

b. Treated as simultaneously determined, with first-stage estimates using (nonendogenous) variables in Table 1.

c. Treated as simultaneously determined, with first-stage estimates using (nonendogenous) variables in Table 1.

Achievement test scores is within-family, so within-family estimates are based on an important share of the variance in individual cognitive achievement that is examined in the previous section. Among the key right-side variables, over a third of the variance in both years of schooling and the Z score for height and almost three-fifths of the variance in preschool ability also is within-family.

A further question might be whether the sibling sample is a selected one. Since having two children in the relevant age range to be in the within-family sample reflects unobserved family characteristics that are controlled in the sibling estimates, the use of such a subsample should not be contaminated by selectivity bias. Nevertheless it is reassuring that the OLS

34. This point also is made by Heckman and MaCurdy (1980), Pitt and Rosenzweig (1990), and Behrman and Deolalikar (1993).
Table 4 (continued): Cognitive Achievement Production Function Estimates for Sibling Data and Family and Community Fixed Estimates, Ghanaian LSMS

<table>
<thead>
<tr>
<th>(3) Easy Tests Cognitive Achievement Scores Only Individuals who Took Easy Tests</th>
<th>(4) Only Those Individuals with Positive Cognitive Achievement Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS Estimates: Height Predetermined</td>
<td>Family and Community Fixed Effects Height Simultaneous</td>
</tr>
<tr>
<td>Height Simultaneous</td>
<td>Predetermined</td>
</tr>
<tr>
<td>(8)</td>
<td>(10)</td>
</tr>
<tr>
<td>Child Characteristics</td>
<td>Parental Schooling</td>
</tr>
<tr>
<td>Height (2 score)</td>
<td>0.24(2.1)</td>
</tr>
<tr>
<td>Years of schooling</td>
<td>0.73(9.5)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.002(0.3)</td>
</tr>
<tr>
<td>Male</td>
<td>0.61(2.4)</td>
</tr>
<tr>
<td>Pre-School Ability</td>
<td>0.29(13.2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parental Schooling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Father</td>
</tr>
<tr>
<td>Mother</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
</tbody>
</table>

| R^2 | 0.58 | 0.55 | 0.54 | 0.49 | 0.79 |
| Root MSE | 2.9 | 2.9 | 8.8 | 9.1 | 7.8 |
| F | 115.2 | 100.4 | 96.5 | 78.4 | 5.3 |
| N | 587 | 587 | 587 | 564 | 564 |

a. See Table 1 for basic statistics for the underlying data. Absolute values of t statistics are given in parenthesis to the right of the point estimate. These are corrected in the fixed effects estimates for the degrees of freedom used for the fixed effects.

b. Treated as simultaneously determined, with first-stage estimates using (nonendogenous) variables in Table 1.

c. Treated as simultaneously determined, with first-stage estimates using (nonendogenous) variables in Table 1.

estimates and the estimates with control for simultaneity for child health (using the same instruments as in Section 3) for the sibling subsample in columns 1 and 2 in Table 4 do not differ significantly from the estimates parallel to those in columns 1 and 3 of Table 2 for the subsample of children co-resident with at least one parent (and, like the latter estimates, differ from the full sample estimates in columns 2 and 3 of Table 2 only with respect to the effect of mother’s schooling being larger).

Column 3 in Table 4 gives sibling estimates for the cognitive achievement production function with child height representing child health. Consider first the coefficient estimates of the variables other than child height. Those for age and being male are much smaller for the sibling than for the individual estimates and are not significantly nonzero in the sibling estimates. Therefore, once there is control for observed and unobserved family and community characteristics, what appears in the individual estimates to be a pattern of older
children and males having higher cognitive achievement evaporates. The coefficient estimates for years of schooling and for pre-school ability, in contrast, are about as precisely estimated in these sibling estimates as in the estimates in Table 2 and are not significantly different for the sibling estimates in Table 4 from those in the individual estimates in Table 2. For these two variables, therefore, there is no evidence that random measurement error causes substantial biases towards zero in their sibling coefficient estimates.

Now, what happens to the coefficient estimate of primary concern, that for height? The results are striking, and quite in contrast to those in Section 4. The sibling coefficient estimate for child height is very imprecisely estimated, not significantly different from zero, and negative. If this is an appropriate estimate for this coefficient, it suggests that the OLS estimates and even more so the simultaneous estimates using the instruments that are used in Section 4 are substantially upward biased and misleading regarding the impact of the observed range of child health on child schooling performance. The failure to control for unobserved heterogeneity in tastes regarding child quality versus parental consumption, unobserved heterogeneity in predetermined family and community endowments that affect the production of child quality, and unobserved variations in capital market access under this interpretation, leads to substantial upward biases in the OLS estimates of the impact of the range of observed child health behavior on child cognitive achievement. And the bias is even larger for the simultaneous estimates with family and community instruments such that we used in Section 4 because of the apparent correlation of such instruments with the disturbance term in the cognitive achievement production function.

Family and Community Fixed Effects Estimates for Other Samples and Definitions of Dependent Variables

Because of the questions concerning impact of the imputations that are discussed in Section 3, it is useful again to explore how robust are these estimates to the same three alternative samples and dependent variable definitions that are explored in Section 4. Columns 4-12 in Table 4 give triplets of estimates parallel to that in columns 1-3 for the basic sample. The results are robust to these alternatives. The coefficient estimates for years of schooling and pre-school ability do not change significantly with control for family and community fixed effects, but those for child height become much smaller, more imprecisely estimated and negative.

35. They also are robust for the further alternative subsample in which in the field "na" was not used to indicate that the individuals were unable to answer any questions (see note 14 above). The coefficient estimates for health and schooling, respectively (with t tests in parentheses) are 0.23 (1.6) and 0.61 (6.4) in the OLS estimates, 0.74 (2.4) and 0.63 (2.5) in the instrumental variable estimates, and -0.087 (0.5) and 0.43 (4.0) in the family and community fixed effects estimates.
Other Aspects of These Estimates

Might random measurement error be accounting for these results? It does not seem likely that this is the case for three reasons: First, the greater noise-to-signal ratio for the sibling than for the individual estimates by itself should not reverse the sign of the estimates, but just result in estimates that are smaller in absolute magnitude. Second, as noted above, the estimates for years of schooling and pre-school ability do not suggest a critical role of such measurement error in the sibling estimates, and we have no reasons to think that the anthropometric indicators are much more contaminated by random measurement error than are these other two variables (and, in fact, it would seem to us likely that the anthropometric measurements have less random measurement error than the pre-school ability measure). Third, we have been able to explore somewhat the measurement error possibility by using earlier anthropometric measures as instruments for those that were observed concurrently with the cognitive achievement tests under the assumption that measurement errors are not correlated over time, but this procedure does not change our sibling estimates.

Might the fact that sibling estimates do not control for important unobserved individual factors underlie our results? The simultaneous estimates reported in Section 4 use family and community instruments that are independent of unobserved child-specific components in the disturbance term. If the biases due to the individual-specific components of the disturbance term are much more important than are biases due to the family and community components of the disturbance term, perhaps the simultaneous estimates are yielding results that are closer to the underlying reality than are the sibling estimates. This would require that there is an unobserved individual component in the disturbance term that is important and that is negatively associated with observed health. Such a possibility is discussed in Section 2 as the fourth reason why OLS estimates might be downward biased with the example that some children have more inherent intellectual curiosity and others have greater innate physical exuberance (so that $I_u$ and $I_{mu}$ are negatively correlated). We have doubts, however, about the relevance of such a possibility since this data set includes information about pre-school ability, which would seem to control for what may be the most important often unobserved individual specific component in the production of child cognitive achievement (and for any correlated attributes). We are able to investigate, nevertheless, somewhat further the possibility that measurement error and unobserved individual components account for our results. We do so by instrumenting the within-family deviations of years of schooling and of child health with family and community level variables. These instruments are, by construction, independent of the disturbance term in the within cognitive achievement production function relation. Therefore they result in consistent estimates of the coefficients of child schooling and child health. The instrumenting relations, hardly surprisingly, are consistent with a small proportion of the variance in the within-family deviations of the child years of schooling and the child height. This results in quite imprecise coefficient estimates in the instrumented within-family deviation estimates, and indeed does. But the point estimates still should be consistent since the use of the within-family procedure eliminates unobserved family and community components in the disturbance term and the use of these instruments eliminates any correlation with the individual specific components. The
(imprecisely estimated) coefficient estimates of years of schooling are somewhat smaller than in the simultaneous level estimates and than in the uninstrumented within-family estimates, but still positive and of the same order of magnitude. The (imprecisely estimated) coefficient estimate of child height, however, is negative as in the uninstrumented sibling estimates and in sharp contrast to the OLS and the instrumented individual estimates. Therefore these instrumented within-family estimates reinforce our conclusion above that appearance of apparent importance of the range of observed child health on cognitive achievement in the OLS estimates and the instrumented individual estimates is the result of not controlling for unobserved family and community components directly in the disturbance term of the cognitive achievement production function or working through unobserved choice inputs.

Also there is the question of how robust are these results to changes in the measurement of child health or to separating the estimates for reading versus mathematics scores. All of the results hold in the same pattern as in Table 4 if BMI is used instead of the Z score for child height. Estimates with reading and mathematics scores separated, moreover, do not appear to be very different in their implications from the estimates for the combined cognitive achievement scores (though they differ from the estimates summarized in Section 4 in that there are not significant differences between the estimates for reading and mathematics). Though years of schooling and pre-school ability have significantly positive effects on both reading and mathematics scores in the sibling estimates that do not differ significantly from those for the OLS estimates, for neither mathematics nor reading is the sibling estimate of height significantly nonzero.

Further, it has been suggested to us that perhaps the birth-order or the gender effects are not simply additive as is assumed in the estimates in Table 4 that have been discussed to this point, and the birth-order or gender effects confound the estimates of the impact of child health. Therefore we have undertaken family and community fixed effects estimates in which the siblings are ordered, respectively, by birth order and by gender. Neither ordering affects the sibling results discussed above. The coefficient estimates of years of schooling and of pre-school ability do not differ significantly in these within-family estimates from those in Table 4, but again the coefficient estimates of the Z score for child height becomes negative and insignificant.

Finally, as noted above, the within-family estimates control for both unobserved family and unobserved community effects. What happens if there is control only for unobserved community effects? If such community effects are controlled, ceteris paribus, the estimates may be either lower or higher (see the third reason for an upward bias and the fourth reason for a downward bias in Section 2). Community fixed-effects estimates suggest that much of the effects of the within-sibling estimates in Table 4 are due to unobserved community effects rather than unobserved family effects beyond the community effects. The coefficient estimates of height are less than 30 percent of those in the OLS estimates, and 10

36. Though with control for birth order, being male has a much smaller and much more imprecisely estimated point estimate.
percent or less of those if height is treated as simultaneous and insignificant at standard levels (though not negative). Thus the implication is that communities with a favorable unobserved health environment also have a favorable unobserved learning environment, the failure to control for which in standard estimates leads to erroneous attribution of the effects of that unobserved community environment on cognitive achievement to health. In contrast to the effects on the coefficient estimates of health, the coefficient estimates for years of schooling and for per-school ability are not affected significantly by control for community fixed effects and are estimated with considerable precision.

However just because community effects are important does not mean that family effects (beyond the community effects) are unimportant. To the contrary, an F test for restricting the coefficients on such family effects to be zero using the sibling data set is rejected ($F = 1.92$ with critical values of 1.19 at the 5 percent level and 1.28 at the 1 percent level). Therefore the preferred estimates are those in Table 4 that control for both unobserved community effects and unobserved family effects beyond the community effects.

37. These community effects are not captured by observed schooling quality characteristics (see discussion in Section 4 of columns 6 and 7 in Table 2).
Might Child Health Affect Cognitive Achievement through Schooling Attainment?

Some previous studies have estimated relations with schooling attainment as the dependent variable and with child health among the right-side variables (e.g., Jamison 1986, Moock and Leslie 1986). But there seems to be no natural production function with such a combination, and the reduced form in relation (5) for child schooling attainment, of course, does not include child health among the right-side variables. If one could substitute child health for one of the right-side variables in this reduced form, a conditional reduced form for child schooling attainment with child health on the right side could be obtained and estimated (identified by the exclusion of the variable that is substituted out). This may be the rationale for the specification used in such previous studies, but it is not made explicit (and, in an important sense, not followed since child health is treated as predetermined). Moreover, the coefficient of child health in such a relation would merely be the ratio of the coefficients of the right-side variable in the reduced forms for child health and child schooling attainment that is substituted out to obtain the conditional reduced form — and thus would vary depending on which right-side variable is substituted out in this process.

Elaboration on this last point is useful since similar relations, often characterized as "conditional demand functions," are not uncommon in the more general empirical literature. Consider the two following linear and simplified forms of the reduced form relation (5) for child health and child schooling:

\begin{align*}
(5A) \quad H_C &= a_{11}P_H + a_{12}P_S + e_1 \text{ and} \\
(5B) \quad S_C &= a_{21}P_H + a_{22}P_S + e_2.
\end{align*}

\(P_H\) can be eliminated from these two relations to obtain:

\begin{align*}
(5C) \quad S_C &= a_{31}H_C + a_{32}P_S + e_3.
\end{align*}

Now it might appear that this relation is an expression in which schooling depends on child health and the variables that has been eliminated to obtain this expression (i.e., \(P_H\)) can be used as an instrument to control for the endogeneity of child health (\(H_C\)). However, what is \(a_{31}\)? It is but the ratio of the coefficient of the variable that has been eliminated in relation (5B) to the coefficient of that variables in relation (5A) (i.e., \(a_{31} = a_{21}/a_{11}\)). This ratio tells us about the relative effect of the eliminated variable on schooling in comparison with that on health, not anything about the effect of health on schooling. To strengthen this point, note that if we had eliminated \(P_S\) instead to obtain a relation similar to (5C), the coefficient of \(H_C\) would have been \(a_{22}/a_{12}\), so the so-called "conditional demand" effect of \(H_C\) depends on what variable has been eliminated from the reduced forms to obtain the "conditional demand relation."

---

38. We have benefitted from discussions with Andrew Foster on this point.
Table 5: Cognitive Achievement Production Function Estimates
For Sibling Data, Ghanaian LSMS, without Schooling*  

<table>
<thead>
<tr>
<th>OLS Estimates:</th>
<th>Family and Community Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height Predetermined (1)</td>
<td>(2)</td>
</tr>
<tr>
<td><strong>Child Characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>Height (Z score)</td>
<td>1.0(3.9)</td>
</tr>
<tr>
<td>Age</td>
<td>0.10(8.0)</td>
</tr>
<tr>
<td>Male</td>
<td>1.9(2.9)</td>
</tr>
<tr>
<td>Pre-School Ability</td>
<td>0.89(14.7)</td>
</tr>
<tr>
<td><strong>Parental Schooling</strong></td>
<td></td>
</tr>
<tr>
<td>Father</td>
<td>0.37(5.4)</td>
</tr>
<tr>
<td>Mother</td>
<td>0.48(5.1)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-25.1(13.6)</td>
</tr>
<tr>
<td>R²</td>
<td>0.49</td>
</tr>
<tr>
<td>Root MSE</td>
<td>8.5</td>
</tr>
<tr>
<td>F</td>
<td>117.2</td>
</tr>
<tr>
<td>N</td>
<td>727</td>
</tr>
</tbody>
</table>

*See Table 1 for basic statistics for the underlying data. Absolute values of t statistics are given in parenthesis to the right of the point estimate. These are corrected in the fixed effects estimates for the degrees of freedom used for the fixed effects.

Therefore it is not clear how we could obtain conditional demand estimates from our data. But if child health is affecting child cognitive achievement through child schooling, then the estimated impact of child health in the cognitive achievement production function should increase if we restrict the coefficient of child schooling to be zero since, in such a case, the coefficient of child health includes both the impact of itself and of the correlated schooling effect. Table 5 presents some relevant estimates for the sibling sample. Column 1 is the OLS estimate (i.e., not controlling for family and community effects) without schooling. The coefficient estimate of child height is about 40 percent greater than in column 1 of Table 4, so child height indeed does pick up part of the effect of the omitted schooling.

---

39. With longitudinal data including past price "shocks" (unanticipated fluctuations in relative prices) it would be possible to obtain an estimate of the impact of the health stock in period t - 1 on schooling attendance in period t.

40. This would seem to give an upper bound on the direct and indirect (through schooling) effects since child health may be correlated with child schooling because both respond with the same sign to other variables (e.g., family or community resources, preferences that favor child quality over parental consumption) even if child health does not affect child schooling directly.
variable. Column 2 in Table 5 uses the sibling data to control for family and community effects, as in Section 5. In this case the estimated effect of child health is positive and about 0.5 greater than in column 3 of Table 4, so child health again seems to pick up some of the effect of the omitted school variable. But the coefficient estimate is less than a third of that in column 1, and not significantly different from zero at the 35 percent level. Thus our estimates do not suggest that child health is operating in some important manner through schooling to affect child cognitive achievement.

41. If the coefficient estimate of pre-school ability also is restricted to zero, the coefficient estimate of child height increases to 0.6, but still is not significantly nonzero at the five percent level.
Conclusions

The previous literature based on socioeconomic survey data that has investigated the impact of child health and nutrition on child schooling success has concluded that such effects are positive and important over the range of child health observed in children who attend school, in addition to effects on who attends school. However, this literature does not control for the probable endogeneity of the determination of child health. The failure to control for such endogeneity may result in biases that are upwards or downwards in the estimated impact of child health on child schooling.

Our estimates with control for such endogeneity by using family and community characteristics as instruments, conditional on the assumption that such instruments are appropriate, suggest that the biases in the previous literature have been big and downward. That is, that the true effects are much bigger — about three times as large — as suggested by using the OLS procedures that were used in the previous literature. This result seems robust to control for selectivity regarding who proceeds at least through the first grade and who resides with at least one biological parent, to use of an expanded instrument set, to the inclusion of school quality indicators, and to use of an alternative shorter-run indicator of child health. On a disaggregated level it is stronger for reading than for mathematics and for males than for females. Some may be tempted to conclude on the basis of these results that the impact of variations of child health (within the range observed for children attending school) on child schooling success is more important than previously realized because of important downward biases in OLS estimates.

However this result is not robust to control for unobserved family and community characteristics using within-family estimates. In fact, such estimates yield negative (though insignificant) estimates for child health. This effect of control for these fixed effects on the child health coefficient contrasts strongly with the lack of an effect of such within-estimates on the coefficient estimates of child schooling and pre-school ability. The coefficient estimates of the latter two variables are not changed substantially by the control for unobserved family and community variables. This general pattern holds if the mathematics and reading scores are considered separately, if siblings are ordered by birth order or gender, and, with less precision, if the within-estimates are instrumented with family and community level variables to eliminate correlations with unobserved individual characteristics. Within-community estimates suggest that much of the impact on the coefficient estimates of within-estimates is due to unobserved community factors rather than to unobserved family factors beyond the community effects. That is, there are important unobserved community characteristics that affect both child anthropometric measures and child cognitive achievement which, if not controlled, tend to lead to the incorrect inference that child health as indicated by anthropometric measures affect child cognitive achievement. But unobserved family effects also have significant effects beyond the unobserved community effects. Finally, child health may be working partly through child schooling at the time of the cognitive achievement test scores, but even allowing for such a possibility the impact of child health on cognitive achievement is not significantly nonzero if there is control for unobserved family and community effects.
Thus, the instrumented 2SLS estimates apparently overstate substantially the impact of child health on child schooling success and cause greater distortions than the OLS estimates.\textsuperscript{42} The problem is that the instruments that are used apparently are not independent of the disturbance term in the cognitive achievement production function, but operate importantly in part through unobserved choice variables. This problem well may be important also in other contexts in which efforts have been made to estimate production functions or other relations with family and community variables as instruments under the assumption that there are no unobserved choice inputs, such as Pitt, Rosenzweig and Hassan (1990), Rosenzweig and Schultz (1983, 1987, 1988), and Schultz and Tansel (1993). In the present context, our best estimates are that the true impact of child health over the relevant range on child schooling success is not significant, despite the appearance of great importance with the use of typical family and community instruments. This has important implications for understanding the child health-child schooling associations, and for undertaking similar research on a wide range of other topics that depend on assumptions that observed family and community characteristics are independent of the disturbance terms in the relations being estimated.

\textsuperscript{42} We would not claim that our results raise questions about all previous studies of the impact of child health on child cognitive achievement. Some experimental studies provide some more persuasive positive evidence. Recent double-blind placebo trials for 159 Jamaican school children, for example, suggest that effective treatment for \textit{Trichuris trichiura} (whipworm) within nine weeks improved cognitive achievement significantly so that the previously moderately to heavily infected children no longer performed significantly different than an unaffected control group (Nokes, et al. 1992a,b). But we do think that our results raise questions about nonexperimental studies that claim to find a significantly positive effect of child health within the ranges observed in school children on child cognitive achievement but which have not explored the estimation problems due to behavioral determinants of child health that we consider in this paper.
References


Florencio, Cecilia A., 1988, "Nutrition, Health and Other Determinants of Academic Achievement and School-Related Behavior of Grades One to Six Pupils," Quezon City, Philippines: University of the Philippines, mimeo.


Olneck, Michael, 1977, "On the Use of Sibling Data to Estimate the Effects of Family
Background, Cognitive Skills, and Schooling: Results from the Kalamazoo Brothers
Study," in Paul Taubman, ed., Kinometrics: Determinants of Socioeconomic Success
within and between Families, Amsterdam: North-Holland, 125-163.

Paxson, Christina H., 1993, "Consumption and Income Seasonality in Thailand," Journal of
Political Economy 101:1 (February), 39-72.

Pitt, Mark M. and Mark R. Rosenzweig, 1990, "Estimating the Behavioral Consequences of
Health in a Family Context: The Intrafamily Incidence of Infant Illness in

Pitt, Mark M., Mark R. Rosenzweig, and Donna M. Gibbons, 1993, "The Determinants and
Consequences of the Placement of Government Programs in Indonesia," The World
Bank Economic Review 7:3 (September), 319-348.

Pitt, Mark M., Mark R. Rosenzweig, and M.N. Hassan, 1990, "Productivity, Health and
Inequality in the Intrahousehold Distribution of Food in Low-Income Countries,"
American Economic Review 80:5 (December), 1139-1156.


Pollitt, Ernesto, Phongian Hathirat, Nittaya J. Kotchabhakdi, Lavon Missell, and Aree
Valyasevi, 1989, "Iron Deficiency and Educational Achievement in Thailand," The

Psacharopoulos, G., 1985, "Returns to Education: A Further International Update and

Observer 3:1 (January), 99-1.

Raven, J.C., 1956, Guide to the Coloured Progressive Matrices (Sets A, Ab, B). London:
Lewis.

Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth

Rosenzweig, Mark R. and T. Paul Schultz, 1987, "Fertility and Investments in Human
Capital: Estimates of the Consequences of Imperfect Fertility Control in Malaysia,"
Journal of Econometrics 36, 3-184.


Argentina
Carlos Hirsch, SRL
Galería de Cunas
Florida 145, 4th Floor-Dolce 453/446
1333 Buenos Aires

Australia, Papua New Guinea, Fiji, Solomon Islands, Vanuatu, and Western Samoa
D.A. Information Services
648 Whitehorse Road
Mitcham 3132
Victoria

Austria
Gerald and Co.
Gartenstrasse 31
A-1011 Wien

Bangladesh
Micro Industries Development Assistance Society (MIDAS)
House 5, Road 16
Dhanmondi 1/R/Area
Dhaka 1209

Bench offices:
Pine View, 1st Floor
100 Agradhun Commercial Area
Chittagong 4100

Belgium
Jean de Lannoy
Av. du Roi 202
1060 Brussels

Canada
Le Diffuseur
151 Avoine, de Montage
Boucherville, Québec J4B 5S6

Renouf Publishing Co.
1294 Algoma Road
Ottawa, Ontario K1B 3W8

Chile
Invertex S.A.
Av. Santa Maria 6400
Edificio INTET, Of. 201
Santiago

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

Colombia
Informaciones 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
Center of Applied Research
Cyprus College
6, Dogenous Street, Engomi
P.O. Box 2006
Nicosia

Denmark
Sambandsforlaget
Rosensverk Allé 11
DK-1970 Frederiksberg C

Dominican Republic
Editora Taller, C. por A.
Restauración e Isabel la Católica 309
Apartado de Correos 2390 2-1
Santo Domingo

Egypt, Arab Republic of
Ali Ahran
Al Galaa Street
Cairo

The Middle East Observer
41, Sheif Street
Cairo

Finland
Akatemiens Kirjakauppa
P.O. Box 128
SF-00101 Helsinki 10

France
World Bank Publications
66, avenue d’Aix
75116 Paris

Germany
UNO-Verlag
Pappelsdorfer Allee 55
D-5000 Bonn 1

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

Hungary
Foundation for Market Economy
Dombovart Ut 17-19
H-1117 Budapest

India
Allied Publishers Private Ltd.
721 Mount Road
Madras - 600 002

Bench offices:
15 J.N. Heretella Marg
Ballard Estate
Bombay - 400 038

13/14 Asaf Ali Road
New Delhi - 110 002

17 Chittaranjan Avenue
Calcutta - 700 072

Jayadeva Hostal Building
5th Main Road, Ghandinagar
Bangalore - 560 009

3-5-129 Kachiguda
Cross Road
Hyderabad - 500 027

Prathama Flats, 2nd Floor
Near Thakore Bungo, Navrangpura
Ahmedabad - 380 009

Patel House
16-A Ambik Marg
Lucknow - 226 001

Central Bazaar Road
60, Bajaj Nagar
Nagpur - 440 001

Indonesia
Pt. Indira Limited
Jalan Sambudra 10
Jakarta 10320

Iran
Kowkab Publishers
P.O. Box 19971-511
Tehran

Ireland
Government Supplies Agency
4-5 Harcourt Road
Dublin 2

Israel
Yoznot Literature Ltd.
P.O. Box 58655
Tel Aviv 61560

Italy
Unione Commissionaria Sanremesi SPA
Via Duca Di Calabria 1/1
Casale Postale 552
50125 Firenze

Japan
Eastern Book Service
Hong Kong 3-Chome, Bunkyo-ku 113
Tokyo

Kenya
Africa Book Service (E.A.) Ltd.
P.O. Box 342-45
Nairobi

Korea, Republic of
Pan Korea Book Corporation
P.O. Box 101, Kwangwhamun Seoul

Korean Stock Book Centre
P.O. Box 34
Yeouido

Malaysia
University of Malaysia Cooperative Bookshop
P.O. Box 1127, Jalan Pantai Baru
50770 Kuta Lumpur

Mexico
INFOTEC
Apartado Postal 22-860
16400 Tlatyop, Mexico D.F.

Netherlands
De Linboonbouw/InOn-Publikaties
P.O. Box 202
7460 AE Hasseltoren

New Zealand
EBCO NZ Ltd.
Private Mail Bag 99914
New Market Auckand

Nigeria
University Press Limited
Three Crown Building Jericho
Private Mail Bag 5903
Ibadan

Norway
Narway Information Center
Book Department
P.O. Box 6125 Etterstad
N-0662 Oslo 6

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

Peru
Editorial De Barril SA
Apartado 3814
Lima 1

Philippines
International Book Center
Suite 1703, Cityland 10
Condominium Tower 1
Ayala Avenue, H.V. dela Costa Extension
Makati, Metro Manila

Poland
International Publishing Service
Ul. Placna 51/37
00-477 Warsaw

For subscription orders:
IPS Journals
Ul. Okresna 3
02-916 Warsaw

Portugal
Livraria Portugal
Rua Do Carmo 70-74
1200 Lisbon

Saudi Arabia, Qatar
Jami Book Store
P.O. Box 3194
Riyadh 11471

Singapore, Taiwan, Myanmar, Brunei
Gower Asia Pacific Pte Ltd.
Golden Wheel Building
41, Kalang Pudding, 02-03
Singapore 1334

South Africa, Botswana
For single titles
Oxford University Press
Southern Africa
P.O. Box 1541
Cape Town 8000

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

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

Libreria Internacional AEDOS
Consell de Cent, 391
08009 Barcelona

Sri Lanka and the Maldives
Lake House Bookshop
P.O. Box 244
100, Sir Chitampanal A.
Gandineri Mawatha
Colombo 2

Sweden
For single titles:
Fritzes Facebookforetaget
Regensplan 12, Box 16356
S-103 27 Stockholm

For subscription orders:
Wengeren-Williams AB
P. O. Box 1305
S-171 25 Solna

Switzerland
For single titles:
Librairie Payot
Case postale 3121
CH 1002 Lausanne

For subscription orders:
Librairie Payot
Service des Abonnements
Case postale 3121
CH 1002 Lausanne

Thailand
Central Department Store
306 Siron Road
Bangkok

Trinidad & Tobago, Antigua, Barbuda, Barbados, Dominica, Grenada, Guyana, Jamaica, Montserrat, St. Kitts & Nevis, St. Lucia, St. Vincent & Grenadines
Systematics Studies Unit
#4 Wata Street
Curepe

Trinidad, West Indies

United Kingdom
Microinfo Ltd.
P.O. Box 3
Alton, Hampshire GU34 2PG
England

Microinfo Ltd.
P.O. Box 3
Alton, Hampshire GU34 2PG
England
LSMS Working Papers (continued)

No. 67  King, Does Education Pay in the Labor Market? The Labor Force Participation, Occupation, and Earnings of Peruvian Women
No. 68  Kozel, The Composition and Distribution of Income in Côte d’Ivoire
No. 69  Deaton, Price Elasticities from Survey Data: Extensions and Indonesian Results
No. 70  Glewwe, Efficient Allocation of Transfers to the Poor: The Problem of Unobserved Household Income
No. 71  Glewwe, Investigating the Determinants of Household Welfare in Côte d’Ivoire
No. 72  Pitt and Rosenzweig, The Selectivity of Fertility and the Determinants of Human Capital Investments: Parametric and Semiparametric Estimates
No. 73  Jacoby, Shadow Wages and Peasant Family Labor Supply: An Econometric Application to the Peruvian Sierra
No. 74  Behrman, The Action of Human Resources and Poverty on One Another: What We Have Yet to Learn
No. 75  Glewwe and Twum-Baah, The Distribution of Welfare in Ghana, 1987–88
No. 76  Glewwe, Schooling, Skills, and the Returns to Government Investment in Education: An Exploration Using Data from Ghana
No. 77  Newman, Jorgensen, and Pradhan, Workers’ Benefits from Bolivia’s Emergency Social Fund
No. 78  Vijverberg, Dual Selection Criteria with Multiple Alternatives: Migration, Work Status, and Wages
No. 79  Thomas, Gender Differences in Household Resource Allocations
No. 80  Grosh, The Household Survey as a Tool for Policy Change: Lessons from the Jamaican Survey of Living Conditions
No. 81  Deaton and Paxson, Patterns of Aging in Thailand and Côte d’Ivoire
No. 82  Ravaillon, Does Undernutrition Respond to Incomes and Prices? Dominance Tests for Indonesia
No. 83  Ravaillon and Datt, Growth and Redistribution Components of Changes in Poverty Measure: A Decomposition with Applications to Brazil and India in the 1980s
No. 84  Vijverberg, Measuring Income from Family Enterprises with Household Surveys
No. 85  Deaton and Grimard, Demand Analysis and Tax Reform in Pakistan
No. 86  Glewwe and Hall, Poverty and Inequality during Unorthodox Adjustment: The Case of Peru, 1985–90
No. 88  Ravaillon, Poverty Comparisons: A Guide to Concepts and Methods
No. 89  Thomas, Lavy, and Strauss, Public Policy and Anthropometric Outcomes in Côte d’Ivoire
No. 90  Ainsworth and others, Measuring the Impact of Fatal Adult Illness in Sub-Saharan Africa: An Annotated Household Questionnaire
No. 91  Glewwe and Jacoby, Estimating the Determinants of Cognitive Achievement in Low-Income Countries: The Case of Ghana
No. 92  Ainsworth, Economic Aspects of Child Fostering in Côte d’Ivoire
No. 93  Lavy, Investment in Human Capital: Schooling Supply Constraints in Rural Ghana
No. 94  Lavy and Quigley, Willingness to Pay for the Quality and Intensity of Medical Care: Low-Income Households in Ghana
No. 95  Schultz and Tansel, Measurement of Returns to Adult Health: Morbidity Effects on Wage Rates in Côte d’Ivoire and Ghana
No. 96  Louat, Grosh, and van der Gaag, Welfare Implications of Female Headship in Jamaican Households
No. 97  Coulombe and Demery, Household Size in Côte d’Ivoire: Sampling Bias in the CILSS
No. 98  Glewwe and Jacoby, Delayed Primary School Enrollment and Childhood Malnutrition in Ghana: An Economic Analysis
No. 99  Baker and Grosh, Poverty Reduction through Geographic Targeting: How Well Does It Work?
No. 100  Datt and Ravaillon, Income Gains for the Poor from Public Works Employment: Evidence from Two Indian Villages
No. 101  Kostermans, Assessing the Quality of Anthropometric Data: Background and Illustrated Guidelines for Survey Managers
No. 103  Beneo and Schultz, Determinants of Fertility and Child Mortality in Côte d’Ivoire and Ghana