Fordham University
Department of Economics
Discussion Paper Series

Long-Term Impact of Investments in Early Schooling – Empirical Evidence from Rural Ethiopia

Subha Mani
Fordham University, Department of Economics
John Hoddinott
International Food Policy Research Institute
John Strauss
University of Southern California, Department of Economics

Discussion Paper No: 2009-09
November 2009

Department of Economics
Fordham University
441 E Fordham Rd, Dealy Hall
Bronx, NY 10458
(718) 817-4048

Long-term impact of investments in early schooling – Empirical evidence from rural Ethiopia

Subha Mani, Fordham University

John Hoddinott, International Food Policy Research Institute

John Strauss, University of Southern California

Acknowledgements:
We thank the following organizations for funding the survey work on which this paper is based: SIDA, USAID and the World Food Programme. We thank John Ham, Jeffrey B Nugent and seminar participants at the PAA meetings for comments and suggestions. We are thankful to Yisehac Yohannes for clarifying several data related questions. Financial support was provided by the College of Letters, Arts and Sciences, University of Southern California and the International Food Policy Research Institute.

Corresponding author:
Subha Mani**,** Department of Economics, Fordham University, 441 East Fordham Road, Bronx, New York 10458. Email – smani@fordham.edu

Abstract:

This paper identifies the cumulative impact of early schooling investments on later schooling outcomes in a developing country context using enrollment status and relative grade attainment as short-run and long-run measures of schooling. Using a child-level longitudinal data set from rural Ethiopia, we estimate a dynamic conditional schooling demand function where the coefficient estimate on the lagged dependent variable captures the impact of all previous periods schooling inputs and resources. We find that this lagged dependent variable indicates a strong positive association between current and lagged schooling. Past history matters more for girls than boys and for children from higher income households compared to the poor.

1. Introduction

In much of the developing world, households reside in risky environments. In the absence of full insurance or other smoothing mechanisms, the realization of these risks, shocks, leads to losses of utility. As Dercon (2005), Alderman, Hoddinott and Kinsey (2006) and other have noted, the importance of these losses from a policy perspective depends partly on whether such shocks induce path dependence. That is, do transitory shocks have permanent consequences? Or put another way, is past history destiny?

In the last ten years, a series of papers have demonstrated that in the context of one dimension of human capital, nutrition, does indeed demonstrate path dependence. Maccini and Yang (2009) have shown that rainfall in the year and district of birth in Indonesia have long-run effects, on both attained adult height for men and women and on completed years of schooling for women. Alderman, Hoddinott and Kinsey (2006) have shown that early childhood health as measured by height, has lasting effects on the level of schooling completed, among children in rural Zimbabwe. Hoddinott and Kinsey (2001) and Mani (2008) find evidence of path dependence in child heights, in rural Zimbabwe and Indonesia respectively.

Schooling outcomes – such as the decision to continue or withdraw from school, or to enroll having previously not enrolled in school – would seem to be intimately linked to past schooling decisions which themselves were influenced by prior community, school and home resources. The “value-added” specification of human capital accumulation, in which the impact of lagged resources is captured using a cumulative measure of lagged schooling outcome [Hanushek (1979), Boardman and Murnane (1979), Todd and Wolpin (2003, 2007), Andrabi et. al (2009)], has been used in the context of developed countries to explore these and related issues. However, in developing countries this is still not well understood. In the most recent *Handbook of Development Economics*, Orazem and King (2008) write, “Longitudinal analysis of cognitive attainment is needed to establish whether lost human capital from transitory increases in child labor or school absences due to adverse income shocks is reversible or permanent” (p 3550).[1]

This paper contributes evidence on this issue by :- (a) using the value-added specification of human capital accumulation to capture the cumulative impact of past schooling inputs and resources on future schooling outcomes - enrollment status and relative grade attainment, short-run and long-run indicators of schooling; (b) estimating a dynamic conditional schooling demand function that replaces the endogenous schooling inputs with exogenous observables and accounts for the problem of missing school inputs, and (c) drawing on estimation strategies that address the potential correlations between lagged schooling outcome and unobserved endowments. It does using data from rural Ethiopia, a poor African country with low (though rising) levels of grade attainment.

We find that a child who is enrolled in the last period is 32 percentage points more likely to be enrolled today compared to his counterpart who was not enrolled in the last period and that past levels of relative grade attainment affect current levels of this outcome. That is, there is path dependence in schooling outcomes. The path dependence

[1] An exception is Behrman et. al (2005) who use experimental data from Mexico to assess the impact of a Conditional Cash Transfer program, *PROGRESA,* on schooling outcomes using a probability transition matrix which specifies the vector of schooling states for the next age. This methodology allows them to capture the association between an individual's enrollment status in the past period and its effect on current period enrollments but does not account for the impact of socioeconomic factors or child level unobservables that affect a child's complete trajectory of current and future schooling outcomes.

also varies with background characteristics, and is much stronger for girls (69 percentage point differential) than boys (21 percentage points), and for children from high income (81 percentage points) compared to low income households (7 percentage points).

The paper is organized as follows. Section 2 describes the data; section 3 outlines the theoretical model which guides the empirical specification estimated in section 4. The empirical results are discussed in section 5 and concluding remarks follow in section 6.

2. Data

The data used are taken from the 1994, 1999, and 2004 waves of the Ethiopian Rural Household Survey (ERHS). The ERHS is a socioeconomic survey administered in selected rural peasant associations of Ethiopia during 1989-2004.[2] The first wave of the ERHS was fielded in 1989 during which households from 7 farming villages in central and southern Ethiopia were surveyed. In 1989, only a narrow set of questions were administrated. In 1994, 6 of the 7 original villages from 1989 and 9 new villages that account for the diverse farming systems practiced in Ethiopia were additionally selected for survey purposes. A total of 15 rural villages were surveyed in 1994 with the aim of constructing a longitudinal data set. In 1994, two waves of the ERHS were administered, the first wave during January-March and the second during August-October. Households were re-interviewed in 1995, 1997, 1999 and 2004 [see Dercon and Hoddinott (2004) and Dercon et. al (2009) for more details on survey design]. The ERHS provides extensive information on household composition, income, consumption expenditure, farm and non-farm assets, ownership and value of land and livestock units, anthropometrics, harvest

[2] The smallest administrative unit in Ethiopia is called a `peasant association', which is sometimes equivalent to one village or a cluster of villages. We use ``villages'' and ``peasant association'' interchangeably.

use and schooling outcomes. In 1997 and 2004, the survey also collected detailed community level information on infrastructure availability, prices of consumption goods, and wage earnings.

In all survey rounds, data was collected on children's school enrollment status and grade attainment. Using these data, we constructed a child-level longitudinal data set that follows children aged 7 to 14 years (i.e., those of primary school age) in 1994 through the 1999 and 2004 waves of the ERHS. This allows us to avoid complications arising from the irregular spacing of the survey rounds; since the 1994, 1999 and 2004 rounds were fielded in approximately the same months, it also avoids seasonality concerns. As with any longitudinal data set, there are always concerns regarding selective sample attrition. Household level attrition is minimal in the ERHS; only 13% of the sample was lost between 1994 and 2004. This partly reflects the relative immobility of the sample (it is difficult to obtain land if households migrate) and partly a high degree of institutional continuity in the development of these surveys [see Dercon et. al (2006)]. Child level attrition related concerns are addressed in the results section.

3. Conceptual framework

We use a dynamic model of the determinants of schooling outcomes to guide our choice of variables that appear in our empirical model. Households are assumed to maximize an expected lifetime utility function, U (1) subject to a lifetime budget constraint (2) and a period specific dynamic child schooling production function (3).[3]

[3] This approach is similar to dynamic models used in the health literature; see Strauss and Thomas (2008) and Mani (2008).

$$Max: U = E_t \sum_{t=0}^{T} \beta^t u_t(C_t, S_t, L_t; \theta_{pt}) \tag{1}$$

$$s.t: A_T = \prod_{t=0}^{T} (1+r_t)A_0 + \sum_{t=0}^{T} \prod_{\tau=t}^{T} (1+r_\tau)[w_t(T_t - L_t) + \pi_t - p_t^c C_t - p_t^n N_t] \tag{2}$$

$$S_t = f(S_{t-1}, N_t, I_t, \theta_{ct}, \theta_c, \mu_{ht}, \mu_h) \tag{3}$$

We assume that – (a) the household's lifetime utility function is additively separable over time [Fedorov and Sahn (2005); Strauss and Thomas (2008)], (b) the one-period lagged schooling outcome in equation (3) is a sufficient statistic capturing the impact of all lagged schooling inputs, environmental factors, and other time-varying characteristics from birth up until the last observed period in the sample[4], (c) the sub-utility functions are quasi-concave and twice differentiable, (d) the household can borrow and or lend against its future in each period t, and (e) household members have common preferences and pool all resources, that is we assume a unitary household model.[5]

Utility depends upon food and non-food consumption goods, C_t, leisure, L_t, child's schooling outcome, S_t. Schooling outcomes are modeled here as a pure consumption good from which the household derives utility. Household utility is also affected by unobserved preference shocks, θ_{pt}. β is the subjective discount factor. E_t is the expectations operator conditional on the information available at time t. p_t^c is a vector of price of food and non-food consumption goods, p_t^n is a vector of price of schooling

[4] A similar assumption is employed in value-added cognitive achievement production functions, see Todd and Wolpin (2007, 2003) and Hanushek (1979, 2003) and in dynamic health production functions [Strauss and Thomas (1995, 2008), Grossman (1972), Cebu Study Team (1992)].

[5] There exists little empirical validation for the existence of a unitary household model. However, with the data available to us, a collective model of the household would not change the empirical specification we can estimate.

inputs, and w_t, is wage rate. T_t, is parents total time endowment. A_0 is assets the households owns at the beginning of period 0. Profit income from farm and non-farm activities and all other sources of non-labor income is captured by π_t.

S_t is written as a function of lagged schooling outcome and current period schooling inputs, community resources, child characteristics and household characteristics. Schooling inputs N_t include books, school uniform, food intake and other home inputs. Environmental characteristics, I_t capture overall resource availability in the community including measures that capture availability of primary schooling, access to electricity and other community infrastructure. θ_c and θ_{ct} include child specific time-invariant and time-varying characteristics such as child's sex and age capturing age and gender specific differences in the accumulation of schooling outcomes. θ_c and θ_{ct} also include time-varying and time-invariant measures of innate ability that capture overall cognitive development and learning potential. μ_h and μ_{ht} capture household demographic characteristics and other time-invariant and time-varying rearing and caring practices, all of which affect schooling outcomes.

The optimal choice of schooling input N_t^* is written as:

$$N_t^* = f(S_{t-1}, I_t, P_t^c, P_t^n, w_t, \lambda, \pi_t, \theta_{ct}, \theta_c, \mu_{ht}, \mu_h, \theta_{pt}, E_t(Z_{t+j})) \quad \text{for j=1,….,T-t} \qquad (4)$$

Z= $I_t, P_t^c, P_t^n, w_t, \pi_t, \lambda, \theta_{ct}, \theta_c, \mu_{ht}, \mu_h, \theta_{pt}$

N_t^* is a function of the one-period lagged schooling outcome, prices of consumption goods, prices of schooling inputs, wage rates, environmental factors, λ (marginal utility of wealth in period 0), a set of time-varying and time-invariant child level and household level characteristics, and household's expectations at date t about all future period -

prices, environmental characteristics, and household demographics as captured by the term Z.

The dynamic conditional schooling demand function (5) is obtained by replacing N_t in equation (3) by N_t^*:

$$S_t^* = f(S_{t-1}, I_t, P_t^c, P_t^n, w_t, \lambda, \pi_t, \theta_{ct}, \theta_c, \mu_{ht}, \mu_h, \theta_{pt}, E_t(Z_{t+j})) \quad \text{for j=1,….,T-t} \qquad (5)$$

$$\text{and} \quad Z = I_t, P_t^c, P_t^n, w_t, \lambda, \pi_t, \theta_{ct}, \theta_c, \mu_{ht}, \mu_h, \theta_{pt}$$

4. Empirical Specification

The empirical counterpart of the dynamic conditional schooling demand function (5) is:

$$S_{it} = \beta_0 + \beta_1 S_{it-1} + \sum_{j=1}^{R} \beta_j X_{jit} + \sum_{j=1}^{S} \beta_j Z_{ji} + \varepsilon_i + \varepsilon_h + \varepsilon_{vt} + \varepsilon_{it} \qquad (6)$$

S_{it} is enrollment status and relative grade attainment of child i at time t. Enrollment status is defined as a dummy variable which takes a value 1 if the child is enrolled in school at the time of the survey, zero otherwise.[6] Relative grade attainment is defined as actual grades divided by potential grades where potential grades is calculated as total number of grades accumulated had the individual completed one grade of schooling by age 7 and continued to accumulate an additional grade of schooling in each subsequent year. Table 1 provides descriptive statistics for these outcomes as well as the regressors used in the empirical specification of (6).

Xs capture time-varying characteristics and Zs capture time-invariant characteristics. At the individual level, we control for age of the child, male dummy,

[6] Some children are enrolled in religious schools. Our interest is limited to measuring human capital accumulated through learning subjects like mathematics, science and social science; none of which is taught in religious schools. For this reason, we treat children enrolled in religious schools as not enrolled.

mother's age and measure of parental schooling. In the dynamic specification, we use lagged age in years which is specified as a spline variable with age cut-off at 15 years. The spline specification allows us to capture non-linearities in age specific differences in schooling. The male dummy equals one if male, 0 if female capturing gender specific differences in schooling outcomes. Age is interacted with the male dummy to capture age-gender specific differences in schooling. The majority of parents in this region have no formal schooling and so we characterize parental schooling using dummy variables, where the dummy variable takes a value one if the mother (father) has at least one grade of formal schooling, zero otherwise. Mother's age is included in the regressions to capture mother's experience and knowledge.

Household level regressors include number of adult (>18 years) males and number of adult (>18 years) females capturing household demographic composition. Age of the head of the household is included to capture household experience and life-cycle position. These demographic composition variables are specified in lags to avoid potential biases associated with treating household demographic composition as exogenous. Current period demographic composition may be correlated with household specific time-invariant unobservables that are correlated with current and lagged period schooling outcomes.

In all specifications we include village by survey round dummy variables. This controls for all time varying shocks and changes in prices and environmental factors, both negative (drought) and positive (improvements in infrastructure), at the village level. In our preferred first-difference specification, this also allows for village specific time trends.

The dynamic specification does not include an explicit measure of household income, except for λ, the marginal utility of wealth at time zero. As λ is time-invariant, it is first-differenced out of our preferred econometric specification. Recall, however, that we assume that households can freely borrow and lend in each period. This is not true in rural areas of Ethiopia where formal and informal credit markets are badly underdeveloped. To ensure that our results are not sensitive to treating λ as time-invariant, we capture borrowing constraints by including a lagged measure of log of household's real per capita consumption expenditure as an additional explanatory variable.

There are four unobservables in equation (6), ε_i, ε_h, ε_{vt} and ε_{it}; ε_i captures individual specific time-invariant unobservables such as child's innate ability to perform well in school; ε_{vt} captures village specific time-varying unobservables such as prices of schooling inputs and home inputs ; ε_h captures household specific time-invariant unobservables such as parental preferences towards schooling and their time preferences; and ε_{it} is a random time-varying unobservable that are unknown to both the individual and the econometrician at date t. We assume that all factors that enter the dynamic schooling demand function through the expectations term $E_t(Z_{t+j})$ are unknown to the econometrician at date t and captured in empirical specification through the time-varying error term (ε_{it}). Note that an OLS estimate of β_1 is likely to be biased due to the presence of time-invariant unobservables such as child's innate ability to perform well in school, parental preferences towards schooling and community's political connections; all of which are likely to be correlated with the lagged schooling outcome, S_{it-1} [Deaton (1997); Blundell and Bond (1998); Wooldridge (2002)]. Given these unobservables, our estimation strategy must be sensitive to violations of the assumption of zero correlation between the lagged dependent variable, (S_{it-1}) and the error term.

5. Results

Dynamic regression results

Results of the dynamic enrollment regressions are reported in Table 2 and results for the dynamic relative grade attainment regressions are reported in Table 3. In addition to the econometric concerns noted above, in both sets of specifications we need to address concerns regarding the exogeneity of lagged per capita consumption. It too may be correlated with household specific unobservables such as preferences and discount rates and is also vulnerable to concerns regarding random measurement error. Given this, in addition to estimating an OLS version of equation (6), we also use two IV strategies. The first is an Arellano-Bond (1991) type estimator where first-differenced lagged schooling enrollment (relative grade attainment) and first-differenced lagged PCE are instrumented using twice-lagged schooling enrollment (relative grade attainment) and twice-lagged per capita consumption. Second, we estimate an Arellano-Bond model where only the first-differenced lagged schooling enrollment (relative grade attainment) is instrumented using two-period lagged enrollment (relative grade attainment) and first-differenced lagged PCE is treated as exogenous. Both sets of IV estimates assume zero first- and second-order serial correlation in the error terms and no measurement error in the school enrollment or relative grade attainment variables.[7]

[7] We do not have enough rounds of data to test the validity of these assumptions.

The OLS estimate of β_1 is 0.34 (table 2, column 1) and is significant at the 1 percent level indicating strong positive association between lagged enrollment and current enrollment. A similar estimate is obtained when we treat lagged enrollment and consumption as endogenous (column 2) or only when lagged enrollment is treated as endogenous (column 3). As a C statistic test does not reject the null that first difference lagged log consumption is exogenous, column (3) represents our preferred specification. These show that if a child was enrolled in the previous period, (s)he is 32 percentage points more likely to be enrolled today relative to a child not previously enrolled.

Table 3 reports regression results of a dynamic conditional schooling demand function for relative grade attainment. The first three columns are comparable to those used for enrollment; an OLS specification and two variants of Arellano-Bond. The OLS estimate of β_1 is 0.62 (column 1) indicating a strong positive association between lagged and current relative grade attainment. The magnitude of β_1, however, falls by more than half when lagged relative grade attainment is treated as endogenous (columns 2 and 3). However, the Arellano-Bond estimation strategies followed here assume that there is zero first-order and second-order serial correlation in the error terms. To relax this assumption, we use a variant of a first-differenced Generalized Methods of Moments (FD-GMM) estimator where first-differenced lagged relative grade attainment is instrumented using twice lagged *enrollment* (not relative grade attainment) as an instrument. This estimation strategy addresses the measurement error bias in relative grade attainment as it allows for random measurement error in lagged enrollment and relative grade attainment, but assumes that the two sources of measurement error are independent. The FD-GMM estimator reported in column 4, table 3 yields an unbiased and consistent estimate on lagged RGA without relying on the assumption of lack of serial correlation in the error terms. Column (4) shows that the FD-GMM estimate of β_1 is 0.31 and statistically significant. Current grade progression depends on past grade progression.

Robustness

The regression results are robust to issues of instrument validity and sample attrition. In the presence of weak correlation between the endogenous regressor and the instruments, the IV estimates will suffer from higher inconsistency and bias compared to the OLS estimate [Murray (2006)]. To test for the presence of weak instruments, we use the Kleibergen-Paap Wald rk F statistic, which is robust to the presence of heteroskedasticity, autocorrelation and clustering [Kleibergen and Paap, 2006]. In the presence of a single endogenous regressor, the Kleibergen-Paap test statistic reduces to the usual F statistic on the excluded instruments.. The F statistic on the excluded instruments reported in our preferred IV regressions is almost always above 10, satisfying the Staiger and Stock (2003) rule of thumb rejecting the null of weak correlation between the instruments and the endogenous regressor. [8]

The coefficient estimate on β_1 is robust to concerns regarding sample attrition. If sample attrition was related to the outcome variable of interest either through observables or through unobservables, then the coefficient estimate on twice-lagged schooling outcome would suffer from attrition bias [Fitzgerald et. al (1998)]. We have 2047 observations on primary school age children in 1994, of which 809 could be followed through the 1999 and 2004 waves of the ERHS. The panel sample has an annual rate of

[8] The Staiger and Stock (2003) rule of thumb is approximately a 5% significance test that the worst relative (IV to OLS) bias would be 10% or less [see table 1, p 39 (Staiger and Stock, 2003)].

attrition at 6%. At first, this might seem large; however, most of this attrition is age related. Given that some our children are of age 12, 13 or 14 when first observed, it is only natural for them to have left their natal household by 2004, when they are in their early 20s. Much of the attrition in our sample is associated with demographic changes and is common among other longitudinal panel data set such as the IFLS. It is the presence of time-invariant unobservables such as child's innate ability that is likely to affect both the decision to migrate and the endogenous covariate, schooling outcome. The preferred FD-GMM estimation strategy used here removes all sources of time-invariant unobservables addressing this potential source of attrition bias.

To determine the extent to which endogenous observables create attrition bias, we also estimate a linear probability model of sample attrition, where the dependent variable is defined as attrition which takes a value 1 if the primary school aged child can be followed through the 1994, 1999, and 2004 waves of the ERHS and 0 otherwise. The regression results on sample attrition are reported in Appendix Table A2. In column 1, attrition is regressed upon enrollment status from 1994 and baseline characteristics from 1994 which include measure of household income, age, mother's schooling, father's schooling, and household composition variables. The regression results reported in column 1(table A2) indicates that sample attrition is negatively associated with age; older children are much less likely to be followed over time compared to younger children. This is consistent with migration patterns in the region [Ezra and Kiros (2001), Fafchamps and Quisumbing (2005)]. The attrition regression results outlined in column 1, table A2 indicate that attrition is unrelated to the endogenous observable, enrollment status. Hence, our preferred estimates of the dynamic enrollment regression are not likely to be confounded by attrition bias.

We also estimate a linear probability model of sample attrition using relative grade attainment. In column 2 of table A2, attrition is regressed upon RGA from 1994, and other covariates as controlled in the attrition regression for enrollment status. The regression results reported in column 2, table A2 indicate that sample attrition is negatively associated with relative grade attainment and is statistically significant that is attrition is related to endogenous factors such as household income and schooling attainment. The potential correlation between attrition and these endogenous covariates is addressed by our preferred FD-GMM estimator where two-period lagged schooling enrollment is used as an instrument, and is uncorrelated with attrition and hence provides us with an unbiased estimate on lagged RGA.

The empirical results on lagged enrollment and lagged RGA are also robust to treating λ as a constant. We estimate our preferred specifications for enrollment (using Arellano-Bond) and relative grade attainment (using FD-GMM) without controlling for two-period lagged per capita expenditures where λ is treated as a constant would get first-differenced from both specifications. We find that our estimates on the lagged dependent variable are not statistically significantly different from the one's reported in column 3 (table 2) for enrollment and column 4 (table 3) for relative grade attainment. Hence, the parameter estimate on the twice-lagged schooling outcome variable is robust to treating lambda as a constant.

Disaggregations

The degree of path dependence in these schooling outcomes may differ by economic or demographic group. We explore such differences here. Table 4 shows that boys and girls have very different degrees of path dependence in school enrollment, with path dependence being much stronger for girls than boys (0.69 compared to 0.21, significant at 1% and 10% respectively). This may surprise some, since if there were strong boy-preference, one might think that if both boys and girls were in school, it is more likely that boys would stay in school. This is not so during this period in these rural Ethiopian villages.

There is also a distinct difference by initial age of the child, with children under 11 years in 1994 having a higher path dependent coefficient than older children (0.41 compared to 0.25, both significant at 5%). This makes sense if the younger children are more likely to stay in primary school over this period instead of potentially graduating to the next level, at which dropout rates are high.

Finally, children from households with higher per capita expenditures have a much higher degree of path dependence than children from poorer households (0.81 compared to 0.07 and not significant even at 10%). Apparently, coming from a higher income household means that if the child is in school, he or she will likely stay, while children from lower income households are more susceptible to period specific shocks.

These results for current enrollments are replicated for our measure of relative grade attainment, as shown in Table 4. Now the percentage point differences are not quite as large, but they still exist and in the same directions as for enrollments.

6. Conclusion

This paper examines the cumulative impact of early schooling investments on later schooling outcomes. We estimate a dynamic conditional schooling demand function where the coefficient estimate on the lagged dependent variable captures the impact of all previous periods' schooling inputs and resources. The dynamic specification is estimated using longitudinal data on primary school children in rural Ethiopia between 7 and 14 years in 1994 followed through 1999 and 2004. We use two measures of schooling outcomes – enrollment status (short-run measure of schooling) and relative grade attainment (long-run measure of schooling attainment) to give a comprehensive view of the impact of investments in early schooling resources on final attainments. The preferred first-difference GMM estimation strategy used here addresses concerns regarding omitted variable bias for enrollment status, and also measurement error bias for relative grade attainment.

Our results indicate that the history of schooling inputs and resources have a strong impact on individual's later schooling outcomes. We find that a child who is enrolled in the last period is 32 percentage points more likely to be enrolled today compared to his counterpart who was not enrolled in the last period. We obtain similar findings using relative grade attainment (RGA); grade progression today is affected by grade progression in the past. Any lags and delays that affected progression in the past will have a permanent impact on final grades accumulated. These results differ by groups of children, being stronger for girls, younger children and children from higher income households.

References

Alderman, H., J. Hoddinott, and B. Kinsey. 2006. Long-term consequences of early childhood malnutrition, Oxford Economic Papers, 58(3): 450-474.

Andrabi, T., J. Das, A.I. Khwaja, and T. Zajonc. 2009. Here today, gone tomorrow? Examining the extent and implications of low persistence in child learning. Working paper series, Pomona College.

Arellano B. and S.R. Bond. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application of employment equations, Review of Economic Studies 58, pp 277-297.

Blundell, R. and S. Bond. 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics, 87: 115-143.

Boardman, A.E. and R.J. Murnane. 1979. Using Panel Data to Improve Estimates of the Determinants of Educational Achievement. Sociology of Education, 52(2):113–121.

Behrman, J. R., Sengupta, P., and Todd, P. 2005. Progressing through PROGRESSA: An Impact Assessment of a School Subsidy Experiment in Rural Mexico. Economic Development and Cultural Change, 54: 237-275.

Cebu Study Team. 1992. A child health production function estimated from longitudinal data. Journal of Development Economics, 38(2):323-351.

Deaton, A. 1997. The analysis of household surveys: A microeconometric approach in development policy. Baltimore: Johns Hopkins University press.

Dercon, S. and J. Hoddinott. 2004. The Ethiopian Rural Household Survey: Introduction. mimeo, Washington D.C: International Food Policy Research Institute.

Dercon, S. 2005. Risk, Poverty and Vulnerability in Africa, Journal of African Economies, vol. 14(4): 483-488.

Dercon, S., J. Hoddinott and T. Woldehanna. 2006. Growth, poverty and chronic poverty in rural Ethiopia: Evidence from 15 Communities 1994-2004. Background paper for The Chronic Poverty Report 2008-09, Manchester, UK: Chronic Poverty Research Centre (CPRC).

Dercon, S., D.O. Gilligan, J. Hoddinott and T. Woldehanna. 2009 The Impact of Roads and Agricultural Extension on Consumption Growth and Poverty in Fifteen Ethiopian Villages, American Journal of Agricultural Economics, vol. 91(4), pp 1007-1021

Ezra, M. and G.E. Kiros. 2001. Rural Out-Migration in the Drought Prone Areas of Ethiopia: A Multilevel Analysis. International Migration Review, 35(3): 749-771.

Fafchamps, M. and A. Quisumbing. 2005. Marriage, Bequest and Assertive Matching in Rural Ethiopia. Economic Development and Cultural Change, 53: 347-380.

Fedorov L. and D. Sahn. 2005. Socioeconomic determinants of children's health in Russia: A longitudinal survey, Economic Development and Cultural Change, pp 479-500.

Fitzgerald, J., P. Gottschalk, and R. Moffitt. 1998. An analysis of sample attrition in panel data. Journal of Human Resources, 33(2): 251-299.

Grossman, M. 1972. On the Concept of Health Capital and the Demand for Health. Journal of Political Economy, 80: 223-55.

Hanushek, E.A. 1979. Conceptual and Empirical Issues in the Estimation of Educational Production Functions. The Journal of Human Resources, 14(3):351–388.

Hanushek, E.A. 2003. The Failure of Input-Based Schooling Policies. Economic Journal, 113(485):64–98.

Hoddinott J. and B. Kinsey. 2001. Child growth in the time of drought, Oxford Bulletin of Economics and Statistics, vol 63 (4): 409-436.

Kleibergen, F. and R. Paap. 2006. Generalized reduced rank tests using the singular value decomposition. Journal of Econometrics, 133 (1): 97-126.

Maccini, S. and D. Yang, 2009. Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall, American Economic Review, vol. 99(3): 1006-26.

Mani, S. 2008. Is there Complete, Partial, or No Recovery from Childhood Malnutrition – Empirical Evidence from Indonesia. Fordham discussion series, Fordham University.

Murray, M. P. 2006. Avoiding invalid instruments and coping with weak instruments. Journal of Economic Perspectives, 20(4): 111-132.

Orazem, P.F. and E.M. King. 2008. Schooling in Developing Countries: The Roles of Supply, Demand and Government Policy. Handbook of Development Economics, Volume 4, edited by T. P. Schultz and J. Strauss.

Staiger, D. J.H. Stock. 2003. Testing for Weak Instruments in Linear IV Regression, working paper series, Harvard University.

Strauss, J. and D. Thomas. 1995. Human Resources: Empirical Modeling of Household and Family Decisions. Handbook of Development Economics, vol. 3, edited by Jere R. Behrman and T.N. Srinivasan.

Strauss, J. and D. Thomas. 2008. Health Over the Life Course, Handbook of Development Economics, vol. 4, edited by T. Paul Schultz and John Strauss eds., Amsterdam: North Holland Press.

Todd, P.E. and K.I. Wolpin. 2003. On the Specification and Estimation of the Production Function for Cognitive Achievement. Economic Journal, 113(485):3–33.

Todd, P.E. and K.I. Wolpin. 2007. The production of cognitive achievement in children: home, school, and racial test score gaps. Journal of Human Capital, 113(485):3–33.

Wooldridge, J. 2002. Econometric Analysis of Cross Section and Panel Data, Cambridge: MIT Press.

Table 1: Descriptive statistics

Variable	Mean Std. dev
Enrollment, Enrollment=1 if currently enrolled in school and 0 otherwise	0.38 (0.48)
Completed grades of schooling	2.26 (2.80)
Relative grade attainment (actual grade/potential grade given age)	0.23 (0.28)
Household size	7.65 (2.68)
Log real per capita household consumption expenditure (PCE)	3.96 (0.76)
Mother's schooling	0.08 (0.27)
Father's schooling	0.25 (0.41)
Male dummy	0.58 (0.49)
Age (years)	15.09 (4.55)
No. of adult males	1.78 (1.12)
No. of adult females	1.78 (1.01)
Mother's age	42.47 (10.15)
Age of the head of the household	51.28 (12.02)

Table 2: Determinants of school enrollment

Covariates	**(1) OLS**	**(2) Arellano-Bond**	**(3) Arellano-Bond**
Lagged enrollment	0.34*** (0.02)	0.32*** (0.09)	0.32*** (0.09)
Lagged log real per capita consumption	0.0029 (0.01)	-0.03 (0.04)	-0.051** (0.02)
Male dummy	0.03 (0.11)		
Lag age in years (in spline): <15 years	-0.024*** (0.008)	-0.03** (0.01)	-0.03** (0.01)
Lag age in years (in spline): >=15 years	-0.034** (0.01)	-0.03** (0.01)	-0.03* (0.01)
Lag age in years (in spline): <15 years *x* Male dummy	0.002 (0.009)	0.0005 (0.01)	0.0006 (0.01)
Lag age in years (in spline): >=15 years *x* Male dummy	-0.008 (0.02)	-0.01 (0.02)	-0.012 (0.02)
Mother's schooling	0.093** (0.04)		
Father's schooling	0.062** (0.03)		
Number of adult males, lagged	0.005 (0.01)	-0.010 (0.02)	-0.011 (0.02)
Number of adult females, lagged	0.030** (0.01)	0.016 (0.02)	0.014 (0.02)
Mother's age	-0.0024* (0.001)		
Age of household head, lagged	-0.0012 (0.001)	-0.004 (0.003)	-0.004 (0.003)
Village *x* survey round dummy variables included	Yes	Yes	Yes
Test of endogeneity of first-differenced lagged per capita consumption		0.25 (0.61)	
Kleibergen-Paap F statistic		111.38*** (0.00)	219.96*** (0.00)

Notes: Robust standard errors in parentheses; * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level. Sample size is 1618 (column 1) and 809 (columns 2 and 3). In column (2), first differenced lagged per capita consumption and first-differenced lagged enrollment are treated as endogenous. In column (3), only first-differenced lagged enrollment is treated as endogenous.

Table 3: Determinants of relative grade attainment

Covariates	(1) OLS	(2) Arellano-Bond	(3) Arellano-Bond	(4) First-difference GMM
Lagged relative grade attainment	0.62*** (0.02)	0.25*** (0.06)	0.25*** (0.06)	0.31*** (0.06)
Lagged log real per capita consumption	0.022*** (0.007)	-0.003 (0.01)	0.005 (0.04)	0.005 (0.007)
Male dummy	0.010 (0.05)			
Lag age in years (in spline): <15 years	-0.009** (0.004)	0.0008 (0.004)	0.00004 (0.004)	-0.001 (0.005)
Lag age in years (in spline): >=15 years	0.014* (0.007)	0.02 (0.006)	0.019 (0.006)	0.019 (0.006)
Lag age in years (in spline): <15 years *x* Male dummy	0.0025 (0.004)	0.002 (0.003)	0.001 (0.003)	0.001 (0.003)
Lag age in years (in spline): >=15 years *x* Male dummy	0.0003 (0.009)	0.001 (0.007)	0.0009 (0.007)	0.00003 (0.008)
Mother's schooling	0.035* (0.01)			
Father's schooling	0.012 (0.01)			
Number of adult males, lagged	0.011** (0.004)	0.011 (0.006)	0.011* (0.006)	0.011* (0.007)
Number of adult females, lagged	0.009* (0.005)	0.015* (0.008)	0.015* (0.008)	0.016* (0.008)
Mother's age	-0.0002 (0.0006)			
Lagged age of household head	-0.0009* (0.0005)	0.001 (0.0009)	0.0009 (0.0009)	0.0008 (0.0009)
Village *x* survey round dummy variables included	Yes	Yes	Yes	Yes
Test of endogeneity of first-differenced lagged per capita consumption		0.72 (0.39)		
Kleibergen-Paap F statistic		144.93 (0.00)	280.51 (0.00)	20.42

Notes: Robust standard errors in parentheses; * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level. Sample size is 1618 (column 1) and 809 (columns 2, 3 and 4). In column (2), first differenced lagged per capita consumption and first-differenced lagged enrollment are treated as endogenous. In columns (3) and (4), only first-differenced lagged relative grade attainment is treated as endogenous.

Table 4: Determinants of enrollment and relative grade attainment by selected disaggregations

	Enrollment		Relative Grade attainment	
Disaggregation	Parameter estimate (standard error)	Sample size	Parameter estimate (standard error)	Sample size
Children 11 years and older, 1994	0.25** (0.13)	281		
Children less than 11, 1994	0.41** (0.16)	528		
Children 12 years and older, 1994			0.13 (0.09)	201
Children less than 12, 1994			0.32*** (0.10)	608
Boys	0.21* (0.11)	474	0.29* (0.17)	474
Girls	0.69*** (0.21)	335	0.39** (0.19)	335
Poor households	0.07 (0.10)	525	0.23* (0.14)	525
Less poor households	0.81*** (0.22)	284	0.40** (0.20)	284
Mother has some schooling	0.21 (0.19)	63	0.26* (0.13)	63
Mother has no schooling	0.35*** (0.11)	746	0.34** (0.13)	746
Father has some schooling	0.17 (0.13)	184	0.38** (0.15)	184
Father has no schooling	0.38*** (0.14)	625	0.17** (0.07)	625

Notes: Robust standard errors in parentheses; * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level. Lowess plots of enrollment and two-period lagged age were used to determine the cut-off for stratifying the sample by age. Poor households have log per capita consumption below 4 in 1994; less poor households have consumption levels above this cut-off. Lowess plots were used between enrollment and two-period lagged PCE to determine the cut-off point at which the sample should be stratified. Lowess plots and full regression results are available on request.

Appendix

Table A1: First-stage regressions for our preferred estimates reported in column 3, table 2 and column 4, table 3

Covariates	(1) Column 3 table 2	(2) Column 4 table 3
Lagged enrollment	-0.685***	-0.149***
	(0.04)	(0.03)
Lagged log real pce (in first-differences)	0.04**	-0.0008
	(0.01)	(0.01)
Lag age in years (in spline): <15 years (in first-differences)	0.09***	0.03***
	(0.009)	(0.005)
Lag age in years (in spline): >=15 years (in first-differences)	0.04***	0.01
	(0.01)	(0.008)
Lag age in years (in spline): <15 years *x* Male dummy (in first-differences)	0.01	0.004
	(0.008)	(0.004)
Lag age in years (in spline): >=15 years *x* Male dummy (in first-differences)	0.02	0.017
	(0.02)	(0.01)
no. of adult males, lagged (in first-differences)	0.017	-0.005
	(0.01)	(0.08)
no. of adult females, lagged (in first-differences)	0.016	-0.008
	(0.02)	(0.01)
age of the head of the household, lagged (in first-differences)	0.0005	0.001
	(0.02)	(0.001)

Notes: Robust standard errors in parentheses; * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level. Sample size is 809 (columns 1 and 2).

Table A2: Determinants of sample attrition for enrollment status and relative grade attainment

Covariates	(1)	(2)
Enrollment	0.03 (0.03)	
Relative grade attainment		0.07*** (0.02)
Log of real per capita consumption	0.03 (0.01)	0.03** (0.01)
Mother's schooling	0.04 (0.04)	0.05 (0.04)
Father's schooling	0.03 (0.02)	0.02 (0.02)
Male dummy	0.01 (0.05)	0.02 (0.05)
dummy =1 if the child is 8 yrs	-0.11** (0.05)	-0.10* (0.05)
dummy =1 if the child is 9 yrs	-0.16*** (0.05)	-0.15*** (0.05)
dummy =1 if the child is 10 yrs	-0.12** (0.05)	-0.12** (0.05)
dummy =1 if the child is 11 yrs	-0.25*** (0.06)	-0.25*** (0.05)
dummy =1 if the child is 12 yrs	-0.29*** (0.05)	-0.28*** (0.05)
dummy =1 if the child is 13 yrs	-0.37*** (0.05)	-0.36*** (0.05)
dummy =1 if the child is 14 yrs	-0.37*** (0.05)	-0.36*** (0.05)
No. of adult males	0.01 (0.009)	0.01 (0.09)
Mother's age	0.0006 (0.001)	0.0006 (0.001)
No. of adult females	0.012 (0.01)	0.013 (0.01)
Age of the head of the household	0.01 (0.0008)	0.001 (0.0008)

Notes: Robust standard errors in parentheses; * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level. Sample size is 809 (columns 1 and 2). Age interacted gender dummies are suppressed. Attrition takes the value 1 if the individual was followed during the subsequent waves and 0 otherwise.