Estimating the impact of women's education on fertility, child mortality, and empowerment when schooling ain't learning

Michelle Kaffenberger[*], Lant Pritchett[†‡], and Justin Sandefur[‡]

March 7, 2018[§]

Abstract

We use literacy data available in the Demographic and Health Surveys (DHS) for 129 survey rounds, across 54 countries, to estimate the impact of female *basic education*–which we define as completing six years of schooling *and* acquiring literacy–on a woman's fertility, survival of her children, and (for the 69 DHS rounds where it is available) a measure of the woman's empowerment. First, our estimates of the impact of basic education on these three outcomes are 3 to 4 times larger than the standard approach that estimates the impact of schooling. For instance, using OLS and data on completed years of schooling produces estimates that female basic schooling (completing six years of schooling) reduces child mortality by 21 percent, whereas our results which use IV techniques and data on both schooling and literacy suggest a reduction in child mortality from female basic *education* of 68 percent. Second, our results suggest that achieving literacy accounts for 36 percent of the child survival improvement, 50 percent of the reduction in fertility, and 80 percent of the increase in female empowerment from basic education. These results suggest that (a) the non-pecuniary returns from female education are even much higher than previously believed and (b) the returns to investing in cost-effective actions to improve learning outcomes, such as literacy acquisition, of girls already in school could be very high, higher than investing in expanding the years of schooling.

Keywords: return to schooling, gender, literacy, child mortality, fertility, empowerment

JEL Classification Numbers: I1, I2, J1, O1

[*]RISE Research Fellow

[†]Harvard Kennedy School

[‡]Center for Global Development

[§]The views expressed here do not necessarily reflect those of the Center for Global Development, its board, or its funders. All errors are ours.

Contents

1 Introduction 4

2 Impact of Schooling versus impact of education 7

2.1 DHS Measures for Schooling and Learning 13

2.2 Schooling and the production of literacy 14

3 Estimates of impact of schooling and literacy: OLS and IV 17

3.1 Outcome and other regression control variables 18

3.1.1 Outcome variables 18

3.1.2 Other control variables included in all regressions 19

3.2 Meta-analysis weighting 20

3.3 Distinguishing partial and total impact of schooling and impact of education with OLS 21

3.4 Understanding the consequences of measurement error on decomposing the relative impacts of schooling and learning 28

3.5 Using Instrumental Variable Estimation Techniques for Literacy and Schooling 31

3.6 Results from Using Instrumental Variable Techniques to Estimate Impact of Literacy and Schooling 32

3.7 Trade-offs from Using Instrumental Variable Techniques 38

4 Individual literacy as a mediator between schooling and outcomes 42

5 Education quality as a moderator of schooling's effects on outcomes 48

6 Conclusion 52

7 Bibliography 53

A Country by country data **56**

B Cumulative distribution function graphs for fertility and female empowerment **58**

1 Introduction

There is a long-standing and broad consensus that women's schooling is one of the most powerful forces for improving well-being in the developing world, with positive impacts in the labor market as well as variety of non-market wage impacts such as child health, fertility, and empowerment[1]. Montenegro and Patrinos (2014) summarize data from 139 different surveys and find the estimated labor market wage returns are even higher for women than for men. There are also enormous empirical literatures showing that women's schooling is associated with a variety of outcomes, over and above labor market and income impacts (e.g. Glewwe (1999)). Mothers with more schooling tend to have higher child survival[2], lower child malnutrition (Christiaensen and Alderman, 2004), lower fertility[3], and greater female empowerment[4].

Yet these empirical literatures are still radically incomplete. They do not address *how* schooling has these impacts. Nearly all empirical studies use only a measure of time served: "years of schooling", "highest level of school attended", "grade attainment". But nearly everyone assumes that *schooling* has an impact on outcomes because it produces *education*: individuals through attending school acquire skills, capabilities, attitudes, beliefs, and dispositions-and it is this *education* that then has long-lasting positive impacts on wages, health, empowerment, and well-being. If schooling (time served) and education (capabilities acquired) were consistently, strongly, and tightly related this casual elision of empirics based on schooling and conclusions about education would be harmless. But in many countries today "schooling ain't learning." (**?**). A recent ASER report of youth aged 14 to 18 in rural

[1]Nearly 25 years ago Lawrence Summers **?**, then head of the World Bank's research division, argued that girl's schooling was arguably the highest rate of return investment available to governments.

[2]For example, the causal link between mothers' schooling attainment and decreased child mortality is "one of the most consistent and powerful findings in public health," with a Lancet review estimating 4.2 million child deaths (51.2 percent) from 1979-2009 can be attributed to increased grade attainment in women of reproductive age (Gakidou et al., 2010).

[3]Behrman (2015) shows female schooling impacts on in reducing women's ideal and actual family size,

[4]There are also empirical literatures at the country level suggesting countries with higher levels of female education (either absolutely or relative to male) tend to be more democratic (XXX), have less corruption (XXX), and XXX

India found that while over 80 percent had completed grade 8,roughly half or less could do simple division, calculate how much a price discount of 10 percent would save them, follow simple instructions, or understood measuring length with a ruler. With a weak connection between shooling and learning no claim can be made about the "impact of education" using only data on schooling–even with a randomized control trial (RCT).

There is a "learning crisis" in developing countries (**?**, **?**, **?**). For example the 2012 UNESCO Global Monitoring Report analyzes the DHS and reports that "many children in poor countries have not become literate even by the time they have completed primary school," while "curricula around the world expect children to learn to read by the end of the second year of primary school" (United Nations Educational and , UNESCO). Further, in some countries the crisis appears to be worsening. The Annual Status of Education Report (ASER) assessed over a half a million children in rural India in 2014, finding that the percentage of grade 5 students who can read a simple story fell from 54 percent to 48 percent from 2010 to 2014, and the percentage of grade 5 students who could do a simple division problem fell from 36 percent in 2010 to just 26 percent in 2014. (Programme, n.d.).

Particularly as most countries in the world near universal enrollment and universal primary school completion a pressing policy question is whether to allocate efforts to raising the learning of those who attend school or extend the duration children spend in school. Unfortunately the vast and expanding literature on the economics of education, including the rapidly burgeoning using of methods that are careful about causal identification, has two branches. One large branch estimating the efficacy of various actions in raising learning of those in school (and some part of that branch estimates cost effectiveness)–but this generally does not produce estimates of the gains in outcomes from the additional learning. Another large branch examines the positive benefits of attending school, but that branch does not distinguish the causal pathways and hence how much of the schooling effect was due to the increased learning and hence whether similar benefits could have been accomplished with more learning while in school rather than more years of schooling. This leaves policy makers and practitioners completely in the dark on an increasingly key issue of addressing the learning crisis.

We use the data on women's directly assessed literacy in 129 survey rounds in 54 countries from the DHS to examine the association of outcomes with schooling attended and learning/literacy. Regression analysis on three outcome variables: women's fertility, child mortality, and women's empowerment allows us to distinguish between the impact of schooling (at whatever learning that schooling happened to produce) and the impact of basic education, which we define as completing six years of schooling during which women, at a minimum, learn to read. We use instrumental variable (IV) estimates using the clustered sampling of the DHS to create enumeration area leave-out-means of schooling and literacy as instruments to address measurement error. We use standard approaches to "meta-analysis" to summarize these country-survey round regression results.

The standard estimates of the impact of schooling *underestimates* the impact of women's *education* (schooling plus learning) on outcomes by a factor of 3 to 4. For example the typical approach shows six years of schooling would reduce average total fertility by one-third of a child (a 10% drop from the average fertility in the sample of 3.4 children) while our improved estimate shows that basic education - six years of schooling plus attaining basic literacy - reduces average fertility by 1.21 children, a 36% reduction. Further, the typical approach suggests six years of schooling reduces child mortality by 22%, while our improved estimates show six years of schooling plus literacy reduces child mortality a whopping 68%. By a standard measure of women's empowerment (scaled to a standard deviation of 1) six years of schooling alone suggests an improvement of .14 units whereas basic education improves empowerment by .59 units.

Our estimates of the gains from basic education can be divided into the gains from schooling *per se* at a given level of literacy and the gains from literacy (our empirical proxy for learning) at a given level of schooling. Our IV results suggest that greater literacy accounts for 36 percent of the gain from basic education in child survival, 50 percent of the reduction in fertility, and 80 percent of the gain from empowerment. Depending on the relative costs of improving learning relative to the cost of an additional year of schooling this implies improving learning could be *orders of magnitude* more cost effective in improving outcomes for women. In Section 4 we use robustness checks to examine the attribution of total impacts to mediating pathways which suggest the OLS findings are not robust to even small degrees of violations of the assumptions of "sequential ignorability" in the treatment effects literature, but these robustness checks cannot be run on our IV regressions.

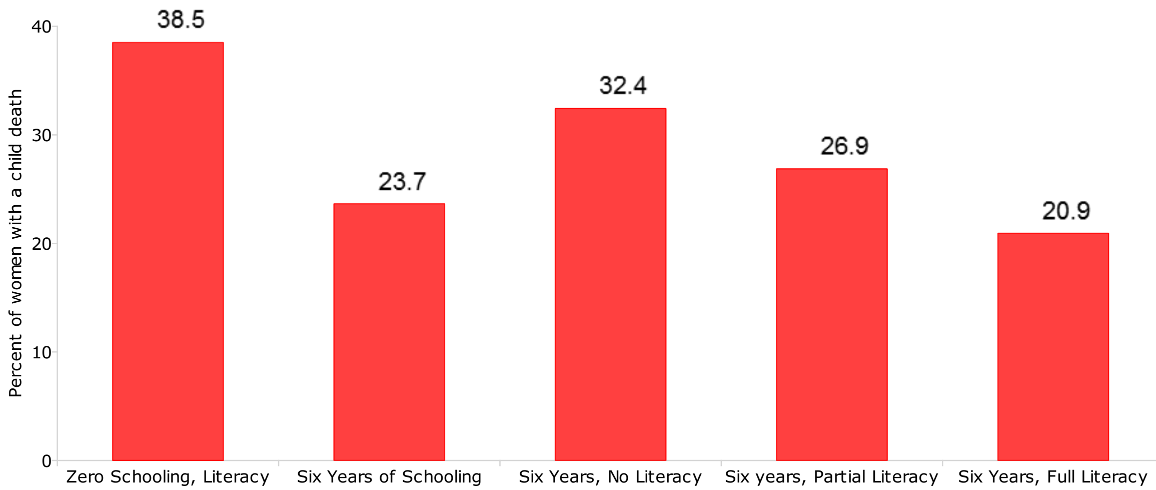
Finally, we use an alternative approach to examine the same core question of learning as the mechanism for schooling's impact: moderation analysis. We compute an aggregate measure of school quality in a given region or country, and examine the role of schooling quality as a moderator for schooling's impact on fertility, child mortality, and women's empowerment. Interacting this school quality measure with individual years of schooling, we find that the returns to schooling in terms of child survival, for instance, are three-quarters larger at the highest level of school quality compared to the lowest.

2 Impact of Schooling versus impact of education

We start with the simplest possible cross tabulation of outcomes and schooling in Figure 1 for two reasons. First, it is easy to hone the intuition for the question we wish to explore and nearly all econometric and experimental (RCT) methods are in some ways just elaborations on this basic cross-tab. Second, because they are easily understood in policy making and public discourse about the impact of schooling, simple graphs like these are widely used.

Figure 1 compares the fraction of women who have ever experienced the death of a child (of women who ever had a child) between women with with no schooling and zero literacy and with six years of schooling complete at various levels of literacy, first at the average literacy level, and then those with none, partial, and full literacy among the over 850,000 women in our DHS sample who have had a child. Of women with no schooling 38.5 percent have had

Figure 1: Fraction of women aged 15-49 who have experienced the death of a child by schooling and literacy



Source: Authors' calculations based on DHS microdata for 54 countries.

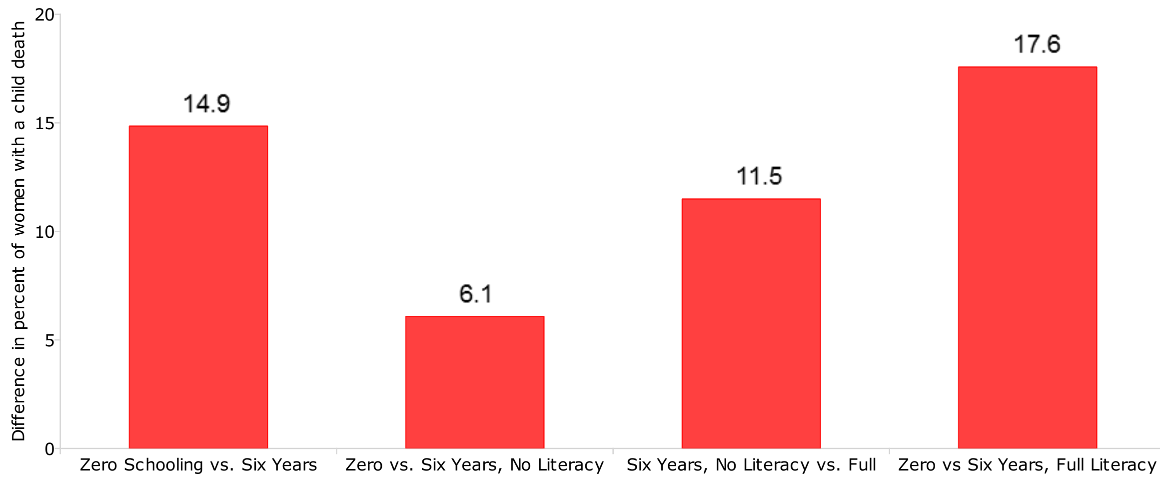
a child die. This is only 23.7 percent among women with six years of school complete. In many studies this difference would be referred to as the "impact of schooling."

But some women with schooling could read a sentence and some could only read parts of a sentence and some could read the full sentence. Among the women with schooling complete but who could not read the sentence 32 percent has experienced a child death, only about 6 percentage points lower than women with no schooling (Figure 2). For women with schooling complete and who could read a sentence only 20.9 percent had experienced a child death, 17.6 percentage points lower than those with no schooling (Figure 2). Gaining literacy therefore nearly triples the outcomes associated with six years of schooling.

So another interpretation is that the "impact of schooling" is mostly mediated by the acquisition of literacy and that the "impact of basic *education*" defined as completing schooling and acquiring literacy is high (17.6 percent reduction–cutting child death by nearly in half from 38 percent) but the impact of school alone, for those who did not acquire literacy, is quite modest, only 6 percent.

This reveals there are two distinct meanings of the "impact of schooling." One is the *total* impact of schooling through all of its mediating causal channels–including the acquisition

Figure 2: Differences in child survival by schooling and literacy



Source: Authors' calculations based on DHS microdata for 54 countries.

of literacy–at the average or typical rate at which those mediators, like the acquisition of literacy, happen with schooling. The other is the *partial* impact of schooling, that is, the impact of schooling except through channels that are accounted for, like literacy or wealth.

We define the "impact of education" as the difference in outcomes between those with no schooling and those with basic schooling (six years) and literacy; the "partial impact of basic schooling conditional on literacy" as the difference between those with and without schooling at the same degree of literacy; and the "total impact of basic schooling" as the difference in outcomes between those with no schooling and those with six years complete (which includes the partial impact of schooling and the impact of schooling through its mechanisms and is itself the weighted average of the outcomes of those with schooling and literacy and those with schooling and no literacy).

This leads to a simple expression for outcome (Y) for a woman as a linear (for simplicity) function of her years of schooling completed (S), her extent of literacy (L) and all other conditioning factors (Z, e.g. urban/rural residence, her age, a wealth index, etc.):

$$Y_i = \alpha + \beta_{S|L,Z} * S_i + \beta_{L|S,Z} * L_i + \theta_{Z|L,S} * Z_i + \epsilon_i \tag{1}$$

We add extra notation for the schooling and literacy coefficients in this equation for clarity. In this notation the *partial* impact of schooling, or how much we would expect an outcome to be better if schooling went up by S years *but literacy and other factors did not change* is:

Partial impact of years of schooling (L, Z fixed) on outcome Y:

$$\Delta Y = (\Delta Y/\Delta S)|_{L=\bar{L},Z=\bar{Z}} * \Delta S = \beta_{S|L,Z} * \Delta S \tag{2}$$

The total impact of schooling then is how much we would expect an outcome to be better if schooling went up by S years and (rather than L being held constant) we include the improvement in outcome Y associated with the learning obtained from S years of schooling:

Total impact of years of schooling including impact on literacy (Z fixed) on outcome Y:

$$\Delta Y = \beta_{S|L,Z} * \Delta S + \beta_{L|S,Z} * (\Delta L/\Delta S) * \Delta S \tag{3}$$

We regard *neither* of these are *conceptually* the same as the *impact of basic education* which is the acquisition of *both* additional schooling *and* the acquisition of literacy:

Impact of basic education:

$$\Delta Y = \beta_{S|L,Z} * \Delta S + \beta_{L|S,Z} * \Delta L \tag{4}$$

Where for the impact of "basic education" ΔS is "completing primary school" (say, six years of schooling) and ΔL is "acquired literacy" (on the scale that is measured).

The "total impact of schooling" and "impact of education" are conceptually distinct and it is not merely that one is an empirically flawed estimate of the other. The total impact of schooling answers the question of how much one would expect outcomes to improve from additional schooling–at whatever degree of additional learning (measured as increased capability acquisition or in this simple case, literacy), that happens to bring. The impact of education asks how much improvement there would be if a person acquired both schooling *and* learning.

Because of the widespread availability of household data sets that include both a measure of schooling and measures of a variety of outcomes there have been literally thousands of empirical studies comparing the outcomes associated with various levels of reported schooling. This includes studies of both pecuniary (e.g. wages, incomes) and non-pecuniary (e.g. child health, empowerment) gains.

None of these studies (whether cross-tabs, regressions, or experimental) using schooling can be regarded as an estimate of the impact of *education* at all, but rather are only estimates of the total impact of *schooling.* The total impact of *education* depends on (a) the relative causal pathways of both schooling *per se* and of various educational attainments on outcomes and (b) the extent of education achievement that schooling produced.

Various pathways are posited to mediate the effect of schooling, including the pure act of girls attending school, building familiarity with new social interactions and networks; specific knowledge actually transferred in school to future mothers; and finally, the acquisition of learning skills, literacy and numeracy, allowing women to accumulate knowledge both in and outside of school (Bongaarts and Watkins, 1996; Christiaensen and Alderman, 2004; Dearden, Pritchett and Brown, 2004; Glewwe, 1999; La Ferrara, Chong and Duryea, 2012). The authors of the Lancet review on child mortality note that "many hypotheses have been proposed for the mechanisms through which increased education could lead to reductions in child mortality rates, including individual level effects through improved use of health services, economic advantages, empowerment and independence of women, and community-level effects" (Gakidou et al., 2010). These are not mutually exclusive or necessarily universal. In an analysis of Malawi, Uganda and Ethiopia, Behrman (2015) finds some pathways to be common across contexts while others are country-specific.

However, a growing body of studies emphasize the role of skills (e.g. literacy) in mediating the association between schooling and beneficial social outcomes. Glewwe (1999) finds that "education improves child health primarily by increasing [mothers'] health knowledge" in Morocco. Though Moroccan schools did not include health knowledge as part of the curriculum, the impact of schooling on health behaviors appeared to be the result of skills learned in school; literacy and numeracy allowed mothers to improve their health knowledge outside of school, for example via the ability to read medicine labels (Glewwe, 1999). Based on survey data and ethnographic research in Nepal, including basic literacy tests, LeVine et al. (2004) found that literate mothers had better comprehension of both print and radio health messages, and were better able to tell an organized health narrative to an interviewer; schooling had no significant effect separate from that mediated by literacy.

To understand the total impact of education, then, we must account for both the causal pathway of schooling itself (the partial impact of schooling) *and* the causal pathway of learning. This will give us the potential impact on outcomes that could be achieved if schooling consistently produced learning.

By analyzing non-pecuniary benefits to education, we have an advantage in disentangling these pathways over the vast literature that studies the connection between schooling, education, and wages, incomes, and economic productivity. The bulk of this literature at the individual level consists of "Mincer regressions" (**?** that estimate the association with schooling, and the finding that individuals (men and women) with higher levels of schooling tend to make more in the labor market is probably the second most replicated and reliable fact in economics (after Engel's law) as there are literally thousands of Mincer regression studies across hundreds of countries, many dating back to the 1960s and 1970s. But with wages or economic productivity there are massive challenges, of two types. First, there is a "micro-macro" paradox as many countries have had upward sloping Mincer curves showing those with more education make higher wages at each point in time–but massive increases in schooling have not led to higher average wages (**?**, **?**). Second, the Spence (1973) model suggests people might make higher wages from more schooling not because it raises their productivity but because having more schooling signals productivity individuals have independently of their schooling. The current conventional wisdom is that cognitive skills gained in school have been much more predictive of countries' economic growth than expansion of schooling itself (Hanushek and Woessmann (2012),?). We argue examining the non-pecuniary gains for women is better for disentangling the relative pathways of schooling and learning as these outcomes are often individual choices (e.g. fertility) or relational (empowerment) and hence are less susceptible to the complications of market mediated outcomes and Spence signalling.

2.1 DHS Measures for Schooling and Learning

Since 2000 the DHS instrument includes a literacy test in the women's questionnaire which makes the analysis of the distinct impacts of schooling, literacy, and hence education possible. We use data from 129 survey rounds from 54 different countries, mostly low-income and lower-middle-income (Demographic and Survey, n.d.). Given the DHS sampling for the administration of the questionnaire containing the literacy assessment the samples are designed to produce estimates that are nationally representative of women of reproductive age (15-44) in each country.

DHS asks each woman whether she attended school, and if so the highest level she attended (primary, secondary, or tertiary). The survey also asks each woman the highest grade she attended. We use this self-reported years of schooling as our schooling measure. For aggregation and reporting, we use six years of schooling as anapproximation for "primary completion" even though the exact cutoff for primary completion differs from country to country.

For the literacy indicator, all respondents who have not attended secondary school are included in the literacy assessment. This truncates our sample to only those women who report their highest level as less than secondary, which has implications we discuss below[5].

[5] In the DHS data, women with secondary schooling or higher are classified as literate without taking the assessment. By limiting our samples to those who took the assessment our calculated literacy rates cannot be compared to official DHS reports.

Women doing the literacy assessment are asked to read a sentence from a card that contains one simple sentence such as the following:

- *Parents love their children.*
- *Farming is hard work.*
- *The child is reading a book.*
- *Children work hard at school.*

The enumerators are provided with cards in the variety of languages they expect to encounter and each woman is allowed to choose the language she wishes to read–so this is not a test of literacy in English or in the dominant national language but of a woman's ability to read in a language of her choosing. The fraction of women without the appropriate language card is very small.

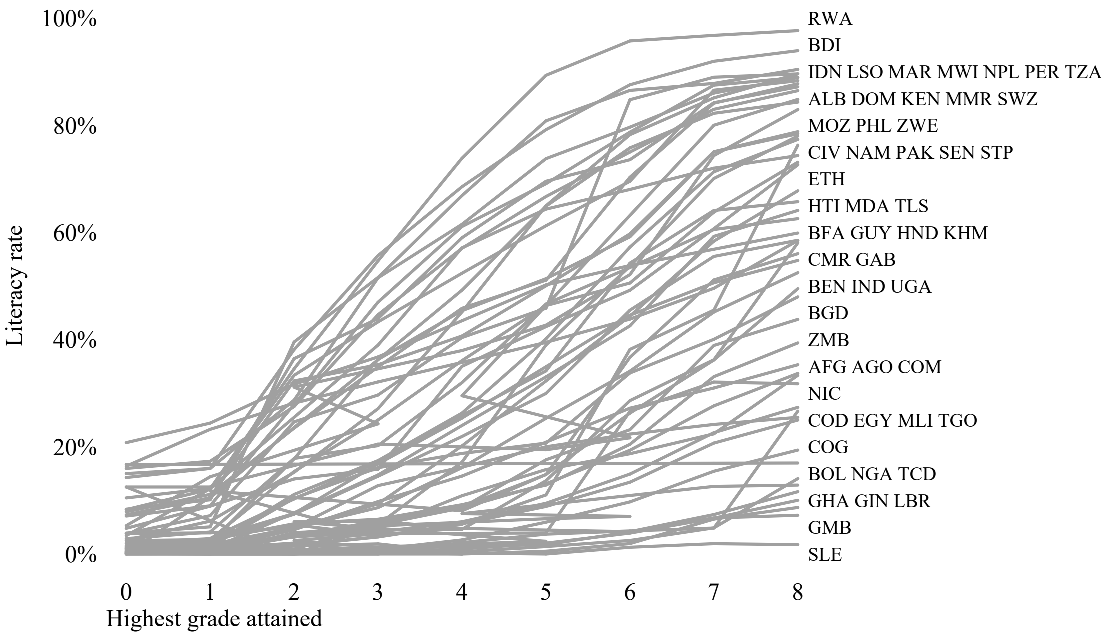
Literacy is measured on a three level scale, with enumerators recording whether the woman 1) could read the full sentence; 2) could read parts of the sentence only; or 3) could not read at all. Because reading "parts" of the sentence can include those who are only able to read a single word, we define "literate" as those who could read the full sentence.

This test provides a very low-bar for literacy. The sentences respondents are asked to read are simple and there is no measure for whether respondents would be able to read a more complicated passage. In addition, the categories do not imply any level of understanding of the sentence, and there is no test for ability to write. No other measure of learning, such as numeracy, is assessed.

2.2 Schooling and the production of literacy

We generate a "learning profile" for each survey round, which is the predicted literacy for young women (aged 25-34) at each grade (using a local polynomial regression rather than cross-tabs to smooth the relationship and account for small cell sizes). Figure 3 gives the learning profile for the most recent survey round for each of the 54 countries, and shows that the relationship between literacy and grade attainment varies enormously across countries.

Figure 3: Literacy among women age 25-34, by schooling level



Source: Authors' calculations based on DHS microdata for 54 countries, based on the most recent available round by country. Lines show fitted values from a local polynomial regression of the literacy score on years of schooling, limited to a sample of women age 25-34 with less than secondary schooling.

In almost all cases, literacy is below 20% and often near zero for women with zero years of schooling and then the relationship fans out. Among women who complete third grade, a third or more have gained literacy in some countries, such as Rwanda and Peru. But in other countries, such as Liberia and Mali, literacy rates remain at approximately zero after three years of schooling. Looking at women with six years of schooling, several West Africa countries still report literacy rates of less than ten percent, including the largest country in the region, Nigeria. These learning profiles suggest a majority of women are illiterate even after six years of schooling in twenty-one of the fifty-four countries[6].

[6] **?** show that across the countries with DHS data on literacy among young adult women with the exact same measured schooling–six years–the fraction who can read a single simple sentence varied from almost zero ( ), to very modest levels (e.g.) to nearly 100 percent ( ), with an average literacy across countries for those with six years of schooling of around 50 percent. Pritchett and Kaffenberger using an entirely different collection of data for 10 countries (the Financial Inclusion Index) that includes both men and women find very similar levels and variability in literacy for those with the same schooling

As we saw above, the expression in the simplest linear model for the total derivative of outcomes w.r.t. to schooling was:

$$dY^{i,c}/dS^{i,c} = \beta_S^c + \beta_L^c * (dL^{i,c}/dS^{i,c}) \tag{5}$$

Which implies that even if the partial effect of schooling (β_S) and the partial effect of literacy (β_L) were constant across countries the "total impact of schooling" would differ across countries if the impact of a year of schooling differed in the production of literacy ($dL^{i,c}/dS^{i,c}$), which Figure 3 strongly suggests is the case. With the DHS data we can run exactly the regression of literacy (on the three level scale) on schooling (highest grade completed) for each of the 129 survey rounds.

$$Literacy_{i,c} = \alpha^c + \beta_S^c * S^{i,c} + \epsilon_{i,c} \tag{6}$$

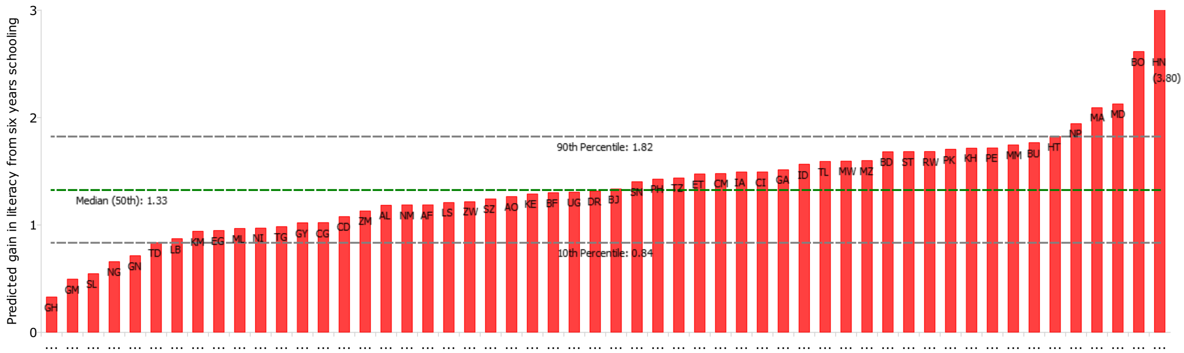
Using the fact that we have repeated observations for most of the countries (40 of 54) we can calculate that the variation in the literacy on schooling coefficients across the 129 rounds is 96 percent *across* countries, as the results for the same country tend to be quite stable across multiple rounds. For instance, for Ghana's four rounds the coefficient only varies from .29 to .38 while Rwanda's varies from 1.63 to 1.73.

The results in Figure 4 show the average observed literacy gain per year of schooling (dL/dS) by country. The gains are low on average (the overall median is that six years of schooling produce just 1.37 units on the 0 (cannot read) to 2 (can read) scale)[7]. And this varies massively across countries,[8] indicating that the total impact of schooling on outcomes

[7] Recall that UNESCO expects children to gain basic literacy by their second year of primary school.

[8] The association between schooling and literacy evident in the figures may be biased estimates of the true causal relationship of literacy and schooling for a variety of reasons. The primary concern that an individual's unobserved learning capability would contribute to both grade progression and observed literacy, which would imply the observed slopes are biased *upward*–making the observed shallow learning profiles all the more troubling. A second and closely related bias is that the proportion of women who go on to secondary schooling varies across countries, and these women are excluded from the literacy measure. If high secondary schooling rates are driven by low thresholds for grade progression, we would expect to find a negative correlation between secondary schooling rates and literacy among those with primary or less. This might be particularly apparent in countries where grade attainment has expanded rapidly, as is the case in

Figure 4: The likelihood of acquiring literacy (a two point gain) from six years of schooling varies widely across countries



Source: Authors' calculations based on DHS microdata for 54 countries.

will also vary massively across countries.

3 Estimates of impact of schooling and literacy: OLS and IV

To disentangle the partial impact of schooling, the total impact of schooling, and the impact of education, we begin by running the typical approach of using OLS to estimate the impact of schooling on our three outcomes with no measure of learning. We next add the literacy variable to these OLS regressions. Third, we instrument for literacy to correct for measurement error, and finally instrument for both schooling and literacy.

The first subsection describes our outcome and control variables from the DHS datasets. Second subsection discusses the use of meta analysis techniques to aggregate across datasets. The third subsection gives results on partial and total impacts of schooling and impact of education from OLS regressions, and the fourth gives consequences of applying instrumental variables techniques.

much of sub-Saharan Africa. In fact, we find the opposite. In the sample of 106 country-years with available data, the correlation between literacy rates among non-completers and the rate of primary completion is 0.73, and when looking at annual changes in both rates across survey rounds, the correlation is 0.69. As more women finish primary school, the literacy rate among those who don't goes up. This is suggestive that sample selection is not biasing the observed learning profiles upward.

3.1 Outcome and other regression control variables

3.1.1 Outcome variables

We analyze three outcome variables. First, we use total fertility, which is the self-reported total number of live births per woman. Second, we compute child survival, which is total number of live children divided by total number of births, the inverse of which is child mortality.

Third we compute a measure of women's empowerment. The DHS includes modules on women's empowerment which are widely used to estimate the impacts of various factors on empowerment outcomes (**?**). We use principle component analysis with the three most commonly used sets of empowerment questions, to create an empowerment index. The questions included in the index are:

- Whether the woman has any say in the following decisions (positive indicators):
  - Her own healthcare
  - Making large household decisions
  - Visiting family or relatives
  - What to do with money her husband earns
- Whether the woman believes a husband beating or hitting his wife is justified if the wife (negative indicators):
  - Goes out without telling him
  - Neglects the children
  - Argues with him
  - Refuses to have sex with him
  - Burns the food
- Whether the woman believes a wife may refuse sex with her husband if he "has other women" (positive indicator)

We run our regressions for each survey round individually and therefore we also create the empowerment index individually for each survey round. For women's empowerment, we use 67 survey rounds that included the empowerment modules (the remainder of the survey rounds did not include the necessary questions). We standardize the indices to have a standard deviation of 1 so that coefficients are easily comparable.

3.1.2 Other control variables included in all regressions

In all our regressions we use standard control variables, including age, age squared, and age cubed; rural or urban residence (dummy); dummies for regions within each country; and the wealth index included in the DHS.

Table 1: Summary statistics for key variables (pooled for all countries)

Row name	Mean	SD	N
Age	30.63	9.68	1055655
Literacy Rate (%)	27.17	44.48	1055655
Schooling Years	2.35	2.68	1055655
Rural Residents (%)	75.74	42.86	1055655
Rich (% of sample in top 40% population)	29.39	45.56	1055655

Table 2: Summary statistics for key health variables (pooled for all countries)

Row name	Mean	SD	N
Children Ever Born	3.37	2.77	1055655
Survival Rate of Children (%)	89.19	22.39	854758
Empowerment Index (normalized)	0	.99	386573

3.2 Meta-analysis weighting

We run each regression specification across our three outcomes of interest, and do this separately for the 129 survey rounds of data collection in our sample (67 for empowerment). Analysis of each survey round could be a study itself – many papers are written analyzing the impact of schooling on outcomes in one or a few countries. Therefore, by analyzing each survey round individually and aggregating the results, we are effectively conducting a self-contained meta-analysis.

We therefore use standard meta-analysis techniques to aggregate regression results. The standard approach is to weight regression coefficients by the inverse of their variance. In this way, more precise estimates receive greater weighting, and less precise estimates receive lower weighting. There are two models for assigning weights, the fixed effects model and the random effects model. Fixed effects models assume there is one true effect size which is shared by all included studies, and therefore the aggregated result is the estimate for this common effect size. The random effects model allows that the true effect could vary from study to study, for example if the underlying samples differ, if the interventions across studies differ in some ways, or if the context for each study differs in ways that could affect the true effect. Under the random effects model, the aggregated result is the mean effect across a distribution of true effects. In this model, estimates are weighted by the inverse of their variance and the variance of the true effects across studies to account for both the within-study and between-study variance **?**.

We use the random effects model to weight and aggregate results across survey rounds. It is reasonable to expect that the true effect of schooling and literacy will differ across survey rounds, given varying country contexts, time periods of data collection, and underlying populations of women with less than secondary schooling. The formula for the random effects weighted sum is:

$$\beta^K = \sum_{i=1}^{N} \frac{\beta_i^K}{var(\beta_i^K + \tau^2)} / \sum_{i}^{N} \frac{1}{var(\beta_i^K + \tau^2)} \tag{7}$$

Where β^K is the weighted sum of betas for either schooling, literacy, or "education" (the linear combination of schooling and literacy). β_i^K is the coefficient from survey round i, $var\left(\beta_i^K\right)$ is the variance of β_i^K, and τ^2 accounts for the variation between studies in the random effects model. Each coefficient β_i^K is weighted by the inverse of its variance plus τ^2 and these weighted coefficients are summed to give the aggregated, weighted average across the survey rounds.

3.3 Distinguishing partial and total impact of schooling and impact of education with OLS

The vast majority of estimates in the literature use observational data on "years of schooling", data on pecuniary (wages, incomes) or non-pecuniary (child health, fertility) outcomes, and some other control variables (e.g. age, region, sex) and interpret the coefficient (loosely) as the "impact of schooling." This approach neither distinguishes cleanly between whether it is estimating the *partial* impact of schooling controlling for learning or the *total* impact of schooling, with the average impact of schooling via learning included into the estimate of schooling, essentially through omitted variables bias, nor can it distinguish between the impact of schooling and learning and hence education.

We begin by running what would be the standard OLS regression which uses years of schooling and control variables to estimate the "impact" on outcomes. We do this for our three outcomes across our 129 survey rounds (67 survey rounds for women's empowerment)[9] and report the random effects meta-analysis weighted results in Table 3. This yields the "total impact of schooling" (the impact of schooling plus the impact of the learning produced by schooling) but *not* the total impact of education. Column (1) shows the results from this OLS that includes schooling but not literacy. Coefficients are scaled by six to represent the impact of completing six years of schooling, an approximation for primary school completion.

Completing six years schooling (or, roughly, primary completion), with this specification,

[9] The empowerment module was not included in some survey rounds

is associated with a reduction in average fertility of one-third of a child, or a 10% reduction. In the child mortality regressions, primary schooling is associated with a reduction in child mortality of 2.3 percentage points; with an average child mortality rate in the sample of 10.7%, this is the equivalent of a 22% drop in child mortality. And, this approach shows that primary completion (not controlling for learning) is associated with a 0.146 standard deviation increase in the women's empowerment index.

We next estimate the OLS regression including literacy. As expected, the coefficient on schooling, which is now the *partial* impact conditional on literacy $\beta_S|L, Z$ is smaller, by about 20-45% as shown in column (2) of Table 3.

We also now have estimates of the impact of literacy, conditional on schooling. Going from illiterate to literate for a given amount of schooling completed ($\beta_{L|S,Z}$) is associated with a reduction in average fertility of 0.11. For child mortality, achieving literacy is associated with a 0.009 percentage point reduction in child mortality. Finally, for women's empowerment, literacy yields a 0.108 standard deviation increase in the empowerment index.

Now we are ready for the first big reveal. Remember that the equation for the total impact of schooling, including through the pathway of augmenting literacy was:

Total impact of years of schooling including impact on literacy (Z fixed) on outcome Y:

$$\Delta Y = \beta_{S|L,Z} * \Delta S + \beta_{L|S,Z} * (\Delta L/\Delta S) * \Delta S \tag{8}$$

If we express the regression of literacy on schooling as:

$$L^{i,c} = \alpha_L^c + \pi_L^c * S^{i,c} + \Theta_L^c * Z_L^{i,c} + \nu^{i,c} \tag{9}$$

Then we now have two distinct estimates of the total impact of completing six years of schooling (at current learning levels in country c). The first is to just not include measures of learning and rely essentially on "omitted variables bias" to estimate the total.

Table 3: Adding literacy to OLS regressions increases estimated impact of education by 13-32%

	OLS with schooling only	OLS with Schooling and Literacy	Ratio ($\beta_{S(1)}/\beta_{S(2)}$)	Ratio impact of education to typically estimated impact of schooling (($\beta_{S(2)}+\beta_{L(2)}$)/ $\beta_{S(1)}$)
	(1)	(2)	(3)	(4)
Fertility				
Schooling	-0.327	-0.257	0.787	1.131
	0.018	0.020		
Literacy		-0.107		
		0.015		
Linear combination		-0.370		
		0.018		
Reduction in Child Mortality				
Schooling	0.023	0.017	0.726	1.185
	0.001	0.001		
Literacy		0.009		
		0.001		
Linear combination		0.028		
		0.001		
Women's empowerment				
Schooling	0.143	0.081	0.567	1.321
	0.013	0.014		
Literacy		0.104		
		0.013		
Linear combination		0.189		
		0.014		

Random effects meta-analysis estimates from 129 DHS surveys in 54 countries. Regressions contain controls for age, age squared, age cubed, wealth (using the DHS wealth index), a rural/urban dummy, and dummies for regions within countries. Schooling coefficients have been scaled by six to reference the effect of completing primary schooling; literacy coefficients have been scaled by two to reference the effect of going from illiterate to literate on the 0 (None),1 (Partial),2 (Full) scale.

That is, if there is a structural model such that:

$$Y = \alpha + \beta X + \gamma W + \epsilon \tag{10}$$

$$X = \pi * W + \nu \tag{11}$$

Where the equation linking X and W in the standard OVB (omitted variable bias) set up is just a reduced form and does not necessarily represent a causal model and the partial correlation coefficient π could be induced because X caused W, W causes X or both are jointly caused/correlated with a third variable. The OLS estimate of the coefficient on X, β, in an estimating equation that excludes W is affected by the omission of W and converges to: $\beta^{OLS(OVB)} \xrightarrow{p} \beta + \pi * \gamma$

In our case if we interpret (provisionally) π as the causal impact on literacy of an additional year of schooling then we have our first estimate of the "total impact of six years of schooling" for each country c. Estimating the OLS regression including schooling and not literacy essentially relies on OVB to estimate the total impact of schooling (at current learning levels) through its various causal pathways, including in β the impact schooling has by way of learning: $(\beta^c_{S|Z} * 6)$.

The second estimate of the "total impact of six years of schooling" is able to articulate the learning pathway separately. When we run our OLS with the literacy variable, we no longer have OVB for literacy and thus we have separate estimates for the partial impact of schooling, β, and the impact of literacy, γ. We also run the regression of literacy on schooling, from equation **??**, and we recover our estimate of π. From this, we have our second estimate of the total impact of schooling: $\beta^c_{S|L,Z} * 6 + \beta^c_{L|S,Z} * \pi^c_L * 6$

Basically all of the previous literature has relied on the OVB bias estimates of the impact of schooling. Turns out, our estimates of the total impact of schooling via OVB and our estimates of the total impact of schooling via the pathway are practically identical. The correlation between the two across countries is .998, .999 and .998 for child survival, fertility, and empowerment.

The advantage of having and using the literacy data is not that we have different estimates of the total impact of schooling (at current learning levels). We can produce exactly the same estimates. But there are three large advantages to having the literacy data.

First, everything that has ever been said empirically that conflated, explicitly or implicitly, the "impact of schooling" with the "impact of education" based on schooling data alone is wrong, and is wrong by an amount we can calculate with our estimates exactly.

The impact on outcomes of six years of schooling at the existing learning profile in country c (π_L^c) is:

$\beta_{S|L,Z}^c * 6 + \beta_{L|S,Z}^c * \pi_L^c * 6$

Whereas we can *define* by education the achievement of a given level of learning (say, full literacy) which, given our 0,1,2 scaling of literacy is:

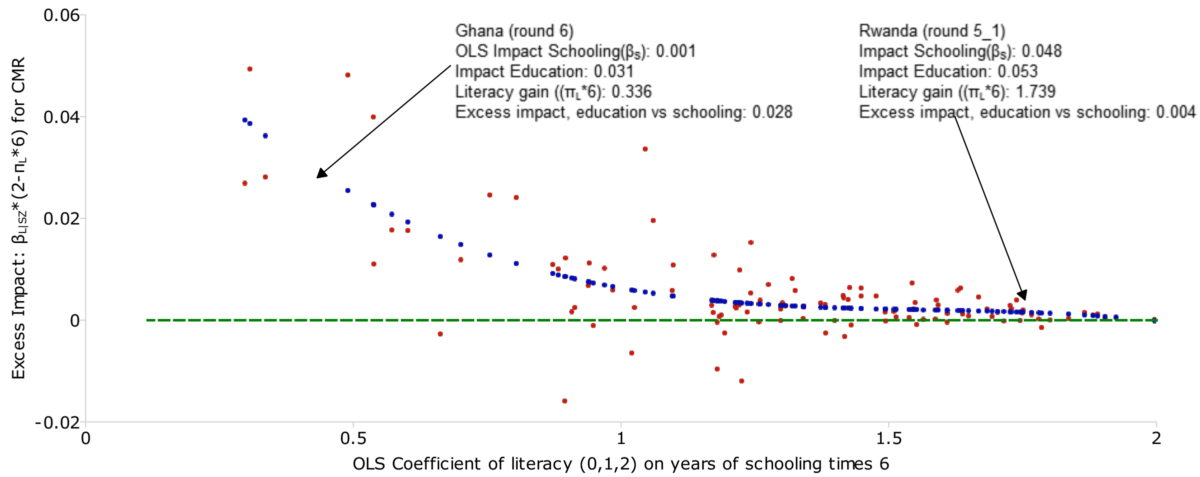
$\beta c_{S|L,Z} * 6 + \beta_{L|S,Z}^c * 2$

Hence in country c the excess of the impact of education (six years, reaching full literacy) and impact of schooling (six years at the actual learning profile) is:

$$\text{Excess of impact of education over schooling} = \beta_{L|S,Z}^c * (2 - \pi_L^c * 6) \tag{12}$$

Second, given the massive attention that the question of causal identification and particularly causal identification using randomized methods, receives these days, it is important to stress that in this case we know, *ex ante* an RCT is not particularly helpful for two reasons. One, even a causally clean estimate of the partial or total impact of schooling does not help resolve the question of the impact of *education*. That is, a "treatment" that increased years of schooling (at whatever level of learning that schooling produced) in a target population would provide at best a cleanly identified estimate of the total impact of schooling that allowed causal attribution. But this still relies on the causal impact being mediated by the level of learning achieved and hence is of no help in identifying the impact of education or in decomposing how much of that impact was due to a "pure" schooling impact versus how

Figure 5: Impact of education on child survival is larger than the impact of school by an amount that depends on how much literacy schooling produces



Source: Authors' calculations based on DHS microdata.

much was mediated by the causal impact of the additional schooling on learning.

Two, in this case we know *for sure* there cannot be "external validity" of estimates of the total impact of schooling across countries. Since observed $\Delta L/\Delta S$ or π_L^c differs by an order of magnitude across countries (see figure 4) and the formula for the total impact of schooling in country c involves that term (see equation 12) rigorous estimates of the total impact of schooling *must* differ across countries. For instance, take the estimates of meta-analysis average partial impact of schooling (.081) and partial impact of literacy (.104) on women's empowerment in table 3. The estimates of π_L^c in Figure 4 differ from a 10th percentile of .66 to a 90th percentile of 1.87. Therefore even if the partial impacts ($\beta_{S|L,Z}$ and $\beta_{L|S,Z}$) were the same in two countries (at the average value) the total impact of schooling–which an RCT should correctly estimate–would be different by over 50 percent ($.12 = 081 + (.104/2) * .66$ in the low learning country versus $.18 = .081 + (.104/2) * 1.87$). Whereas the impact of education would be, by construction, exactly the same.

For instance, Duflo (2000) and Breierova and Duflo (2004) exploit variation from a nationwide school construction program in Indonesia to recover causal estimates of the impact of increased parental schooling on child mortality. But this causal impact of schooling is mediated by the average efficacy of schooling in producing literacy in that context and cannot be used to estimate the impact of schooling in either lower (e.g. Ghana, Nigeria) or higher (e.g. Bolivia) learning countries. Similarly, introduction of Universal Primary Education in Nigeria in 1976 and Uganda in 1997 provided researchers with a source of exogenous change; based on this analysis Osili and Long (2008) suggests that increasing female schooling by one year reduces early fertility in Nigeria while Keats (2014) finds that women in Uganda with more schooling prefer to have fewer children, delay having their first child, and reduce overall fertility at any age, while investing more in their children's health. Similarly recent work in Ghana of a program to extend girl's enrollment in school estimates the impact on XXX **?**. But without having and using additional information on what was learned in school these studies do no more to estimate impact of *education* than do observational studies and *cannot*, in and of themselves, provide evidence with any external validity about the impact of schooling independently of estimates of learning profiles.

This helps us distinguish between "external validity" of estimates of what is conceptually the same, e.g. comparing the impact of literacy on female empowerment across countries, and comparing "impact" estimates that are known to have different causal pathways, e.g. the total impact of schooling including the impact mediated by learning. An advantage of having multiple rounds for countries is that we can decompose the observed variance in estimates that is between countries (using the average estimate for each country) and the variation across rounds in the same country. For instance, take the coefficient of literacy on fertility. While the estimate from Table 3 show the *average* coefficient (using standard error weights) of -0.107 is estimated with considerable precision, a standard error of .015 (t-statistic over 7), there is considerable variation across countries. The standard deviation across the country means is .136 and this is about 60 percent of the total variation in literacy coefficient estimates across countries.

Third, these estimates provide some indication of the channels whereby schooling has its impact on outcomes, which is central to informed decisions about priorities in the education sector. As we pointed out above all of the existing evidence on the impact of schooling is completely uninformative on key decisions facing policymakers in the sector about the allocation of priorities between extension of schooling and investments and actions to improve the learning profile.

3.4 Understanding the consequences of measurement error on decomposing the relative impacts of schooling and learning

Before launching into reports of the IV estimates it is worth honing intuition of what we should expect. We have two simple textbook challenges that anyone with a moderate exposure to econometrics should understand: measurement error and omitted variables bias. Here we just want to hone intuition of what happens when those combine.

Start with the simple bivariate example of measurement error and assume the true model is:

$y_i = \beta_L * L_i + \epsilon_i$

but that the variable L is measured with error so that what is observed is L*:

$L_i^* = L_i + \nu_i$

Then we know that OLS estimates will suffer from attenuation bias (will be closer to zero than the true value) and we have an exact formula for the degree of measurement error bias:

$$\beta_L^{OLS} \longrightarrow \beta_L * \frac{\sigma_L^2}{\sigma_\nu^2 + \sigma_L^2} \tag{13}$$

Which has a very simple and clear intuition: the OLS coefficient is the true coefficient times the ratio of signal to the noise plus true signal in the variable. This implies that as noise goes to zero ($\sigma_\nu^2 \to 0$) OLS converges to the true coefficient whereas if the signal goes to zero ($\sigma_L^2 \to 0$) OLS converges to zero irrespective of the value of the true coefficient. This also implies that if there is an IV estimate that is consistent for (converges in probability to) the true coefficient the ratio of the two is itself an estimate of the signal to noise plus signal ratio:

$$\beta_L^{IV} / \beta_L^{OLS} \longrightarrow \frac{\sigma_L^2}{\sigma_\nu^2 + \sigma_L^2} \tag{14}$$

But what happens if there are two variables and one has measurement error? We know that we can recover the exact OLS estimate with two variables with repeated OLS, so that if the true model is:

$y_i = \beta_S * S_i + \beta_L * L_i + \epsilon_i$

Then we can regress Y on L and S on L and regress the residuals of those regressions and recover numerically exactly the coefficient estimate on S from a multivariate regression of Y on S and L.

$(y_i - \hat{\beta}_L * L_i) = \beta_S * (S_i - \hat{\pi}_L * L_i) + \eta_i$

Where "$\hat{\beta}$" is the OLS estimate.

This is the intuition behind the formula for omitted variables bias, as the simple bivariate regression of y on S is the equivalent of the above procedure but where β_L and π_L are forced to equal 0 rather than their OLS values. But the simple bivariate case is just a special case, think of estimating the following equation:

$$(y_i - \tilde{\beta}_L * L_i) = \beta_S * (S_i - \tilde{\pi}_L * L_i) + \eta_i \tag{15}$$

where the *tilde* represents any arbitrary value. If $0 < \tilde{\beta}_L < \beta_L$ then the estimate of S will suffer from what we call "partial omitted variables bias." Suppose that L is included in the regression (not omitted) but measured with error (and for now that S is not measured with error) then repeated least squares will not produce an estimate for S that converges to the true value but will, if π_L is positive, be *overestimated*. When L is measured with error, *part* of its signal is omitted, and with omitted variables bias, *some* of that signal is heaped onto S. The conclusion is that OLS multivariate regression of y on S and L when L has measurement error and S and L are correlated will produce estimates on S ($\beta_{S|L}$) that are too large and estimates on L ($\beta_{L|S}$) that are too small as the estimates of L will suffer from the standard attenuation bias and estiamtes of S will suffer from partial omitted variables bias induced by the attenuation bias from measurement error in L.

The consequences of this for attempting to decompose the total impact of education on outcomes through the schooling conditional on learning and learning conditional on schooling as in equation 16 can be severe.

$$\text{Fraction of impact of education due to learning} = \frac{\beta_{L|S} * \Delta L}{\beta_{L|S} * \Delta L + \beta_{S|L} * \Delta S} \tag{16}$$

Suppose the true value of the "pure" impact of schooling was zero and the entire impact of education was through learning ($\beta_{S|L=0}, \beta_{L|S} > 0$) so the true fraction in equation 16 is 1. But suppose the the ratio of noise to noise plus signal for the measure of learning was .5 and the correlation of S and L was .9. Then the estimated fraction of the impact of education due to learning would be only .065 even though we know (by construction) the true fraction is 1.

Our situation is the perfect storm of multicollinearity and measurement error. One part of the storm is that we can see from the data and from the OLS results that schooling and learning are highly correlated and the covariance of the coefficient estimates is hence very high. XXX. Hence any measurement error in school or literacy will very strongly affect *both* estimates, making one too low (attenuation bias) and the other too high (partially omitted variable bias) and hence the estimate of the ratio doubly wrong.

The second part of the perfect storm is that there are good reasons to believe that assessing whether a woman can read a single arbitrary sentence is a very noisy measure of literacy and literacy is a very noisy proxy for learning, much moreso than self-reports by women of their years of schooling. In this case the *relative* measurement error will strongly affect the attempt to decompose the causal channels.

This is why we need to use estimation techniques that can potentially reach adequate estimates even in the presence of measurement error.

3.5 Using Instrumental Variable Estimation Techniques for Literacy and Schooling

Measurement error in literacy could come from at least four main sources. First, we are using one single question to assess an entire domain and hence perhaps a woman could have read three other sentences but not the particular one presented. In most assessments of say, math ability, a respondent would be asked several questions in a particular domain (e.g. several problems using division) and then these scores would be weighted to produce an overall score. So this 0,1,2 indicator is noisy relative to a woman's true reading ability or literacy. Second, the literacy tests are administered by enumerators who must judge how well a respondent could or could not read. These judgments are not perfectly consistent across enumerators. Third, a literacy test captures only one aspect of learning – reading simple sentences – and so there is measurement error in using this as a proxy for learning, as other aspects, such as numeracy, are not included. Finally, categories are truncated such that those at the higher end of the literacy test are all grouped in the same top category, so there is error as there is no variance in the higher ability levels.

To instrument for literacy we take advantage of the fact the sampling in the DHS is clustered. In order to save time and expense the sampling frame first chooses small geographic areas (often census tracts) as enumeration areas (EA) and then samples women within that EA. Women in the same EA are therefore geographic neighbors. We create a "enumeration area leave-out-mean" (EALOM) for each individual i which is the average literacy level of everyone else in enumeration area j except respondent i (this way the respondent's literacy level is not included in calculating the instrument):

$$\bar{L}_{i,j} = \sum_{k=1}^{N_j, k \neq i} L_{i,j}/(N_j - 1) \tag{17}$$

where $L_{i,j}$ is the literacy of the i^{th} woman in the j_{th} EA and N_j is the total number of respondents in enumeration area j.

The measures of schooling might also contain measurement error, which could come from women mis-reporting their schooling attainment, enumerators mis-recording it, or other sources. We use a similar approach and use the schooling completion EALOM as an instrument for schooling.

A valid instrument must meet two criteria.

First, the instrument must be correlated with the variable being instrumented, which is the "inclusion" criteria. Failure of this condition leads to imprecise IV estimates and incorrect standard errors. It is plausible that a respondent's literacy level would be correlated with that of her neighbor's as they may have gone to the same school or similar quality schools or by way of geographic sorting of similarly educated people into neighborhoods. A weak instrument is indicated by a low R-Squared or F-test on the instrument in the "first stage" regression, which is the woman's literacy on her EALOM. The standard tests for instrument inclusion show this instrument performs adequately as the F-statistics are typically above 10.

The second criteria is that the instrument must satisfy the exclusion restriction; the instrument must not have a direct causal impact on the outcome of interest and can therefore be properly excluded from the original equation. This is discussed further in Section 3.7.

3.6 Results from Using Instrumental Variable Techniques to Estimate Impact of Literacy and Schooling

The results when the EALOM instrument is used only for literacy only are in column (1) of Table 4 and illustrate the sensitivity of the estimates of schooling and literacy to method. If measurement error is addressed through IV only for literacy the estimate of the literacy impact goes up substantially and hence, due to the multicollearity discussed above, drives the estimate of the coefficient of schooling down leading to "wrong signed" estimates of the partial schooling effect ($\beta_{S|l,Z}$) for all three outcome variables.

Table 4: Method matters: Instrumenting for schooling and learning yields an estimated impact of education (schooling + learning) 3 times higher than that estimated by OLS

	IV for Literacy (1)	Ratio of IV estimate in column (1) to OLS estimate (column (2) of Table 3) (2)	IV for Literacy and Schooling (3)	Ratio of IV estimate in column (3) to OLS estimate (column (2) of Table 3) (4)
Reduction in child mortality				
Schooling	-0.040	-2.350	0.036	2.108
	0.011		0.009	
Literacy	0.098	10.916	0.020	2.183
	0.018		0.009	
Education (S and L)	0.044	1.607	0.073	2.654
	0.004		0.007	
Fertility				
Schooling	0.816	-3.172	-0.527	2.047
	0.122		0.122	
Literacy	-1.818	16.989	-0.529	4.939
	0.192		0.116	
Education (S and L)	-0.764	2.064	-1.206	3.261
	0.050		0.090	
Women's empowerment				
Schooling	-0.427	-5.271	0.125	1.541
	0.097		0.106	
Literacy	0.937	8.975	0.474	4.546
	0.167		0.123	
Education (S and L)	0.431	2.283	0.595	3.147
	0.053		0.078	

Random effects meta-analysis estimates from 129 surveys in 54 countries. Regressions contain controls for age, age squared, age cubed, wealth (using the DHS wealth index), a rural/urban dummy, and dummies for regions. Schooling coefficients have been scaled by six to reference the effect of completing primary schooling; literacy coefficients have been scaled by two to reference the effect of going from illiterate to literate on the 3-point scale.

Column (3) of Table 4 are estimates that use the EALOM as an instrument for both schooling and literacy. These estimates have three important features.

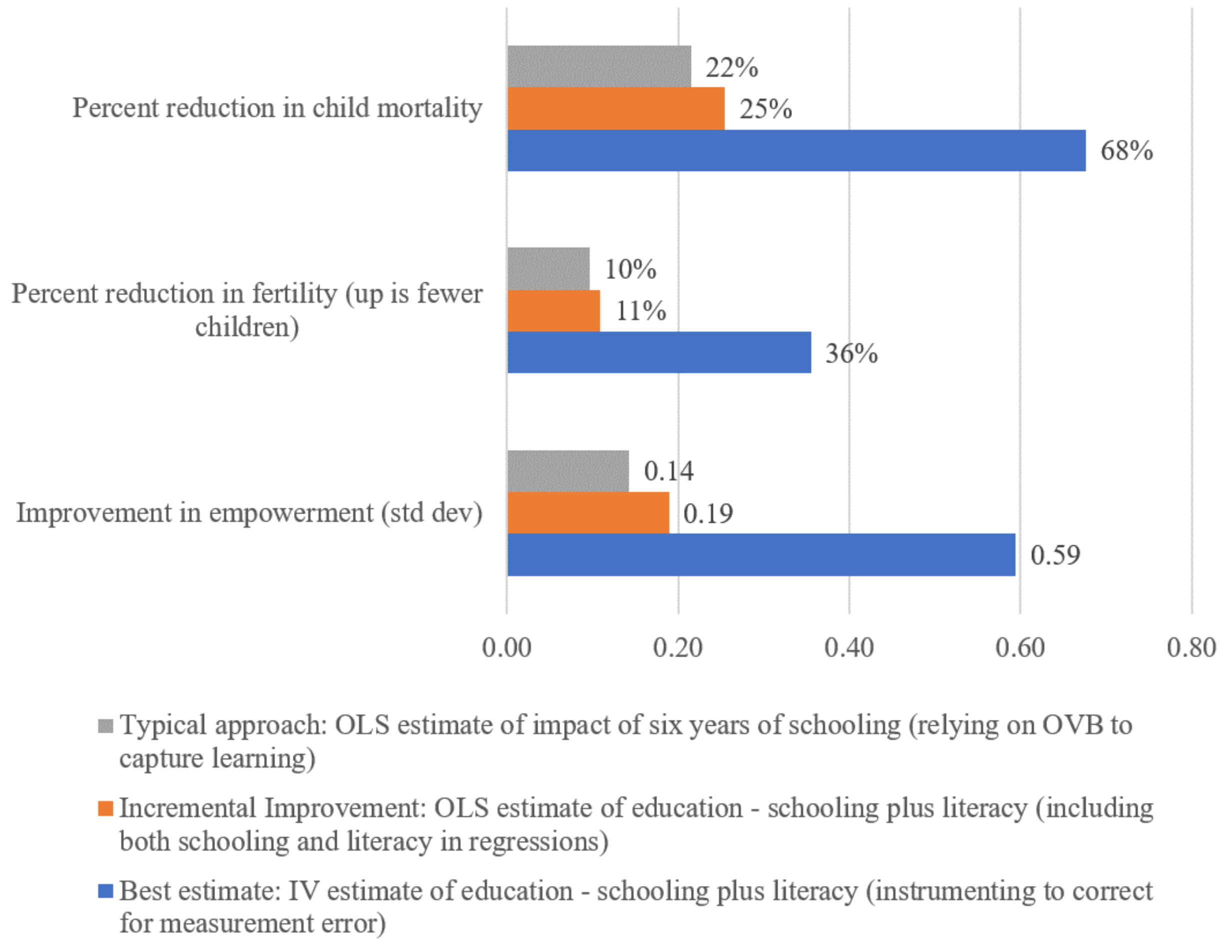
First, the estimate of the impact of basic education is much higher than using OLS with both schooling and literacy, which was in turn higher than the estimate of the impact of schooling when learning is not included. The impact of basic education increases child survival by .073, which given that child survival in the sample was already .89 (Table 2), implies a two thirds reduction in child mortality. The standard estimate of the impact of schooling from cross-tabs or regressions on schooling alone is .023 (Table 3, column 1), three times smaller.

For fertility, the total impact of basic education using IV is 1.2, off an average of 3.37 (Table A.2). The cross-tab/simple regression approach would have estimated an "impact of schooling" of only .33 (3, column 1). Basic education increases female empowerment by .59 of a standard deviation, compared to the estimated .14 impact of schooling from the traditional approach.

The most important finding of the paper is that using data on both schooling and learning and using instrumental variables techniques to account for the attenuation bias from measurement error suggests that the impact of women's education on the well-being of women (child health, fertility, empowerment) *three to four times* bigger than the standard methods suggest.

Second, it appears that the estimates of the impact of literacy are much more affected by measurement error than are estimates of schooling, so the proportion of education's impact coming from literacy versus the partial schooling effect are much larger using IV as the attenuation bias was worse for literacy than schooling (Table 4 vs. Table 3. In Table 4 we see the ratio of the IV to OLS estimates is 4.94 for fertility, 4.54 for female empowerment and 2.18 for child survival. In contrast, this ratio of IV to OLS is 3.26, 3.15 and 2.11 for fertility, empowerment, and child survival. Recall from equation 14 that the ratio of OLS to IV estimates (which is the inverse of the ratios above) is an estimate of the signal to

Figure 6: Typical estimates of the impact of schooling *underestimate* the impact of women's *education* on outcomes by a factor of 3 to 4



Source: Random effects meta analysis estimates from 129 surveys in 54 countries signal plus noise ratio. This suggests that the variability in the DHS literacy indicator is only about 20 percent signal (1/4.94) for the purpose of measuring the impact on fertility, and 45% (1/2.18) for survival.

This might seen like an excessive degree of measurement error, but two points. First, ratios of OLS to IV this small are not uncommon. **?** show the ratio of OLS to IV estimates of the impact of household consumption per person on child school enrollment were .15 for Pakistan, .16 for Indonesia, and .46 for Nepal. Second, the measurement error is not the measurement error of literacy as a measure of literacy alone but also of literacy as a proxy for all other learning that may affect the outcomes. One can easily imagine the signal of a simple 0,1,2 indicator of literacy is associated with, but only weakly, the extent to which learning affects fertility choices.

Third, the implication of the IV estimates is that the fraction of the impact of education that is due to increased learning is much higher than the OLS results would suggest. This then returns us to the key policy question on which the existing literature provides limited guidance. Suppose there was a set of policy actions that increased the literacy gain per year from its current level to a level such that, on average, a woman with grade 6 complete was fully literate ($\pi = 2/6 = .333$; e.g. each year of schooling increased literacy by 0.333 on three-point scale). Contrasted with that there might be a different set of actions that increased a woman's time in school by 2 additional years. Using the random effects weighted average coefficients[10] for outcome gains plus the learning profile to calculate for each country:

$$\text{Gain in outcome from increased literacy for six years of primary school} = (.333-\pi^c)*\beta_{L|S,Z})*6 \tag{18}$$

Alternatively, we can calculate the gain from two more years of schooling at the existing pace of learning as:

[10] We use the average rather than the country by country estimates for the reasons that become clear below.

Table 5: The impact on outcomes of raising literacy over the six years of primary schooling compared to the gains from expanding schooling by two additional years

Fertility	Fertility Lit gain	Fertility Sch gain	CS	CS lit gain	CS sch gain	Emp	Emp lit gain	Emp sch gain
Average	-0.164	-0.297		0.006	0.016		0.149	0.150
Low learning								
Ghana(7)	-0.429	-0.209		0.016	0.013	GH6	0.394	0.068
Nigeria(6.2)	-0.369	-0.229	Nigeria(6.2)	0.014	0.014	Nigeria(6.2)	0.331	0.089
Moderate								
Egypt(6)	-0.280	-0.258	EG6	0.010	0.015			
India(5)	-0.133	-0.307	IA5	0.005	0.017	IA5	0.120	0.160
High								
Rwanda(6.2)	-0.094	-0.320	RW6_2	0.004	0.017	Rwanda(7)	0.087	0.170
Cambodia(6)	-0.078	-0.326	KH6	0.003	0.017	KH6	0.070	0.176

Notes: Ghana(7) indicates Ghana, round 7 of DHS.
The gain for outcome i from a learing profile of .333 gain per year versus existing country level for primary school is: $\beta^i_{L|S,Z} * (.333 - \pi^c) * 6$
The gain for outcome i from two additional years of schooling is: $\beta^i_{S|L,Z} * 2 + \beta^i_{L|S,Z} * (\pi^c) * 2$

$$\text{Gain in outcome from two more years of schooling at existing learning profile} = \beta_{S|L,Z} * 2 + \beta_{L|S,Z} * \pi^c * 2 \tag{19}$$

Table 5 shows the results. On average a country that increased its learning profile to this moderate and achievable level (this level is achieved by countries like Bolivia, Honduras, Morocco) would achieve as much gain to female empowerment as expanding schooling by two years. The gain from a steeper learning profile for fertility is about half as large as schooling expansion, and for child survival 37 percent as large. These are averages, though, and we know from the learning profiles there is no "external validity" for these results; they vary widely across countries. In low learning countries like Ghana the fertility gain would be twice as large from improving learning of women in primacy school than expanding the years completed. In a moderate-learning-level country such as Egypt, the gains from literacy versus schooling would be about the same. Obviously for countries where learning is already high, like Cambodia, the gains from moving towards a moderate learning profile are small.

So the choice of investing in improving learning versus expanding schooling (beyond what countries consider a basic right–we are in no way suggesting anything less than universal primary completion is an acceptable policy or goal) depends on relative costs. If, as many are suggesting, there are massively cost effective ways of improving learning then in low learning environments it may be that investments in improving learning are *orders of magnitude* more cost effective than spending that expands attendance in low learning schools.

3.7 Trade-offs from Using Instrumental Variable Techniques

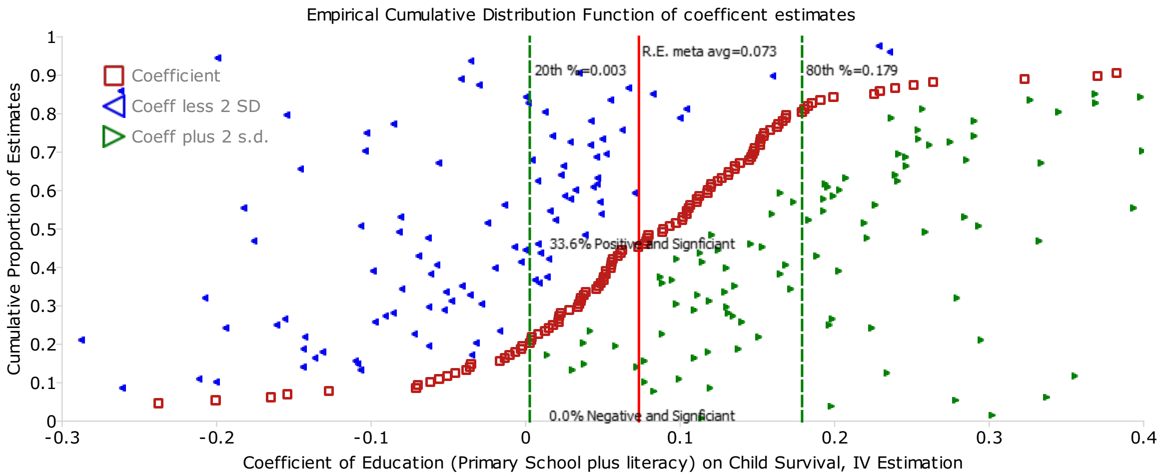
The IV estimates come at an econometric sacrifice. While the estimated effects are much larger using instrumental variables, the precision is lower in each country and the variation in the estimates across countries increases as well. Figure 7 shows the empirical cumulative distribution function of the IV estimates of the impact of education and the two standard error confidence intervals around those estimates. As in Table 4 the IV estimates of the impact of education are much higher than their OLS counterparts (the weighted average is .073 vs .028). But Figure 7 shows that the variance of the IV estimates in each country and the variability across countries is large. While the t-test for the null hypothesis that the mean of the impact of education is zero is around 10 because the standard error of the estimate of this average from meta-analysis is low, the variability of estimates is large. The 20th percentile of the estimates is essentially 0 and the 80th percentile is .179.

Figure 8 shows the empirical cumulative distribution function for the 129 estimates of the impact of schooling (β_S) and impact of education estimated with OLS and with IV for child survival (these same graphs for fertility and female empowerment are in an appendix. [11]. As we saw in section 3 the impact of education (estimated with OLS or IV) is consistently higher than the impact of schooling by an amount the depends on learning.

Figure 9 compares the empirical cdf of the IV results and shows for each country the OLS

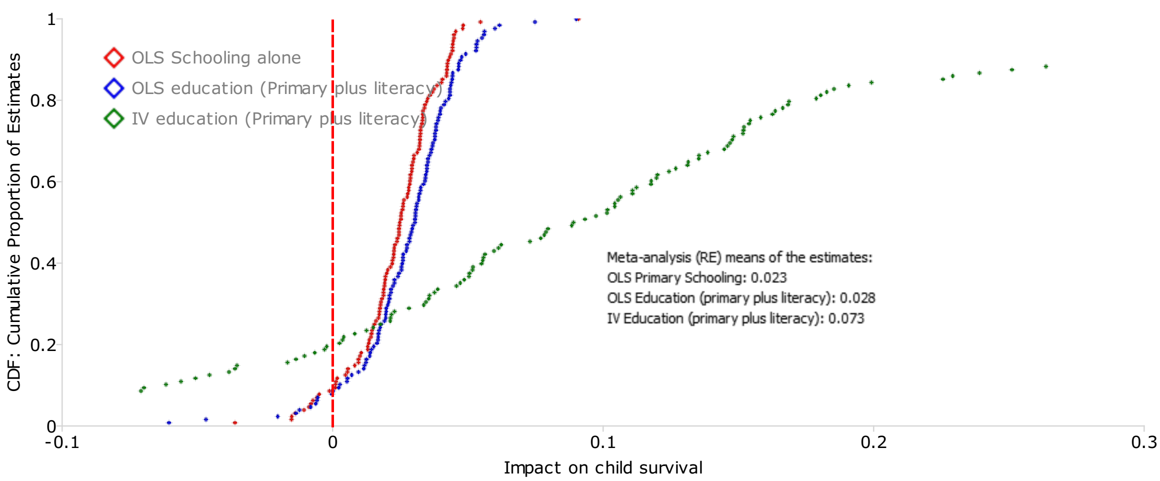
[11] In this graph each of the estimates is sorted from lowest to highest and each "row" of the graph is the nth largest estimate and hence these are not for the same country. That is the 10^{th} largest estimate for OLS schooling, OLS education and IV education can be three different countries.

Figure 7: Empirical cumulative distribution functions of the IV estimates of the impact of education on child survival with standard error bounds



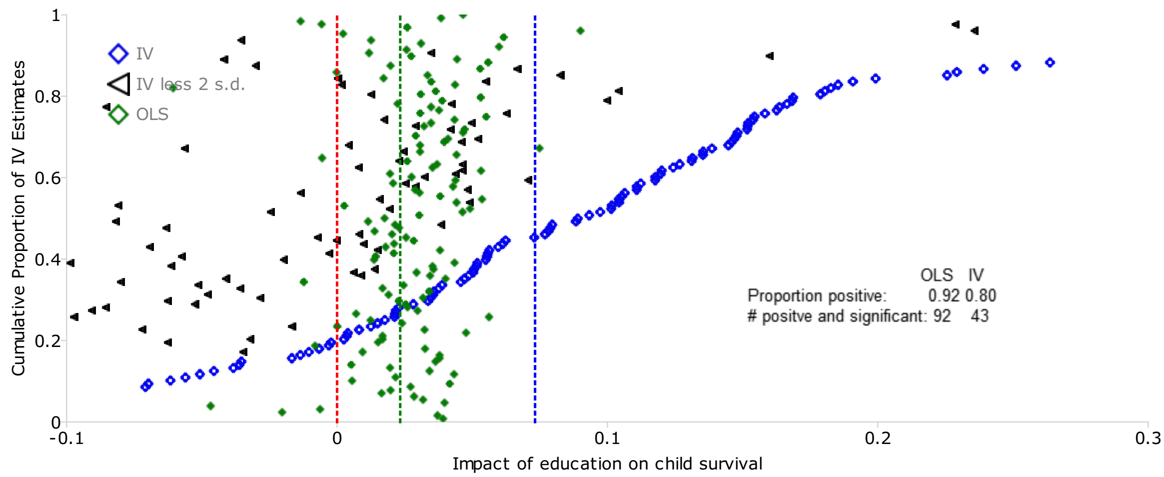
Source: Authors' regression estimates on DHS microdata for 54 countries.

Figure 8: Empirical cumulative distribution functions of the estimates of the impact of schooling (OLS) and education (OLS and IV) for child survival



Source: Authors' regression estimates on DHS microdata for 54 countries.

Figure 9: Comparing the level and precision of OLS and IV estimates of the impact of basic education on child survival



Source: Authors' regression estimates on DHS microdata for 54 countries.

estimate[12] and shows the IV estimate less 2 standard errors and hence the lower end of the confidence interval that an individual IV estimate is greater than zero. This graph shows that, while the average of the IV estimates is much higher there are many countries where the IV estimate is less than the OLS estimate and in fact, nearly 20 percent of IV estimates are negative. And the greater imprecision implies that, while for the OLS estimates 93 of 114 were positive and statistically significant for IV only 43 of 114 are.

A related cost of the much higher imprecision of the IV estimates is that this imprecision, combined with multicollinearity and hence high covariance of the estimates of schooling and literacy is that the decomposition of the total impact of education into the schooling channel and the literacy channel is unreliable country by country. Essentially when the estimated impact of literacy is high this induces a low estimate of schooling and hence the estimates of schooling and literacy are negatively correlated across countries for each outcome. This includes driving the effects to negative values. For instance, 24 percent of the IV estimates of the partial effect of schooling ($\beta_{S|L,Z}$) on child survival are negative and 31 percent of the

[12] Hence, in contrast to Figure 8, here each "row" on the vertical axis is a country and they are sorted by the magntude of the IV estimate.

estimates for fertility are positive, both of which are implausible. So our confidence in the estimates of the relative impacts of literacy and schooling comes from the aggregation across many countries/periods that smooths over this induced variability in the individual terms.

Finally, since our estimates are "just identified" we cannot test the exclusion restriction. But it is possible that the level of literacy of other women in a woman's neighborhood directly affects her outcomes. For instance, it is possible that having more literate neighbors leads to women having better health information as they learn from their neighbors (for some weak evidence to this effect see Dearden, Pritchett and Brown (2004)). Or, it is possible that living in a neighborhood with more literate women, who themselves have fewer children, causes women to reassess their own preferences for children. Two points. There is literature suggesting that true peer effects are often overstated **?** as many findings of peer effects are just the result of the exclusion of local variables and it is nearly impossible to disentangle "locality" effects and peer effects. **?**, using the same type of EALOM instrument with different datasets found similar results in rural and urban areas where one might *a priori* suspect the peer effects are smaller as one is much much less likely to be in direct social contact with EA neighbors in urban than rural areas which suggests (weakly) that the peer effect channel is not the dominant case of the IV results. But more deeply, if there are true peer effects then this makes our IV estimates of the direct effect of a woman's literacy on her own outcomes "biased" but this would mean the total aggregate effect of raising women's schooling on outcomes is higher than the sum of the effects for individual women. If there are true peer effects then there are positive externalities of schooling (of at least some geographic scope) and the impact on say, child survival, of increasing the literacy of one woman is the impact on her own outcomes plus the sum of the impact she has on all her connected peers. So if our IV estimates of the direct effect is "too high" because of peer effects the direct effect is "too low" as an estimate of the total impact of increased education.

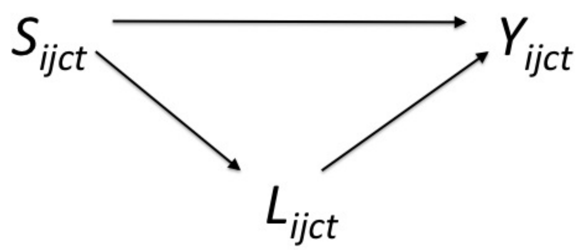
4 Individual literacy as a mediator between schooling and outcomes

So far in this paper we have been explicitly using the word "impact" loosely as a synonym for a variety of the estimates of the difference in outcomes for women with different levels of schooling and literacy and we do not imply this correctly identifies a causal impact. These procedures from cross-tabs to OLS regressions to IV regressions do produce facts. But we do need to address whether these regression facts are consistent with different causal models. In this section we examine two distinct models of the relationship between schooling, S_{ijct} and learning, L_{ijct}, where subscripts denote inidvdual i in region j and country c in period t. In model we have been using up to this point, the literacy of individual women mediates the relationship between schooling and outcomes for individual women (Figure 10). A different approach is that average literacy rates conditional on schooling in a given population are used as a metric of school quality, which we treat as a potential moderator of the return to schooling.

Figure 11 is a more sophisticated version of the simple cross-tabs in Figure 1 and compares the outcomes across women who completed various grades and who have varies degrees of literacy. The solid line is all women, irrespective of their literacy, and hence traces out the association of outcomes with schooling. The short dashed line shows results just for women who had no literacy and the line is much less steep. The graphs in Figure 11 provide prima facie evidence for literacy mediating the return to schooling as defined in terms of child survival, fertility, and empowerment outcomes.

We now write our previous equations in the terminology of the treatment effects literature. The average treatment effect of schooling on a given outcome can be decomposed into the average causal mediation effect (ACME) of learning, and the average direct effect (ADE) of schooling unrelated to learning. Following **?**, we can estimate the ACME in a linear regression framework as the coefficient, $\beta_{L|S,Z}$ on the mediator in an outcome regression

Figure 10: Causal pathways directly from schooling and mediated via literacy



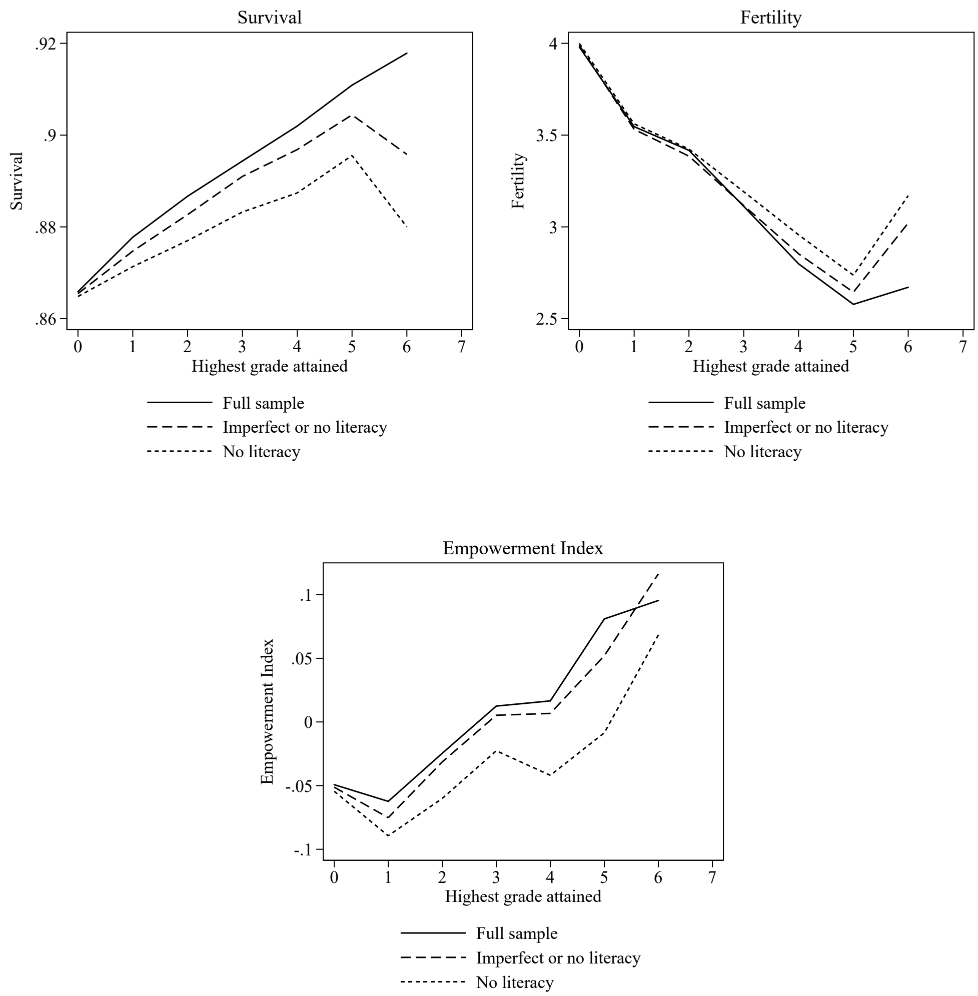
controlling for schooling:

$$L_{ict} = \alpha_L + \pi_L^{ct} * S_{ict} + \theta_L Z_{ict}^L + u_{ct} \tag{20}$$

$$Y_{ict} = \alpha_Y + \beta_{S|L,Z} * S_{ict} + \beta_{L|S,Z} * L_{ict} + \Theta_Y Z_{ict}^Y + v_{ct} \tag{21}$$

While this approach to testing the causal mediation is widespread in the social sciences, it rests on a set of assumptions which **?** refer to as "sequential ignorability." The first element of sequential ignorability is a standard exogeneity assumption about schooling, known in the treatment effects literature as the "ignorability of treatment." The second step in sequential ignorability is an assumption that the mediator is ignorable conditional on the treatment status. These assumptions cannot be directly tested, or controlled for, even in an RCT context. That is, even if an RCT had assigned people to a treatment group that increased their participation in schooling this, in and of itself, would not be sufficient to recover estimates of the ACME and ADE separately. Therefore we follow the method proposed by **?** to assess the sensitivity of our results to violations of these assumptions. We define ρ as the correlation between the errors in equations (20) and (21). **?** show how for a given estimate

Figure 11: Does literacy mediate the return to schooling?



Note: Lines represent the sample average for each outcome at each level of schooling, averaging over all survey years.

of $\beta_{L|S,Z}$ the true ACME depends on ρ, and suggest reporting cut-off values of ρ at which one can no longer reject the null that the ACME is zero.

These sensitivity tests are only currently possible for OLS, not IV, so we revert to our OLS model which includes both schooling and literacy. Table 6 shows results from these pooled regressions. The scaled coefficients for schooling and literacy and for the total impact are very close to the meta-analysis random effects standard error weighted results reported in Table 3[13]. In table 6 scaling the coefficient on S_{ijct} to reflect primary schooling (six years) produces estimates of 2.1 percentage point reduction in child mortality, similar to the reduction of 1.7 percentage points estimated through the meta-analysis techniques. Similarly, scaling L_{ijct} by two to reflect going from illiterate to literate reduces child mortality 0.8 percentage points, comparable to the 0.9 percentage points estimated through the meta-analysis. Across the three outcomes, these simple regressions suggest literacy accounts for about one-third to one-half of "education's" impact. And, not surprisingly for regressions with hundreds of thousands of observations, all of these are statistcally significant at any p-level.

Table 6 also contains a set of regressions that, instead of examining the data for individual women aggregates the data into averages across regions, producing 1471 distinct regional observations from the 54 countries and 129 country rounds[14]. Aggregation is a different way of dealing with measurement error, as if there is measurement error for each individual then smoothing across observations is a way of reducing the influence of measurement error by increase signal to noise. Interestingly, the literacy coefficients are 5.7 times higher in the regional regressions for fertility and 4.3 times higher for empowerment–which are simular to the IV to OLS ration in Table 4 of 4.9 and 4.5. But this is not observed for child survival. And the coefficients on schooling fall to statistically insignificant levels for all three outcomes.

Turning to the sensitivity tests, the **?** procedure produces a threshold for ρ above which a correlation between errors in the literacy regression and the outcome regression would

[13]This is reassuring and not surprising as the pooled OLS results are just a different weighting on the same observations.

[14]827 regions for empowerment, as the questions were not included in every country.

Table 6: Mediation analysis: Does literacy mediate the association between schooling and women's outcomes?

	Child Survival	Fertility	Women's Empowerment
	(1)	(2)	(3)
Woman level:			
S_ijct	0.003***	-0.049***	0.017***
	(0.000)	(0.001)	(0.001)
L_ijct	0.004***	-0.057***	0.048***
	(0.000)	(0.003)	(0.003)
W_ijct	0.014***	-0.285***	0.069***
	(0.000)	(0.002)	(0.002)
Scaled schooling (6 years)	0.021	-0.295	0.105
Scaled literacy (2 levels)	0.008	-0.114	0.097
Total effect of schooling at current learning levels	0.026	-0.372	0.170
Percent mediated at current learning levels	21.00	21.00	38.00
Total effect of Education (schooling + literacy)	0.029	-0.409	0.201
P-value (H0: ACME=0)	0	0	0
Sensitivity analysis: threshold ρ	0.03	-0.02	0.06
Obs. (women)	847653	1055701	386573
Country-year cells	129	129	129
Region level:			
S_ijct	0.002	-0.015	-0.029
	(0.002)	(0.022)	(0.026)
L_ijct	0.005	-0.324***	0.209***
	(0.006)	(0.075)	(0.080)
W_ijct	0.022***	-0.329***	0.058
	(0.003)	(0.035)	(0.039)
Scaled schooling (6 years)	0.012	-0.088	-0.171
Scaled literacy (2 levels)	0.011	-0.648	0.417
Total effect of schooling at current learning levels	0.019	-0.518	0.106
Percent mediated at current learning levels	37.00	83.00	262.00
Total effect of Education (schooling + literacy)	0.023	-0.736	0.246
P-value (H0: ACME=0)	0	0	1
Sensitivity analysis: threshold ρ	0.02	-0.11	0.09
Obs. (regions)	1471	1471	827
Country-year cells	129	129	129
Enumeration area (EA) level:			
	0.005***	-0.066***	0.033***
	(0.000)	(0.004)	(0.004)
L_ijct	0.004***	-0.113***	0.084***
	(0.001)	(0.013)	(0.012)
W_ijct	0.016***	-0.419***	0.108***
	(0.001)	(0.006)	(0.006)
Scaled schooling (6 years)	0.028	-0.397	0.199
Scaled literacy (2 levels)	0.008	-0.226	0.168
Total effect of schooling at current learning levels	0.032	-0.543	0.307
Percent mediated at current learning levels	15.00	27.00	35.00
Total effect of Education (schooling + literacy)	0.035	-0.623	0.366
P-value (H0: ACME=0)	0	0	0

overturn significant results. In the bottom row of each part of the table, it can be seen that the ρ values are quite small, suggesting that any substantial positive correlation of the error terms would indicate that the ACME is statistically indistinguishable from zero.

These sensitivity tests apply only to the OLS regressions and may not produce the same results in the IV case, for two reasons. First, the impact of literacy on outcomes is 2 to 5 times larger when estimated with instrumental variables. Since the sensitivity essentially examines a downward sloping relationship between the ACME and ρ (where the OLS results are ACME at $\rho = 0$) and asks at what value of ρ ACME becomes zero this just mechanically implies a larger ACME implies a large ρ (all else equal). This however would still produce modest levels of ρ. The second point is subtler but potentially more powerful. One model of why the learning conditional on schooling and the outcomes conditional on schooling and learning regressions would have correlated errors is essentially omitted variables bias. Say there was some underlying degree of "mental adeptness" and women with more of that learned more from a given exposure to school and also had better outcomes. Then the OLS coefficient on learning in the outcome equation would be biased upward for the usual omitted variables bias reasons and the results of estimating the two equations without that measure included could not distinguish between a causal mediation effect of learning and this unobserved ρ from the missing variable in both equations–and intuitively, the lower the R-squared in the equations the larger the fraction of the variance in learning and outcomes is in the error term (by definition) and hence the smaller the ρ would have to be. But, omitted variables bias depends on a correlation between the included and excluded variables (in this hypothetical example between a woman's "mental adeptness" and learning). But the IV estimates essentially project the learning into the space of the EALOM and use only the portion of the variability across women in learning that is correlated with the EALOM to estimate the impact of learning. But there is no reason to believe that a woman's "mental adeptness" is correlated with her enumeration area neighbor's measurement literacy. So it is at least possible that there are cases in which the OLS is not robust to cross-equation error (which is a violation of the assumptions of sequential ignorability) but the IV results are nevertheless robust.

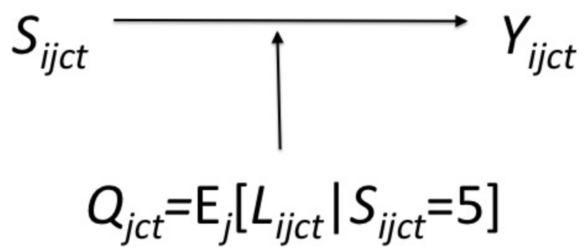
5 Education quality as a moderator of schooling's effects on outcomes

The previous section tested whether individual literacy mediates the relationship between girls' schooling and adult outcomes; this section tests whether school quality – as measured by the average propensity of schooling to generate literacy – moderates the relationship between girls' schooling and adult outcomes. This is an alternative approach to answering the same core question: is learning the mechanism linking schooling to reduced fertility and child mortality, and increased women's empowerment? If the social return to schooling is significantly higher where school quality is better, this is indicative of a learning channel as illustrated in Figure 12 where Q_{jct} in the figure is the quality of schooling in the jth country/region as proxied by the average literacy level of women who completed grade 5 in region j.

Concretely, we use the same pooled DHS microdata as in Section 4 to now regress child survival at the individual-woman level on geographic and demographic controls and years of schooling, and we replace our individual-level literacy measure with an interaction of years of schooling with an aggregate school quality measure.

Let Y_{ict} denote the woman-specific child-survival rate for individual i in country c at time t. We regress this measure on years of schooling, S_{ict}, and its interaction with our aggregate school quality variable, for which we use the average literacy level of a woman with five years

Figure 12: Causal pathway from a woman's school to her outcomes moderated by regional school quality



of schooling, $\bar{L}_{ct}$.[15]

$$Y_{ict} = \alpha S_{ict} + \beta(\tilde{S}_{ict} \times \tilde{L}_{ct}) + \gamma X_{ic} + \mu_{ct} + \varepsilon_{ict} \tag{22}$$

where the tilde denotes that the variable has been de-meaned by subtracting the country-year average in the case of individual level variables (such that $\tilde{S}_{ict} \equiv S_{ict} - \bar{S}_{ct}$) and aggregate variables are de-meaned by subtracting the overall sample average (such that $\tilde{L}_{ct} \equiv L_{ct} - \bar{L}$). This adjustment is purely to aid interpretation and ensure that the α coefficient is not changed by the inclusion of interaction terms. Additionally, because the school quality indicator is measured at the country level, in the analysis we cluster the standard errors at the country level.

To allow for non-linearity in the impact of schooling, we also estimate a version of equation (22) replacing the linear schooling term with a spline function, with a single knot at five years

[15]Because the literacy test is only administered to women with primary schooling or less, and some countries consider primary school complete at five years (and thus those with more schooling are not tested), we use average literacy level at five years of schooling to provide a measure of school quality at a grade level where it is reasonable to expect students to have learned basic literacy and to maximizes the number of survey rounds included in the sample.

Table 7: Moderator analysis: Does school quality moderate the relationship between schooling and child survival?

	Linear schooling				Spline (knot at $S = 5$)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Schooling								
S_ic	0.006***	0.006***	0.006***	0.006***				
	(0.000)	(0.000)	(0.000)	(0.000)				
S_ic^{0-4}					0.006***	0.006***	0.006***	0.006***
					(0.000)	(0.000)	(0.000)	(0.000)
S_ic^{4-8}					0.005***	0.006***	0.006***	0.006***
					(0.001)	(0.001)	(0.001)	(0.001)
Moderators								
$\tilde{S}\_ic \times \tilde{L}\_c$		0.002	0.002**	0.002**				
		(0.001)	(0.001)	(0.001)				
$\tilde{S}\_ic^{0-4} \times \tilde{L}\_c$						0.003***	0.004***	0.004***
						(0.001)	(0.001)	(0.001)
$\tilde{S}\_ic^{4-8} \times \tilde{L}\_c$						-0.011***	-0.007**	-0.007**
						(0.003)	(0.003)	(0.003)
Obs. (women)	845985	845985	845985	845985	845985	845985	845985	845985
Country-year cells	125	125	125	125	125	125	125	125

The dependent variable is the woman-specific child survival rate, i.e., total number of living children over total live births. The sample is restricted to women with less than secondary schooling. Regressions are weighted so that each country-year receives equal weight. Standard errors are clustered at the country level. Variables marked with a tilde are demeaned: individual-level variabls (e.g., $\tilde{S}_{ic}$) are demeaned at the country level while country-level varialbes (e.g., $\tilde{L}_c$) are demeaned at the global level.

of schooling, such that

$$
\begin{aligned}
S_{ict}^{0-4} &= \min(S_{ic}, 4) \\
S_{ict}^{4-8} &= \max(S_{ic}, 4) - 4.
\end{aligned}
$$

These two spline terms are also interacted with our school quality measure in later specifications.

Averaging over all countries and primary grade levels, the results in Table 7 show a statistically robust association between years of schooling and child survival, when controlling for both country fixed effects and a cubic polynomial of the mother's age and with no learning measure included in the regression. Six years of schooling (at existing learning levels) is associated with an increased probability of child survival of roughly 3.3% (column 1), relative to an average (unweighted) survival rate of 90.5% in the sample, representing a 35% reduction in child mortality.

The spline regressions (column 5) show a slightly larger relationship in the first four years of schooling (0.6% per year) than in the latter four years (0.5% per year).

Turning to the coefficient of primary interest, β on the interaction term in equation (22), we see a positive but statistically insignificant coefficient across all years of primary schooling (column 2). However, this interaction term is larger and significant for the first four years of primary schooling and negative (and significant) in the latter four years (column 6).

Note that so far we have only controlled for characteristics of the individual woman and fixed effects at the country level, allowing for no other determinants of the variance in the return to schooling across countries besides our index of school quality. Equation (22) could be re-cast as a bivariate cross-country regression where the independent variables are the country-specific coefficients on S_{ict}. Controlling for other basic factors in this interaction space increases the β coefficient on the schooling-quality interaction. For instance, once we control for the interaction of schooling and the average child survival rate in a country-year cell, the coefficient of interest become statistically significant in the linear specification (column 3), and increases for lower-primary in the spline specification, while the coefficient for upper primary becomes smaller (though is still negative, column 7).

6 Conclusion

An enormous literature demonstrates both the economic and broader social returns to girls' schooling. Yet nearly every study (and there have been literally thousands) that attempts to estimate the "impact" of girl's education on some outcome has used years of schooling as a proxy for education.

The first implication of our results is that every study that has taken this approach of using "years of schooling" as a measure of "education" has drastically *underestimated* the real impact of an education that includes learning. Including a measure of learning and correcting for measurement error increases the estimated impact of *education* by a factor of three to four.

Second, estimates based only on schooling cannot distinguish the causal pathways of schooling versus learning, even with an RCT. Therefore, they also can give no information on the relative benefits of increasing learning or expanding schooling duration at existing learning levels. Our best estimates suggest learning accounts for 35-80% of education's impact. Given that learning interventions can be substantially less expensive than schooling expansion, investing in learning could be *orders of magnitude* more cost effective at improving outcomes.

Finally, these new estimates have implications for comparing investments in education with other development interventions. Higher returns to education than what have typically been estimated suggests that investing in education could have a greater impact per dollar spent, relative to other interventions, than previously believed.

7 Bibliography

Behrman, Julia Andrea. 2015. "The effect of increased primary schooling on adult women's HIV status in Malawi and Uganda: Universal Primary Education as a natural experiment." *Social science & medicine*, 127: 108–115.

Bongaarts, John, and Susan Cotts Watkins. 1996. "Social interactions and contemporary fertility transitions." *Population and development review*, 639–682.

Breierova, Lucia, and Esther Duflo. 2004. "The impact of education on fertility and child mortality: Do fathers really matter less than mothers?" National Bureau of Economic Research.

Christiaensen, Luc, and Harold Alderman. 2004. "Child Malnutrition in Ethiopia: Can Maternal Knowledge Augment the Role of Income?*." *Economic Development and cultural change*, 52(2): 287–312.

Dearden, Kirk, Lant Pritchett, and Jeff Brown. 2004. "Learning from neighbors: Social learning about child feeding during diarrheal episodes." Annual meeting of the Population Association of America, Boston, MA.(Abstract# 454.2).

Demographic, and Health Survey. n.d.. "DHS Program." *http://dhsprogram.com*, Accessed: 2016-04-01.

Duflo, Esther. 2000. "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment." National Bureau of Economic Research.

Gakidou, Emmanuela, Krycia Cowling, Rafael Lozano, and Christopher JL Murray. 2010. "Increased educational attainment and its effect on child mortality in 175 countries between 1970 and 2009: a systematic analysis." *The Lancet*, 376(9745): 959–974.

Glewwe, Paul. 1999. "Why does mother's schooling raise child health in developing countries? Evidence from Morocco." *Journal of human resources*, 124–159.

Hanushek, Eric A, and Ludger Woessmann. 2012. "Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation." *Journal of Economic Growth*, 17(4): 267–321.

Keats, Anthony. 2014. "WomenâĂŹs Schooling, Fertility, and Child Health Outcomes: Evidence from UgandaâĂŹs Free Primary Education Program." Mimeo, Wesleyan University.

La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea. 2012. "Soap operas and fertility: Evidence from Brazil." *American Economic Journal: Applied Economics*, 1–31.

LeVine, Robert A, Sarah E LeVine, Meredith L Rowe, and Beatrice Schnell-Anzola. 2004. "Maternal literacy and health behavior: a Nepalese case study." *Social science & medicine*, 58(4): 863–877.

Montenegro, Claudio E., and Harry Anthony Patrinos. 2014. "Comparable estimates of returns to schooling around the world." *World Bank Policy Research Working Paper*, , (WPS 7020).

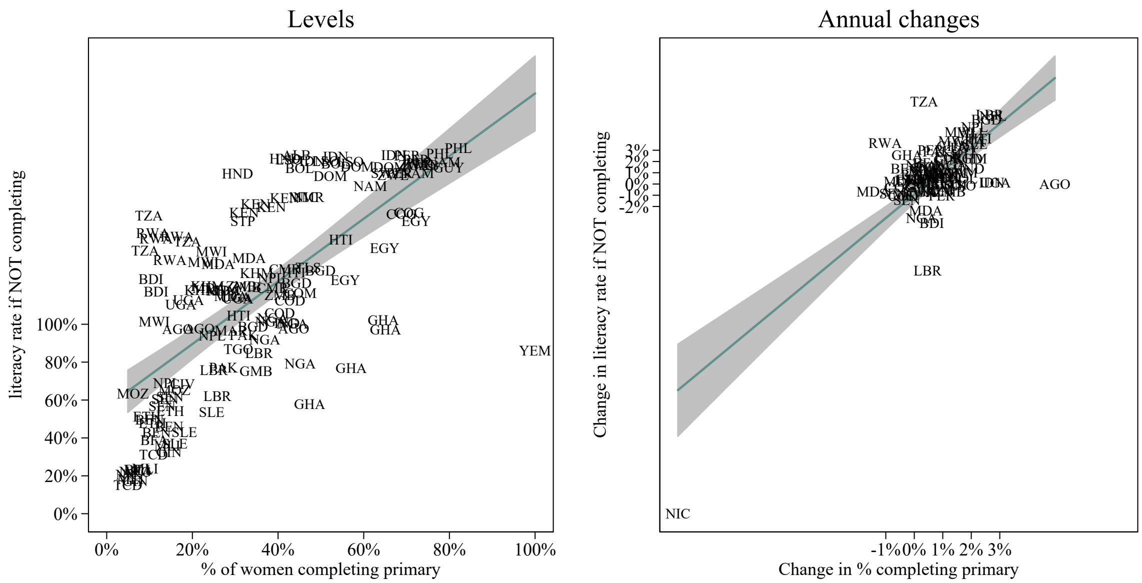
Osili, Una Okonkwo, and Bridget Terry Long. 2008. "Does female schooling reduce fertility? Evidence from Nigeria." *Journal of Development Economics*, 87(1): 57–75.

Programme, RISE. n.d.. "RISE Vision Document 2: The Pivot from Schooling to Education." *http://www.riseprogramme.org/sites/www.rise.ox.ac.uk/files/RISE_Vision_document-2.pdf*, Accessed: 2016-07-20.

Spence, A. 1973. "Job Market Signaling." *The Quarterly Journal of Economics*, 87(3): 355–374.

United Nations Educational, Scientific, and Cultural Organization (UNESCO). 2012. *Youth and skills: putting education to work.* UNESCO, Paris, France.

Figure 13: Evidence on the direction of sample selection bias in literacy rates measured among women with incomplete primary



The vertical axis measures the level or change in the percentage of women who are literate among those who did not complete primary school. The sample is restricted to women age 25-34. Each country appears up to three times, corresponding to waves 4-6 of the DHS. Survey dates vary by country; median survey year for wave 4 is 2004, 5 is 2007, and 6 is 2013. Annual changes are calculated as the percentage point change between waves divided by the (country-specifc) timespan between waves.

A Country by country data

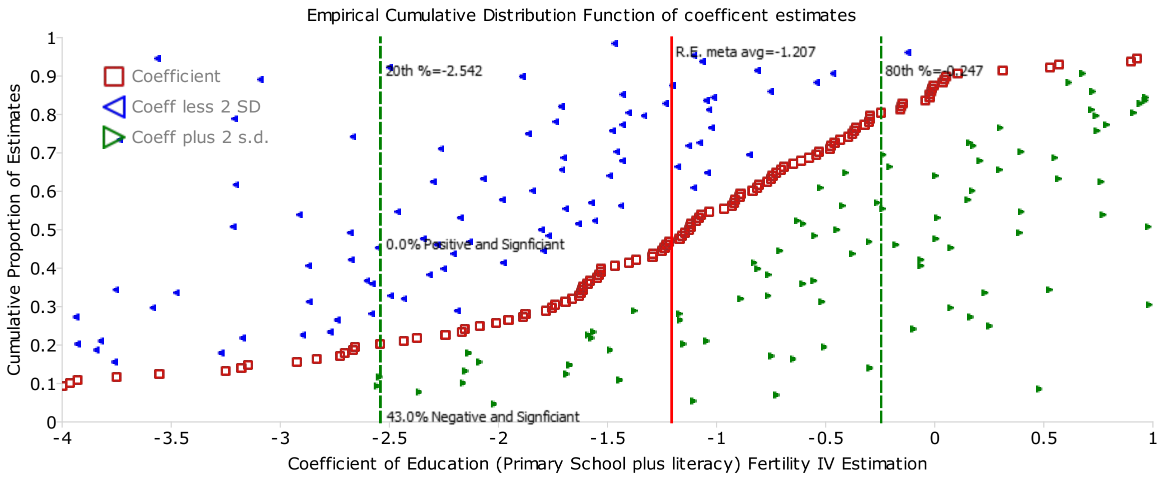
Table A.1: Summary statistics for key variables

Country	Age (Avg.)	Literacy Rate (%)	Schooling Years (Avg.)	Rural Residents (%)	Rich (% of those in sample in top 40
					Afghanistan
31.37	1.63	.3	78.47	33.82	
Albania	32.01	80.09	7.77	74.09	18.21
Angola	28.09	21.35	2.73	58.14	22.5
Bangladesh	33.43	18.05	1.87	70.73	29.48
Benin	30.05	5.7	.87	63.62	34.14
Bolivia	36.77	17.02	.89	71.92	9.49
Burkina Faso	29.53	5.12	.66	79.57	36.94
Burundi	28.33	48.56	2.13	86.63	37.38
Cambodia	31.92	31.07	2.71	78.82	32.85
Cameroon	29.52	27.22	3.2	66.8	23.48
Chad	28.79	2.24	.81	75.18	42.03
Comoros	30.82	8.39	1.72	70.72	28.41
Congo	29.19	11.12	3.49	72.88	12.34
Cote D'Ivoire	29.36	14.31	1.28	62.91	33.27
Democratic Republic of Congo	29.44	12.32	2.65	79.2	20.69
Dominican Republic	32.07	60.8	4.7	54.42	13.55
Egypt	35.35	5.93	1.15	75.42	15.57
Ethiopia	28.5	15.19	1.57	78.24	40.62
Gabon	31.77	38.98	3.97	48.04	8.58
Gambia	29.45	2.1	.99	70.03	24.61
Ghana	31.57	4.03	1.71	72.27	16.7
Guinea	29.95	1.39	.6	73.56	34.5
Guyana	34.61	53.51	4.08	89.11	18.22
Haiti	30.08	29.45	2.42	67.52	24.87
Honduras	33.6	31.85	.85	87.73	4.94
India	31.61	11.77	1.19	70.04	29.7
Indonesia	35.19	63.47	4.28	71.97	18.25
Kenya	29.21	56.66	5.18	71.95	28.12
Lesotho	29.73	79.33	5.73	82.82	27.2
Liberia	29.76	5.53	1.57	65.67	23.09
Malawi	28.79	49.8	4.15	88.41	33.71
Mali	29.11	1.68	.52	74.72	37.14
Moldova	29.15	39.89	2.26	82.65	29.77
Morocco	31.02	20.7	.94	59.94	24.37
Mozambique	29.3	26.16	2.45	65.68	42.49
Myanmar	34.33	56.74	2.97	85.01	23.79
Namibia	33.45	50.34	3.84	70.75	18.71
Nepal	31.75	26.12	.92	79.42	29.44
Nicaragua	29.1	2.88	.6	74.53	46.67
Nigeria	30.55	4.3	1.71	77.95	18.89
Pakistan	33.03	12.17	.76	67.11	28.14
Peru	34.48	63.09	3.81	67.77	6.97
Philippines	33.49	65.19	4.57	72.34	12.49
Rwanda	29.09	62.89	3.5	83.59	36.07
Sao Tome and Principe	30.75	42.59	3.75	60.09	26.2
Senegal	28.65	10.41	1.1	68.11	26.6
Sierra Leone	30.34	1.5	.84	72.55	31.89
Swaziland	29.21	63.8	4.37	80.9	27.34
Tanzania	29.53	59.12	4.69	78.06	37.57
Timor Leste	32.35	21.43	2.01	83.33	25.56
Togo	31.53	8	1.93	72.53	28.46
Uganda	28.82	30.31	3.55	86.67	30.78
Zambia	30.05	25.87	4.37	69.86	24.11
Zimbabwe	31.02	56.28	5.51	86.74	19.72

Table A.2: Summary statistics for key health variables

Country	Children Ever Born (Avg.)	Children By Age 25 (Avg.)	Child Survival Rate (%)	Empowerment Index (normalized)
Afghanistan	4.41	2.37	93.3	0
Albania	2.04	.99	96.59	-.01
Angola	3.32	1.67	90.97	.01
Bangladesh	3.22	2.24	90.2	
Benin	3.46	1.88	89.53	0
Bolivia	5.2	2.27	85.02	-.09
Burkina Faso	3.67	1.6	86.43	0
Burundi	2.99	1.04	89.59	0
Cambodia	2.63	1.27	91.11	0
Cameroon	3.6	2.1	85.53	.01
Chad	4.11	2.38	85.42	-.02
Comoros	3.51	1.86	94.34	0
Congo	3.33	1.89	89.27	
Cote D'Ivoire	3.22	1.83	87.96	0
Democratic Republic of Congo	3.7	1.95	85.86	0
Dominican Republic	3.02	2.08	94.5	0
Egypt	3.82	2.13	93.01	
Ethiopia	3.2	1.79	87.3	0
Gabon	3.83	2.08	91.92	.01
Gambia	3.47	1.89	92.62	0
Ghana	3.57	1.45	90.39	-.01
Guinea	3.57	1.95	83.55	0
Guyana	3.33	1.99	95.13	0
Haiti	3.03	1.44	89.1	0
Honduras	4.31	2.28	93.4	-.16
India	3	2.04	89.91	0
Indonesia	2.99	1.73	92.04	0
Kenya	3.26	1.76	89.72	0
Lesotho	2.5	1.4	90.16	-.01
Liberia	3.67	2.06	84.45	0
Malawi	3.36	1.87	88.27	0
Mali	3.85	1.75	86.74	0
Moldova	3.28	1.13	86.41	0
Morocco	2.41	1.22	92.63	
Mozambique	3.21	1.87	84.27	0
Myanmar	2.5	1.14	90.74	0
Namibia	3.14	1.48	92.08	-.01
Nepal	3.06	2.92	135.53	0
Nicaragua	4.17	2.38	83.67	0
Nigeria	4.22	2.17	80.72	0
Pakistan	4.22	2.11	90.16	
Peru	3.67	2.03	93.26	0
Philippines	3.71	1.8	93.39	
Rwanda	2.64	1.05	88.07	0
Sao Tome and Principe	3.52	2.02	92.9	.01
Senegal	3.11	1.53	92.66	0
Sierra Leone	3.7	1.57	75.22	0
Swaziland	3.09	1.83	87.69	-.01
Tanzania	3.33	1.82	89.19	
Timor Leste	3.76	1.48	89.96	0
Togo	3.57	1.73	89.53	0
Uganda	3.94	1.86	79.95	0
Zambia	3.93	2.19	88.32	0
Zimbabwe	3.21	1.85	91.97	-.01

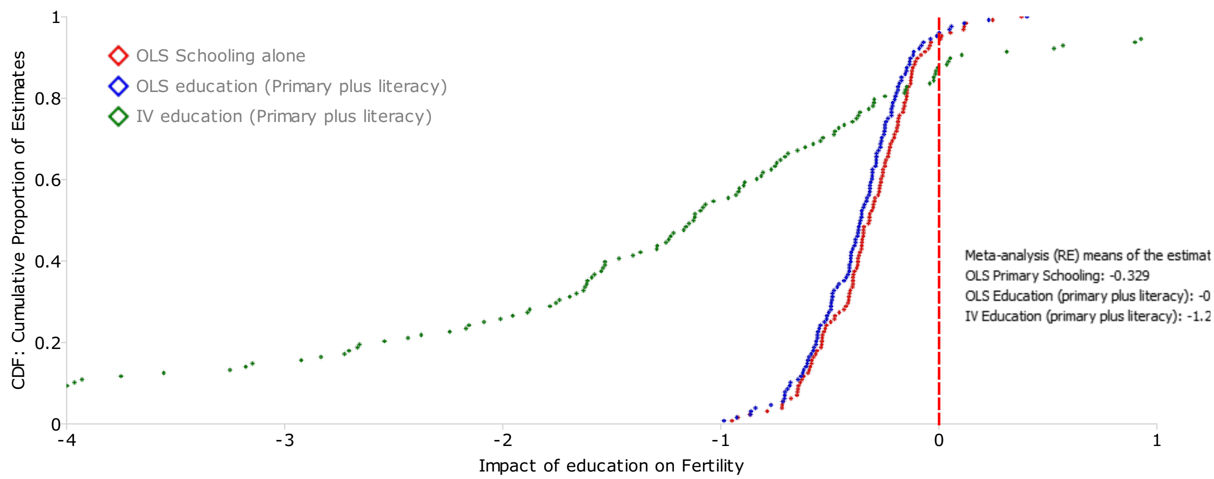
Figure B.1: Empirical cumulative distribution function for IV estimates of the impact of education on fertility



Source: Authors' calculations based on DHS microdata for 54 countries.

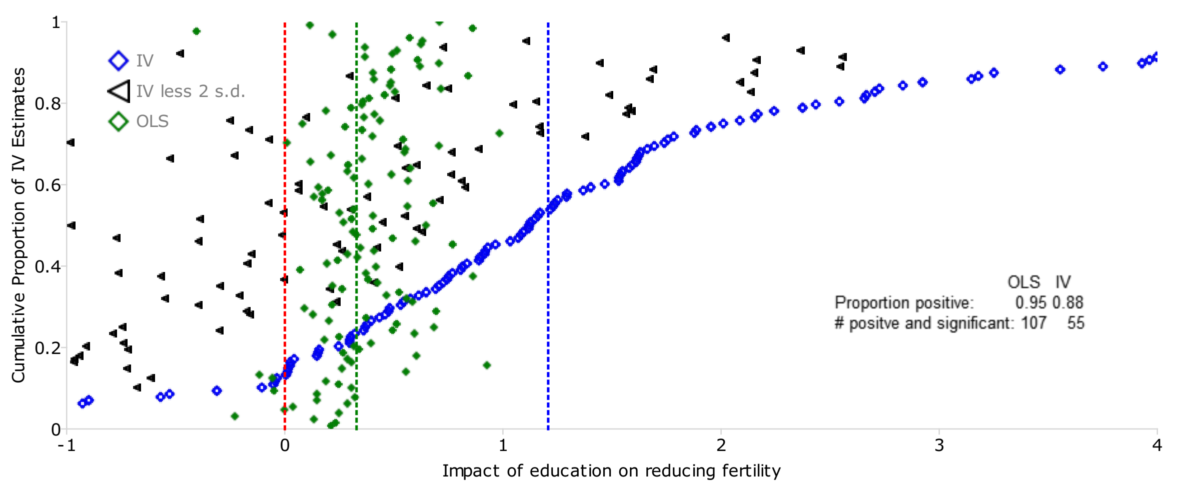
B Cumulative distribution function graphs for fertility and female empowerment

Figure B.2: Empirical cumulative distribution functions for OLS estimates of the impact of schooling and OLS and IV estimates of the impact of education on fertility



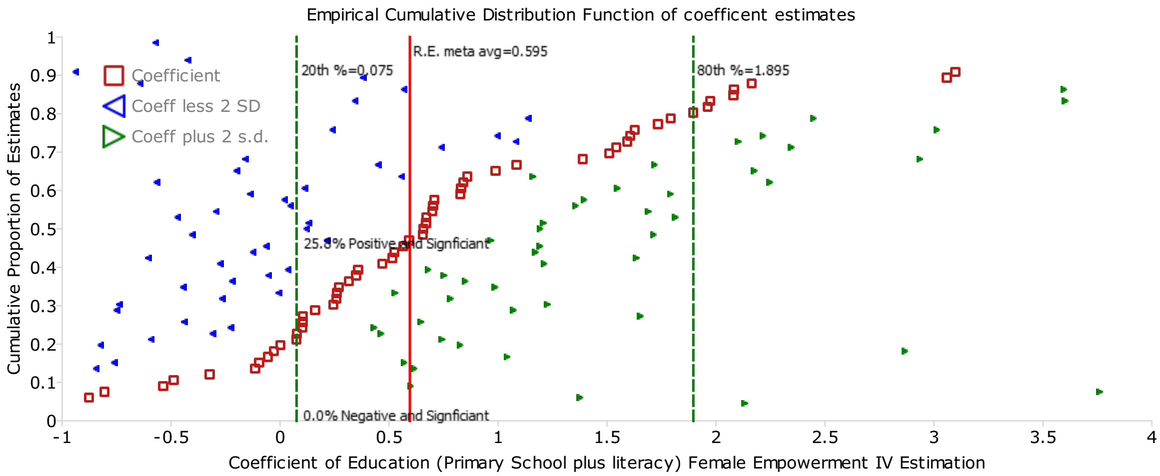
Source: Authors' calculations based on DHS microdata for 54 countries.

Figure B.3: Comparing IV and OLS estimates of the impact of education on fertility



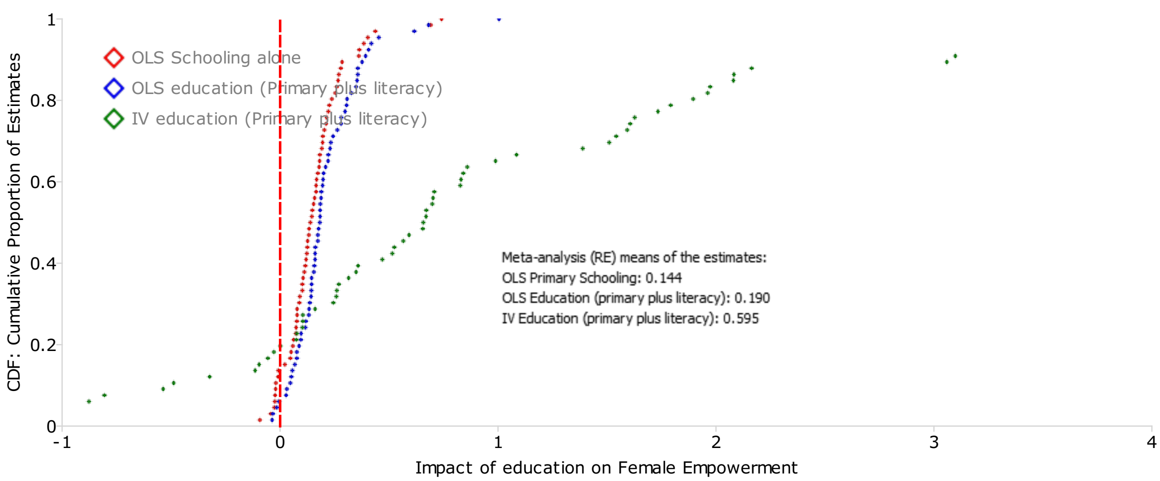
Source: Authors' calculations based on DHS microdata for 54 countries.

Figure B.4: Empirical cumulative distribution function for IV estimates of the impact of education on female empowerment



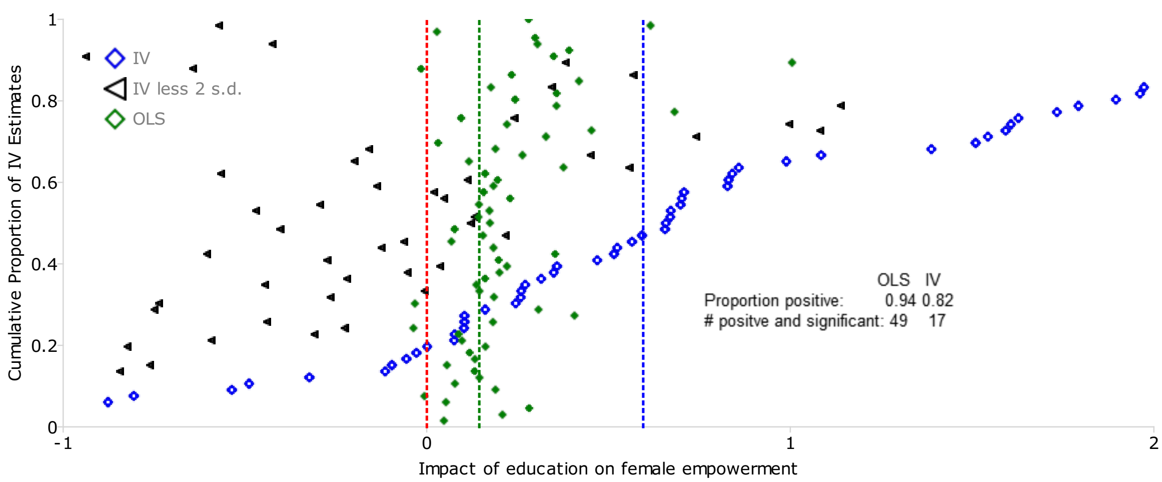
Source: Authors' calculations based on DHS microdata for 54 countries.

Figure B.5: Empirical cumulative distribution functions for OLS estimates of the impact of schooling and OLS and IV estimates of the impact of education on female empowerment



Source: Authors' calculations based on DHS microdata for 54 countries.

Figure B.6: Comparing IV and OLS estimates of the impact of education on female empowerment



Source: Authors' calculations based on DHS microdata for 54 countries.