Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19103-6297

pier@ssc.upenn.edu
http://www.econ.upenn.edu/pier

PIER Working Paper 97-023

"Child Health and Schooling Achievement: Association, Causality and Household Allocations"

by

Jere. R. Behrman and Victor Lavy

CHILD HEALTH AND SCHOOLING ACHIEVEMENT:

ASSOCIATION, CAUSALITY AND HOUSEHOLD ALLOCATIONS

by

Jere R. Behrman and Victor Lavy*

April 28, 1997

Direct correspodence to: Jere R. Behrman, Professor of Economics, McNeil 160, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297, USA; telephone 215 898 7704, fax 215 573 2057, e-mail: jbehrman@econ.sas.upenn.edu

*Behrman is the William R. Kenan Jr. Professor of Economics at the University of Pennsylvania, Philadelphia PA, 19104-6297, USA and Lavy is Professor of Economics, Hebrew University, Jerusalem, Israel. Both initiated work on this paper as consultants to the World Bank. Behrman's subsequent input was funded by NIH 5-RO1-HD30907, NIH 1 R01-AG-11725-01, and NSF SBR95-11955. We thank for useful comments Harold Alderman, Simon Appleton, Nancy Birdsall, Anil Deolalikar, Paul Glewwe, Estelle James, Dean Jolliffee, 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). The authors alone, and not the World Bank, the NIH nor the NSF, are responsible for the contents of this paper.

Abstract

Child health is widely perceived to affect strongly schooling. But evidence is quite limited because numerous studies based on socioeconomic surveys fail to consider the endogeneity of child health, measurement error, and the impact of unobserved fixed and choice inputs. This paper shows that a priori the resulting biases may be positive or negative depending on which of a number of household allocation behaviors dominate and what is the nature of any unobserved choice inputs in educational production. Then illustrative empirical analysis, using rich data from Ghana, is presented, with the following results: (1) IV estimates based on observed family and community characteristics similar to those used in other studies suggest a downward bias in OLS. (2) Family and community fixed effects estimates 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 not significant despite the strong association that leads to the appearance of an effect in standard OLS or IV estimates using family and community variables. (3) The usual assumption that there are no unobserved choice inputs in educational production probably leads to an upward bias in the estimated impact of child health on schooling even if there is good control for the endogeneity of child health and measurement error. (4) Child health also does not significantly affect child cognitive achievement through schooling attainment; consideration of the relations that usually have been used to investigate such a possibility, moreover, suggests that the coefficients that are usually estimated are <u>not</u> coefficients that represent the impact of child health on child schooling. (5) The preferred estimates control for unobserved family and community fixed effects and are robust to other estimation problems, so the standard estimates overstate the impact of child health in the observed range on child schooling success, which should strike a cautionary note about the interpretation of the many production function estimates in the literature, including, but not limited to those that focus on household production.

Section 1. Introduction

Common sense and casual observation suggest that very 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 that include the levels observed among children actually attending schools (i.e., surveys in Pollitt 1990, Miller and Korenman 1993, and World Bank 1993). In part because of this perception there is strong advocacy for improving child health because improvements are claimed to have substantial fairly immediate effects on child education and, through this channel, important longer-run effects on labor productivity.

But 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.[1] 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 and possible measurement problems not only regarding child health but also inputs into educational success that are not observed in the data.

A priori it would seem that child health and child schooling are determined simultaneously by households given relevant observed and unobserved (by analysts) characteristics of households and of the communities in which they reside. 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 clear *a priori*. On one hand, this bias may arise primarily from unobserved (to researchers) household or community characteristics that contribute both to better child education and to better child health -- for example, how warm and nurturing is the general environment in which children are being raised. If so the standard estimates give upward-biased estimates of the impact of child health on child education. On the other hand the bias may arise largely from unobserved parental and community endowments for child health that are inversely related to those for educational investments -- for example, the tradeoff between rural communities with healthy fresh air and open spaces but not much intellectual stimulation versus large polluted cities with more stimulating intellectual environments. In this case, the standard estimates may give downward-biased estimates of the impact of child health on child education.

The basic estimation problem is to isolate the effect of child health on child education through assuring that the representation of child health used is independent of the disturbance term in the relation being estimated. This disturbance term, in turn, may include unobserved family, and community factors that may operate directly

---

[1]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 Gopladas 1989, Nokes et al. 1992a,b). However these experimental 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.

on child cognitive achievement or whatever indicator of child schooling success is used. Further, even if the estimation procedure used assures that the indicator of health is not correlated with the disturbance term, if the estimation is of a production function such as for cognitive achievement and there is an unobserved behavioral input into that production process (e.g., parental time devoted to homework), the estimates are biased in an indeterminate direction.

For these reasons the correlations between anthropometric indicators of child health and child schooling achievement in studies such as Freeman, Klein, Townsend and Lechtig (1980), Wilson (1981), Wolfe (1985), Chutikul (1986), Moock and Leslie (1986), Jamison (1986), Florencio (1988), Sigman, Neumann, Jansen and Bwibo (1989), Gomes-Neto, Hanushek, Leite, and Frota-Bezzera (1992), and Miller and Korenman (1993) are <u>not</u> compelling evidence of the extent of the impact of child health on schooling success despite the fact that this interpretation is quite widespread.[2] <u>Associations</u> between indicators of child health and indicators of child schooling achievement do not demonstrate that child health <u>causes</u> child schooling achievement to the same degree. The true effects may be much smaller or much larger.

In this paper we investigate the impact of child health on child schooling success, given the endogenous determination of child health and possible measurement problems. 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 remaining sections explore the empirical relevance of the endogenous determination of child health on child schooling success and of measurement problems for this data set, which leads to five major conclusions. (1) Instrumental variable (IV) 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 OLS estimates is <u>downward</u> -- i.e., that the unobserved health endowments are inversely associated with the unobserved education endowments and/or random measurement error is important. (2) Family-and-community- fixed effects (FC-FE) estimates suggest that the direction of the bias in standard estimates is <u>upward</u> and that the true effects of the range of observed child health on school success is not significant despite the strong association that leads to the appearance of an effect in standard OLS or IV level estimates using common family and community variables. That is, there are unobserved family and community

---

[2]Glewwe and Jacoby (1995) is an exception among published studies that is sensitive to possible problems of simultaneity in their investigation of the impact of anthropometric indicators on delayed schooling enrollment and schooling duration in Ghana. However, they do not adequately justify their specification from the perspective of the discussion in Section 7 below of relation (13C). Two recent preliminary studies -- Glewwe, Jacoby and King (1996) and Alderman, Behrman, Lavy and Menon (1997) -- also attempt to deal with this general estimation problem using longitudinal data. But the former relies on the maintained assumption that child anthropometric measures through age two are exogenous instruments for subsequent health and schooling decisions and the latter limits its analysis to delayed schooling enrollment and not more direct indicators of schooling success.

effects -- such as the general nature of the child development environment -- that cause upward biases in the standard OLS and IV level estimates. These FC-FE estimates, moreover, seem robust with respect to a number of sample, specification, and estimation alternatives. (3) The usual assumption that there are no unobserved choice inputs in educational production also probably (though our estimates are imprecise) leads to an upward bias in the estimated impact of child health on schooling health even if there is good control for the endogeneity of child health and measurement error. (4) 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. (5) The preferred estimates control for unobserved family and community fixed effects and unobserved choice inputs into educational production and are robust to other estimation problems, so that the standard estimates overstate the impact of child health in the observed range on child schooling success. More generally, our explorations illustrate that production function estimates may be very sensitive to assumptions regarding the importance of unobserved allocated inputs and unobserved fixed factors, which should strike a cautionary note about the interpretation of the many production function estimates in the literature, including, but not limited to those that focus on household production (e.g., Pitt, Rosenzweig and Hassan 1990, Rosenzweig and Schultz 1983, 1987, 1988).

Section 2. 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 cognitive achievement test scores (that have been shown for the data set that we use to have a significant positive relation with adult wages).[3] But the OLS estimate of the impact of child health in a production function for cognitive achievement may be biased due to a number of factors that relate to the underlying household behaviors and what is observed and not observed in the data. The direction of the bias, moreover, may be upward or downward. Therefore it is difficult to know how good are previous estimates of the impact of child health on schooling success -- or even if they have established upper or lower bounds for the true effects.

To clarify these points we model within a simple one-period log-linear system the joint determination by the household of child health and child schooling success separable from other household behavioral decisions,

---

[3]See Glewwe (1994, 1996). This result also is reported for the U.S. (e.g., Bishop 1989, 1991, Blackburn and Neumark 1993 and references therein) and in the other four data sets from developing countries of which we are aware for which such relations have been explored (for urban Kenya and Tanzania in Boissiere, Knight and Sabot 1985; for rural Pakistan in Alderman, Behrman, Ross and Sabot 1996b; for Morocco in Lavy, Spratt and Leboucher 1997; and for South Africa in Case and Deaton 1996 and Moll 1996).

indicating what variables are observable and unobservable with reference to the data that we use for our estimates below. With these explicit functional forms, behavioral assumptions, and assumptions about what is observable, we able to clarify the nature of the bias problem and to derive empirically tractable relations that we estimate below.

Assume that the fth family in the cth community has a CES welfare function that is defined over the child quality (Q) of the n children in the family ($i = 1, n$) and that enters separably from other arguments into the overall family objective function:

(1) $W_{fc} = \sum w_{ifc}(Q_{ifc})^{z}$, where sum is over $i = 1, n$.

Assume that child quality is a log-linear weighted average of child education (E) and child health (H):[4]

(2) $Q_{ifc} = (E_{ifc})^{q}(H_{ifc})^{1-q}$.

In the welfare function in relation (1), the weights on Q for each child in the family are given by w, and the preference tradeoffs among Q for different children in the family are given by z. This formulation of the welfare function allows for the full range of inequality aversion regarding the distribution of Q among children in the family: the welfare function becomes Rawlsian (infinite inequality aversion) with fixed coefficients as $z \to -\infty$ at one extreme and linear (zero inequality aversion as $z = 1$ at the other. With the log-linear production technology below, moreover, $z = 0$ is critical for intrahousehold allocations because welfare-maximizing households compensate for innate differences (i.e., by investing more in children who are less able) if z is less than zero and reinforce innate differences (i.e., by investing more in children who are more able) if z is greater than zero. Therefor $z = 0$ represents "neutrality," with neither compensation for or reinforcement of innate differences among children in the household. If the weights w are the same for all children within a family, the welfare function exhibits "equal concern" and is symmetrical around the 45° ray from the origin. (For more discussion of this welfare function, see Behrman, Pollak and Taubman 1982.)

The welfare function is maximized subject to production function constraints for E and H and a resource constraint for total resources devoted to human resource investments in the children (R):

(3) $E_{ifc} = a(S^{c}_{ifc})^{a1}(H_{ifc})^{a2}(X^{e}_{ifc})^{a3}(A^{e}_{ifc})^{a4}(S^{p}_{fc})^{a5}(I^{e}_{ifc})^{a6}(F^{e}_{fc})^{a7}(C^{e}_{c})^{a8}\exp(u_{ifc})$,

(4) $H_{ifc} = b(X^{h}_{ifc})^{b1}(S^{p}_{fc})^{b2}(I^{h}_{ifc})^{b3}(F^{h}_{fc})^{b4}(C^{h}_{c})^{b5}\exp(v_{ifc})$, and

(5) $R_{fc} \geq \sum(P^{s}S_{ifc} + P^{xe}X^{e}_{ifc} + P^{xh}X^{h}_{ifc})$, where the sum is over $i = 1,n$.

Both production functions are posited to be log-linear. The inputs into production of education for the ith child in the fth family in the cth community are posited to include: two observed current choice variables (time in schooling

[4]Both education and health may be of interest because of their impact on productivity and/or in themselves. The parental interest in their productivity effects may be motivated by altruism, expected returns from intergenerational investments and transfers in the presence of market imperfections, and/or interest in the family dynasty.

S^c and health H for that child);[5] one unobserved choice variable X^e (e.g., parental time devoted to education); one observed fixed child variable A^e (ability); observed fixed parental schooling S^p; unobserved fixed individual child, family and community variables that affect child education (I^e, F^e, C^e); and a stochastic term (u). The estimate of interest for this study is that for a2, the elasticity of child education with respect to child health. The inputs into production of health for the ith child in the fth family in the cth community are posited to include: one unobserved choice variable X^e (e.g., nutrients); observed fixed parental schooling S^p; unobserved fixed individual child, family and community variables that affect child health (I^h, F^h, C^h); and a stochastic term (v). The resource constraint states that the fth household's allocation to children R is greater than or equal to expenditures on observed and unobserved current choice inputs in the production of education and health. (Throughout the superscripts e and h refer respectively to education and health.)

Constrained maximization under assumptions necessary for an interior solution, implies:

(6) $S^c_{ifc} = (w_{ifc}*z*q*a1)(E_{ifc})^{qz}(H_{ifc})^{z-zq}/(\lambda_{fc}P^s)$,

(7) $X^e_{ifc} = (w_{ifc}*z*q*a3)(E_{ifc})^{qz}(H_{ifc})^{z-zq}/(\lambda_{fc}P^{xe})$, and

(8) $X^h_{ifc} = (w_{ifc}*z*b1)(a2*q + 1-q)(E_{ifc})^{qz}(H_{ifc})^{z-zq}/(\lambda_{fc}P^{xh})$,

where λ_{fc} is the Lagrangian multiplier for the resource constraint for human resource investments in children in the fth household in the cth community.

Bias occurs in the OLS coefficient estimate of the impact of H in relation (3) if H is correlated with the disturbance term (including unobserved components) in that relation, if H is measured with error or if there are unobserved choice inputs into educational production. To learn more about possible bias, consider two cases that are distinguished by whether or not there is an unobserved choice input in the education production function.

(1) <u>No unobserved choice educational input</u>: If there is no unobserved choice input in the education production function, then relation (3) (with a3=0) is estimated and the question of bias reduces to whether H in that relation is correlated with the unobserved variables in that relation beyond the stochastic term -- i.e., I^e_{ifc}, F^e_{fc}, and C^e_c and any random measurement error in H.[6] To explore this question, use relations (3) (with a3=0), (4), and (8) to obtain an expression for H without X^h or E:

(9) $H_{ifc} = [b((w_{ifc}*z*b1)(a2*q + 1-q)([a(S^c_{ifc})^{a1}(A^e_{ifc})^{a4}(S^p_{fc})^{a5}(I^e_{ifc})^{a6}(F^e_{fc})^{a7}(C^e_c)^{a8}\exp(u_{ifc})]^{qz}/(\lambda_{fc}P^{xh}))^{b1}(S^p_{fc})^{b2}(I^h_{ifc})^{b3}(F^h_{fc})^{b4}(C^h_c)^{b5}\exp(v_{ifc})]^{(1-[a2+z-zq)b1}$.

Whether the OLS estimate of the elasticity of H in relation (3) is biased upwards or downwards depends on the sign of the correlation between the right side of relation (9) and I^e_{ifc}, F^e_{fc}, and C^e_c and the measurement error in H

---

[5]We distinguish between education (E) as measured by cognitive achievement and time spent in schooling (S) which is an input into the production of education, not education itself.

[6]If H is has random measurement error, the error component also is in the disturbance term and causes classical measurement error bias even if there are no unobserved fixed or choice variables. Because the effect of such measurement error is parallel to that of the unobserved fixed variables, in order to avoid further complicating the notation we do not explicitly include the measurement error in what follows.

(which we put aside until the end of this paragraph). Generally there is a bias because these three unobservables appear explicitly in relation (9). Under the assumption that they are not correlated with other right-side variables in relation (9), the sign of the correlation of each of these unobservables with the right side of relation (9) is positive, causing an upward bias in estimates of a2 that do not control for this correlation.[7] But the assumption that they are not correlated with other right-side variables almost surely is too strong. Among these other right-side variables, for example, are the unobserved endowments that directly affect health: I^h_{ifc}, F^h_{fc}, and C^h_c. If these unobserved health endowments are positively correlated with the unobserved education endowments, then this correlation reinforces the positive bias due to the direct representation of the unobserved educational endowments in relation (9). But the unobserved health endowments need not be positively correlated with the unobserved education endowments. Individuals with strong unobserved health endowments may tend to have lesser unobserved (net of measured ability) educational endowments. Households with unobserved role models, knowledge and environments that promote child health more than average may tend to have less than average unobserved characteristics that support child education. Communities with unobserved endowments that support child health more than average (e.g., fresh air, open spaces, relatively cheap food, labor markets that reward physical strength) may tend to be more limited than average in their support of education (e.g., with less communication with rest of world, less general learning environment, labor markets that do not reward education much). If the unobserved health endowments are sufficiently negatively correlated with the unobserved education endowments, the overall correlations between the unobserved education endowments in relation (3) and H as determined in relation (9) can be negative. In such a case, estimates of the impact of child health and education in relation (3) that do not control for the behavioral choices that determine child health are downwardly biased. Random measurement error in H also tends to yield a downward bias. Thus, the failure to control for the determination of child health in the estimation of the impact of child health on school success and for random measurement error in H leads to a biased estimate with the bias of indeterminate sign.

(2) <u>Unobserved choice educational input</u>: First use relations (3) and (7) to eliminate X^e and to solve for E:

(10) $E_{ifc} = [a(S^c_{ifc})^{a1}(H_{ifc})^{a2+a3(z-zq))}((w_{ifc}*z*q*a3)/(\lambda_{fc}P^{xe}))^{a3}(A^e_{ifc})^{a4}(S^p_{fc})^{a5}(I^e_{ifc})^{a6}(F^e_{fc})^{a7}(C^e_c)^{a8}\exp(u_{ifc})]^{1-a3zq}$.

If there is an unobserved choice input in the education production function, relation (10) is the expression that effectively is estimated rather than relation (3). Relation (10) makes explicit the critical role of an unobserved choice input X^e: the exponent of H generally is not a2 but (a2+a3(z-zq))(1-a3zq). This exponent for H reduces to a2 for positive values of a3 only if a3 = (z(1-q)-zqa2)/(1-q), a condition among parameters of the education production function and the household welfare function that is not likely. Generally the exponent of H is not a2

[7]Because they are multiplied by a function of z in relation (9) and z can be negative, it might appear *prima facie* that the correlations can be negative. But z also appears in the exponent of this expression, which offsets the multiplicative effect on the sign of the correlations.

and, in fact, can be greater than or less than a2. Whether this exponent is greater than or less than a2 depends in a complicated way on the same parameters for the education production function and the household welfare function that are noted above (i.e., whether z(1-q) -a2zq-a3z(1-q) > or < 0). If z >0 (so there is reinforcement of innate differences by household actions) this expression is larger the less are q (the share of education in quality), a2 (the elasticity of education with respect to health) and a3 (the elasticity of education with respect to the unobserved choice input into education production) -- and vice versa if z < 0 (so there is compensation of innate differences by household actions). This exponent is larger the greater is z if (1-q)(1-a3)>a2q (i.e., the less important is education in child quality and the lesser is the elasticity of education with regard to the unobserved choice input relative to the elasticity of education with regard to health) -- and vice versa if (1-q)a3<a2q. Note that if there is an unobserved choice educational input, even if H is not correlated with the unobserved variables on the right-side of relation (10) and is not measured with error, the estimates are biased with an indeterminant sign.

Relation (10) indicates that even if H can be purged of whatever is correlated with the disturbance term, what is being estimated is the desired coefficient a2 in combination with a3, q and z. From the direct estimation of relation (10), moreover, there is no way to identify the production function parameter a2 from the other parameters.

It is possible, however, to identify the welfare function parameters q and z from estimation of intrahousehold allocations using data on the ith and jth children in the same family. Consider the ratio of relation (6) for two such children, under the assumption that the true schooling S^c is measured with error so that observed schooling $S^{c*}=S^c\exp(y)$ where y is the measurement error in schooling.[8]

(11) $S^{c*}_{ifc}/S^{c*}_{jfc} = (w_{ifc}/w_{jfc})(E_{ifc}/E_{jfc})^{qz}(H_{ifc}/H_{jfc})^{z-zq}\exp(y_{jfc}/y_{ifc})$.

Unbiased estimates of this relation enable identification of q and z (though <u>not</u> of any of the production function parameters). To obtain such estimates it must be recognized that, within the model specified, E and H are endogenous, so this basically is a relation among endogenous variables. Conditional on the model, identification of q and z requires variables that differ between children i and j in the same household but do not enter directly into relation (11). With such variables (e.g., the sib difference in A), relation (11) can be estimated and the estimates obtained used together with those of relation (10) to help to identify the production function parameter of interest, a2.

Knowledge of q and z, however, is not sufficient to identify a2. But there are *a priori* bounds on a3 that limit the ranges of possibilities. The elasticity of the unobserved choice input in educational production a3 is

[8]Recent literature has revived interest in the importance of measurement error in schooling, particularly in within-family estimates, though the possibility of measurement error being important was noted at least as long ago as Bishop (1976) and Griliches (1977). See for example, Ashenfelter and Krueger (1994), Behrman, Rosenzweig and Taubman (1994, 1996), and Miller, Martin and Mulvey (1995, 1997).

bounded below by 0 and is unlikely to exceed 1.[9] Therefore below we are able to summarize how sensitive the estimate of a2 is to different assumptions regarding the value of a3 given the estimates of q and z from relation (11) and the estimates of relation (10). Thus, though we can not identify a2 in the presence of an unobserved choice educational input, we can establish bounds on a2 -- which is an advance on the previous literature that basically assumes that there is no unobserved choice input.

In addition to the problem of identifying a2 from other parameters even if child health is not correlated with the disturbance term in relation (10), of course, there may be biases because child health is correlated with unobserved variables on the right side of relation (10). Relations (4), (8) and (10) can be used to obtain an expression for H without X^h or E in the case in which there is an unobserved choice input in relation (3):

$$(12) \qquad H_{ifc} = [b((w_{ifc}*z*b1)(a2*q + 1-q)([a(S^c_{ifc})^{a1}((w_{ifc}*z*q*a3)/(\lambda_{fc}P^{xe}))^{a3}(A^e_{ifc})^{a4}(S^p_{fc})^{a5} (I^e_{ifc})^{a6}(F^e_{fc})^{a7}(C^e_c)^{a8}\exp(u_{ifc})]^{1-a3qz)qz}/(\lambda_{fc}P^{xh}))^{b1}(S^p_{fc})^{b2}(I^h_{ifc})^{b3}(F^h_{fc})^{b4}(C^h_c)^{b5}\exp(v_{ifc})]^{(1-[(a2+a3(z-zq))(1-a3qz)qz+z-zq]b1}.$$

Relation (9) above is just the special case of relation (12) in which a3 = 0. Whether the OLS estimate of the impact of H in relation (10) is biased upwards or downwards beyond the effect of not having an exponent equal to a2 as in relation (3) depends on the sign of the correlation between the right side of relation (12) and the unobserved variables on the right-side of relation (10) -- which in this case includes not only I^e_{ifc}, F^e_{fc}, and C^e_c and the random error in H as in the case above, but also the price of unobserved choice inputs into education P^{xe}, family welfare parameters (w_{ifc}, z, q), and the Lagrangian on family resources devoted to children λ_{fc}. Though the relations underlying the correlations are somewhat more complicated in this case, the basic result is parallel: generally there is a bias of *a priori* indeterminant sign.

Thus, this framework suggests that biases may occur in estimates of the elasticity of H in relation (3) in either direction. Moreover, if there are unobserved choice educational inputs, there is a bias in an *a priori* indeterminate direction even if estimation procedures can be used to assure that H is not correlated with the disturbance term in the relation estimated. 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 major progress with the data that we use by exploring the first-order question of what direction of bias dominates and then explore some of the possible reasons for that dominance.

Section 3. Data

[9] If the returns to scale for the choice inputs is r, then r-a1-a2 gives an estimate of a3. For some production function technologies a value of r quite close to 1 appears appropriate (e.g., Griliches and Ringstad 1971). For the production of cognitive achievement for the ranges considered in this study, however, there are claims of considerable returns to scale. Therefore, though we are not aware of any empirical estimates, a value of r significantly above 1 is possible. Were reliable estimates of r available from other sources, we could use such estimates and the estimates of a1 and a2 to estimate a3. In the absence of reliable (or any) estimates from other sources, we think it likely that a3 < 1, particularly in light of the magnitudes of our estimates of a1 and a2 below.

The data that we use for our illustrative empirical analysis of the considerations discussed in Section 2 are from the Ghanaian Living Standard Measurement Study (LSMS) data. These data include child anthropometric measures to represent child health, cognitive achievement test scores to represent schooling success, a test of ability that is designed to be independent of schooling to control for that dimension of child endowments, a fairly rich range of household and community characteristics to use for IV estimation of relation (3) similar to related estimates in the literature, 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 910 observations in our sample. We now discuss briefly the major variables that enter directly into our cognitive achievement production function estimates.[10]

Cognitive Achievement Test Scores: The cognitive achievement tests are on reading and on mathematics, with a basic and an advanced subtest for each. The cognitive achievement tests were adapted from tests designed for use in East Africa by the Educational Testing Service in Princeton, NJ as part of the project summarized in Knight and Sabot (1990) in which also are published sample test questions. The original intention was to administer these tests (and the Raven's test of ability described below) only to individuals age nine or older with three or more years of schooling. In the field work, 1467 individuals age 9-17 took the Raven's test and 910 individuals took at least the basic subtests. We use this latter group of 910 individuals as our basic sample.[11] 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 those who are recorded in the data files as having cognitive achievement scores was low on both test scores, with fairly large variances. In our estimates below we translate

---

[10]Our knowledge of details of the data collection procedures is based on documentation prepared by Glewwe (e.g., Glewwe 1994) and conversations with Glewwe (World Bank project head for the data collection) and Dean Jolliffee (Glewwe's research assistant at the data preparation stage). Also see Glewwe (1994, 1996) and Glewwe and Jacoby (1994, 1995).

[11]About a sixth of the eligible children did not take any tests primarily because they were absent. Absent children may not be a random draw of the population. But (i) selective absence also may be a problem with other studies of cognitive achievement, (ii) to the extent that such absence reflects household or community characteristics our FC-FE estimates control for such sample selectivity, and (iii) explicit control for such selectivity through a two-stage procedure does not change materially the estimates presented below. 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 cognitive achievement tests because they were not able to answer any of the basic subtest questions (these individuals were imputed a score of zero in the data files). The advanced subtests, as is standard practice for this type of test, were supposed to be given only to those who answered correctly at least four questions on the basic tests -- about half of those who took the basic reading test and two-thirds of those who took the basic mathematics test. The limitation on giving the advanced subtests only to those who took the basic test generally was followed, though there were some exceptions. Those who did not take the advanced tests because of their poor performance on the basic tests were imputed a score of zero on the advanced tests.

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

Endogenous Variables	Entire Sample		Sibling Subsample	
	Mean	Standard Deviation	Mean	Standard Deviation
Cognitive achievement test				
overall	11.9	12.9	12.3	13.0
reading	4.7	8.2	4.7	6.1
mathematics	7.3	6.0	7.6	8.1
height (centimeters)	93.3	5.5	93.2	5.6
years of schooling	5.5	2.7	5.6	2.6
Predetermined Child Characteristics				
male	0.57	0.50	0.57	0.49
age (years)	13.1	2.5	13.1	2.4
pre-school ability	18.7	6.2	18.6	6.1
Family Background				
log (per-capita expenditures	10.7	0.32	10.7	0.32
father's years of schooling	7.4	5.9	7.6	5.8
mother's years of schooling	3.9	4.9	3.7	4.9
mother's age (years)	26.8	20.9	27.7	20.5
mother's height[a]	61.5	46.5	63.7	45.9
father's height[a]	41.5	47.3	44.3	47.7
mother's height missing	0.36	0.48	0.34	0.47
father's height missing	0.56	0.50	0.54	0.50
head of household age (years)	49.2	13.6	49.6	13.1
head of household sex male	0.62	0.49	0.64	0.48
Community Characteristics				
minutes of nearest middle school	27.5	47.1	20.6	30.2
minutes to nearest primary school	17.0	23.7	14.3	20.5
% of households in cluster with				
private water	0.20	0.34	0.24	0.36
standpipe	0.06	0.20	0.03	0.14

water from vender	0.03	0.11	0.04	0.10
well with pump	0.10	0.26	0.06	0.18
well without pump	0.18	0.30	0.15	0.27
flush toilet	0.05	0.14	0.05	0.15
latrines	0.54	0.39	0.60	0.38
pan/bucket	0.14	0.23	0.15	0.24
health facilities				
cluster 0 to 6 miles from facility	0.78	0.42	0.88	0.32
ampicillin in stock	0.41	0.49	0.40	0.49
number working medical doctors	1.1	2.3	1.0	2.2
number beds in facility	20.6	44.3	16.3	36.0
price per consultation	71.1	104.9	74.8	111.2
hours open per week	63.9	48.1	69.6	52.1
postnatal services	0.72	0.44	0.78	0.41
public facility	0.67	0.47	0.73	0.44
laboratory in facility	0.37	0.48	0.38	0.48
offer immunizations	0.61	0.49	0.66	0.47
geographical area				
forest zone	0.57	0.50	0.58	0.49
Savannah zone	0.09	0.28	0.08	0.27
rural	0.43	0.50	0.46	0.50

[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.

the origin for the total cognitive achievement test score by adding one to each observation so that there are not zero observations. In Table 2 we also present alternative estimates that use only the subsample of 877 individuals who had positive cognitive achievement test scores and positive years of schooling.[12]

Child Health: We use 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). We normalize these scores by the means for each age by sex so as not to confound age and sex effects with health effects. The mean value for our sample for height is 93.3 and the standard deviation is 5.5. Translated into Z scores for comparison with usual international standards, the mean 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.[13] In some alternative estimates we also have explored the impact of children's body mass index (BMI), a shorter-run index of health.[14] because some of the 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 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: Average completed schooling was 4.3 grades, with a standard deviation of 2.5 grades. The distribution of the completed schooling variable is concentrated within a limited range, as are the distributions of our cognitive achievement and child health variables. The use of the 9-17 age group means that the observed completed grades of schooling variable is right-censored as a representation of completed schooling for those still in school. But in our estimate of the cognitive achievement production function in relation (3) we want schooling completed at the time that the cognitive achievement test was taken, not eventual completed schooling.

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 (see Knight and Sabot 1990 for some examples from this test). The test is designed so that formal schooling is not supposed to influence performance. Out of a maximum possible score of 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 ability in estimates of the determination of

---

[12]We also have undertaken estimates for other samples, such as the 1205 individuals who are recorded in the data set as having cognitive achievement test scores (including imputations of zero for those who were unable to answer questions on the basic tests) and for whom the other data that we use are available. We do not present in this paper all the estimates that we have undertaken because the flood of estimates would obscure the basic point. But we report here that the basic pattern in the estimates concerning the estimated impact of child health is robust to the use of other subsamples that we have explored (as well as to other aspects of the specification that we note in the text below).

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

[14]The BMI is weight in kilograms /square of height in meters. Cole (1991) surveys the use of BMI.

cognitive achievement in some other studies in the economics and other literatures (e.g., Boissiere, Knight and Sabot 1985, Glewwe and Jacoby 1994, Nokes et al. 1992b). Because this test result may be endogenous, we also have estimated 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 the impact of child health on cognitive achievement across the different assumptions that are of interest in this study (though it affects the level of estimated elasticities, particularly for schooling, in any particular relation).

Parental Schooling: The Ghanaian LSMS includes information on 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 (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). The mean years of schooling for fathers is 5.8 years, almost twice the mean of 3.0 years for mothers.

Child Age: because the critical data on cognitive achievement test scores are available only for children nine years or 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 First-Stage Estimates for Child Health and Child Schooling Attainment When Cognitive Achievement Test 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). 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^e), constitute a plausible set of first-stage instruments.[15]

Section 4. Estimates of Cognitive Achievement Production Function with no Unobserved Choice Educational Input and with Instrumental Variable Control for Simultaneity and Measurement Error

Table 2 presents alternative estimates of the overall cognitive achievement production function in relation (3) in the case in which there is no unobserved choice input so a3 *a priori* is equal to zero. To focus on the point

---

[15] There also 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 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 our cognitive achievement production function estimates, but also do not include them in our basic set of instruments because they *a priori* would seem to be correlated with the disturbance term in relation (3). We note that 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 are not changed significantly.

Table 2. Cognitive Achievement Production Estimates with No Unobserved Choice Educational Inputs, Ghanaian LSMS: OLS and Instrumental Variable Estimates[a]

Right-Side Variables	Full Sample (n = 910)			Positive Values of Schooling and Test Scores (n = 877)		
	(1)	(2)	(3)	(4)	(5)	(6)
Child Characteristics						
Height	0.89	2.98[b]	6.66[c]	1.05	3.88[b]	8.85[c]
	(1.96)	(2.44)	(3.83)	(2.03)	(2.51)	(4.26)
Years of Schooling	0.93	1.15[b]	1.42[c]	0.82	0.70[b]	1.22[c]
	(11.69)	(4.28)	(4.46)	(10.30)	(2.51)	(3.41)
Ability	1.05	0.92	0.71	1.15	1.10	0.77
	(12.66)	(7.96)	(5.05)	(12.30)	(8.58)	(4.53)
Age	-0.083	-0.314	-0.50	0.011	0.435	-0.272
	(0.43)	(0.69)	(0.95)	(0.05)	(0.74)	(0.38)
Male	0.12	0.20	0.34	0.09	0.17	0.38
	(2.40)	(3.13)	(4.10)	(1.49)	(2.35)	(3.74)
Parental Schooling						
Father	0.069	0.061	0.049	0.066	0.067	0.041
	(2.89)	(2.37)	(1.73)	(2.44)	(2.29)	(1.19)
Mother	0.041	0.028	0.009	0.071	0.071	0.038
	(1.63)	(1.02)	(0.30)	(2.52)	(2.28)	(1.03)
Intercept	-6.42	-14.78	-30.48	-7.71	-22.37	-41.31
	(2.64)	(2.33)	(3.60)	(2.70)	(2.99)	(4.10)
$\bar{R}^2$	0.459	0.403	0.359	0.449	0.399	0.341
Root MSE	0.727	0.742	0.822	0.807	0.820	0.942
F	111.35	88.73	73.81	103.08	84.03	65.76

a. See Table 1 for basic statistics for the data used. Absolute values of t statistics are given in parentheses beneath the point estimates. All variables are in natural logarithms except for the male dummy (which equals 1 for males and 0 for females).

b. Predicted with family background and community characteristics in Table 1 used as first-stage instruments. F tests for the significance of identifying first-stage instruments for column 2 are F(35,869) = 3.96 for height and F(35, 869) = 2.34 for schooling and for column 5 are F(35, 836) = 3.89 for height and F(35, 836) = 2.10 for schooling (all of which are significant at least at the 0.0002 level).

c. Predicted with community characteristics in Table 1 used as first-stage instruments. F tests for the significance of identifying first-stage instruments for column 3 are F(19, 885) = 2.61 for height and F(19, 885) = 3.24 for schooling and for column 6 are F(19, 852) = 2.36 for height and F(19, 852) = 2.48 for schooling (all of which are significant at least at the 0.001 level).

of interest in this paper, we keep the specification simple, with the variables that are indicated as observed in relation (3) -- child height, child schooling at the time the cognitive achievement test was taken, child ability, parental schooling -- plus controls for child age and child sex.

OLS Estimates: The first column in Table 2 gives OLS estimates using the basic sample of 910 individuals who took the basic subtests. These estimates treat child height and child schooling attainment at the time of the test as independent of the disturbance term as in most previous literature. These estimates are consistent with *a priori* expectations regarding signs. They indicate significant positive effects of child health, child schooling and ability, all with elasticities close to one.[16] The sum of the elasticities of the two choice inputs -- child health and years of schooling -- is 1.82. As noted in Section 2, many observers think that there are increasing returns to scale in the range of schooling observed in this sample. But this value implies considerable returns to scale. There also are significant effects of father's schooling with a much smaller elasticity of about 0.07, and a significantly better performance by males than females of 12%. One surprise, given the widespread emphasis on mother's schooling in child development, is that the elasticity of mother's schooling is small at 0.04 and not significantly nonzero at the 10% level. The child's age also does not have a significant effect even at the 10% level.

IV estimates with child health and child schooling treated as endogenously determined: Columns 2 and 3 are IV estimates in which child health and child schooling are treated as endogenously determined or child health is treated as measured with random error. If the first-stage instruments are independent of the disturbance term in relation (3) and are correlated with child health and child schooling, these estimates are consistent. The instruments used for column 2 are a set of household and community variables similar to those used in a number of previous studies (e.g., Alderman *et al.* 1996a,b, Glewwe and Jacoby 1995, Haddad and Bouis 1991, Lavy *et al.* 1996, Schultz and Tansel 1996, Thomas *et al.* 1996). For the column 3 estimates only the community variables are used as instruments.

The parameter estimates in column 2 generally do not differ significantly from those in column 1 with one exception -- i.e., the elasticity estimate for child health, which more than triples. The sum of the elasticities for child years of schooling and child health also increases substantially to 4.13, which implies considerable -- and probably not credible -- returns to scale. Conditional on the assumption that the instruments are valid, the estimates in column 2 suggest that OLS underestimates the impact of child health on child cognitive development considerably. This may reflect the inverse association between the unobserved educational endowments and health endowments noted as a possibility in Section 2 and/or measurement error in child health that is controlled with instrumenting. If the residuals from the cognitive achievement production function are introduced into the reduced-form relations for child height and cognitive achievement, the coefficient estimates are opposite in sign in the child health reduced-form relation from that in the reduced-form relation for the cognitive achievement test score; the coefficient estimates (with absolute t values in parentheses) for the residuals in the respective reduced form

[16]We use the term "significant" to refer to the standard 5% level unless otherwise qualified.

relations are 0.88 (46.4) for the cognitive tests and -0.014 (6.0) for height. This suggests that the inverse association between unobserved education and health endowments is an important part of the story.

The substantial effects of the IV procedures on the estimated impact of child health, moreover, contrast sharply with those on the estimated impact of child schooling. Though child schooling is treated as determined by the same set of instruments as is child health, this treatment does not have a significant impact on the estimated elasticity of years of child schooling (though the magnitude of the point estimate does increase by about a fifth).

The estimates in column 3 differ somewhat more from those in column 1 than do those in column 2. Three of the elasticity estimates are significantly larger than those in column 1. The one for being male is almost three times as large, the one for years of schooling is about 50% bigger, and the one for child health is about 7.5 times larger. Thus the choice of instruments seems to make some difference, though the basic pattern remains the same, with the estimated impact of child health and the estimates of the returns to scale particularly strongly affected by instrumenting.

If the assumption that the family and community instruments are independent of the unobserved family and community endowments that are in the disturbance term of relation (3) were plausible, the appropriate next step would seem to be to explore further exactly what instruments are best by conducting so-called exogeneity tests. But this assumption does not seem plausible, so the IV estimates are quite suspect. Therefore we prefer the estimates to which we turn in Section 5.

<u>Variations on these Estimates to Explore Robustness</u>: We discuss in Section 3 that there are 33 children in the sample who have either no schooling or scores of zero on the cognitive achievement test. That raises the question of whether the estimates are sensitive to excluding such children for whom arguably the model in Section 3 is not appropriate because they are at corner solutions. Therefore, in columns 4-6 in Table 2, we present additional sets of estimates for the sample of 877 individuals who have positive values for both schooling and cognitive achievement in the original data. There are some differences between the two sets of estimates in Table 2 (e.g., the effect of mother's schooling). But the basic patterns that we discuss above remain the same -- if child health is treated as simultaneous and/or measured with error by using IV instead of OLS the estimated elasticity with respect to health and the returns to scale increase substantially (by factors of three or more for the former) and significantly, while that for schooling does not change significantly or changes much less depending on which estimator is used.

We also have undertaken other estimates for the cognitive achievement production function in relation (3) with OLS and IV procedures to explore the robustness of the results in Table 2. These variations include (1) selectivity control for children not completing first grade, (2) including school quality indicators in the instrument set, (3) including school quality in the production function, (4) limiting sample only to children co-resident with at least one biological parent, (5) using BMI instead of height, (6) disaggregating cognitive achievement scores into mathematics versus reading components, (7) estimating separately by gender. The basic results are robust to all these variations. Many of the details are given in Behrman and Lavy (1994).

Section 5. Estimates of Cognitive Achievement Production Function with Control for Unobserved Family and Community Effects

The first-stage variables in the IV procedure used in Section 4 are family and community characteristics that *a priori* would appear to be associated with F^e_{fc}, and C^e_c, respectively. 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 -- e.g, average family genetic endowments and the general household learning environment.[17] 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.[18] Thus, while the IV estimates in Section 4 are striking regarding their implications for the strong positive impact of child health on child schooling success and on returns to scale conditional on the maintained assumption about the instruments used, this maintained assumption is suspect.

Therefore in this section we present FC-FE estimates that control for unobserved family and community characteristics. These estimates eliminate from the disturbance term in relation (3) all unobserved family and community characteristics and thus 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 unobserved predetermined family endowments that affect the production of child cognitive achievement and child health. However, these estimates do not eliminate biases due to unobserved individual child characteristics. Also these estimates are more subject to biases towards zero due to random measurement error because the noise-to-signal ratio is increased by focusing on deviations from family means.[19] We return to these problems below.

Family and Community Fixed-Effect (FC-FE) Estimates for Basic Sample: For the FC-FE estimates we use the subsample of 587 individuals in the 238 families in our data for which there are at least two siblings in the

---

[17]A number of studies have claimed that unobserved family characteristics affect importantly child education (e.g., Behrman *et al* 1980, Behrman, Rosenzweig and Taubman 1994, 1996, Behrman and Wolfe 1984, 1987, Foster and Roy 1995, Foster and Rosenzweig 1996, Olneck 1977, Rosenzweig and Schultz 1987, Rosenzweig and Wolpin 1994).

[18]The importance of unobserved community characteristics in understanding various dimensions of human resource investments is emphasized in several recent studies (e.g., Behrman and Deolalikar 1993, Foster and Rosenzweig 1996, Foster and Roy 1995, Gertler and Molyneaux 1994, Pitt, Rosenzweig and Gibbons 1993, Rosenzweig and Wolpin 1986).

[19]The increased bias towards zero in such estimates due to random measurement error has been emphasized at least since Bishop (1976) and Griliches (1977). Ashenfelter and Krueger (1994) 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. In contrast, Behrman, Rosenzweig and Taubman (1994, 1996) and Miller, Mulvey and Martin (1995, 1997) find that endowments are significant even with control for measurement error by using reports of others as an instrument as suggested by Ashenfelter and Krueger.

same family. Of course to undertake such estimates there must be some within-family variance in the relevant variables. Table 3 presents the share of the total variance that is within-family rather than between-family. Over half of the variance in cognitive achievement test scores is within-family, so FC-FE 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, about half of the variance in both years of schooling and ability and over three-fifths of the variance in height also is within-family.

A further question is whether the FC-FE sample is a selected one. Because having two children in the relevant age range to be in the FC-FE sample reflects unobserved family characteristics that are controlled in these estimates, the use of such a subsample should not be contaminated by selectivity bias (as note in previous studies, e.g., Heckman and MaCurdy 1980, Pitt and Rosenzweig 1990, and Behrman and Deolalikar 1993). Nevertheless it is reassuring that the OLS and IV estimates for the FC-FE subsample in columns 1-3 in Table 4 do not differ significantly from the estimates in columns 1-3 of Table 2.

Column 4 in Table 4 gives FC-FE estimates for the cognitive achievement production function with child height continuing to represent child health. The coefficient estimates for years of schooling and for ability, in particular, are about as precisely estimated in these FC-FE estimates as in the OLS estimates in Table 2 and are not significantly different for the FC-FE estimates in Table 4 from those in the OLS estimates in Table 2. For these variables, therefore, there is no evidence that random measurement error causes substantial biases towards zero in their FC-FE coefficient estimates.

However, for the coefficient estimate of primary concern -- that for height -- the results are striking, and quite in contrast to those in Section 4. The FC-FE coefficient estimate for child height is imprecisely estimated, not significantly different from zero even at the 40 percent level and only about 40 percent of the magnitude of the OLS estimate in column 1 and about 20 percent of the IV estimate in column 2. If this is the appropriate estimate for this coefficient, it suggests that the OLS estimates and even more so the IV estimates in Section 4 are substantially upward biased and misleading regarding the impact of the observed range of child health on child schooling performance. The estimates in Section 4 also substantially overestimate the extent of increasing returns to scale for the choice inputs, for which the point estimate from column 4 in Table 4 is 1.22 with substantial imprecision as compared with 2.27 for the OLS estimates in column 1, 4.66 for the IV estimates in column 2, and 8.52 for the IV estimates in column 3. The failure to control for unobserved heterogeneity in predetermined family and community endowments that affect the production of child quality 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 extent of returns to scale. And these biases are even larger for the IV estimates with family and community instruments such that we use in Section 4 because of the apparent correlation of such instruments with unobserved family and community variables in the disturbance term in the cognitive achievement production function.

Robustness of Patterns in Estimates: Might random measurement error be accounting for these results? It does not seem likely that this is the case for three reasons: (1) The estimates for years of schooling and ability

Table 3. Percentages of Total Variance That Is Within Family for Key Child Characteristics Ghanaian LSMS: Sibling Sample (n = 587)

Child Characteristics	Percent of Total Variance That is Within Family
Cognitive Achievement	58
Height	62
Years of Schooling	47
Ability	52

Table 4. Cognitive Achievement Production Estimates with No Unobserved Choice Educational Inputs, Ghanaian LSMS: Sibling Sample[a]

Right-Side Variables	Sibling Sample (n = 587)				
	OLS	Height and Schooling Instrument. Variables	Height and Schooling Instrument. Variables	Family & Comm. Fixed Effects	Comm. Fixed Effects
	(1)	(2)	(3)	(4)	(5)
<u>Child Characteristics</u>					
Height	1.37	2.93[b]	5.99[c]	0.56	0.53
	(2.50)	(2.00)	(3.08)	(0.77)	(0.93)
Years of Schooling	0.91	1.72[b]	1.92[c]	0.66	0.80
	(8.50)	(4.82)	(4.18)	(4.60)	(7.07)
Ability	1.03	0.76	0.61	0.91	0.90
	(10.43)	(5.55)	(3.58)	(7.41)	(8.47)
Age	0.092	-1.07	-1.18	0.48	0.37
	(0.37)	(1.76)	(1.60)	(1.61)	(1.47)
Male	0.18	0.27	0.39	0.11	0.14
	(2.83)	(3.21)	(3.84)	(1.50)	(2.20)
<u>Parental Schooling</u>					
Father	0.047	0.016	0.002		
	(1.59)	(0.48)	(0.04)		
Mother	0.063	0.032	0.014		
	(2.11)	(0.95)	(0.37)		
<u>Intercept</u>	-9.35	-11.32	-24.54		
	(3.11)	(1.42)	(2.43)		
$\overline{R}^2$	0.480	0.416	0.379	0.775	0.613
Root MSE	0.703	0.754	0.828	0.605	0.653
F	78.17	60.68	50.55	22.68	10.02

a. See Table 1 for basic statistics for the data used. Absolute values of t statistics are given in parentheses beneath the point estimates. All variables are in natural logarithms except for the male dummy (which equals 1 for males and 0 for females).

b. Health and schooling predicted with family background and community characteristics in Table 1 used as first-stage instruments.

c. Health and schooling predicted with community characteristics in Table 1 used as first-stage instruments.

do not suggest a critical role of such measurement error in the FC-FE estimates, as noted above, and we have no reason to think that the anthropometric indicators are much more contaminated by random measurement error than are these other two variables. In fact, it would seem to us likely that the anthropometric measurements have less random measurement error than the ability measure. (2) The use of 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 for the subsample for which these data are available as in Pitt, Rosenzweig and Hassan (1990) does not change significantly our FC-FE estimates.[20] (3) For the IV estimates, that the residuals have estimated coefficients with opposite signs when they are included in reduced-form estimates for child health and child education (as noted above) suggests that the differences among the estimates with and without instrumenting in Table 4 is not due to measurement error alone.

Might the fact that FC-FE estimates do not control for important unobserved individual factors underlie our results? This would require that there is an unobserved individual component in the disturbance term of relation (3) that is important and that is negatively associated with observed health. We have doubts about the relevance of such a possibility because we control for ability (and for any correlated attributes), which would seem to be the most important often unobserved individual-specific component in the production of cognitive achievement. We do not have the possibility of exploring this possibility through instrumenting, however, because we do not have individual-specific instruments beyond those that already appear in our estimated relations. Also there is the question of how robust are these results to the other variants mentioned in Section 4, such as using BMI to measure child health or separating the estimates for reading versus mathematics scores. The answer is that the results with these variants hold in the same pattern as in Table 4 (we give some examples in Behrman and Lavy 1994).

Community Fixed Effects (C-FE) Alone: The FC-FE 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. C-FE estimates in column 5 of Table 4 suggest that much of the effects of the FC-FE estimates in column 4 of Table 4 are due to unobserved community effects rather than unobserved family effects beyond the community effects. The elasticity estimates with respect to height are less than 40 percent of those in the OLS estimates, and less than 20 percent of those if height is instrumented and insignificant at standard levels. 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

[20]The data include earlier observations on height for 330 of the 587 members of our FC-FE sample. For this subsample, the correlation between current and earlier height is 0.99, suggesting that there is little measurement error that is independent over time. If the earlier measure is used as an instrument to control for measurement error for this subsample, the coefficient estimate of height is practically zero with a large standard error.

environment on cognitive achievement to health.[21] In contrast to the effects on the coefficient estimates of health, the coefficient estimates for years of schooling and for 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 necessarily mean that family effects (beyond the community effects) are unimportant. An F test for restricting the coefficients on such family effects to be zero using the FC-FE sample is rejected (F = 1.51 with critical values of 1.22 at the 5 percent level and 1.32 at the 1 percent level). Therefore the preferred estimates are those in column 4 of Table 4 that control for both unobserved community effects and unobserved family effects beyond the community effects. But note that control for community effects alone leads to basically the same (relatively small, insignificant) estimate for the elasticity of cognitive achievement with respect to height as does control for both family and community effects. The overall significance of family effects beyond community effects is manifested primarily in changes in other estimates.

Section 6. Effect of Including the Possibility of An Unobserved Choice Educational Input

The usual production function estimates assume (usually implicitly, but in some cases explicitly) that all choice inputs are observed. But, as is demonstrated with regard to relation (11) in Section 2, if there is an unobserved choice input, the production function parameter estimates are biased in an indeterminant direction even if the observed production inputs -- including child health in this case -- are treated so as to control their correlation with the disturbance term. We now implement the strategy outlined with regard to relation (11) in Section 2 to explore the possible importance of an unobserved choice input.

Panel A in Table 5 gives alternative estimates of relation (11), in column 1 with OLS and in column 2 with IV procedures using the individual-specific variables (i.e., ability, age, sex) and their interactions with the observed family and community variables in Table 1 as instruments. Panel B gives the estimates for the welfare parameters in relations (1) and (2) that are implied by the estimates in Panel A. The implied point estimates for inequality aversion (z) imply little (OLS) or no (IV) concern about inequality in cognitive achievement among children in a household, with human capital investments allocated more to children with greater endowments so as to reinforce their impact on inequality. The implied point estimates for the weight of cognitive achievement in the child quality production function (q) imply that child health directly (<u>not</u> indirectly through cognitive achievements) is weighed about twice as much as cognitive achievement in the composite child quality index.

Panel C gives alternative estimates of the true value of a2, the elasticity of cognitive achievement with respect to child health (height) in the cognitive achievement production function in relation (3). These alternatives are conditional on: (i) the estimates of relation (11) in Table 5 (OLS versus IV), (ii) the estimates of the exponent of the ratio of child height in relation (10) as presented in Table 4 (OLS is from column 1, IV from column 2, FC-

[21]These community effects are not captured by observed schooling quality characteristics (see discussion in Section 4).

Table 5. Welfare Function Parameter Estimates (Relation 11), Ghanaian LSMS: Sibling Sample Using Each Family as an Observation (n = 493)[a]

	OLS			Instrumental Variable Estimates[b]		
A. Summary of Estimates	(1)			(2)		
Cognitive Achievement	0.24 (13.22)			0.31 (10.65)		
Heights	0.52 (1.89)			0.75 (1.32)		
Intercept	-0.23 (11.37)			-0.20 (8.56)		
$\overline{R}^2$	0.271			0.187		
Root MSE	0.740			0.752		
F	92.43			57.70		
B. Implied Welfare Parameters						
Inequality Aversion (z)	0.76			1.06		
Weight of Cognitive Achievement in Child Quality (q)	0.32			0.29		
C. Implied Value of a2, conditional on value of a3 and on estimated parameter for child health in prod. function	Estimate from Table 4 for Child Health Parameter			Estimate from Table 4 for Child Health Parameter		
	OLS (1.37)	IV (2.93)	FE (0.56)	OLS (1.37)	IV (2.93)	FE (0.56)
a3 = 0.33	1.32	3.68	0.44	1.28	3.02	0.38
a3 = 0.67	1.28	3.51	0.32	1.23	3.20	0.20
a3 = 1.00	1.28	3.33	0.22	1.24	3.50	0.06

a. See Table 1 for basic statistics for the data used. Absolute values of t statistics are given in parentheses beneath the point estimates. All variables are the natural logarithms of ratios for siblings in Panel A, which gives two alternative estimates of relation (11). Panel B gives the values of the welfare parameters in relations (1) and (2) that are implied by these estimates. Panel C gives the estimated elasticities of cognitive achievement with respect to child health (height) that are implied by the estimates of relation (11) in this table, the estimates in Table 4 (OLS refers to column 1 in Table 4, IV to column 2, and FE to column 4), and the alternative assumptions for a3 that are indicated in the first column.

b. Cognitive achievement test scores and height predicted with first-stage instruments including age, sex and ability and interactions of these three individual variables with family and community variables in Table 1.

FE from column 4), and (iii) the true value of the elasticity of cognitive achievement with respect to the unobserved input a3 (for which values of 0.33, 0.67 and 1.00 are given as illustrations). The estimates in Panel C indicate that the true value of a2 may differ a fair amount from the value estimated as if there is no unobserved choice input and that these differences may be biased towards or away from zero. In the examples given in this table the ratio of the estimated value of a2 under the assumption that there is no unobserved choice input to the true value range from 80% to 933%. Thus the existence of unobserved choice inputs may mean that production function parameters for observed inputs are mis-estimated considerably. For the FC-FE estimates that we argue in Section 6 are preferable, Panel C implies lower elasticities of cognitive achievement with respect to child height than would be the case if there were no such unobserved choice input. For the preferred IV estimates of relation (11) in Table 5, for example, they imply an elasticity of 0.38 if a3 = 0.33 (and lower if a3 is greater), which is two-thirds of the magnitude implied in Table 4 if a3 = 0.

Section 7. 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 for child schooling attainment that underlies the model in Section 2, 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 if 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 because similar relations, often characterized as "conditional demand functions," are not uncommon in the more general empirical literature.[22]

Consider the two following log linear and simplified forms of the reduced form relations for child health and child schooling:

(13A) $\ln H = a_{11}\ln P_H + a_{12}\ln P_S + e_1$ and

(13B) $\ln S = a_{21}\ln P_H + a_{22}\ln P_S + e_2$.

$\ln P_H$ can be eliminated from these two relations to obtain:

(13C) $\ln S = a_{31}\ln H + a_{32}\ln P_S + e_3$.

Now it might appear that relation (13C) is an expression in which schooling depends on child health and the

[22]We have benefited from discussions with Andrew Foster on this point.

Table 6. Cognitive Achievement Production Estimates without Schooling, Ghanaian LSMS: Sibling Sample[a]

Right-Side Variables	Sibling Sample (n = 587)		
	OLS	Height and Schooling Inst. Variables	Family and Community Fixed Effects
	(1)	(2)	(3)
Child Characteristics			
Height	2.68 (4.82)	4.63[b] (3.24)	1.33 (1.81)
Ability	1.22 (11.95)	1.15 (10.24)	1.02 (8.20)
Age	1.57 (8.53)	1.69 (8.32)	1.56 (8.03)
Male	0.19 (2.90)	0.27 (3.20)	0.14 (1.86)
Parental Schooling			
Father	0.073 (2.37)	0.068 (2.18)	
Mother	0.086 (2.71)	0.078 (2.42)	
Intercept	-21.72 (7.80)	-31.00 (4.53)	
$\bar{R}^2$	0.416	0.403	0.761
Root MSE	0.745	0.753	0.623
F	70.52	67.03	4.52

a. See Table 1 for basic statistics for the data used. Absolute values of t statistics are given in parentheses beneath the point estimates. All variables are in natural logarithms except for the male dummy (which equals 1 for males and 0 for females).

b. Health and schooling predicted with family background and community characteristics in Table 1 used as first-stage instruments.

variable that has been eliminated to obtain this expression (i.e., $\ln P_H$) can be used as an instrument to control for the endogeneity of child health (H). However, what is a_{31}? It is but the ratio of the coefficient of the variable that has been eliminated in relation (13B) to the coefficient of that variable in relation (13A) (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, <u>not</u> 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 (13C), the coefficient of H would have been a_{22}/a_{12}, so the so-called "conditional demand" effect of H depends on what variable has been eliminated from the reduced forms to obtain the "conditional demand relation."

Therefore it is not clear how we could obtain conditional demand estimates from our data.[23] 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 because, in such a case, the coefficient of child health includes both the impact of itself and of the correlated schooling effect.[24] Table 6 presents some relevant estimates for the FC-FE sample. Column 1 is the OLS estimate without schooling. The coefficient estimate of child height is about 95 percent greater than in column 1 of Table 4, so child height indeed does pick up part of the effect of the omitted schooling variable. Column 2 is the IV estimate using family and community characteristics as instruments, without schooling. The coefficient estimate of child height is almost 60 percent greater than in column 2 of Table 4, so again child height picks up part of the effect of the omitted schooling variable. Column 3 in Table 6 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 over twice as large as in column 4 of Table 4 and is significantly nonzero at the 10 percent level, so child health again seems to pick up some of the effect of the omitted school variable. But the coefficient estimate is only half of that in column 1 in Table 6, and not significantly different from zero at the 5 percent level. Thus our estimates suggest somewhat, but not very strongly, that child health is operating through schooling to affect child cognitive achievement and that this effect is substantially overestimated if there is not control for unobserved fixed family and community determinants.

---

[23] 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. Alderman, Behrman, Lavy and Menon (1997) follow this estimation strategy for estimating the impact of child anthropometrics on age of school enrollment.

[24] This would seem to give an upper bound on the direct and indirect (through schooling) effects because 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.

Section 8. 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). This is an important result because of the perception that schooling is a critical human capital investment that has substantial effects on a wide range of outcomes, including economic productivity and income distribution. However, this literature generally does not control for the probable endogeneity of the determination of child health, measurement error or the possibility of unobserved choice inputs -- which *a priori* may result in biases that are upwards or downwards in the estimated impact of child health on child schooling.

Our estimates with control for health endogeneity and measurement error by using family and community characteristics as instruments, conditional on the assumption that such instruments are appropriate, suggest that the biases in the estimated impact of child health on schooling success and in the estimated returns to scale in education production in the previous literature have been big and downward. This result seems robust to a number of alternative approaches regarding the sample, the dependent variable and the observed variables. 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 and returns to scale in education production are more important than previously realized because of important downward biases in OLS estimates.

However the IV results depend on the implausible assumption that the family and community characteristics used as instruments are independent of the unobserved family and community endowments that directly affect child cognitive development. Moreover, these results regarding the greater impact of child schooling and greater returns to scale are <u>not</u> robust to control for unobserved family and community characteristics using family and community fixed-effects estimates. In fact, such estimates yield insignificant estimates for child health that are only 40 percent of the magnitude of OLS estimates and 20 percent of the magnitude of the IV estimates and estimates of the returns to scale for variable inputs that only are about half of the OLS estimates and about a quarter of the IV estimates. This effect of control for these fixed effects on the child health coefficient contrasts strongly with the lack of an effect of such FC-FE estimates on the coefficient estimates of child schooling and ability, two variables that also (like child health and cognitive achievement) have limited variance in the data and possibly are measured with error. The elasticity estimates of the latter two variables are not changed substantially by the control for unobserved family and community variables. This general pattern is robust to use of different samples, different dependent variables and exploration of measurement error. Within-community estimates suggest that much of the impact on the coefficient estimates of FC-FE estimates is due to unobserved community factors rather than to unobserved family factors beyond the community effects. Apparently there are important unobserved community characteristics that affect both child anthropometric measures and child cognitive achievement that, 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. Further, allowing for unobserved choice inputs in education production in an innovative procedure suggests that the usual estimates are further biased, probably upward though our efforts to explore this possibility are characterized by considerable imprecision. 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 at the standard 5 percent level if there is control for unobserved family and community effects.

We argue that the family and community fixed-effects estimates with control for the possibility that there are unobserved choice inputs are preferred because there would always seem to be possibilities of important unobserved family and community effects and unobserved allocated inputs that affect behavioral data. If there are such factors, the IV estimates with instruments such as in Table 1 overstate substantially the impact of child health on child schooling success and the extent of returns to scale and cause greater distortions than the OLS estimates. The problems are that the instruments that generally are used apparently are not independent of the disturbance term in the cognitive achievement production function and that generally there is likely to be unobserved choice variables. These problems raise questions about the previous empirical estimates of the impact of child health on child school success based on behavioral data. Similar problems 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 and returns to scale are limited (if they exist), 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 in order to control for behavioral choices or random measurement error and that there are no unobserved choice production inputs.

References

Alderman, Harold, Jere R. Behrman, Victor Lavy, and Rekha Menon, 1997, "Child Nutrition, Child Health, and School Enrollment: A Longitudinal Analysis," Washington, D.C.: World Bank, PRWP Paper 1700.

Alderman, Harold, Jere R. Behrman, David Ross, and Richard Sabot, 1996a, "Decomposing the Gender Gap in Cognitive Skills in a Poor Rural Economy," *Journal of Human Resources* 31:1 (Winter), 229-254.

Alderman, Harold, Jere R. Behrman, David Ross, and Richard Sabot, 1996b, "The Returns to Endogenous Human Capital in Pakistan's Rural Wage Labor Market," *Oxford Bulletin of Economics and Statistics* 58:1 (February), 29-55.

Ashenfelter, Orley and Alan Krueger, 1994, "Estimates of the Economic Return to Schooling from a New Sample of Twins," *American Economic Review* 84:5 (December), 1157-74.

Behrman, Jere R., 1993, "The Economic Rationale for Investing in Nutrition in Developing Countries," *World Development* 21:11 (November), 1749-1772.

Behrman, Jere R. and Anil B. Deolalikar, 1993, "Unobserved Household and Community Heterogeneity and the Labor Market Impact of Schooling: A Case Study for Indonesia," *Economic Development and Cultural Change* 41:3 (April 1993), 461-488.

Behrman, Jere R., Z. Hrubec, Paul Taubman, and T.J. Wales, 1980, *Socioeconomic Success: A Study of the Effects of Genetic Endowments, Family Environment and Schooling*, Amsterdam: North-Holland Publishing Company.

Behrman, Jere R. and Victor Lavy, 1994, "Child Health and Schooling Achievement: Association, Causality, and Household Allocations," Washington, D.C.: World Bank, World Bank Living Standards Measurement Study Paper No. 104.

Behrman, Jere R., Robert A. Pollak, and Paul Taubman, 1982, "Parental Preferences and Provision for Progeny," *Journal of Political Economy* 90:1 (February), 52-73.

Behrman, Jere R., Mark R. Rosenzweig, and Paul Taubman, 1994, "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment," *Journal of Political Economy* 102:6 (December), 1131-1174.

Behrman, Jere R., Mark R. Rosenzweig, and Paul Taubman, 1996, "College Choice and Wages: Estimates Using Data on Female Twins," *Review of Economics and Statistics* 73:4 (November), 672-685.

Behrman, Jere R. and Barbara L. Wolfe, 1984, "The Socioeconomic Impact of Schooling in a Developing Country," *Review of Economics and Statistics* 66:2 (May), 296-303.

Behrman, Jere R. and Barbara L. Wolfe, 1987, "How Does Mother's Schooling Affect the Family's Health, Nutrition, Medical Care Usage, and Household Sanitation?" *Journal of Econometrics* 36, 185-204.

Blackburn, McKinley, L. and David Neumark, 1993, "Omitted-Ability Bias and the Increase in the Return to Schooling," *Journal of Labor Economics* 11:3 (July 1993), 521-544.

Bishop, John, 1976, "Reporting Errors and the True Return to Schooling," Madison, WI: University of Wisconsin, mimeo.

Bishop, John H., 1989, "Is the Test Score Decline Responsible for the Productivity Growth Decline?" *American Economic Review* 79 (March), 178-197.

Bishop, John H., 1991, "Achievement, Test Scores, and Relative Wages," in Marvin H. Kosters, ed., *Workers and Their Wages: Changing Patterns in the United States*, Washington, D.C.: American Enterprise Institute Press.

Boissiere, M., J.B. Knight, and R.H. Sabot, 1985, "Earnings, Schooling, Ability and Cognitive Skills," *American Economic Review* 75, 1016-1030.

Case, Anne and Angus Deaton, 1996, "School Quality and Educational Outcomes in South Africa," Princeton, NJ: Princeton University, mimeo.

Chutikul, Sirilaksana, 1986, *Malnourished Children: An Economic Approach to the Causes and Consequences in Rural Thailand*, East-West Population Institute, East-West Center, Honolulu, Hawaii, Number 102.

Cole, T. J., 1991, "Weight-Stature Indices to Measure Underweight, Overweight, and Obesity," in *Anthropometric Assessment of Nutritional Status*, Wiley-Liss, Inc, 83-111.

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.

Foster, Andrew D. and Mark R. Rosenzweig, 1996, "Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution," *American Economic Review* 86:4 (Sept.), 931-953.

Foster, Andrew D. and Nikhil Roy, 1995, "The Dynamics of Education and Fertility: Evidence from a Family Planning Experiment," Philadelphia: University of Pennsylvania, mimeo.

Freeman, H.E., R.E. Klein, J.W. Townsend, and A. Lechtig, 1980, "Nutritional and Cognitive Development Among Rural Guatemalan Children," *American Journal of Public Health* 70:12, 1277-1285.

Gertler, Paul and John W. Molyneaux, 1994, "How Economic Development and Family Planning Programs Combined to Reduce Indonesian Fertility," *Demography* 31:1 (February), 33-63.

Glewwe, Paul, 1994, "Are Rates of Return to Schooling Estimated from Wage Data Relevant Guides for Government Investments in Education? Evidence from a Developing Country," LSMS Working Paper No. 76, Washington, D.C. World Bank.

Glewwe, Paul, 1996, "The Relevance of Standard Estimates of Rates of Return to Schooling for Education Policy: A Critical Assessment," *Journal of Development Economics* 51:2 (December), 267-290.

Glewwe, Paul and Hanan Jacoby, 1994, "Student Achievement and Schooling Choice in Low Income Countries: Evidence from Ghana," *Journal of Human Resources* 29:3 (Summer), 842-64.

Glewwe, Paul and Hanan Jacoby, 1995, "An Economic Analysis of Delayed Primary School Enrollment and Childhood Malnutrition in a Low Income Country," *Review of Economics and Statistics* 77:1 (February), 156-169.

Glewwe, Paul, Hanan Jacoby, and Elizabeth King, 1996, "Estimating an Economic Model of Nutrition and Learning using Longitudinal Data," Washington, D.C.: The World Bank, mimeo.

Gomes-Neto, Joao Batista, Eric A. Hanushek, Raimundo Helio Leite, and Roberto Claudio Frota-Bezzera, 1992, "Health and Schooling: Evidence and Policy Implications for Developing Countries," Rochester: Rochester University Center for Economic Research Working Paper No. 306, mimeo.

Griliches, Z., 1977, "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica* 45, 1-22.

Griliches, A. and V. Ringstad, 1971, *Economies of Scale and the Form of the Production Function*, Amsterdam: North-Holland, Contributions to Economic Analysis No. 72.

Haddad, Lawrence and Howarth Bouis, 1991, "The Impact of Nutritional Status on Agricultural Productivity: Wage Evidence from the Philippines," *Oxford Bulletin of Economics and Statistics* 53:1 (February), 45-68.

Heckman, J.J. and T.E. MaCurdy, 1980, "A Life Cycle Model of Female Labor Supply," *Review of Economic Studies* 47: 47-74.

Jamison, Dean T., 1986, "Child Malnutrition and School Performance in China," *Journal of Development Economics* 20:2, 299-310.

Knight, J.B. and R.H. Sabot, 1990, *Education, Productivity and Inequality: The East African Natural Experiment*, New York: Oxford University Press.

Lavy, Victor, Jennifer Spratt and Nathalie Leboucher, 1997, "Patterns of Incidence and Change in Moroccan Literacy," *Comparative Education Review* 41:2 (May).

Lavy, Victor, John Strauss, Duncan Thomas, and Philippe de Vreyer, 1996, "Quality of Health Care and Survival and Health Outcomes in Ghana," *Journal of Health Economics* 15, 333-357.

Miller, Jane E. and Sanders Korenman, 1993, "Poverty, Nutritional Status, Growth and Cognitive Development of Children in the United States," Princeton, NJ: OPR Working Paper No. 93-5.

Miller, Paul, Charles Mulvey and Nick Martin, 1995, "What Do Twins Studies Tell Us About The Economic Returns to Education? A Comparison of US and Australian Findings," *American Economic Review* 85:3, 586-599.

Miller, Paul, Charles Mulvey and Nick Martin, 1997, "Family Characteristics and the Returns to Schooling: Evidence on Gender Differences from a Sample of Australian Twins," *Economica*.

Moll, Peter G., 1996, "Primary Schooling, Cognitive Skills and Wages in South Africa," Chicago: mimeo.

Moock, Peter R. and Joanne Leslie, 1986, "Childhood Malnutrition and Schooling in the Terai Region of Nepal," *Journal of Development Economics* 20:1, 33-52.

Nokes, C., S. M. Grantham-McGregor, A. W. Sawyer, E.S. Cooper and D.A.P. Bundy, 1992a, "Parasitic Helminth Infection and Cognitive Function in School children," *Proceeding of the Royal Society London*

B 247 (22 February), 77-8 .
Nokes, C., S. M. Grantham-McGregor, A. W. Sawyer, E.S. Cooper, B.A. Robinson and D.A.P. Bundy, 1992b, "Moderate to Heavy Infections of *Trichuris trichiura* Affect Cognitive Function in Jamaican School Children," *Parasitology* 104, 539-547.
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.
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 Indonesia," *International Economic Review* 31:4 (November), 969-989.
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, E., 1990, *Malnutrition and Infection in the Classroom*, Paris: Unesco.
Pollitt, Ernesto, Phongjan Hathirat, Nittaya J. Kotchabhakdi, Lavon Missell, and Aree Valyasevi, 1989, "Iron Deficiency and Educational Achievement in Thailand," *The American Journal of Clinical Nutrition* 50, 687-697.
Raven, J.C., 1956, *Guide to the Coloured Progressive Matrices (Sets A, Ab, B).* London: Lewis.
Rosenzweig, Mark R. and T. Paul Schultz, 1983, "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight, *Journal of Political Economy* 91:5 (October), 723-746.
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.
Rosenzweig, Mark R. and T. Paul Schultz, 1988, "The Stability of Household Production Technology: A Replication," *Journal of Human Resources* 23:4, 535-549.
Rosenzweig, Mark R. and Kenneth J. Wolpin, 1986, "Evaluating the Effects of Optimally Distributed Public Programs," *American Economic Review* 76:3 (June), 470-487.
Rosenzweig, Mark R. and Kenneth J. Wolpin, 1994, "Are There Increasing Returns to Intergenerational Production of Human Capital?" *Journal of Human Resources* 29:2 (Spring), 670-693.
Schultz, T. Paul and Aysit Tansel, 1996, "Wage and Labor Supply Effects of Illness in Cote D'Ivoire and Ghana: Instrumental Variable Estimates for Days Disabled," New Haven, CT: Yale University, mimeo.
Seshadri, Subadra and Tara Gopaldas, 1989, "Impact of Iron Supplementation on Cognitive Functions in Preschool and School-Aged Children: The Indian Experience," *The American Journal of Clinical Nutrition* 50, 675-586.
Sigman, Marian, Charlotte Neumann, Ake A.J. Jansen, and Nimrod Bwibo, 1989, "Cognitive Abilities of Kenyan Children in Relation to Nutrition, Family Characteristics, and Education," *Child Development* 60, 1463-1474.
Soemantri, A.G., 1989, "Preliminary Findings on Iron Supplementation and Learning Achievement of Rural Indonesian Children," *The American Journal of Clinical Nutrition* 50, 698-702.
Soemantri, A.G., E. Pollitt, and I. Kim, 1985, "Iron Deficiency Anemia and Educational Achievement," *The American Journal of Clinical Nutrition* 42, 1221-1228.
Thomas, Duncan, Victor Lavy, and John Strauss, 1996, "Public Policy and Anthropometric Outcomes in the Cote d'Ivoire," *Journal of Public Economics*.
Wilson, A.B., 1981, "Longitudinal Analysis of Diet, Physical Growth, Verbal Development and School Performance," in J.B. Balderston, A.B. Wilson, M.E. Freire and M.S. Simonen, eds., *Malnourished Children of the Rural Poor*, Boston: Auburn House.
Wolfe, Barbara, 1985, "The Influence of Health on School Outcomes: A Multivariate Approach," *Medical Care* (October), 1127-1138.