PN-ACA-317

Scott Grosse
June 7, 1995

Maternal education and child heights in rural Rwanda

The association of maternal education with child growth in sub-Saharan Africa bears closer examination. Elsewhere, maternal education is one of the most powerful predictors of child health status, including child heights. In sub-Saharan Africa, on the other hand, most statistical studies using standardized child heights as dependent variables have failed to find a significant association with maternal schooling or literacy. This includes studies previously carried out in Rwanda (Csete, 1989; Schnepf, 1992; Ogden, 1993). Elsewhere, we have speculated as to why the association may be weaker in African societies than elsewhere.

Our goal in this paper is twofold. First, we want to clarify whether a significant relationship between maternal schooling and child growth status in Rwanda has emerged in recent years. Second, to the extent that a consistent association is found, we would like to be able to decompose the effect into different pathways of influence. In particular, we would like to know to what extent education has an effect on efficiency in production of child health independently of economic conditions and availability of material inputs. This in turn can in principle be decomposed into effects on income and on expenditure decisions. Controlling for income level, educated mothers may be influential in allocating more resources to health-related inputs. Furthermore, it would be of interest to know to what extent increased efficiency in household production, to the extent that it is observed, operates chiefly through household hygiene and feeding practices or through use of preventive and curative health services.

We have the advantage of two data sets to analyze the effect of maternal schooling on child growth status in rural Rwanda. The National Nutrition and Food Security Survey conducted by the Division des Statistiques Agricoles (DSA) of the Ministry of Agriculture is a sample of roughly 1300 farm households with children under 60 months of age conducted during 1991/92. This sample was derived from a cohort of 2500 farm households identified and followed up since 1988. The Demographic and Health Survey (DHS) was conducted during 1992 on a rural sample of 5400 households and an urban sample of 1100 households, not all of which had young children present. In the following discussion, we restrict our focus to children ages 24-59 months, roughly 1100 in number in the DSA data and 2000 in number in the DHS data.

DSA sample

We begin with the DSA data set. The cohort of farm households identified in 1988 for follow-up was divided into two halves, one half to be followed on a regular basis, the "intensive" sample, and one half reserved for occasional use, the "extensive" sample. Agronomic and economic surveys carried out by DSA were administered only to the intensive sample. The division between intensive and extensive samples was ostensibly made on a random basis. To complicate matters, the average standardized child height in the 1991/92 nutrition survey is significantly lower in the intensive sample than in the extensive sample.

For the whole DSA sample, a curve-fitting exercise revealed a marked nonlinearity, with essentially no variation in HAZ up to 6 years of schooling and a strongly negative association beyond 6 years. I created two variables, one the number of years of schooling up to 6 years, which is currently the level of completion of primary schooling in Rwanda, and another for years of schooling beyond 6 years. I refer to the latter as years of postprimary schooling, although this is not strictly correct. In a regression equation with child age in months as well as years of primary and postprimary schooling, we find that the coefficient on the latter is highly significant, and the former is quite insignificant, supporting the initial observation of nonlinearity.

Regression sample:	Whole DSA sample		
Variable	Regression coefficient	t-statistic	Probability level
Age in months	-.019380	5.452	.0000
Years of schooling up to 6	.010486	0.642	.5207
Years of schooling after 6	.048259	2.377	.0176
Adjusted $R^2 = .028$	df = 1119		

As a next step, I control for maternal height, since previous studies have found that controlling for this variable reduces the coefficient of maternal schooling by anywhere from one quarter to one half, because taller women tend to have more years of schooling and also to have taller children. As expected, the coefficient on maternal schooling (post-primary) was reduced by roughly one quarter in magnitude, with a greatly reduced level of statistical significance.

3

Regression sample:	Whole DSA sample		
Variable	Regression coefficient	t-statistic	Probability level
Age in months	-.017597	5.033	.0000
Years of schooling up to 6	.007770	0.489	.6248
Years of schooling after 6	.035505	1.789	.0739
Maternal height (in cm)	.054749	9.166	.0000
Adjusted R^2 = .100	df = 1072		

The next step was to control for economic status, which combines both confounding and maternal income as a pathway of influence. At a first level of analysis, I included only measures of economic status from the nutrition survey available for the whole sample, including housing quality and ownership of farm animals. Inclusion of these variables reduced the coefficient on postprimary schooling still further and made it quite insignificant (results not reported). Subsequently, I decided to include economic variables from the agricultural surveys, available only for the intensive sample. To simplify, I report here the results for the inclusion of only one such variable, per capita family income. I was surprised to find that the association with postprimary schooling became weaker and entirely insignificant, while the association with primary schooling conversely became much stronger and significant.

Regression sample:	Intensive DSA sample		
Variable	Regression coefficient	t-statistic	Probability level
Age in months	-.020866	4.086	.0001
Years of schooling up to 6	.059405	2.477	.0136
Years of schooling after 6	.029106	0.851	.3953
Maternal height (in cm)	.058437	6.538	.0000
Logarithm per capita income	.230454	2.795	.0054
Adjusted R^2 = .140	df = 501		

The obvious next step was to investigate to what extent sample composition was having this unexpected effect on the pattern of coefficients for maternal schooling. Using the split file feature in SPSS, I ran the same models for the extensive and intensive samples, necessarily excluding the income variable. I found quite different patterns of association between the two subsamples. In both samples the positive effect of postprimary schooling was insignificant, while the effect of primary schooling was statistically significant at the .05 level. The discrepancy is that the significant effect was positive in one sample and negative in the other sample. Six years of

primary schooling is found to raise HAZ by 0.43 SD in the intensive sample compared to no schooling and to reduce it by 0.27 SD in the extensive sample.

Regression sample:	Intensive DSA sample		
Variable	Regression coefficient	t-statistic	Probability level
Age in months	-.020426	3.980	.0001
Years of schooling up to 6	.071012	3.014	.0027
Years of schooling after 6	.053975	1.624	.1051
Maternal height (in cm)	.062629	7.081	.0000
Adjusted R^2 = .127	df = 503		

Regression sample:	Extensive DSA sample		
Variable	Regression coefficient	t-statistic	Probability level
Age in months	-.015268	3.242	.0013
Years of schooling up to 6	-.046716	2.186	.0292
Years of schooling after 6	.016095	0.657	.5112
Maternal height (in cm)	.049398	6.115	.0000
Adjusted R^2 = .086	df = 568		

The situation gets even more complicated. I ran the same equations for the two samples without the maternal height variable. I was surprised to find that controlling for maternal height increased the absolute magnitude and significance level of the coefficient on primary years of schooling in both samples. This implies that in the intensive sample uneducated women are taller than women with primary schooling, which is surprising and contrary to what has been reported in the literature. In the extensive sample uneducated women appear to be shorter and controlling for maternal height makes the effect of schooling less positive and more negative. The effect of inclusion of maternal height on the postprimary schooling coefficient was less consistent. In the extensive sample, the coefficient was reduced in magnitude by more than 50%, while in the intensive sample the coefficient was reduced by only 5%, which is trivial. This indicates that the positive correlation between maternal stature and postprimary schooling is almost entirely restricted to the extensive sample.

Regression sample:	Intensive DSA sample		
Variable	Regression coefficient	t-statistic	Probability level
Age in months	-.022677	4.270	.0000
Years of schooling up to 6	.058236	2.381	.0176
Years of schooling after 6	.056954	1.693	.0911
Adjusted R^2 = .041	df = 519		

Regression sample:	Extensive DSA sample		
Variable	Regression coefficient	t-statistic	Probability level
Age in months	-.016692	3.529	.0004
Years of schooling up to 6	-.035349	1.492	.1363
Years of schooling after 6	.035252	1.402	.1616
Adjusted R^2 = .024	df = 599		

To check whether a similar disparity in results would occur with randomly selected subsamples, I randomly split the sample into two halves and reran the analysis on both groups. I found a a significant coefficient on postprimary education in one subsample but not in the other. However, since the difference in coefficients was not statistically significant, this does not appear to be important. Based on this comparison, one would not reject the hypothesis that the same model fits both subsamples.

The effect of education may be contingent upon contextual variables. For example, there may be a positive synergism between the education of the household head and that of the child's mother, or in economic terms they may be complements in household production. To examine this, we examined for interactions within the intensive and extensive samples. In the intensive sample, the effects of maternal literacy and literacy of the household head are additive and highly significant. In the extensive sample, both effects are insignificant, whether additive or interactive. An interaction appears, such that the tallest children have illiterate parents, while the second-tallest group are those with both parents literate, but there does not seem to be any logical way to account for this pattern of observations within the extensive sample.

The disparity in results between the intensive and extensive samples may derive from a number of factors. One possibility is that the division between samples was not entirely random. The fact that both the mothers and children in the intensive sample are significantly shorter than those in the extensive sample suggests that there may have been a systematic factor by which certain households were excluded from the intensive sample. Since ethnicity was not recorded in

the DSA survey, we cannot know whether Tutsis were more likely to be excluded from the intensive sample. In terms of socioeconomic status, the two samples do not differ significantly, so there was no evident bias towards the exclusion or inclusion of more affluent households.

Another, more hopeful explanation is that the maternal education data are more reliably collected from the intensive sample than from the extensive sample. The interviewers had been following the households in the intensive sample on a frequent basis, and were well-known to the respondents. The women in the extensive sample households had never met the interviewers and may have been less inclined to give accurate responses to personal questions such as level of schooling and ability to read.

Finally, we decided to run some analyses for the intensive sample using a community fixed effects model in which each survey cluster save one is represented by a dummy variable. This has the effect of purging the variables of their association with community characteristics. If more educated mothers live in better-off communities on average, part of the apparent effect of schooling may reflect the effect of community residence. The coefficient on maternal schooling is 0.032 when only age is controlled in addition to cluster dummies ($p = .1621$). Controlling for community effects in a model with just age and years of schooling reduces the magnitude of the coefficient on years of schooling by roughly half compared to random effects model and greatly reduces its level of statistical significance. The coefficient is increased to 0.045 when maternal height is added to the regression ($p = .0455$). This type of result has never been reported in the literature and is counterintuitive. Finally, the coefficient is reduced to 0.036 when the logarithm of per capita income is included as well, and is no longer significant ($p = .1203$).

DHS sample

The next phase of the examination of the robustness of the association of maternal schooling with child heights in rural Rwanda involves the DHS survey data, which lacks maternal anthropometric measurements. The analysis of the entire sample revealed an association which was similar to that found within the intensive DSA sample not controlling for maternal height, namely an apparently linear association which is highly significant. The magnitude of the coefficient on years of schooling is somewhat lower, 0.40 as opposed to 0.57 in the intensive DSA sample based upon random effects estimation, although more congruent with the 0.32 estimate from a community fixed effects model.

7

Regression sample:	DHS sample		
Variable	Regression coefficient	t-statistic	Probability level
Age in months	-.017276	6.500	.0000
Years of schooling up to 6	.040560	3.093	.0020
Years of schooling after 6	.039415	2.226	.0261
Adjusted R^2 = .026	df = 2026		

While we are not able to control directly for maternal stature in the DHS data, we can control for maternal ethnicity, which is associated with adult stature. Of the three ethnic groups in Rwanda, the pygmoid Twa are by far the shortest and the Tutsis are known to be generally taller. When we include ethnic dummies as controls, we find that the coefficients on both primary and primary years of schooling decrease in magnitude by less than 10% and remain significant. Because of an apparent linearity, we replace the two schooling variables by a single years of schooling term.

Regression sample:	DHS sample		
Variable	Regression coefficient	t-statistic	Probability level
Age in months	-.017379	6.539	.0000
Years of schooling	.037442	3.297	.0010
Maternal ethnicity -- Tutsi	.274591	2.360	.0183
-- Twa	-2.099680	1.692	.0907
Adjusted R^2 = .029	df = 2021		

To control for confounding by socioeconomic status, we include several variables which are unlikely to be causally influenced by maternal schooling. One of these is ownership of a bicycle, which is generally purchased and used primarily by men. Another is housing quality, defined as an ordinal variable categorized by whether the house has 3 or more sleeping rooms or a cement floor. The two others are dummy variables indicating that the male partner is employed in commerce or a civil service job ("good job"), both of which pay well in Rwanda compared to farming or other rural occupations, including unskilled labor and artisanry ("bad job"). Inclusion of these four socioeconomic variables reduces the coefficient on maternal schooling by roughly one third. The coefficient on Tutsi ethnicity is also reduced, which is not surprising since Tutsis in Rwanda were substantially more affluent, as well as better educated than the majority Hutus.

Regression sample:	DHS sample		
Variable	Regression coefficient	t-statistic	Probability level
Age in months	-.017429	6.571	.0000
Years of schooling	.025727	2.252	.0245
Maternal ethnicity -- Tutsi	.289131	2.499	.0125
-- Twa	-2.007775	1.641	.1009
Ownership of bicycle	.250607	2.343	.0192
Housing quality	.246188	4.959	.0000
Partner has good job	.313595	1.979	.0480
Partner has bad job	-.380719	2.828	.0047
Adjusted R^2 = .052	df = 1982		

To test for robustness of the results, the sample was split randomly into two halves, with the same model estimated for each half. The results were somewhat different between the two subsamples, but the coefficients on maternal schooling, 0.212 and 0.311, did not differ significantly from each other. The hypothesis that the two samples are homogenous cannot be rejected. Unstable results on ethnicity s may reflect correlation with area of residence. To control for effects of area of residence, we adopt a community fixed effects model, in which a dummy variable is included for each sample cluster, save the first. We report the results for the variables other than the cluster dummies, which are of no intrinsic interest. The coefficient on maternal schooling is strengthened as a result of including these cluster dummies. The apparent effect of Tutsi ethnicity is greatly reduced in magnitude, which is consistent with our interpretation of the disparity in results between the two random half-samples.

Regression sample:	DHS sample, community fixed effects		
Variable	Regression coefficient	t-statistic	Probability level
Age in months	-.015696	5.912	.0000
Years of schooling	.031183	2.597	.0095
Maternal ethnicity -- Tutsi	.209178	1.669	.0952
-- Twa	-2.950117	2.337	.0196
Ownership of bicycle	.281944	2.447	.0145
Housing quality	.232682	4.561	.0000
Partner has good job	.295876	1.831	.0673
Partner has bad job	-.303759	2.187	.0289
Adjusted R^2 = .102	df = 1962		

a

Discussion and exploration of pathways of influence

Does maternal schooling favor child growth in rural Rwanda? One would hope that the multiplicity of data sets would allow a conclusive answer to this question. Instead, we find conflicting results both within and between data sets. The most coherent results come from the DHS data set. These results indicate that maternal schooling has a very strong and roughly linear positive association with child heights in preschool-aged children. This is true regardless of whether one controls for community fixed effects or for potential confounding by household socioeconomic status. In the DSA data, a significant positive effect of maternal schooling emerges only in the intensive sample and only in random effects estimation. In the extensive sample and in the pooled sample, no significant association with maternal education is observed. Further, when the effect of area of residence is controlled in the intensive DSA sample, the apparent effect of maternal schooling is greatly reduced and is no longer statistically significant. In the DHS data set, controlling for community fixed effects raises the coefficient on maternal schooling.

The next step toward addressing this question is attempting to model the pathways by which maternal education can positively influence child growth status. Because of the greater robustness in results for maternal schooling in the DSA data, this objective could perhaps be pursued with DHS data. Maternal education can influence child growth through several pathways, including higher incomes, improved feeding practices, better hygiene and environmental sanitation, child spacing, and more effective use of preventive and curative health care. Unfortunately, the absence of any information on income and expenditure levels and the lack of any detailed feeding, hygiene, and environmental sanitation questions in the DHS questionnaire render this exercise necessarily incomplete. The DSA survey also lacks the second set of variables, but has much more complete economic data available and hence is more suited to controlling for economic mechanisms and confounding.

The first step is to determine to what extent the influence of maternal schooling appears to occur through the achievement of functional literacy as opposed to some socialization process which may be unrelated to literacy. The DHS questionnaire classified maternal literacy as an ordinal variable with 3 levels: cannot read, can read with difficulty, and can read easily. In preliminary analyses, we found that including the ordinal scale works just as well as two separate dummy variables, with the coefficients on the two levels of literacy almost equal in magnitude. Including the literacy scale variable, the effect of years of schooling is completely explained, with

the residual coefficient negative and quite insignificant. One other slight change we have made from the forgoing tables is to treat age as a linear spline, with one piece from 24 to 36 months and the other from 37 to 59 months. Excluding years of maternal schooling, the coefficient on the level of maternal literacy is reduced from 0.174 to 0.150. This implies that achieving full literacy is associated with a mean HAZ score higher by 0.30 SD compared to maternal illiteracy.

Regression sample:	DHS sample		
Variable	Regression coefficient	t-statistic	Probability level
Age spline 24-36 months	-.028013	2.910	.0037
Age spline 37-59 months	-.015846	2.311	.0209
Years of schooling	-.010656	0.552	.5813
Level of literacy	.173660	3.235	.0012
Adjusted R^2 = .029	df = 2026		

A very similar finding appears when the data from the intensive DSA sample are analyzed. Controlling for literacy, years of schooling has no effect, while controlling for years of schooling literacy still has a significant association with HAZ. Eliminating the schooling years variable reduces the difference in height between children of literate and illiterate mothers to 0.38 SD, which is modestly larger than the total effect of literacy in the DHS survey before control for confounding. The significance level is also greatly increased, to p = .0006.

Regression sample:	Intensive DSA sample		
Variable	Regression coefficient	t-statistic	Probability level
Age spline 24-36 months	-.047667	2.241	.0255
Age spline 37-59 months	-.025177	1.839	.0665
Years of schooling	-.010349	0.274	.7839
Literacy	.429857	2.112	.0352
Adjusted R^2 = .056	df = 524		

With the DSA data we can control for potential confounding by two variables not present in the DHS dataset. These are maternal stature and literacy of the household head. Introducing just household head literacy reduces the coefficient on maternal literacy from 0.383 to 0.351, while adding maternal height has no additional effect on maternal literacy. Curiously, while maternal stature has no effect on the maternal literacy coefficient, its introduction substantially

reduces the coefficient on literacy of the household head, from 0.241 to 0.185, and reduces its significance level, from p = .0298 to p = .0838. This presumably reflects a tendency for literate households to have taller children, who in turn make taller adults.

The next step is control for confounding by household socioeconomic status and by area of residence. We begin with the DHS data. In results not reported, the coefficient on the literacy variable is reduced to 0.126 when just community fixed effects are incorporated. When individual and household-level confounders are included as well, the coefficient is reduced to 0.098. This indicates that roughly one third of the observed association of child heights with maternal literacy appears to be due to confounding and assortative mating. The variables appearing in that equation also appear in the subsequent tables, so we will not repeat them here.

Similarly, control for confounding greatly reduces the magnitude of the coefficient on maternal literacy in the DSA data. Controlling for community fixed effects, as well as household head literacy and maternal stature, the coefficient is reduced from 0.352 to 0.171, and the significance level is only p = .1592. Controlling in addition for three other confounders: multiple births, housing quality, and value of livestock owned reduces the coefficient still further, to 0.104 and the significance level to p = .4057. The coefficient on household head literacy is also reduced by the addition of these variables, from 0.197 to 0.118, and the significance level rises from .0732 to .2827.

Adding two additional variables, logarithm of per capita water usage and a scale reflecting the occurrence and severity of food insecurity episodes, reduces the coefficient on maternal literacy in the intensive DSA sample still further to 0.086, and p = .4818. The coefficient on household head literacy is reduced to an even greater extent, from 0.118 to 0.078, and the significance level is comparably low to that of maternal literacy, p = .4728. These two variables primarily reflect relative household economic position, but they also result from resource allocation decisions which may be influenced by parental education and hence one cannot assume that these function strictly as confounders. Nonetheless, it is clear that they do not reflect any special role of female literacy as opposed to that of other household adult decision makers.

One potentially important pathway of influence is use of mass media to receive health-related information. The Rwanda DHS survey did not ask about general use of mass media, but it did inquire as to whether the respondents recalling seeing family planning messages in the newspaper or hearing them on the radio within the month preceding the survey. Presumably, those women who did are also likely to have previously recalled specifically health and nutrition messages as well. These variables pick up both the use of mass media and the ability to decode

and remember specific information, both of which should be strongly positively related to functional literacy. Curiously, we find that by itself reporting having read a family planning message in the newspaper (only 3% of the sample) has no significant effect, which reflects a mixed positive effect from having both read the newspaper and listened to the radio and an almost equally strongly negative effect of reading the newspaper but not listening to the radio.

Regression sample:	DHS sample, community fixed effects		
Variable	Regression coefficient	t-statistic	Probability level
Age spline 24-36 months	-.029433	3.022	.0025
Age spline 37-59 months	-.010975	1.597	.1105
Level of literacy	.085837	2.594	.0095
Maternal ethnicity -- Tutsi	.180961	1.455	.1457
-- Twa	-2.84106	2.267	.0235
Twin birth	-.771309	2.877	.0041
Ownership of bicycle	.255493	2.219	.0266
Housing quality	.219912	4.323	.0000
Partner has good job	.239457	1.497	.1355
Partner has bad job	-.300367	2.179	.0294
Heard message on radio	.207407	2.801	.0052
Read message in newspaper	-.946481	2.837	.0046
Interaction of radio and paper	1.322591	3.399	.0007
Adjusted R^2 = .117	df = 1982		

The coefficient on literacy is reduced from .098 to .086 by controlling for mass media usage. In a logistic regression predicting likelihood of recalling hearing a family planning message, basic literacy has a significant positive effect while full literacy has no additional effect. When ownership of a radio is controlled, which is the overwhelming determinant of having heard a message on the radio, the effect of literacy is no longer significant. Basic literacy is associated with a significantly greater likelihood of owning a radio, controlling for household socioeconomic status. Full literacy is not related to the probability of radio ownership. Maternal literacy has a much larger positive effect on newspaper readership than on radio listening, as would be expected, and since this has offsetting effects on child heights, these variables collectively explain relatively little of the causal association.

13

One of the major postulated mechanisms of influence is that literate or educated mothers are more likely to practice good hygiene, thereby protecting their children from risk of infection. The DHS data set has very limited and inadequate indicators of hygiene. These include whether the household has a pit latrine and the type of water source. Neither of these variables is significantly related to child heights, and so we do not include them in the regressions. An alternative approach is to use logistic regression on risks of specific types of disease symptoms reported in the DHS survey for the 2 weeks prior to the interview. These are diarrhea, without or with blood in stools (dysentery), fever, and cough, without or with rapid, indrawn breathing (acute lower respiratory infection or ALRI). Neither maternal literacy nor level of schooling are close to being significantly associated with risks of any of these types of symptoms. Consequently, we conclude that hygiene is probably not an important mechanism linking maternal education to child health and growth in rural Rwanda.

Another set of potentially intervening variables relate to reproductive patterns. Children who are first born, last born, born at least 36 months following the preceding birth, or have a succeeding birth within 24 months are significantly taller, controlling for other variables. The interpretation of the first three of these variables is straightforward, relating to reduced competition from other young siblings for maternal attention. The positive effect of having a short following birth interval has previously been remarked upon in Senegal (Goldberg and M'Bodji, 1985), but the explanation is unclear. Maternal age, parity, and birth order as a continuous variable seem to have no independent influences and are left out of the regression so as to avoid a problem of collinearity among the reproductive variables. One of the more interesting things to note in comparing the following and preceding tables is that the negative effect of age is substantially attenuated when the birth variables are included. As children age, they are more likely to have a younger sibling, which appears to have a sizeable negative effect.

Regression sample:	DHS sample, community fixed effects		
Variable	Regression coefficient	t-statistic	Probability level
Age spline 24-36 months	-.020872	2.102	.0357
Age spline 37-59 months	-.006551	0.938	.3483
Level of literacy	.084302	2.535	.0113
Maternal ethnicity -- Tutsi	.149655	1.213	.2254
-- Twa	-2.701759	2.174	.0298
Twin birth	-.777902	2.922	.0035
Ownership of bicycle	.284598	2.488	.0129
Housing quality	.211044	4.161	.0000
Partner has good job	.192625	1.210	.2265
Partner has bad job	-.283246	2.072	.0384
Heard message on radio	.228133	3.102	.0020
Read message in newspaper	-.982083	2.968	.0030
Interaction of radio and paper	1.354186	3.509	.0005
First birth	.262139	3.356	.0008
Last birth	.190672	2.930	.0034
Preceding interval > 36 mo	.287884	4.255	.0000
Adjusted $R^2 = .132$	df = 1982		

Surprisingly, controlling for reproductive outcomes does not seem to have any appreciable effect on the maternal literacy coefficient. The coefficient is reduced in magnitude only from 0.086 to 0.084. The explanation is that maternal literacy has no clear or consistent positive association with child spacing patterns which are beneficial for child health. Literate mothers are indeed more likely to have fewer children, so that the probability of any given child being the first born child is significantly greater if the mother is literate. This effect is attenuated when maternal age is controlled, but is still significant. While children of fully literate mothers are significantly more likely to be last-born children than are children of barely literate mothers, the same is not true relative to illiterate mothers, revealing a U-shaped curve. Finally, long preceding birth intervals are significantly more likely among illiterate mothers. The explanation may be that literate mothers are more efficient in terms of compressing childbearing into a restricted number of years.

Another possible mechanism of influence from maternal education and literacy to child growth is through greater use of preventive health care services. The DHS survey provides information for four such variables: family planning services, prenatal care, type of delivery care,

and receipt of child immunizations. The problem for this hypothesis is that none of these variables is significantly related to child heights. Indeed, none of the four coefficients has a probability level lower than 0.50. Breaking down the prenatal care variable into different levels of usage improves the fit slightly, since women with medical complications are likely to receive an excessive number of visits, and women with just one or two visits are as likely to have poor outcomes as those with no visits. However, the coefficients are still quite insignificant. Consequently, the fact that literate mothers are more likely to use preventive health services and to use them effectively does not seem to be relevant to the determination of risk of child growth retardation in rural Rwanda.

Another set of variables which are potential candidates for intervening variables are retrospective infant feeding variables. Exploratory analysis revealed that the only association with duration of total breast-feeding is a positive height differential for the small number of children (n = 37) completely weaned from the breast before 9 months of age. It would be unwise to interpret this finding in a causal sense, since it may reflect self-selection bias. Early introduction of animal milk or solids (before 2 months of age) is found to be associated with poorer growth status in preliminary analyses, which is consistent with most previous studies. Finally, children who do not receive any animal milk before 12 months of age are significantly shorter past 24 months of age. This last finding is quite consistent with results of our analysis of the DSA data which show that ownership of milk-producing animals (cows or goats) is associated with significantly better linear growth status in children. The coefficient on this last variable is greatly reduced in magnitude and significance when the community fixed effects are included. We hypothesize that living in a community where drinking milk is common may have a positive effect on child height because of incomplete maternal reporting of milk feeding practices.

Regression sample:	DHS sample, community fixed effects		
Variable	Regression coefficient	t-statistic	Probability level
Age spline 24-36 months	-.018828	1.987	.0579
Age spline 37-59 months	-.006410	0.921	.3571
Level of literacy	.081546	2.459	.0140
Maternal ethnicity -- Tutsi	.151510	1.230	.2187
-- Twa	-2.690519	2.173	.0299
Twin birth	-.697906	2.571	.0102
Ownership of bicycle	.295025	2.585	.0098
Housing quality	.205363	4.056	.0001
Partner has good job	.180877	1.130	.2585
Partner has bad job	-.279839	2.053	.0402
Heard message on radio	.228792	3.118	.0019
Read message in newspaper	-.836790	2.512	.0121
Interaction of radio and paper	1.149654	2.958	.0031
First birth	.243024	3.112	.0019
Last birth	.208305	3.191	.0014
Preceding interval > 36 mo	.284580	4.220	.0000
Short duration breast-feeding	.455511	2.196	.0282
Early introduction of solids	-.808770	2.282	.0226
Early introduction of milk	-.367549	2.189	.0288
No animal milk by 12 months	-.118473	1.802	.0717
Adjusted R^2 = .138	df = 1982		

The effect of the feeding variables on the maternal literacy coefficient is quite modest. The coefficient is reduced by not quite 3 hundredths of a standard deviation, from 0.084 to 0.082. The lack of a stronger effect which might have been expected results from the ambiguous association of maternal literacy with feeding practices. Maternal literacy is unrelated to the probability of early weaning or early introduction of animal milk. Literate mothers appear to be more likely to introduce solid foods too early, but since only 13 children were so affected, the association is not strong. Finally, the only beneficial impact on feeding practices associated with maternal literacy is an increased likelihood of offering animal milk at some point during infancy (after 2 months of age). An inverse U-shaped association was present, with fully literate mothers less likely to do so than those who were only partially literate.

17

Another possible mechanism is maternal income earning potential. In the DHS data set, we have data on maternal occupation but not earnings. Hardly any of the mothers failed to report an occupation. In addition, among family farmers, who are the vast majority of the women respondents in the rural sample, the survey inquired as to whether the women received cash income or not and whether they considered themselves to be self-employed (in control) or not. We have ranked these occupations and categories into a single scale with five categories: farm laborers, unpaid family farmers, farmers receiving cash income, and women who are either self-employed farmers or employed outside agriculture. Rather than report the detailed results of the regression including this variable, we will list only the salient results, namely those for the maternal literacy and maternal work variables.

Regression sample:	DHS sample, community fixed effects		
Variable	Regression coefficient	t-statistic	Probability level
Level of literacy	.072832	2.187	.0289
////////////////// (suppressed)	//////////	///////	///////
Maternal occupational level	.189531	3.541	.0004
Adjusted R^2 = .144	df = 1968		

Maternal occupational level is an intervening variable which appears to account for a part of the association of maternal literacy with child heights. Including this variable reduces the coefficient from .082 to .073. The explanation would seem to be obvious, that more educated women have greater income-earning opportunities. However, only among the 20 children of women with high levels of income-earning opportunities (1% of the sample) does literacy seem to significantly discriminate, and this is true only for full literacy. Partially literate women have occupational patterns virtually indistinguishable from illiterates. It should also be noted that including this variable in the regression reduces the maternal literacy coefficient only in the community fixed effects model; without the cluster dummies the effect is to slightly raise the coefficient on literacy.

Finally, we can also control for possible effects of maternal literacy operating through unspecified risk factors during pregnancy for fetal growth retardation. The DHS questionnaire asked all mothers whether they recalled their children being large, average, or small at birth. This variable is available for the entire sample. In addition, birth weight as recorded on health cards for institutional deliveries is available for roughly one fifth of the rural sample. We report results here only for the subjective estimate of birth size. This variable has a very strong association with

attained height in preschool-age children, one of the strongest associations of any variable in the data set. Nonetheless, the association of birth size with HAZ appears to be orthogonal to virtually all of the included predictor variables, including maternal literacy. Indeed, the literacy coefficient increases slightly, from .073 to .075. The reason is that maternal literacy is associated with smaller birth size, although only weakly and insignificantly.

Regression sample:	DHS sample, community fixed effects		
Variable	Regression coefficient	t-statistic	Probability level
Level of literacy	.074464	2.243	.0250
////////////////// (suppressed)	//////////	///////	///////
Birth size	.163216	3.998	.0001
Adjusted R^2 = .151	df = 1966		

The DSA data set does not offer as much in the way of intervening variables compared to the DHS survey, although this does not seem to be a critical matter since so little of the association remains to be explained. Health services utilization variables, including immunization, prenatal care, distance to health center, and receipt of visit from a community health worker, are not related to child heights. A number of hygiene variables also proved to be without any influence on child heights once economic status is controlled. These include the presence and type of toilet facility, the cleanliness and use of the latrine based on interviewers' observations, presence of soap in the home, kitchen hygiene, and water storage practices. This set of results does not mean that hygiene is not important, but it may be that community-level hygiene is more important than that of the household, as has been reported in a number of other studies. In a community covariates analysis, the proportion of households in a cluster with a covered pit latrine proved to be a highly significant predictor of individual child heights. Also, as in true with the DHS survey, maternal literacy was found to not be significantly related to the presence of any type of disease symptom in children. Thus, we conclude that maternal literacy in rural Rwanda does not protect children against risk of infection.

The one set of significant intervening variables found in both the DHS and DSA data sets relates to reproductive patterns. In the DSA data, the two variables with the strongest associations with child heights are being the first-born child of a given mother or the only child under five in a household. The total number of children under 5 in a household is essentially equivalent to the latter variable. Controlling for reproductive patterns reduces the coefficient on maternal literacy from 0.086 to 0.050, although both coefficients are highly insignificant.

19

Regression sample:	intensive DSA sample, community fixed effects		
Variable	Regression coefficient	t-statistic	Probability level
Age spline 24-36 months	-.040649	1.868	.0626
Age spline 37-59 months	-.019693	1.347	.1789
Literacy of mother	.049842	0.392	.6952
Literacy of household head	.123843	1.132	.2584
Maternal height	.044626	4.765	.0000
Twin birth	-.945201	2.199	.0285
Housing quality logarithm	.191858	1.334	.1829
Livestock value logarithm	.050015	3.559	.0004
Per capita water use logarithm	.226154	1.781	.0757
Food crisis scale	-.174495	2.594	.0099
First birth	.406071	2.395	.0171
Only child under 5 years	.321500	2.642	.0086
Adjusted R^2 = .292	df = 460		

Conclusion

Despite our worries about the lack of robustness of results in the DSA data set, we are reassured to find parallel results between the DHS and intensive DSA samples. In both data sets, we observe an association of child heights with years of maternal schooling at both the primary and postprimary levels. This finding is unusual among studies of child anthropometry in sub-Saharan Africa, including Rwanda, and is worthy of attention. Furthermore, in both data sets we find that the effect of maternal schooling appears to be entirely explained by achieved literacy. Controlling for literacy, additional schooling appears to have no effect.

We are much more successful in accounting for the association of female literacy and child heights in the intensive DSA sample than in the DHS sample. In large part this reflects the much greater explanatory power of the cluster dummies compared to the DHS data. That is, much more of the variation in child heights in the DSA sample is explained by between-area variations. Within-area variations in maternal literacy and in household economic status are greatly reduced compared to the random effects coefficients. We were surprised to find that controlling for maternal height had essentially no effect, because in this sample literate mothers are not taller on average.

Household and community variables appear to account for the largest share of the association of maternal literacy with child heights. In the DHS dataset, control for these variables together with community fixed effects reduces the coefficient on the level of maternal literacy from 0.150 to 0.098. Controlling for various intervening variables, including mass media usage, reproductive patterns, feeding patterns, maternal occupational level, and size at birth, reduced the coefficient still further by less than one quarter, to 0.075, which remained statistically significant. Since this variable is a 3-level scale, the differential in adjusted child heights is 0.15 SD between children of illiterate and fully literate mothers. Since the DHS data set has no control for livestock ownership, food security, most dimensions of housing quality, level of water usage, and other economic variables, one cannot exclude the possibility that much of this differential reflects association with these omitted variables.

In the intensive DSA sample, control for community fixed effects and partial control for household economic status reduces the coefficient on maternal literacy from 0.383 to 0.104. Controlling still further for food security and water consumption reduces the coefficient to 0.086. Controlling for reproductive patterns reduces it finally to 0.050. Thus, over 85% of the association between maternal literacy and standardized child heights is accounted for by the variables included in the model. The residual association is very small in magnitude and far from approaching statistical significance.

To conclude, we find consistent evidence that primary-level maternal schooling has a positive influence on child growth in rural Rwanda, contrary to prior reports from both Rwanda and other African societies. This influence appears to operate through the achievement of functional literacy rather than through an influence of socialization. However, contrary to the widespread belief that the effect of literacy should occur primarily through increased efficiency in the household production of health, we find that other mechanisms appear to be far more important. These include the effect of living in a more literate community, with different standards of hygiene and health-related practices. Literacy is associated with better hygiene and use of health services, but these do not seem to be very important predictors of child growth. We find very weak or nonexistent associations of child heights with use of preventive health care services and with hygienic practices at the household level. The only feeding variable which appears to have a strongly positive influence is the provision of animal milk to young children, which in rural Rwanda is almost entirely a function of the presence of milk-producing animals on the family farm.

Most importantly, our analyses suggest that economic variables have a dominant influence on the risk of child growth retardation in rural Rwanda. This result runs counter to the

conclusion of the major preceding economic analysis of child anthropometric data from rural Rwanda by von Braun et al. (1991). Furthermore, controlling for community fixed effects, the effect of literacy of the household head, over 80% of whom are male, appears to be just as important as maternal literacy. Parental literacy appears to influence child growth almost entirely by influencing access to material resources and, inferentially, through influencing allocation of resources in ways favoring health and nutrition. These include most notably making sure that the household has a milk-producing animal (cow or goat) to provide for young children. The other economic variable of overwhelming importance is food security. Households which experience food shortages, and the degree of severity of the shortage as indicated by household responses, are far more likely to have children with stunted growth. Ownership of livestock is an important factor in helping reduce the severity of food crises, since the sale of livestock and other non-land assets is associated with less adverse outcomes in terms of child growth.