Consultation draft

The long-term decline of school quality in the developing world

Alexis Le Nestour* Laura Moscoviz* Justin Sandefur*

September 2021†

Abstract

This paper documents the evolution of school quality over time in the developing world. We use repeated cross-section surveys from 87 countries to model age, period, and cohort effects in literacy conditional on years of schooling for men and women born between 1950 and 2000. We find little evidence of human capital accumulation over the life-cycle: literacy declines with age, independent of occupation. Cohort effects show long-run stagnation in school quality in all regions, and a steep secular decline in both South Asia and Sub-Saharan Africa. Changing patterns of selection into school appear to explain some but not all of this decline: the fall in observed quality is greater where enrollment grew faster, and after the abolition of user fees, but smaller for women despite experiencing bigger enrollment gains.

JEL classification: I25, N37, O15
Keywords: literacy, school quality, access to education

*Center for Global Development

†This project benefited from discussions with Girin Beeharry, Asyia Kazmi, and Lant Pritchett, and comments from Maryam Akmal, Lee Crawfurd, David Evans, Susannah Hares, Kim Lehrer, Isaac Mbiti, Jack Rossiter, Francis Teal, Andrew Zeitlin, and seminar participants at RISE and CSAE Oxford. The views expressed here should not be attributed to the Center for Global Development or its funders. All remaining errors are our own.

Contents

List of Tables

List of Figures

1 Introduction

International organizations like the World Bank and UNESCO have declared a "learning crisis" in the developing world, with many school systems failing to reliably produce even basic literacy and numeracy skills [World Bank Group, 2018, UNESCO, 2013]. Pupils in India score, on average, at about the 5th percentile for pupils at a similar grade in advanced economies on international learning assessments, while nearly 80 percent of South African children cannot read for meaning in any language by fourth grade [Das and Zajonc, 2010, Howie et al., 2017]. The economic consequences of improving education quality are poorly understood, but potentially large. Recent work in the macroeconomics of development has suggested that school quality may outperform quantity measures in explaining long-run growth in econometric tests, and accounting for quality differentials roughly doubles the explanatory power of education in a development accounting framework [Hanushek and Kimko, 2000, Hanushek and Wossmann, 2006, Schoellman, 2012].[1]

The idea of a learning "crisis" implies something new. But we lack reliable, long-term measures of school quality over time, particularly for the developing world, to evaluate competing explanations for this crisis.[2] For instance, one prominent narrative is that the expansion of mass schooling, e.g., after the abolition of user fees in many African countries in the 1990s and 2000s, led to a quality decline [Ruto and Mugo, 2005, Pritchett, 2013, Taylor et al., 2013, Atuhurra, 2016].[3] Another complementary narrative is that countries at the top of the league table in international learning assessments (e.g., Finland or Vietnam) hold reform lessons for their neighbors [Sahlberg, 2007, Dang and Glewwe, 2018]. This argument often attributes current success to current educational policies, and elides the question of timing. Is, say, Vietnam's success a feature of reforms to its education system during its recent liberalization period, or a longer-standing national characteristic? Are the lagging

[1] This view is not universal: Caselli [2005] questions the explanatory power of test scores in development accounting, based on the weak observed Mincerian returns to test scores for individual incomes in U.S. micro data. Kraay [2018] suggests a different approach, eschewing the Mincerian literature, and attributing a bigger role to school quality.

[2] While a number of developing countries have participated in multiple rounds of international or regional learning assessments, the time span covered is generally limited. For instance, 14 African countries participated in two rounds of the SACMEQ study from 2000 to 2007 while the PASEC, which covers 18 African countries, started in 1995 but has no strict comparability across countries before 2014 and no comparable time series. 19 Latin American countries participated in the Laboratorio SERCE and TERCE tests from 1997 to 2019, and 41 developing countries participated in multiple rounds of TIMSS, PIRLS, or PISA, generally spanning just a single decade or less.

[3] Though see also Blimpo et al. [2019] for a contradictory finding from The Gambia.

learning outcomes in India, or the truly abysmal outcomes we document in Nigeria, the result of a contemporary crisis or do they have deeper historical roots?

To answer these questions, we construct new, comparable time series of education quality for 88 developing countries, stretching from cohorts born in the 1950s up to the late 1990s. We define education quality as the literacy rate for an adult with a given level of schooling, typically five years of primary school. This is roughly analogous to major international learning assessments such as PISA, TIMSS, PIRLS, etc., which measure school quality with standardized tests in a specified grade. To estimate literacy across age cohorts, we draw on repeated cross-sections from the Demographic and Health Surveys (DHS) run by USAID and Multiple Indicator Cluster Surveys (MICS) run by UNICEF, which provide direct, rather than self-reported, tests of literacy for adult women and, in a subset of survey rounds, men, aged 15 to 49, mostly measured between the early 2000s and late 2010s. We also disaggregate results for some larger countries, i.e., India, Nigeria, and Indonesia, to look for patterns of divergence or convergence among sub-national units.

The core analytical challenge throughout the paper is to identify cohort effects in repeated cross-sectional data. As is well known, age, period, and cohort effects are perfectly multi-collinear in a simple linear model, and identification rests on auxiliary assumptions about the underlying data generating process [Fosse and Winship, 2019]. To overcome the identification issue, we follow Deaton and Paxson [1994] and impose the restriction that period effects sum to zero and orthogonal to a trend, which assumes that period effects are shocks and do not have a trend. This method has been widely used in the economic literature dealing with age, period and cohort models. Thus, Lagakos et al. [2018] use the Deaton-Paxson normalization to estimate experience-wage profiles using repeated cross-sectional surveys. Although we believe that literacy conditional on schooling can vary across the life cycle or differ from one cohort to another, it seems unlikely that there are periods where the adult population would learn or forget how to read (period effects). Moreover,the fact that literacy rates are close to 0 and exhibit no trend over time for individuals with no schooling reinforces our conviction that literacy acquisition happens while in school and not at the adult age and we interpret systematic differences between surveys that cannot be accounted by age or cohort effects as survey effects.

Results from our age-period-cohort model of literacy rates reveal a picture of overall stagnation in education quality in the developing world, and fairly stable gaps in quality across countries. We find virtually no case worldwide of dramatic improvements in education quality over a fifty-year time horizon, and for most of the 88 countries in our sample we see flat lines, with little movement up or down. This flat trend holds over multiple decades for some notable high-performers like Burundi and Vietnam, and low-performers like Nigeria. However, we also find evidence of significant secular decline in education quality for a small number of large developing countries, notably India and Ethiopia.

Changing patterns of selection into school are an obvious potential explanation, which we explore in detail in Section 6. We follow the convention in the literature by defining 'school quality' as a simple conditional average — the literacy rate conditional on a certain number of years of schooling. Our primary interest lies in documenting trends in this statistic. Nevertheless, it bears emphasizing that school quality defined this way may decline (as we observe in practice in a plurality of countries) when the causal effect of schooling deteriorates, or when enrollment expands (as we observe in almost all countries) if marginal students have lower scholastic ability. Indeed, we find that on average the selection of pupils into schooling on the basis of height, a characteristic largely determined prior to schooling decisions, weakens over time. Furthermore, in countries and time periods where enrollment expanded more rapidly, school quality declined more quickly, consistent with (though clearly not definitive proof of) a selection effect. Controlling for pupil height or local enrollment trends significantly dampens the measured downward trend in school quality for women but not for men in our overall sample, and leaves a significant negative trend in both cases. We also find, on average, no significant trend break in observed school quality when countries abolished user fees for primary schooling, a policy change which putatively expanded access to schooling for marginal pupils.

Our results add a new dimension (i.e., the evolution of quality over time) to the large literature measuring human capital around the world, with relevance for a variety of empirical applications. The most widely used human capital measure in the economics, due to Barro and Lee [2013], focuses exclusively on quantity rather than quality. More recent work has incorporated richer measures of education quality, but this has come at the cost of either limited coverage for the developing world, or short time series [Yue Yin, 2018, Sulis et al., 2020, Gustafsson, 2016, Van Damme and Bellens, 2017, Kraay, 2018, Patel and Sandefur, 2020]. By combining a quality dimension with long time series and broad developing country coverage, the results here would potentially enable future work to revisit earlier cross-country findings on the determinants of test-score performance, and conversely, the role of education quality in promoting economic development [Hanushek and Wossmann, 2006, Wossmann and Hanushek, 2007]. As an illustration, we find a strong correlation of quality with subsequent economic growth in our developing country sample, though this finding is not robust to controlling for unobserved country effects.

The following section describes the coverage of our survey data sources in more detail and in Section 3 we describe and present some basic validation tests of our measure of school quality. We then turn to our core empirical strategy to overcome the challenge of disentangling age, period, and cohort effects in repeated cross-sectional data on literacy in Section 4. Having found mostly downward trends in observed quality, Section 6 then examines whether changing patterns of selection into schooling drive these trends. Finally, Section 5 presents various stylized facts based on the trends in school quality that we calculate, and section 7 concludes.

2 Data

Our dataset consists in the universe of Demographic and Health (DHS) and Multiple Indicator Cluster surveys (MICS) as of January 2021 with information on literacy and schooling. DHS and MICS collect nationally representative information on literacy and schooling of women, and in some cases men, aged 15 to 49.[4] Since 2000, reading abilities are tested with a short literacy test consisting on reading four simple sentences out loud. Interviewers then score individuals on a three point scale: "cannot read at all", "able to read only parts of the sentence" and "able to read whole sentence". We use a 0-1 scale for literacy with individuals "able to read only parts of the sentence" scoring 0.5 point. Prior to 2000, literacy was self-reported by individuals on a similar three point scale and surveys with self-reported data on literacy are excluded to keep a consistent measure of literacy over time.

The surveys do not test individuals with secondary schooling and, in a few MICS surveys, also lack literacy data on individuals with no schooling. In the latter case, we make the assumption that individuals with no schooling are illiterate.[5] We follow the DHS and MICS

[4]While statistics for women are nationally representative, those for men are often not perfectly so, because men are sampled only from households with adult women present.

[5]Where these observations are not censored, the median literacy rates for individuals with zero schooling are 2 percent for women and 5 percent for men, based on all surveys with at least 10 percent of the population reporting zero schooling. In only 34 surveys out of 242 for women and 22 of 93 for men, the rate of literacy is higher than 10 percent for individuals with no schooling.

convention of assuming that individuals with secondary education are literate to construct trends of literacy.[6]

Note this missing data problem has little effect on our analysis. While the unconditional literacy rates reported below rest on the assumptions we make about literacy at the top and bottom of the schooling distribution, the school quality measures that are the focus of the paper are much less fragile to this missing data problem, as they focus on individuals with five years of primary schooling, all of whom are tested.

Individuals chose the language they want for the literacy test. In most countries, reading cards have been translated into all local languages and across the whole woman sample there are only 0.25 percent of tests that could not take place because of the lack of availability of reading cards in the local language. In a small share of interviews (0.17 percent in the woman sample), individuals could not be tested because they are visually impaired. These observations are excluded from the analysis and the final woman sample includes 99.5 percent of the observations.

Overall, data coverage for men is lower than for women, with fewer countries covered and a total sample size five times smaller (4.6 million observations for women and 1 million for men). Samples of women are nationally representative, although in seven countries only ever-married women have been interviewed [7] but samples for men are not perfectly representative, as men are selected only from households with women respondents, excluding men living alone.

All available DHS and MICS datasets for a country are used for our estimates, with the exception of pre-2000 surveys that collected self-reported literacy. To the exception of Panama and Trinidad-and-Tobago, no country in our dataset is classified as high-income countries,[8] and so our estimates should be seen as representative of non high-income countries. Data coverage is excellent in the woman and man samples for Sub-Saharan Africa and

[6] This assumption is questionable in some countries such as Ghana where literacy rates at the end of primary school are still low, and a few recent DHS surveys that test all adults allow us to test (and reject) this assumption. The numbers of surveys with full coverage is too low, and too recent, to serve as the basis of our main analysis.

[7] Only ever-married women were interviewed in Afghanistan, Bangladesh, Egypt, Indonesia, Jordan, Nepal and Pakistan. Given that most women eventually marry in these countries, potential selection biases should be minimal.

[8] World Bank classification as of 1st of July 2020 https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

South Asia[9] with 99 percent of the population covered for women and men in South Asia and 97 and 86 percent of women and men respectively in Sub-Saharan Africa (see table 1). Coverage is also good for Middle East and North Africa for women (75 percent of the population excluding high income countries) but in this region men are not surveyed and there no available survey. Similarly, in Latin America and Caribbean the coverage is much better for women (41 percent) than for men (10 percent) and the coverage of the region suffers from the absence of data for Brazil, whose surveys pre-date 2000 and only includes self-reported literacy measures, and the fact that there is only one survey available for Mexico. Because there is no available data for China, the coverage for the East and Asia Pacific region is 30 percent for women and 26 percent for men but goes up to and 91 and 77 percent for women and men respectively when China is excluded. Finally, there are only few countries with available data in the European and Central Asia region making up 5 and 1 percent of the population of the region for women and men respectively. In this region, some countries are excluded (e.g. Tajikistan, Ukraine) because virtually all individuals reach secondary school making it impossible to estimate our measure of school quality on individuals with primary schooling or less. Because of the low coverage in the Europe and Central Asia region, we do not report regional averages for this region. Similarly, no regional averages are reported for men for the Middle East and North Africa and the Latin America and Caribbean regions.

There are a large number of individuals in the samples who are still in primary school between the age of 15 and 19. 21 (17) percent of women (men) are still in primary school at age 15 and rates decrease to 2 and 1 percent for women and men respectively by age 20. Including individuals who are still in primary school could bias our estimates as the composition of cohorts might change between two surveys. For this reason, samples are restricted to individuals aged 20 to 49.

DHS and MICS surveys provide information on the number of years of education an individual has reached. We only consider formal education in our estimates. In the few cases where the length of primary schooling has changed between two surveys (e.g. Angola), we set the length of primary to the lowest value of the two surveys and considered that individuals who exceeded this value in the other survey have reached secondary school and are literate. Age is computed as the mean year of interview minus year of birth in all surveys. We do so to avoid any variance in age for individuals born the same year and interviewed during the same survey. Indeed, in some cases, surveys take a long time to administer and

[9] Region classification follows World Bank classification. See here for the complete list: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

places surveyed first may have different characteristics than later-surveyed places, which might create some spurious age effects.

3 Measuring education quality

Although measuring the *quantity* of education of a population is usually straightforward using administrative data or household surveys, measuring the *quality* of this education remains challenging. Large national or international assessment programs make considerable efforts to measure learning in school in a reliable way via standardized tests of representative student samples. However, because of the cost and complexity of collecting such data, standardized tests happen only infrequently, and very few low- and lower-middle income countries participate in internationally comparable assessments whatsoever. In order to measure educational quality in a large number of developing countries and over time, we propose a measure of educational quality based on data available in household surveys, following Oye et al. [2016], and Gollin et al. [2014]. We use the adult literacy rate as our proxy measure of human capital, and conceive of the quantity and quality of schooling as inputs in the literacy production function. The main advantage of literacy as an outcome is that it is available across a wide range of countries and time periods and is measured in a comparable way. While literacy alone is clearly a low bar against which to judge education systems, ceiling effects are less worrisome in our context, as only a handful of countries report rates above 90 percent in any year.

We define educational quality as the tendency for schooling to generate learning. More precisely, we measure quality as the expected level of literacy conditional on five years of schooling.[10]. Primary schooling lasts between four to eight years in sampled countries, with six years being the most common length (63 percent of countries in our sample). Five years is thus close to the end of primary in most countries.

More formally, let $\mathsf{L}_{ijt} \in \{0, 1, 2\}$ be the level of literacy and G_{ijt} the maximum grade attained of an individual i in country j at time t. Our indicator of educational quality faced

[10]This measure, essentially capturing the strength of the cross-sectional correlation between schooling and literacy serves as a substitute for a more ideal measure, where longitudinal panel data is available, tracking learning as students progress through grades. See Muralidharan et al. [2019], Andrabi et al. [2011], Singh [2019]

by an individual in j born in year t is defined as the predicted value $\hat{\mathsf{L}_{jt}}$ at $\mathsf{G} = 5$ from the following regression:

$$\mathsf{L}_{ijt} = \alpha_{jt} + \beta_{jt}\mathsf{G}_{ijt} + u_{ijt} \tag{1}$$

Obvious alternative definitions include using the estimated $\hat{\beta}$ or the simple average of literacy conditional on five years of schooling. The former has a weaker correlation with independent measures of school quality discussed in the next section. The latter performs quite similarly to our preferred measure. One advantage of our regression-based approach is that it allows us to make use of the full dataset, whereas the number of individuals with exactly five years of schooling can become small for some country-sex-birth year cells.

3.1 Comparison with international test scores

To assess the validity of our measure, we briefly examine its cross-country correlation with alternative measures of school quality, as well as per capita GDP and education spending, before turning to our primary analysis of trends over time.

Across countries, our measure of average education quality shows a significant and roughly linear, positive relationship with independent measures of student performance based on more detailed standardized tests – though the correspondence is far from perfect.

No single standardized test covers a large share of developing countries. To examine correlations across a relatively large number of countries, we draw on results reported by Alt [2018] who provide rough correspondences to link scores across international and regional tests on a common scale. The resulting scores, produced by the World Bank under the rubric of Harmonized Learning Outcomes (HLO), combine national averages from various underlying assessments (PASEC, LLECE, EGRA, PISA, SACMEQ and TIMSS/PIRLS) spanning 163 countries. As time-series variation is limited, we focus here on cross-country patterns at a point in time.

We match the birth cohorts in our data to the corresponding survey year and estimate age for each of the tests underlying HLO estimates. Based on the grade at which the tests are administered, we assume the average age of test takers for PASEC, SACMECQ and LLECE is 12 years, for EGRA is 8 years, 10 for TIMSS/PIRLS, and 15 for PISA. We can match 117 country, sex and survey-year cells from our data to the HLO estimates, spanning 40 countries.

Figure 1a shows the comparison between our measure of expected literacy at grade five and HLO estimates. We observe a relatively linear relationship, with one standard deviation in HLO (100 points) equivalent to about 18 percentage points of expected literacy at grade five. The Spearman correlation is 0.41 for the whole sample with no meaningful difference between men and women.

3.2 Comparison with income per capita and educational spending

We also see that richer countries and countries that spend more on education tend to have better school quality by our measure, though again there are some notable exceptions. The estimated correlation coefficients between school quality and GDP per capita or education spending, respectively, are high and positive (0.4 and 0.5). Figure 1b also shows that a group of countries perform well despite low GDP per capita and education spending such as Burundi, Rwanda, Madagascar and Lesotho in Africa; Myanmar, Nepal, and Vietnam in Asia.

4 Empirical strategy for disentangling age, period, and cohort effects

The primary goal of the empirical analysis that follows is to measure the evolution of the quality of schooling over time — i.e., trends in literacy conditional on years of schooling. Recall that our survey data reports literacy and schooling for repeated cross-sections of women and men of various ages, observed at multiple survey years in each country. Thus our core challenge is to disentangle the role of age, period, and cohort effects in these repeated cross-sections.

Age-period-cohort (APC) models of this sort are widely used in demography, sociology, political science and, to a lesser degree, economics (see Fosse and Winship [2019] for a recent review). In our setting, age captures life-cycle effects on literacy, which may evolve even after an individual leaves school if they acquire human capital throughout their career, or instead it decays as time passes. Period effects allow for the possibility that contemporaneous events may affect measured literacy, e.g., if economic booms increase formal employment which contributes to literacy through on-the-job learning. Cohort effects, which are our primary object of interest, capture differences in literacy conditional on schooling for individuals born (and hence educated) in different years. These cohort effects capture secular trends in schools' value added, as well as trends in other non-school inputs into learning such as early-childhood nutrition and – as discussed at length below – changes in selection into schooling based on scholastic ability.

The basic model poses a fundamental identification challenge. If we posit that age, period, and cohort effects are linear and additive, there is an obvious, perfect collinearity between the three factors (age = period − cohort), making it impossible to estimate the model without invoking additional functional-form assumptions. The following sub-sections briefly motivate the modeling assumptions we invoke, and our reasons for eschewing alternative approaches.

4.1 Assumptions underlying our modeling approach

In most of what follows we adopt one of the most common approaches to estimating APC models in the economic literature: the Deaton-Paxson (DP) normalization [Deaton and Paxson, 1994]. The DP approach imposes a parametric restriction on period effects. Specifically, the model assumes that the period effects do not contain any linear time trends. In practical terms, the variables are detrended and the period effects are restricted to be orthogonal to a linear time trend. This approach provides a solution to the APC identification problem and the model is just identified. OLS regressions or other estimation methods can then be used to estimate coefficients for age, period, and cohort. However, the DP method requires a minimum of three surveys for the model to be identified. A shorter cohort cycle can yield biased estimates as the method would attribute any variation in literacy between the first and second survey to age and cohort effects.

The choice of the best method to estimate our APC model depends on what we think the true age, period and cohort effects are. First, we do not want to impose restrictions on *cohort effects*, whose estimation is our main research goal.

Second, we want to allow that there might be *age effects*, although we do not have a strong prior on their direction. The estimation of age effects is also a secondary research goal for us. Finding evidence for negative age effects would demonstrate that some of the education is wasted but finding positive age effects would show that the consequences of low quality education can be overcome as individuals age. Because of the limited number of large longitudinal datasets including information on adult cognitive skills, the process of literacy retention or acquisition after students leave school in developing countries is still not very well known. There is some limited evidence that literacy skills continue to evolve after individuals leave school and that changes in adult literacy vary by sex, grate attained or personal circumstances. For instance, Gorman and Pollitt [1997], using a longitudinal dataset of teenagers in Guatemala, find that literacy continues to increase after students leave school but that the increase was stronger for students who completed fewer grades. In a longitudinal study of teenagers in Malawi, Soler-Hampejsek et al. [2018] find that literacy levels in English tend to deteriorate after students leave school but skills in local language remain the same. Moreover, they find that grade attained, rather than life experience or personal characteristics, is the main determinant of variability in literacy retention, such as the ones who achieved more grades were more able to keep their level of literacy after leaving school. They also find that girls tend to lose literacy skills faster than boys. In another study using the same data from Malawi and longitudinal data from Zambia and Bangladesh, Psaki et al. [2019] find that early child-bearing is associated with larger losses in literacy skills for adolescent girls, with stronger effects for girls with low attainment (grade seven or below). The three previous studies are focused on learning retention of adolescents but, to our knowledge, there is no longitudinal study on literacy changes for adults. It has been found that adults who take part in adult literacy programs often relapse to illiteracy, which might be due to low proficiency levels, lack of use of literacy skills or difficulty in teaching people how to read at an age where the brain is less malleable [Abadzi, 2003].

Third, to achieve identification we assume that *period effects* do not exhibit a trend, that is there is no societal change that allows adults to learn (or forget) how to read. The roll out of adult literacy programs could be a factor of a positive trend in period effects (or negative if they tend to be closed down) but their effectiveness has been questioned and they often do not reach a large share of the population. Adoption of mobile phones could also be a factor in a positive trend in period effects but the relative short span of our data (on average 10 years and maximum 20 years) may not be sufficient to see a change. Note that we allow for the possibility of systematic differences between different periods because of survey effects. Indeed, although DHS and MICS surveys are of good quality, challenges in collecting data in developing countries might lead to some sampling variance between surveys and changes in reading cards or interviewer training could slightly change the grading of on the literacy scale. However, we see no reason to anticipate a systematic trend in these survey effects across multiple rounds. For these reasons, using a DP normalization, which assumes no period trends, seem the most appropriate but we will first turn a visual description of our data to see how our assumptions hold.

4.2 Review of alternative APC modeling approaches

Numerous other models have been developed to deal with APC identification. For instance, Mason et al. [1973] suggests imposing the restriction that at least two age, cohort, or period groups have identical effects in order to estimate the model. Yang et al. [2004] expand on this approach, developing a constrained generalized linear model which has been widely used in the literature. Yang and Land [2008] recommend estimation using cross-classified random-effects two-level models (CCREM) to account for the possibility individuals belonging to the same cohort or survey year may have similar outcomes because of random cohort or survey year components. They also advocate modeling age effects with a linear and quadratic term and grouping birth cohorts (usually in bins of three or five years) to break the identification problem. Clearly these choices are fairly arbitrary, and their justification depends on the question at hand.[11]

More complex methods that allow for a trend in period effects have also been developed: the Intrinsic Estimator (IE) and the Maximum Entropy (ME) method [Yang et al., 2004, Browning et al., 2012]. The IE model allows simultaneous modelling of age, period, and cohort effects without subjective choices of model constraints. The IE method imposes constraints that are unrelated to the subject studied and the variables in the model. The constraints are a function of the number of periods [Yang et al., 2008]. The IE approach is a special form of ridge regression that accounts for the multicollinearity between age, period, and cohort through singular value decomposition of matrices. In other terms, the correlated variables: age, period, and cohort variables are orthogonally transferred into linear uncorrelated variables [Luo, 2013]. Abrahamsen [2015] provides a thorough explanation of the ME approach. This method accounts for the idea that information provided by the

[11] In a setting similar to ours with U.S. data,Yang and Land [2008] compare the performance of a fixed-effect model with a CCREM to analyse 15 repeated cross section surveys on vocabulary. They show that the CCREM performs better than fixed-effect models by performing a Haussman specification test even in the case where they restrict their sample to the last five surveys. However, CCREM models have also be criticized on the grounds that, in the relatively common case where the number of categories available for cohort exceeds the number of categories for periods, the cohort random effects will have a greater variance than period random effects for a the same slope, pushing the linear cohort effect close to zero [Fosse and Winship, 2019]. Moreover, Bell et al. [2014], using Monte Carlo simulations, show that CCREM results are sensitive to the grouping of cohorts and that, in the likely case that cohorts are not grouped in bins where there is no true cohort effects, estimated cohort and period effects will be biased. When testing this model in our data, results are somewhat sensitive to the way cohort bins are defined.

data is limited and cannot provide one unique solution to solve the identification problem [Lassassi and Tansel, 2020]. The ME approach conceptualizes the uncertainty in the model and estimates the most likely solution [Browning et al., 2012].

These two methods have been shown to give similar results the one by Browning et al. [2012] and have been used in several recent papers [Lassassi and Tansel, 2020, Chłoń-Domińczak et al., 2017]. In theory, they are more flexible than the DP normalization because they allow linear period effects. In practice, Luo [2013] has shown that the IE imposes an implicit linear constraint between the slopes of the age, period and cohort effects and that this constraint depends on the number of age, period and cohort groups available in the estimation. Moreover, they are not recommended when surveys are conducted at uneven intervals, as it is often the case with DHS and MICS surveys. Using Monte Carlo simulations, Luo [2013] shows that results obtained with the IE are biased and not consistent when the data generating process does not match this linear constraint. Pelzer et al. [2015] also demonstrates the non-uniqueness of the IE method and shows that results differ significantly depending on the model parametrization.

4.3 APC decomposition using the Deaton-Paxson normalization

In this section, we present the results of APC decompositions of literacy using the DP normalization based on equation 1. All estimates have been performed separately by country and sex with at least three available surveys and 500 observations and include individuals with some primary schooling. We do not include individuals with zero years of schooling as age effects for them might be different and their inclusion could bias our predictions. Indeed, individuals with no schooling have on average very low levels of literacy (5.5 percent and 9.5 percent for 20 year-old women and men respectively) and so it can be expected that age effects for them will be either positive (if they manage to gain literacy skills as they age) or flat (if they fail to acquire new skills) but cannot be negative. An APC decomposition of literacy using the DP normalization for people with no schooling shows small and positive age effects (median effect: 0.08 pp per year or 2.4 pp over 30 years) and very small cohort effects (median effect 0.03 pp per year) (see figure 9 in Appendix). Interestingly, it also means that people with no schooling usually do not manage to learn how to read later in life; there is rarely a second chance for people who missed out on schooling when they were young. The model includes fixed effects for age, cohort and period (survey). The DP normalization can be written as:

$$\mathsf{L}_{iban} = \sum_{b=min+1}^{max} (\alpha_b + \beta_b \mathsf{G}_{ib})\mathsf{C}_b + \sum_{a=16}^{49} (\delta_a \mathsf{A}_a) + \sum_{n=3}^{N} (\theta_n \mathsf{S}_n) + u_{iban} \tag{2}$$

Where a is the age of an individual interviewed in survey n and born in cohort b and G_{ib} being the number of completed years of primary schooling. A_a takes the value one if an individual is aged a. Age fixed-effects estimated in model 2 reflect systematic differences in literacy rates for individuals born the same year but interviewed at different ages. S_n is a set of $N-2$ period dummies that have been transformed to be make period effects effects orthogonal to a time-trend and add to zero. As discussed in Deaton (1994), the transformation is:

$$\mathsf{S}_n^* = \mathsf{S}_n - [(n-1)S_2 - (n-2)S_1] \tag{3}$$

We also tested a more parsimonious model where age fixed effects are replaced by age and age squared as described in equation 4. Indeed, the large number of fixed effects can make models difficult to estimate, especially when using a Tobit estimator.

$$\mathsf{L}_{iban} = \sum_{b=min+1}^{max} (\alpha_b + \beta_b \mathsf{G}_{ib})\mathsf{C}_b + \gamma \mathsf{A}_a + \delta \mathsf{A}_a^2 + \sum_{n=3}^{N} (\theta_n \mathsf{S}_n) + u_{iban} \tag{4}$$

Equations 2 and 4 have been estimated for 24 countries for men and 47 for women with an Ordinary Least Square (OLS) estimator and a Tobit estimator. Literacy rates are bounded in the interval [0;1] and a Tobit estimator will avoid predictions outside this interval and limit the magnitude of age and cohort effects. A summary of results is of the trends in age and cohort effects are presented in figure 10 in Appendix.It shows that OLS and Tobit estimates give similar results but Tobit estimates avoid predictions outside the [0;1] interval.

4.4 A note on countries with only one or two survey rounds

As mentioned earlier, there are several country/sex cases with only two available surveys (38 cases) or even one (36 cases). In these cases, the standard DP normalization cannot work as it requires at least three periods. When two surveys are available, it is possible to fit an Age-Cohort model (AC) ignoring the effect of period or survey. In essence, it makes the difference between surveys entirely attributable to age effects, which given our assumptions of zero or small period effects seems reasonable. The risk is that if the differences between two surveys actually reflects a survey effect, spurious age effects could be estimated, which would bias cohort effects. In order to estimate an AC model, only controls for age and cohort effects are included. However, it is not possible to control for both age and cohort fixed effects with only two available surveys as the number of unique age and cohort combination is lower than the total number of parameters to be estimated. The model is estimated by including age and age squared instead of age fixed effects. In the case of only one available survey, the fitted model is a pure Cohort model (C), as no separate age and period effects can be estimated. Since, in most cases, age effects tend to be negative not controlling for it could bias cohort trends upward.

To test the sensitivity of our results to the inclusion of age and period effects, we run an AC and C models on country/sex cases with at least three available surveys and compare it to the APC results. Results can be seen in figure 11 in Appendix. Ignoring period effects does not change the estimated age and cohort effects in a meaningful way and makes it possible to obtain reliable estimates when only two surveys are available. However, the cohort-only model produces trends in cohort effects which tend to be flatter. Nevertheless, we report results from countries with only one available survey using this cohort-only approach, as the trends across recent cohorts remain informative, but it seems likely that the long term-trends here may underestimate the true decline in school quality.

5 Main results: cohort effects in school quality around the world

Estimating the age-period-cohort model described above provides us with measures of cohort effects in school quality for each country in our sample: 87 countries in the case of women, and 58 for men. In most countries the trends span about 40 years, from the cohort born in the mid 1950's to the cohort born in the mid 1990's. All the results discussed in this section are based on the Tobit model with age and age squared for countries presented above, but the qualitative conclusions of our discussion are not driven by the choice of the specification or by the inclusion of cases with one or two available surveys.

5.1 Graphical analysis for selected large countries

Following Yang and Land (2013), we first present a visual description of the data to detect potential age or period effects and see if our assumptions on age and period effects might hold. Figure 2 shows the literacy rate across cohorts and surveys for women with four to six years of primary schooling in the eight largest countries in our dataset with at least four surveys, with the exception of India where only two surveys are available.[12] If there were no age, period or survey effects, we would expect that literacy rates for women born the same year but interviewed at different dates would be on average the same and literacy rate trends estimated with different surveys would overlap. This tends to be the case in countries such as Indonesia, Nigeria or Egypt. In these three countries, it is unlikely that there are large period or age effects as women born the same year tend to have roughly the same literacy rate when they are interviewed at different dates.[13] In the cases of Bangladesh, Tanzania or Kenya, there are relatively large differences between surveys but no obvious pattern of a decrease or increase in literacy across surveys, suggesting that differences might be due to survey effects. Finally, in the case of Ethiopia and India, there tends to be systematic differences with women interviewed in more recent surveys having lower rates of literacy, suggesting negative period or age effects. This seems particularly large in Ethiopia, where women born the same year seem to have lost about 30 percentage points of literacy between the 2000 and 2016 survey. The 10 percentage-point loss in India in 10 years is also relatively large but, as only two surveys are available, cannot be unambiguously attributed to either period or age effects. Based on additional data analysis (see Appendix A.5), however, we argue that the difference in literacy rates between the two surveys in India is mainly due to age effects. Overall, this descriptive analysis suggests that we should expect moderate age or period effects in most cases. Period effects seem to matter and could create spurious age or period effects, especially in cases with a low number of surveys.

Systematic differences in literacy rates between surveys could also be due to differences in sampling. If this were the case, we would also expect to see differences across survey rounds in other indicators, including time invariant characteristics like schooling. However, we show in figure 3 that the share of women born in the same cohort accessing grade 5 is very stable across surveys. For instance, while literacy conditional on schooling for women

[12]Data for India are also shown in the graph because Indian population makes up about one third of the total population of the countries in our sample and Indian trend can have a large impact on the global trend.

[13]We cannot rule out the rather implausible case of large and opposite age and period effects cancelling each other.

in a given cohort in Ethiopia appears to decline between survey rounds systematically from 2000 and 2019, schooling does not. This pattern is consistent with the idea that literacy may decay over the lifecycle, but schooling, by definition, cannot.

We turn now from 'raw' data to the core estimation results. Figure 4 shows the estimated age, period and cohort effects for women for the eight large countries discussed previously. The estimates are based on the Tobit estimator using the Deaton-Paxson normalization. In all countries, period effects tend to be very small, suggesting that there are no large systematic differences in literacy rates between survey rounds once we control for age and cohort effects.

Age effects, shown in the first column of Figure 4 tend to be either flat or declining. This suggests that for this subset of countries – especially India, Bangladesh, and Ethiopia – rather than acquiring human capital in the form of literacy over their life cycle, women tend to lose it through a process of decay during their adult lifetimes. Kenya and Nigeria also show slight evidence of such decay, with only Egypt showing any sign of human capital acquisition over women's adult lifetimes.

The trends in cohort effects, which are our main object of interest, show a fairly sharp downward trend in India, Nigeria, and Ethiopia. Taken at face value, the estimates for India suggest the probability a woman with five years of schooling would become literate was nearly 100 percent for the 1960s birth cohorts, but had fallen to roughly 40 percent for the mid-1990s cohorts. The India results – again, based on just two survey rounds – are somewhat fragile to our modeling assumptions though, which we explore in detail in Appendix A.5. Ethiopia has experienced a similar decline in school quality as India, though with a slightly less dramatic downward slope. Nigeria, for its part, started off at a lower baseline level of school quality, and has reached a lower point today: with only about 20 percent of adult women educated in the late 1990s able to read.

The other five countries in Figure 4 show relatively flat trends in school quality, as measured by cohort effects. Their quality levels though, are far from uniform: while five years of schooling are a virtual guarantee of literacy in Indonesia, that is true for only less than half of educated women in Egypt. The remaining large countries (Bangladesh, Kenya, and Tanzania) fall somewhere in between, and all of them outperform Nigeria.

5.2 On average, school quality in the developing world has declined steadily over multiple decades, albeit with significant heterogeneity in levels and trends

Now we turn to a summary of findings for a broader set of countries. There are too many countries in the sample to discuss each in detail, so we focus on regional averages, leaving country-by-country statistics for the appendix (Tables 7 and 8 for women and men respectively). Regional averages are a simple average of the countries in the sample, without population weights.

We focus on three main indicators to document changes in educational outcomes: share of birth cohort accessing grade five, as well as the literacy rate and average school quality (i.e., estimated literacy for an adult with five years of schooling) where the latter two statistics are both corrected for age and period effects per the Deaton-Paxson normalization described above.

Several broad patterns stand out in the regional trends shown in Table 3.

First, most of the gains in literacy are associated with school expansion (rather than any improvement in school quality, as we show below). Schooling figures show consistent progress in access to education for both men and women across all regions. From the cohort born in the 1960's to the cohort born in the 1990's, women's access to five or more years of schooling has increased by 42 percentage points in South Asia (from 34 to 76 percent), 25 percentage points in Sub-Saharan Africa (from 41 to 66 percent) and reached above 90 percent in East Asia and Pacific and Latin America regions. Improvements for men were less impressive as they started from higher levels but were still large (+12 pp in East Asia, +22 pp in South Asia and +17 pp in Sub-Saharan Africa). These large gains were accompanied by similar increases in literacy rates across cohorts.

Second, there is large variance across countries in the level of school quality. In Ghana, Guinea, the Gambia, Sierra Leone and Nigeria 20 percent or less of women are able to read at grade five whereas in Vietnam, Rwanda, Burundi, El Salvador, Guatemala, Honduras, Costa Rica and Bolivia more than 95 percent can read by the end of grade five.

Third, despite big cross-country differences in school quality, there are very few cases of dramatic improvements in school quality. In most cases, cross-country gaps in school quality seem to have emerged at least several decades ago. Trends over time are generally flat or declining. In two thirds of countries in our sample, the difference between the school quality faced by the oldest and youngest decadal cohorts is less than 10 percentage points. However, where trends in school quality exist, the trend in school quality tends to be negative. In 56 countries in the sample of women, school quality decreased between the cohort born in the 1960's and the cohort born in the 1990's, while it increased in 14 countries and did not change by more than +/- 5 percentage points in the remaining 17 countries (see Figure 6). Although some countries that were already performing well have managed to improve their school quality (e.g. Vietnam, Pakistan, Peru), no country that was doing poorly in the past has managed to improve its school quality meaningfully.

Finally, despite the general picture of stagnation, regional trends show strikingly different patterns in the evolution of school access and – the more novel contribution of this paper – school quality. To highlight the latter point, Figure 5 shows the trends in school quality rates for women and men on a consistent sample of countries from the birth cohort born in 1955 to the birth cohort born in 1998. We observe five regions that have followed different trends in educational outcomes; East Asia and Latin America show large improvements in access to schooling and literacy rates, and relative stability in school quality over multiple decades; The Middle East and North Africa, as well as South Asia, have experienced the largest gains in literacy and access to schooling, but their trends in school quality have both declined. In the Middle East and North Africa, school quality rates started at a low level and the decrease has been moderate across the whole period. In South Asia, school quality for the 1960s cohorts was among the highest in the world, but since then the region has experienced the largest drop in school quality of all regions (-31 pp for women and -35 pp for men). Finally, Sub-Saharan Africa has experienced modest progress in literacy and access to schooling compared to the two previous regions; however, it has experienced the second largest drop in school quality (-24 pp for women and men).

5.3 The gender gap in school quality has decreased over time (in conjunction with a declining gender gap in school access)

School quality for men and women can be compared for a sub-sample of countries and cohorts (see Table 4). In Sub-Saharan Africa, the number of countries with a positive gender gap in favor of men has remained stable between the 1960's to the 1990's with 12 countries out of 35 countries with data for men and women. The gender gap in school quality has decreased slightly but remained very high (+9 percentage points for men). Unlike Sub-Saharan Africa, the gender gap has decreased in South Asia and school quality for men and women has converged over the past three decades and school quality for the last coohort born in the 1990's is almost identical. In East Asia and Pacific, the gender gap has remained positive for women between the 1960s and the 1980s.

5.4 School quality is correlated with long-term growth

Our estimates allow us to revisit the relationship between schooling, school quality and economic growth [Mankiw et al., 1992, Wossmann and Hanushek, 2007]. Indeed, our estimates of school quality and schooling covers long trends allowing us to estimate the stock of human capital available at different dates and see if it explains subsequent economic growth. We followed the literature and specified the growth equation as a function of initial income per capita and the stock of human capital measured by the quantity of schooling and school quality. We use our measure of school quality, expected literacy at grade five, as a measure of school quality and two separate measures of school quantity: the number of years of schooling of the adult population following Barro and Lee [2013], and the share of the population accessing grade five from our own estimate such as:

$$\mathsf{Y}_i = \alpha + \beta I_i + \gamma S_i + \delta Q_i + u_i \tag{5}$$

where Y_i is the economic growth per year from year i to the latest available data (2019), I_i is the GDP per capita in year i in logarithm, S_i is the quantity of schooling of the adult population either measured as number of years of schooling of the adult population aged 20 to 40 from Barro and Lee [2013] or measured with the share of the female population aged 20 to 40 (or the highest available age) with at least five years of schooling at year i and Q_i the average expected literacy rate at grade five of the population aged 20 to 40 at year i.

Results are presented in Table 5. They show that for our two specifications, there is a correlation between school quality and future economic growth. For instance, economic growth from 2010 to 2019 was 0.7 percentage points higher in countries with one standard deviation (sd=0.23) higher school quality.

6 Accounting for changing patterns of selection into school

So far we have followed the convention in the literature by defining 'school quality' as a simple conditional average — the literacy rate conditional on five years of schooling. Obviously, however, changes in school quality defined this way may reflect changes in schools' value added, or changes in the composition of pupils. The expansion of schooling under free primary education reforms in the 1990s and 2000s, for instance, incorporated a growing share of children from more disadvantaged backgrounds who might have more difficulty learning to read [Deininger, 2003, Bold et al., 2015]. More speculatively, as cost barriers to secondary schooling are lowered, progression into secondary may hinge more and more on academic ability rather than financial resources, rendering those with only primary schooling more negatively selected and less likely to be able to read. As we rely on survey data on adult respondents, we have very little information on the socio-economic backgrounds of individuals during childhood to control for these factors.

In a study using panel data from Vietnam, Peru, India and Ethiopia, Singh [2019] finds that differences in test scores between countries are mainly due to differences in school quality rather than differences in child's endowments, suggesting that differences in composition of students may not be driving our results. The literature on the Mincerian wage returns to schooling is also potentially informative here. While we are not concerned with earnings, there is a general consensus that OLS estimate of the returns to schooling, while theoretically subject to the same upward bias due to unobserved ability that concerns us here, appear quite similar to well-identified estimates from natural experiments Card [1999]. Nevertheless, there is recent evidence from secondary schooling in Ghana that (upward) selection bias is sometimes real and significant Ware, E. O., Kornu, D. D. K., & Adusei [2018]. Estimates of school quality are still informative even if they are partly determined by the composition of students, as they reflect the ability of school systems across countries and over time to teach literacy to students currently attending. But if trends are mostly driven by changes in patterns of student selection, it changes how we should interpret them. Our results show that school performance trends are relatively flat in most countries, in spite of massive changes in the population of students attending school. This could show that selection does not play a major role in changes in school quality or that countries have been able to slowly adjust their school performance over time in order to keep it roughly constant.

In the absence of a valid instrumental variable for access to schooling, it is hard to estimate precisely the impact of selection on school quality. In this section, we explore in more details the possible magnitude of selection and how the cohorts of women and men have changed over time.

6.1 Comparing height between women who did and did not go to school as an indicator of selection

While most selection concerns suggest a downward bias in our estimates of school quality over time, a countervailing mechanism is that as most countries have experienced considerable economic growth and increased schooling over the period covered here, a larger share of parents will have been to school and living conditions would have improved. Therefore, the endowments of the average student will not necessarily drop over time, even if access to the school system has expanded. One way to see how the endowments of students have changed over time is to look at the height of adults, which is recorded for women in 64 countries in our dataset. Adult height is correlated with nutritional status and living conditions during childhood [Garenne, 2011] and can be an indicator of absolute and relative endowments of individuals when they were growing up. Changes in absolute height over time can reflect that living conditions have improved and that new generations are better off than older generations. Also, since height is linked to living conditions, differences in adult height should reflect differences in parental income and, therefore, differences in height between women with some schooling and women with no schooling should reflect the wealth gradient in access to school. Figure 7 shows the trends in female adult heights by schooling status.

Trends in height for all women tend to be positive, with women gaining on average 0.2 centimeters per decade, suggesting that new generations of women tend to be better off. However, in 20 countries, mainly in Sub-Saharan Africa, trends in female adult height are actually negative and the latest cohorts of women in these countries might have actually benefited from a less favorable environment when they were growing up than the first born cohorts.The trends in height for women with some schooling are highly correlated with the trends in height for all women (r=0.94) but are much lower on average (0.07 centimeters per decade) and the trends are even negative in about half of the countries. So, endowment of women going to school measured by their adult height seems to have been relatively stable in most cases. This is due to the fact that the composition of women attending school has changed over time. For women born in the 1960's, height of women going to school was on average 0.47 taller than the general population but the difference shrunk to 0.27 centimeter for women born in 1990's.

Adult height data show an overall improvements in living conditions, although not universal. The average adult height of women with some schooling has actually been relatively stable due to the expansion of school systems and a greater access to schooling from relatively worse-off students (as measured by adult height). So, although absolute endowments of student might not have changed dramatically over time, there is evidence of a reduction in the wealth gradient in school access and the latest cohorts of students should include more students from a relatively poorer background.

6.2 Does measured school quality fall when access to schooling expands?

The correlation between the share of individuals accessing school and school quality is positive for men and women (r=0.23 for men and women) across the whole sample. The positive correlation is mainly driven by the correlation between countries (r=0.43 for men and r=0.38 for women), while the correlations within countries are negative (r=-0.3 for men and r=-0.38 women). The results suggest that countries with higher access to schooling have better school quality (likely they are richer countries with better school systems) but, as the share of individuals accessing school increases in a country, school performance decreases.

However, these correlations could be biased by the secular trends in education expansion and school quality. The association between selection into school and school quality can be tested more formally using results from cohort estimates that provide long time series for each country. Table 6 shows the results of country fixed-effects models for women and men respectively where the dependent variable is school quality and independent variables the year of birth, to estimate secular trends, the share of a cohort with some schooling and the share with the secondary schooling, to capture selection into school. Also, administrative data on repetition rates, available from the 1970's to test if changes in school quality can be explained by changes in repetition practices[14], have been matched with birth cohorts when students were 10 year old. The indicator for repetition rate is the cumulative repetition rates from grade one to five, that is the number of extra years of schooling a student would have at grade five if she had experienced her country repetition rates when she was 10. Finally, for a sub-sample of countries, we control for the height of women with some schooling.

[14]Source: UIS data. Data are matched for men and women separately. Gaps in data are interpolated

For women and men, columns (1) show secular decreasing trends of about -4.6 and -5.9 percentage points per decade respectively in school quality. Controlling for the share of the cohort that has accessed school or has reached secondary schooling (column (2)) has little impact in the man sample but reduce by two the decreasing trend for women. Also, the share of the cohort with no schooling is correlated with expected literacy at grade five such as in a case where all individuals access school compared to the case when none access it expected literacy at grade five decrease by 11.5 percentage points for women and, similarly, as more women reach secondary school school quality decreases. This drop could be due changes in selection of students or to the fact that as school systems expand it is harder and harder for governments to keep quality constant. Assuming that the association between the share of students with schooling and school quality is entirely due to selection would mean that trends of school quality should be flatter keeping the endowments of students constant. For women, the 2018 population weighted average of women born in 1960's not going to school was 42.7 percent decreasing to 17 percent for the ones born in the 1990's while the ones accessing secondary school increase from 31.5 percent to 65 percent. In the meantime, school quality decreased by 23 percentage points from 82 percent to 59 percent. Results suggest that increase in schooling access explains 6.9 percentage points of this decrease in school quality (21 percent).

Columns (3) show that changes in repetition rates are associated with changes in school quality for both men and women such as an extra year of schooling at grade five due to repetition is associated with a 12.1 and 6.8 percentage points increase in school quality respectively. This is a relatively large effect; our estimates show that school quality rates are about 73 percent for men and women in the whole sample, that is about 14.6 percentage points per year of schooling. It would suggest that an extra year due to repetition is worth about half to more than three quarters of a year of schooling. Controlling for repetition rates has little effects on the estimated secular trend. Finally, in column (5), we control for the height of women with schooling, which is positive and statistically significant at 5 percent.

Overall, selection into school and changes in repetition practices may explain some of the downward trend in school quality but not all of it. Interestingly, changes in repetition practices seem to have large effects in school quality. Can differences in repetition rates explain the large differences observed between countries? Although some of the best performing countries have very high levels of repetition (e.g. Madagascar, Rwanda or Burundi), the magnitude of the association between repetition and school quality is too small to change the ranking of countries. Our estimates suggest that a total ban of repetition would drop expected literacy at grade five by 10-15 percentage points for Burundi, Madagascar and Rwanda putting their school quality measure around 80 percent, still way above countries like Ghana or Sierra Leone.

6.3 Free primary education as an exogenous shift in access

Finally, the impact of selection and large school expansion can be tested by looking at the impact of free primary education (FPE) policies. Over the past 40 years, many developing countries have passed reforms to make primary education free. FPE reforms have been credited with an improvement in access to schooling (REF). However, there have been concerns that the sudden influx of new students has led to a deterioration in school quality. We test this hypothesis using our long term trends in school quality. For all FPE reforms we identified in developing countries, we estimated an interrupted time series model with a break for the birth cohort that was eight when the FPE reform was introduced.[15] Models were estimated when at least ten years of data were available before and after the reform and estimates and limited to 20 years before and after the reform. Estimates were weighted by the size of the sample of the cohort to give more weight to more precise estimates.

51 estimates of FPE reforms could be estimated. Changes in trends in school quality before and after FPE reforms are shown in figure 8 for women and men separately. In 27 cases, we observe no statistically significant change in the trend of school quality before and after the implementation of the FPE reform. There are more statistically significant negative effects than positive ones (20 vs 3) and the mean effects are -0.4 percentage point per year for women and -0.9 for men with only few cases of a very rapid drop in school quality after the introduction of FPE reforms. Although the general trends in school quality were already decreasing before the FPE reforms (-0.2 percentage points for women and men), average declines in school quality have accelerated substantially after the reforms (-0.6 and -1.1 percentage points per year for women and men respectively) and could explain a large share of the drop in quality in the sub-sample of countries having introduced FPE.

It is worth noting that the impact of FPE reforms on school quality may not only be due to the change in the composition of the cohorts attending school. A large and sudden

[15] We tested different ages for which the FPE reform was introduced (6, 8, 10 and 12) and our results stand.

expansion of the number of students attending school,[16] could be hard to manage for school systems and reduce available resources for students. Thus, the negative impacts of FPE reforms are probably an upper bound of the effects of selection.

Overall, the data show the composition of cohort of students has changed over time and that a larger share of students accessing schooling is associated with a drop in quality. FPE reforms are associated with an acceleration of the negative trend in school quality but school quality was decreasing before the introduction of these reforms and FPE reforms might also have been associated with a reduction of resources per students.

7 Conclusion

We use literacy tests in survey data to construct long-term trends in literacy for 87 developing countries, spanning birth cohorts from the 1950s to 2000. We show that over this period literacy rates have increased substantially, but virtually all progress has been due to the increase in access to school rather than any improvement school quality, which we define as the propensity for schooling to generate literacy after five years of schooling.

Overall, school quality is low in developing countries with about 70% of women able to read after grade five and quality has been declining over time. From the 1950s to 2000, school quality deterioration implied the probability that a woman with five years of schooling could read a sentence fell by roughly two to four percentage points per decade and about six percentage points for men.

Although the negative trends in school quality is concerning, a more generous reading of the results is that most school systems have managed to dramatically expand their education offer without very large drops in school quality. Trends in school quality are relatively stable over time, and there are few if any identifiable cases of large and rapid improvements in school quality at the national level. These patterns cast doubt on the effectiveness of education reforms as conceived to date to dramatically raise the ability of school systems to teach students how to read, and highlights the central role in expanding school *access* in historical gains in literacy.

[16] Interrupted time series estimates of the share of students accessing grade five before and after the FPE reforms show that the mean trend increased from 0.7 percentage point per year to 1.2.

Two limitations of our approach bear repeating. First, our reliance on simple literacy tests as a measure of human capital clearly ignores important higher-order skills, not to mention numeracy and other domains. Nevertheless, we argue that literacy remains a useful proxy for learning levels in many school systems, where literacy is far from universal, and puts our focus on the lower parts of the learning distribution which are often ignored in favor of policy attention to students passing competitive exams for elite secondary and tertiary schools.

A second limitation relates to the possibility of unobserved selection into schooling (or into higher grades) which may bias our estimates of school quality which are based on the association between schooling and literacy. Of particular concern is that changing patterns of selection, e.g., enrollment of more low-income or otherwise under-prepared students as schooling systems have expanded, might create a spurious impression of declining school quality. We have shown that despite the overall trend toward higher enrollment and lower measured quality, countries and periods with faster enrollment growth did not experience significantly faster deterioration in school quality. Furthermore, correcting for differences across countries and periods in grade repetition rates, or the height of individuals (an indicator of socio-economic status that is mostly determined prior to schooling) does little to change the picture of declining quality.

In terms of policy implications, the fact that countries with nearly universal literacy got there by increasing average years of schooling rather than school quality may be instructing for lagging countries. But the impact of expanding enrollment clearly hinges on quality levels. The large variance in school quality levels across countries coupled with an apparent lack of catch-up by poor performers suggests an opportunity for greater knowledge transfer and even policy replication within the region. An obvious question is why this hasn't happened to date. The paucity of data on school quality may have made it difficult to reliably identify top performers. In this case, additional efforts to measure the quality of school systems in a reliable way might be a way to help low performers to improve. However, the relative stickiness of school quality over time could alternatively be seen as evidence that factors that affect school quality are deeply embedded in the system and not easily portable to new contexts.

Turning to implications for future work, our hope in embarking on this project was that by constructing long-term trends in school quality, we could identify countries who have made transformational changes, and risen to the top of the ranks in school quality in the region. Such exemplars appear not to exist. Instead, the rank-ordering of countries' school quality has been relatively stable over time. Nevertheless, moderate improvements over the period of a decade or more do exist, and future research could usefully complement the large literature on policy evaluations by taking a historical lens to try to understand these successful reformers. Future research could also usefully shed light on possible explanations for the large gaps in school quality between East Africa (e.g., Rwanda, Burundi, etc.) and West Africa (e.g., Ghana, the Gambia, Niger, etc.), though time-series data may be of little help here as these gaps appear somewhat invariant over several decades. Finally, estimated school quality appears to be a good proxy of human capital as can been seen in the link between economic growth and school quality and our indicator of school quality can be used in future work aiming to explain the role of human capital in long term economic growth.

References

Global Data Set on Education Quality (1965-2015). 2018. URL https://doi.org/10.1596/1813-9450-8314.

Helen Abadzi. *Improving Adult Literacy Outcomes Lessons from Cognitive Research for Developing Countries*, volume 0932. 2003. ISBN 0821354930.

Camilla Karto Abrahamsen. Labor supply over the life cycle-an age-period-cohort analysis. Master's thesis, 2015.

Tahir Andrabi, Jishnu Das, Asim Ijaz Khwaja, and Tristan Zajonc. Do value-added estimates add value? Accounting for learning dynamics. *American Economic Journal: Applied Economics*, 3(3):29–54, 2011. ISSN 19457782. doi: 10.1257/app.3.3.29.

Julius F Atuhurra. Does community involvement affect teacher effort? Assessing learning impacts of Free Primary Education in Kenya. *International Journal of Educational Development*, 49:234–246, 2016.

Robert Barro and Jong-Wha Lee. International Trade, Distortions and Long-Run Economic Growth. *IMF Working Papers*, 92(90):i, 2013. ISSN 1018-5941. doi: 10.5089/9781451851335.001.

Bethany A Bell, Grant B Morgan, Jason A Schoeneberger, Jeffrey D Kromrey, and John M Ferron. How low can you go? *Methodology*, 2014.

Moussa P Blimpo, Ousman Gajigo, and Todd Pugatch. Financial constraints and girls' secondary education: Evidence from school fee elimination in The Gambia. *The World Bank Economic Review*, 33(1):185–208, 2019.

Tessa Bold, Mwangi Kimenyi, Germano Mwabu, and Justin Sandefur. Can free provision reduce demand for public services? evidence from kenyan education. *The World Bank Economic Review*, 29(2):293–326, 2015.

Martin Browning, Ian Crawford, and Marike Knoef. The age-period cohort problem: set identification and point identification. Technical report, Centre for Microdata Methods and Practice (cemmap), jan 2012. URL http://cemmap.ifs.org.uk/wps/cwp021212.pdf.

David Card. the Causal Effect of Education on Earnings David. *Handbook of Labor Economics*, 3:1801–1863, 1999. URL http://linkinghub.elsevier.com/retrieve/pii/S1573446399030114.

Francesco Caselli. Accounting for cross-country income differences. *Handbook of economic growth*, 1:679–741, 2005.

Agnieszka Chłoń-Domińczak, Paweł Strzelecki, and Wojciech Łątkowski. Pension reforms or consequences of the economic crisis? measuring the changes of pension incomes in selected eu countries using apc approach. *The Journal of the Economics of Ageing*, page 100131, 2017.

Hai Anh H. Dang and Paul W. Glewwe. Well Begun, but Aiming Higher: A Review of Vietnam's Education Trends in the past 20 Years and Emerging Challenges. *Journal of Development Studies*, 54(7):1171–1195, 2018. ISSN 17439140. doi: 10.1080/00220388.2017.1380797. URL https://doi.org/10.1080/00220388.2017.1380797.

Jishnu Das and Tristan Zajonc. India shining and Bharat drowning: Comparing two Indian states to the worldwide distribution in mathematics achievement. *Journal of Development Economics*, 92(2):175–187, 2010. ISSN 03043878. doi: 10.1016/j.jdeveco.2009.03.004.

Angus Deaton and Christina Paxson. Intertemporal Choice and Inequality. *Journal of Political Economy*, 102(3):437–467, 1994. doi: https://www.jstor.org/stable/2138618?seq. URL https://www.jstor.org/stable/2138618?seq.

Klaus Deininger. Does cost of schooling affect enrollment by the poor? Universal primary education in Uganda. *Economics of Education Review*, 22(3):291–305, 2003. ISSN 02727757. doi: 10.1016/S0272-7757(02)00053-5.

Ethan Fosse and Christopher Winship. Analyzing Age-Period-Cohort Data: A Review and Critique. 2019. doi: 10.1146/annurev-soc-073018. URL https://doi.org/10.1146/annurev-soc-073018-.

Michel Garenne. Trends in nutritional status of adult women in sub-Saharan Africa. *DHS Comparative Reports*, (27):xv–pp, 2011. URL www.measuredhs.com.

Douglas Gollin, David Lagakos, and Michael E Waugh. The agricultural productivity gap. *The Quarterly Journal of Economics*, 129(2):939–993, 2014.

Kathleen S Gorman and Ernesto Pollitt. The Contribution of Schooling to Literacy in Guatemala. Technical Report 4, 1997.

Jan Eric Gustafsson. Lasting effects of quality of schooling: Evidence from PISA and PIAAC. *Intelligence*, 57:66–72, 2016. ISSN 01602896. doi: 10.1016/j.intell.2016.05.004.

Eric A. Hanushek and Dennis D. Kimko. Schooling, labor-force quality, and the growth of nations. *American Economic Review*, 90(5):1184–1208, 2000. ISSN 00028282. doi: 10.1257/aer.90.5.1184.

Eric A. Hanushek and Ludger Wossmann. Does education tracking affect performance and inequality? Differences-in-differences evidence across countries. *The Economic Journal*, 2006. ISSN 1098-6596.

S J Howie, C Combrinck, K Roux, M Tshele, G M Mokoena, and McLeod P N. PIRLS Literacy 2016 : South African Highlights Report What is PIRLS ? Participation in PIRLS cycles Two types of PIRLS Achievement Assessments Questionnaires (Contextual) Grade 4 PIRLS Literacy Attained Sample Grade 4 Achievement in PIRLS Literacy 2016. *South African Highlights Report.*, (December):1–12, 2017. doi: 10.13140/RG.2.2.11110.73282.

Aart Kraay. Methodology for a world bank human capital index. *World Bank Policy Research Working Paper*, (8593), 2018.

David Lagakos, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. Life cycle wage growth across countries. *Journal of Political Economy*, 126(2):797–849, 2018.

Moundir Lassassi and Aysit Tansel. Female Labor Force Participation in Five Selected Mena Countries: An Age-Period-Cohort Analysis (Algeria, Egypt, Jordan, Palestine and Tunisia). 2020. ISSN 2365-9793. URL www.iza.org.

Liying Luo. Assessing validity and application scope of the intrinsic estimator approach to the age-period-cohort problem. *Demography*, 50(6):1945–1967, 2013.

N. G. Mankiw, D. Romer, and D. N. Weil. A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics*, 107(2):407–437, may 1992. ISSN 0033-5533. doi: 10.2307/2118477. URL https://academic.oup.com/qje/article-lookup/doi/10.2307/2118477.

Karen Oppenheim Mason, William M Mason, H H Winsborough, and W. Kenneth Poole. Some Methodological Issues in Cohort Analysis of Archival Data. *American Sociological Review*, 38(2):242, 1973. ISSN 00031224. doi: 10.2307/2094398.

Karthik Muralidharan, Abhijeet Singh, and Alejandro J Ganimian. Disrupting education? Experimental evidence on technology-aided instruction in India. *American Economic Review*, 109(4):1426–60, 2019.

Mari Oye, Lant Pritchett, and Justin Sandefur. Girls schooling is good, girls schooling with learning is better. *Background Paper for The Learning Generation. The Education Commission*, 2016.

Dev Patel and Justin Sandefur. A Rosetta Stone for Human Capital. *Center for Global Development Working Paper 550*, 2020. URL www.cgdev.org/publication/rosetta-stone-human-capital.

Ben Pelzer, Manfred Te Grotenhuis, Rob Eisinga, and Alexander W Schmidt-Catran. The non-uniqueness property of the intrinsic estimator in apc models. *Demography*, 52(1): 315–327, 2015.

Lant Pritchett. The Rebirth of Education Why Schooling in Developing Countries is Flailing; How the Developed World is Complicit; and What to do Next. *Center for Global Development Brief*, (September 2013):1–4, 2013.

Stephanie R. Psaki, Erica Soler-Hampejsek, Jyotirmoy Saha, Barbara S. Mensch, and Sajeda Amin. The Effects of Adolescent Childbearing on Literacy and Numeracy in Bangladesh, Malawi, and Zambia. *Demography*, 56(5):1899–1929, oct 2019. ISSN 15337790. doi: 10.1007/s13524-019-00816-z.

Sara J Ruto and John K Mugo. Free primary education in Kenya . The elusive EFA dream - about to be reached ? *Zeitschrift fur internationale Bildungsforschung und Entwicklungspadagogik 28 (2005) 1, S. 21-23*, 28, 2005.

Pasi Sahlberg. Education policies for raising student learning: The Finnish approach. *Journal of Education Policy*, 22(2):147–171, 2007. ISSN 02680939. doi: 10.1080/02680930601158919.

Todd Schoellman. Education quality and development accounting. *Review of Economic Studies*, 79(1):388–417, 2012. URL https://EconPapers.repec.org/RePEc:oup:restud:v:79:y:2012:i:1:p:388-417.

Abhijeet Singh. Learning More with Every Year: School Year Productivity and International Learning Divergence. *Journal of the European Economic Association*, 2019. ISSN 1542-4766. doi: 10.1093/jeea/jvz033.

Erica Soler-Hampejsek, Barbara S. Mensch, Stephanie R. Psaki, Monica J. Grant, Christine A. Kelly, and Paul C. Hewett. Reading and numeracy skills after school leaving in southern Malawi: A longitudinal analysis. *International Journal of Educational Development*, 59:86–99, mar 2018. ISSN 07380593. doi: 10.1016/j.ijedudev.2017.08.011.

Isabella Sulis, Francesca Giambona, and Mariano Porcu. Adjusted indicators of quality and equity for monitoring the education systems over time. Insights on EU15 countries from PISA surveys, 2020. ISSN 00380121.

Stephen Taylor, Nicholas Spaull, and Stephen Taylor. The effects of rapidly expanding primary school access on effective learning: The case of Southern and Eastern Africa since 2000. *Working Papers*, 2013.

UNESCO. The global learning crisis: Why every child deserves a quality education. Technical report, 2013.

Jan Van Damme and Kim Bellens. *Countries Strive Towards More Quality and Equity in Education: Do They Show Success or Failure? Evidence from TIMSS 2003 and 2011, for Grade 4*. 2017. doi: 10.1007/978-3-319-43473-5_7.

B. M. Ware, E. O., Kornu, D. D. K., & Adusei. Enhancing the entrepreneurial skills among Ghanaian students, the bedrock of job creation (entrepreneurship education). *Journal of Emerging Trends in Economics and Management Sciences*, 9(3):169–179, 2018.

World Bank Group. Learning: To Realize Education's Promise, 2018. ISSN 0305750X. URL https://openknowledge.worldbank.org/bitstream/handle/10986/.../9781464810961.pdf.

Ludger Wossmann and Eric Hanushek. The Role of Education Quality in Economic Growth The Role of School Improvement in Economic Development. *World Bank Policy Research Working Paper 4122*, pages 1–94, 2007. doi: 10.2139/ssrn.960379.

Yang Yang and Kenneth C. Land. Age-period-cohort analysis of repeated cross-section surveys: Fixed or random effects? *Sociological Methods and Research*, 36(3):297–326, feb 2008. ISSN 00491241. doi: 10.1177/0049124106292360.

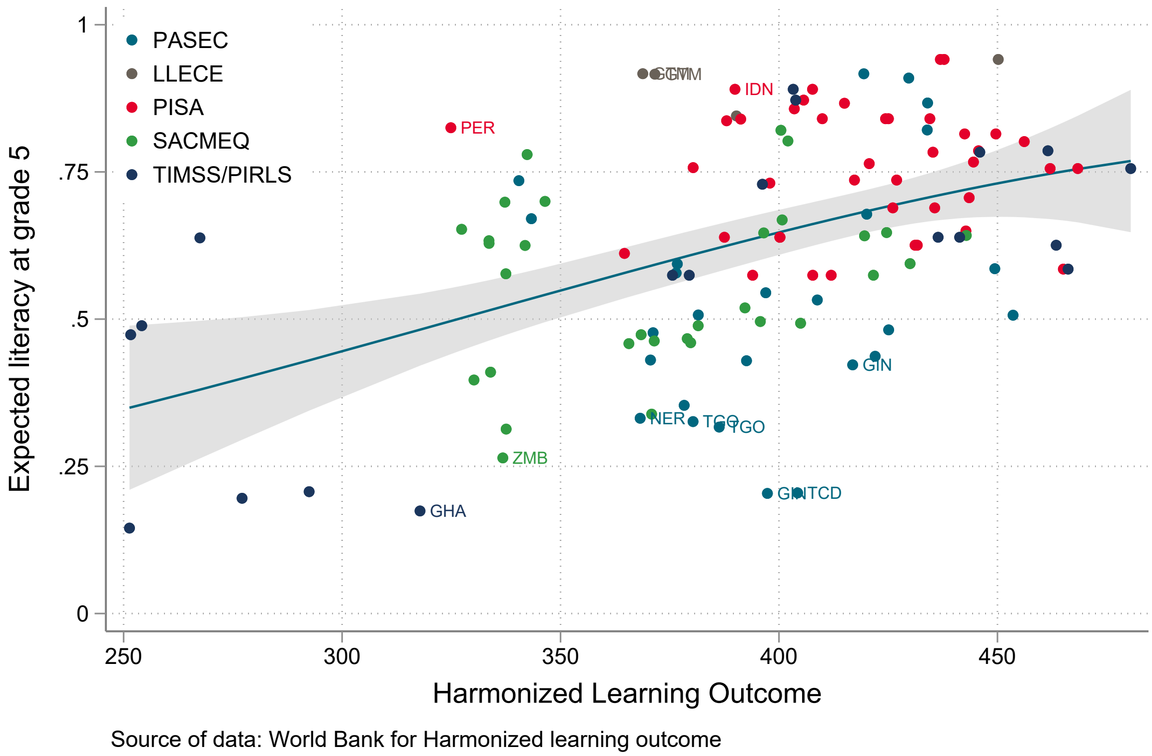
Yang Yang, Wenjiang J Fu, and Kenneth C Land. A methodological comparison of age-period-cohort models: the intrinsic estimator and conventional generalized linear models. *Sociological methodology*, 34(1):75–110, 2004.

Yang Yang, Sam Schulhofer-Wohl, Wenjiang J Fu, and Kenneth C Land. The intrinsic estimator for age-period-cohort analysis: what it is and how to use it. *American Journal of Sociology*, 113(6):1697–1736, 2008.

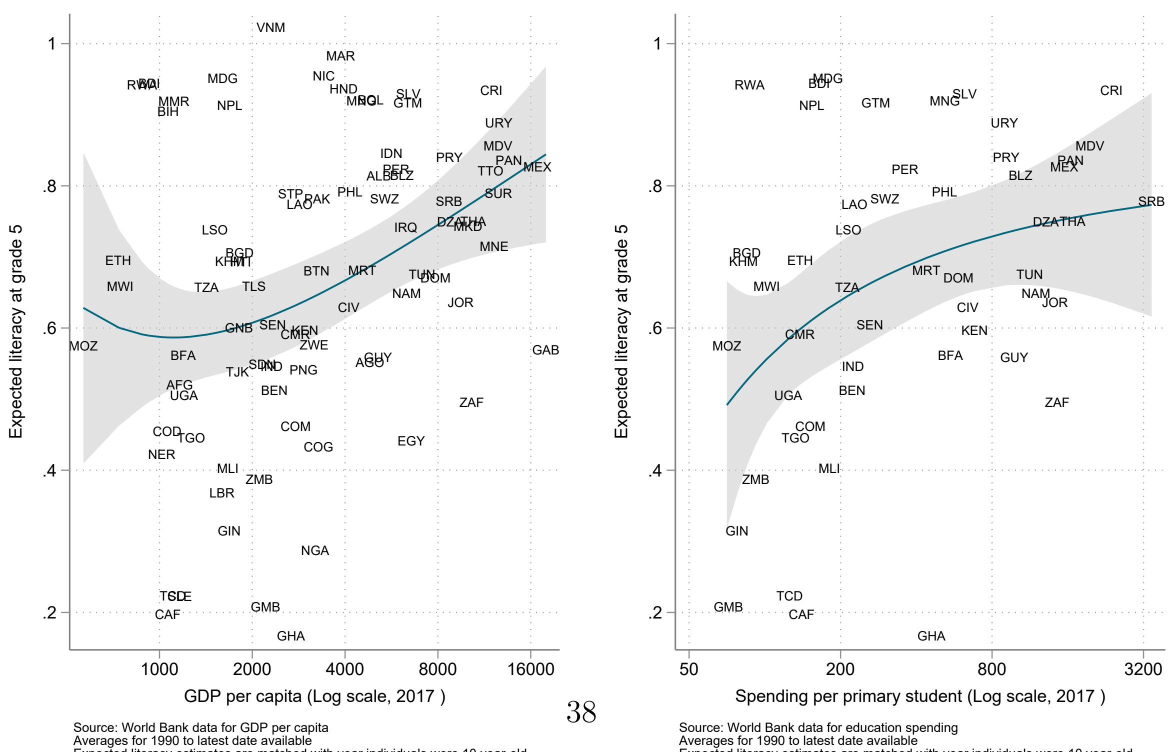
Yue Yin. Reflection on the Quality and Equity of Chinese Basic Education — From PISA 2018 Results : Where All Students Can Succeed. *Journal of Contemporary Educational Research*, 4, 2018.

Figure 1: Cross-sectional validation of school quality measure

(a) School quality and harmonized learning outcomes

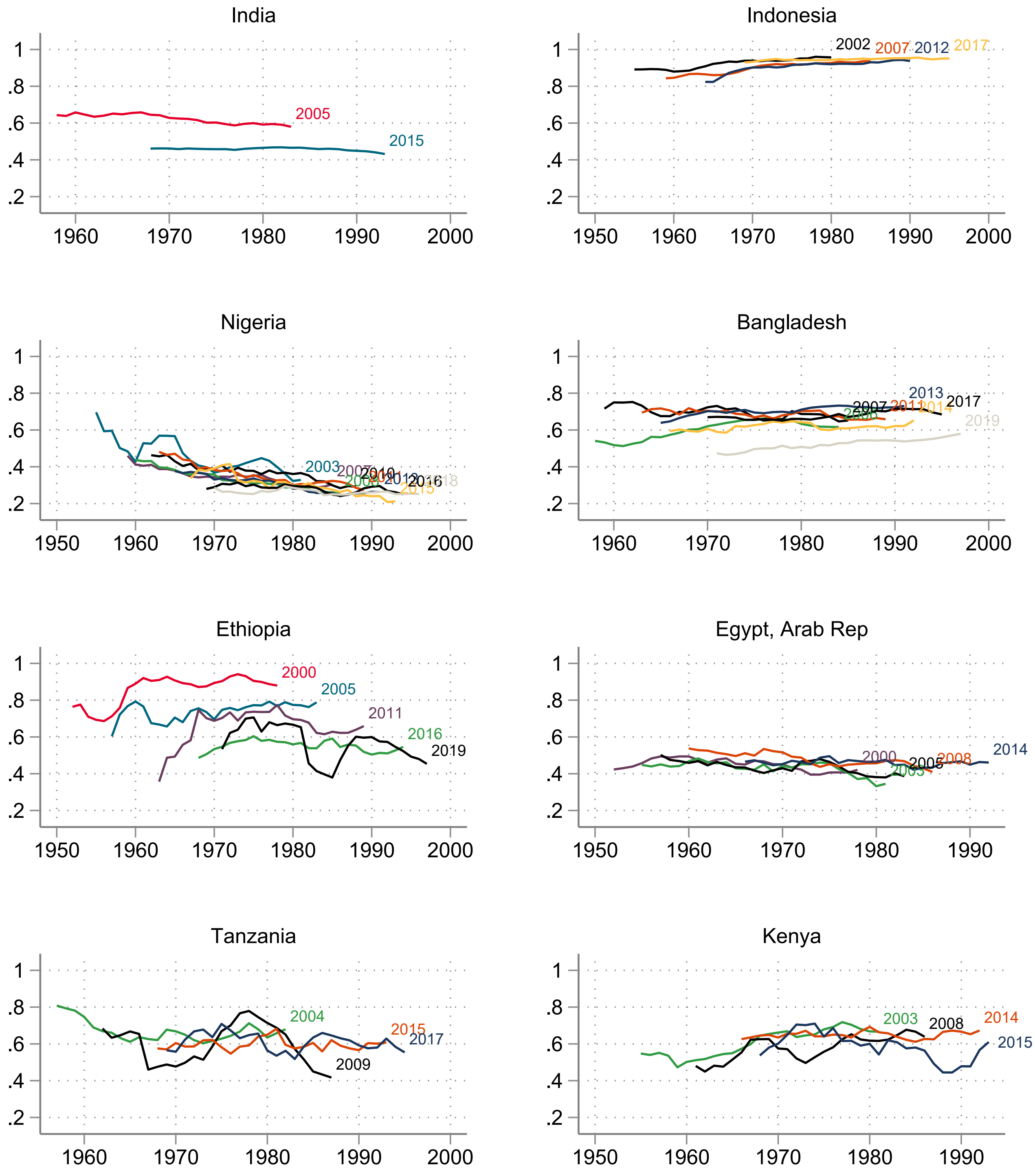


(b) School quality, GDP per capita, and education spending



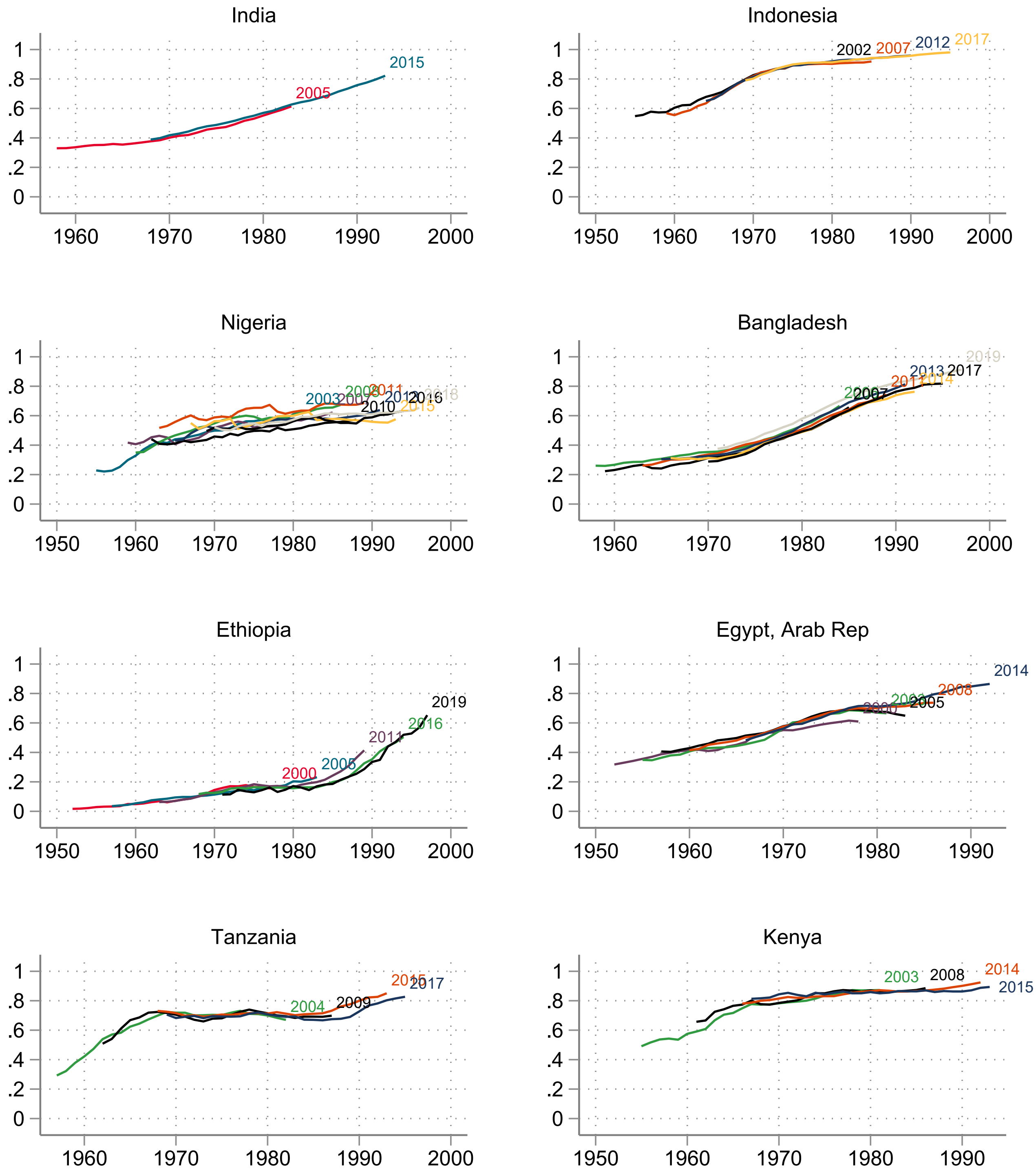
Note:

Figure 2: Literacy rates of women with four to six years of schooling (selected countries)



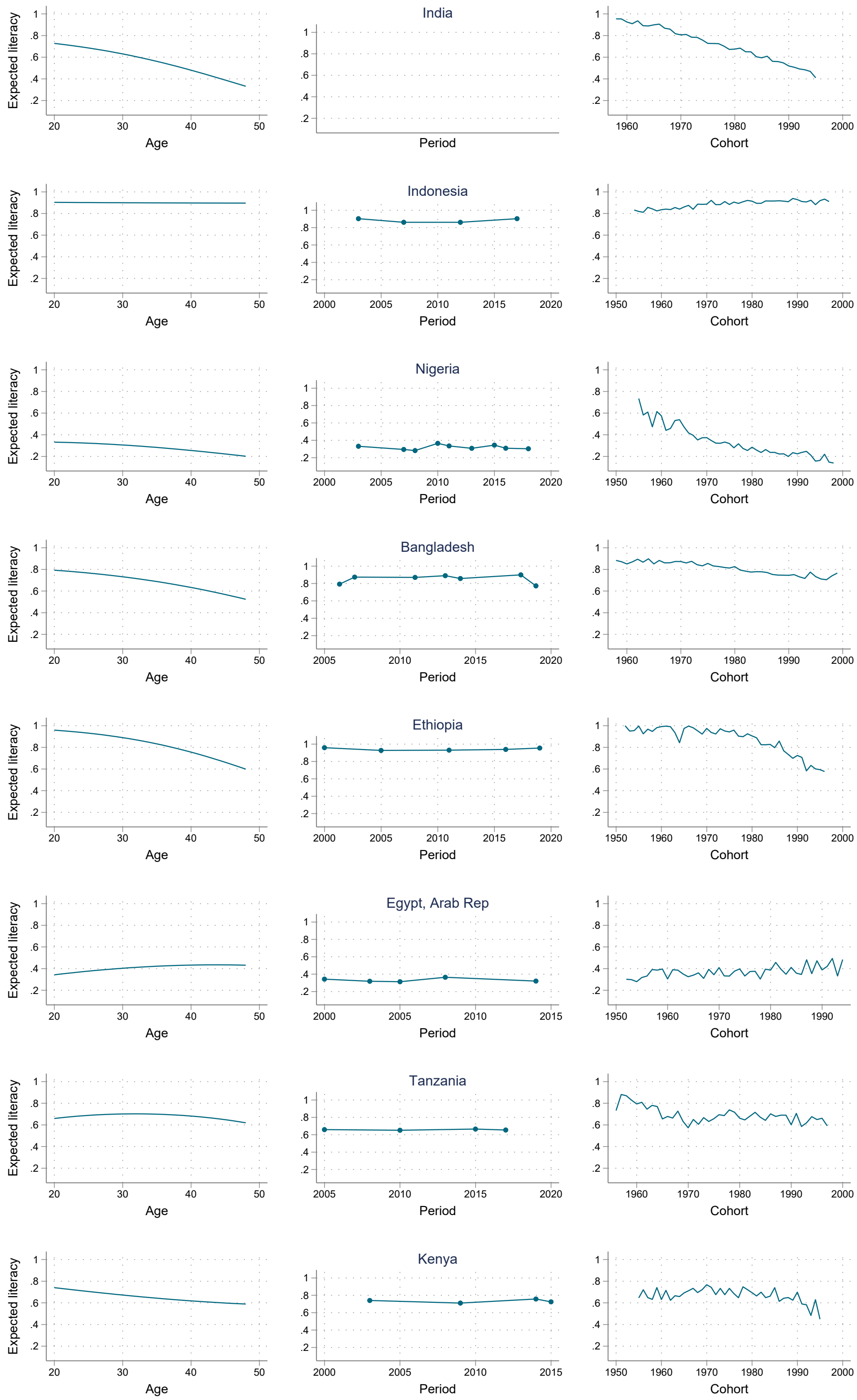
The vertical axis represents literacy rates 5-year moving average for literacy rates

Figure 3: Access to grade five for women (selected countries)



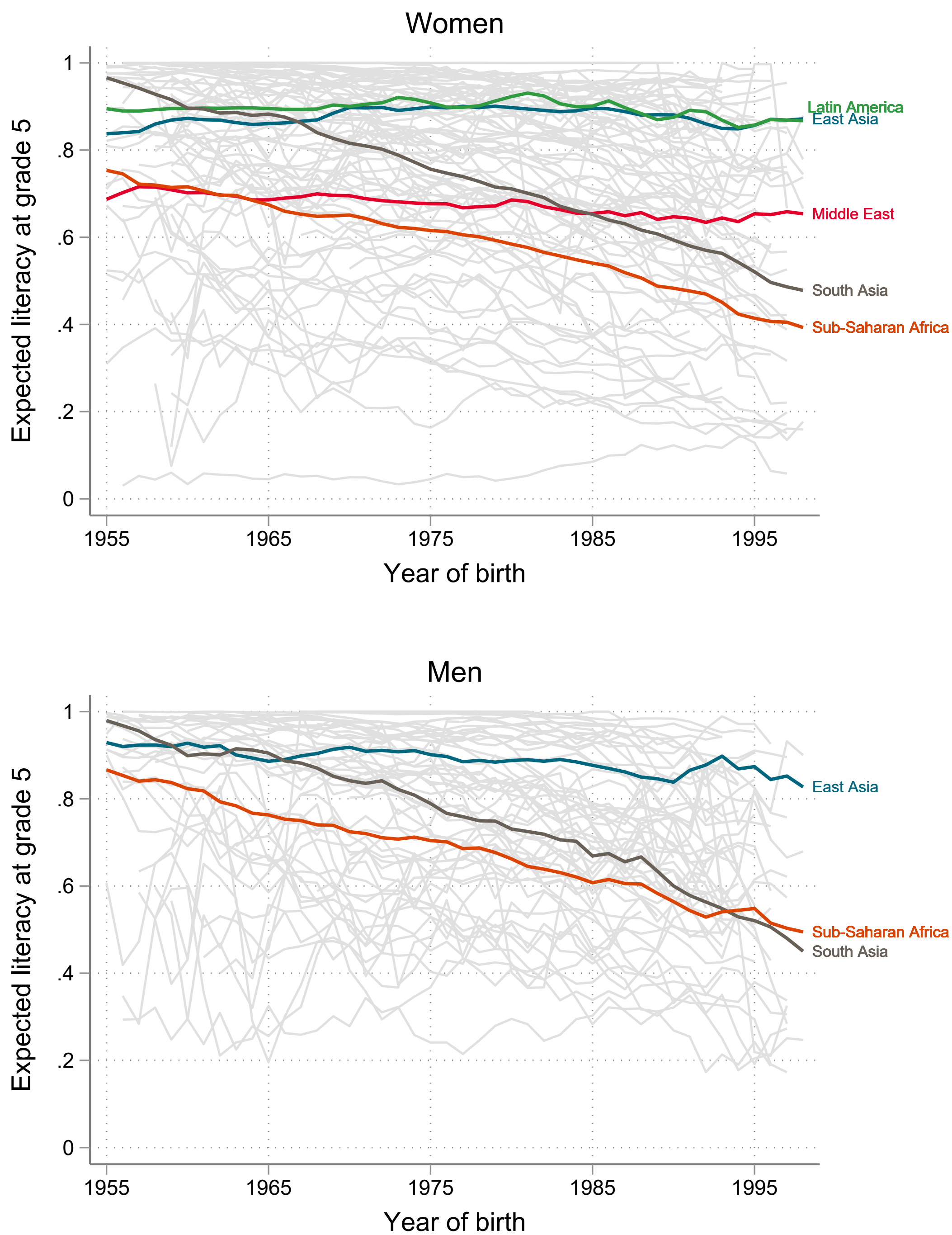
The vertical axis represents 5-year moving average of access to grade 5 or more

Figure 4: Estimated age, period and cohort effects in school quality (selected countries-women)



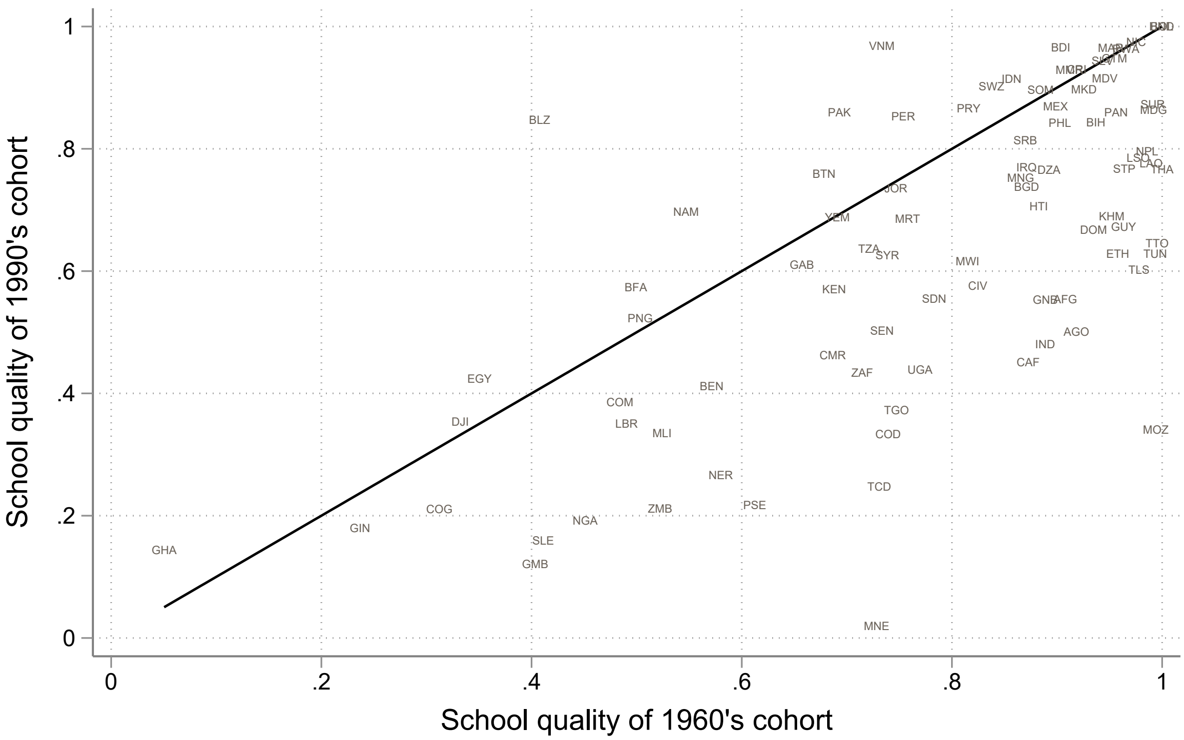
Outcome in all graphs is school quality, i.e. literacy conditional on five years of schooling. Age effects estimated at grade 5 for cohort born in 1975 and no period effect. Period effects estimated at age 20, grade 5 and 1975 cohort. Cohort effects estimated at age 20, grade 5 and no period effect.

Figure 5: National and regional trends in school quality



Note: Light gray lines show national estimates. To ensure regional trends reflect a consistent sample of countries, in cases of missing data for specific country-birth cohorts, national school quality estimates have been extrapolated backward and forward using the average, country-specific rate of change over the whole period.

Figure 6: School quality for the 1960's cohort and 1990's cohort (women)



Note: 1980's cohort instead of 1990's for BIH, BOL, CAF, DJI, MAR, MKD, NIC, PHL, SOM and SYR.

Figure 7: Trends in height for all women and women with some schooling

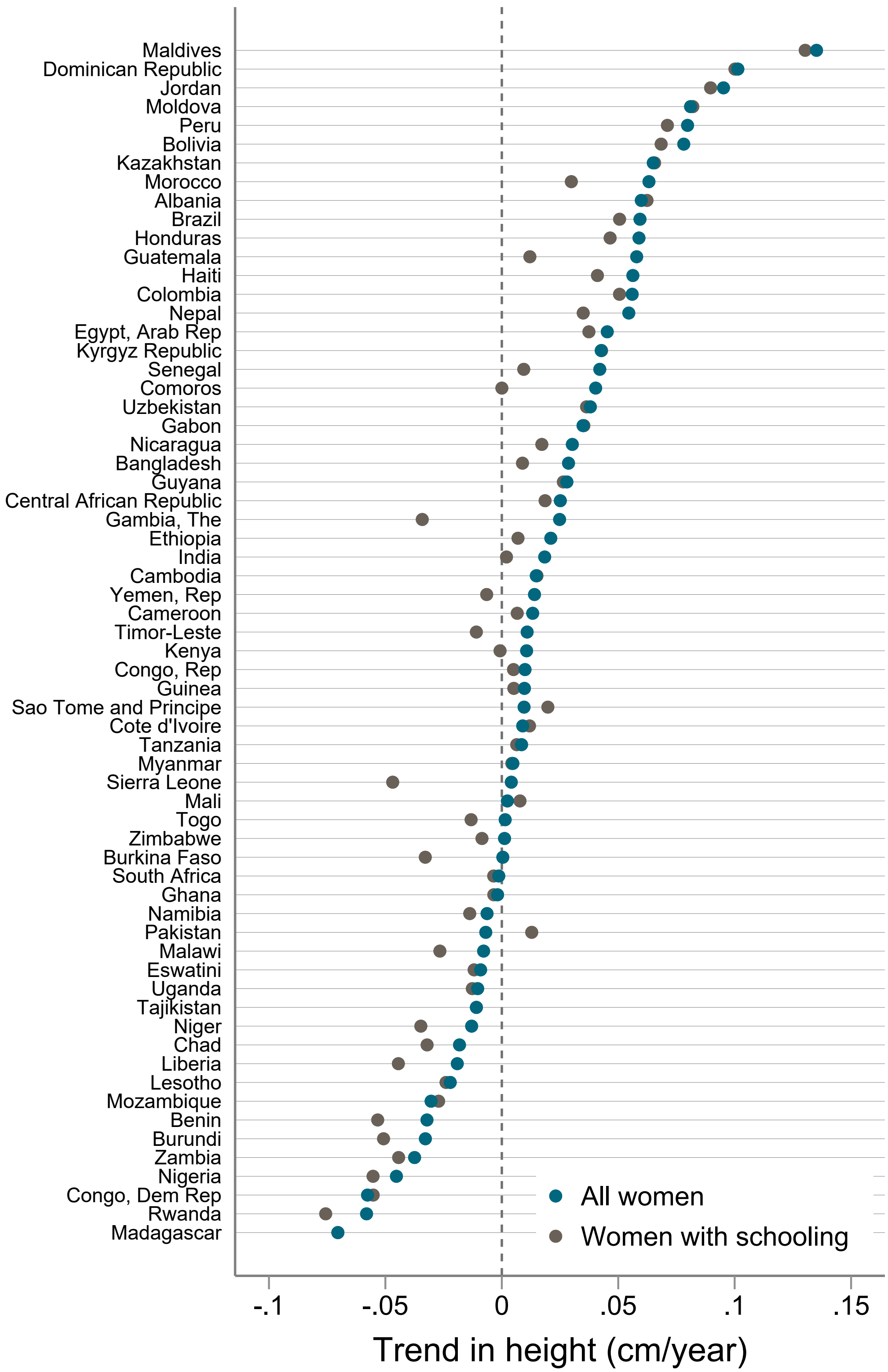


Figure 8: Changes in school quality trends after free primary education (FPE) reforms

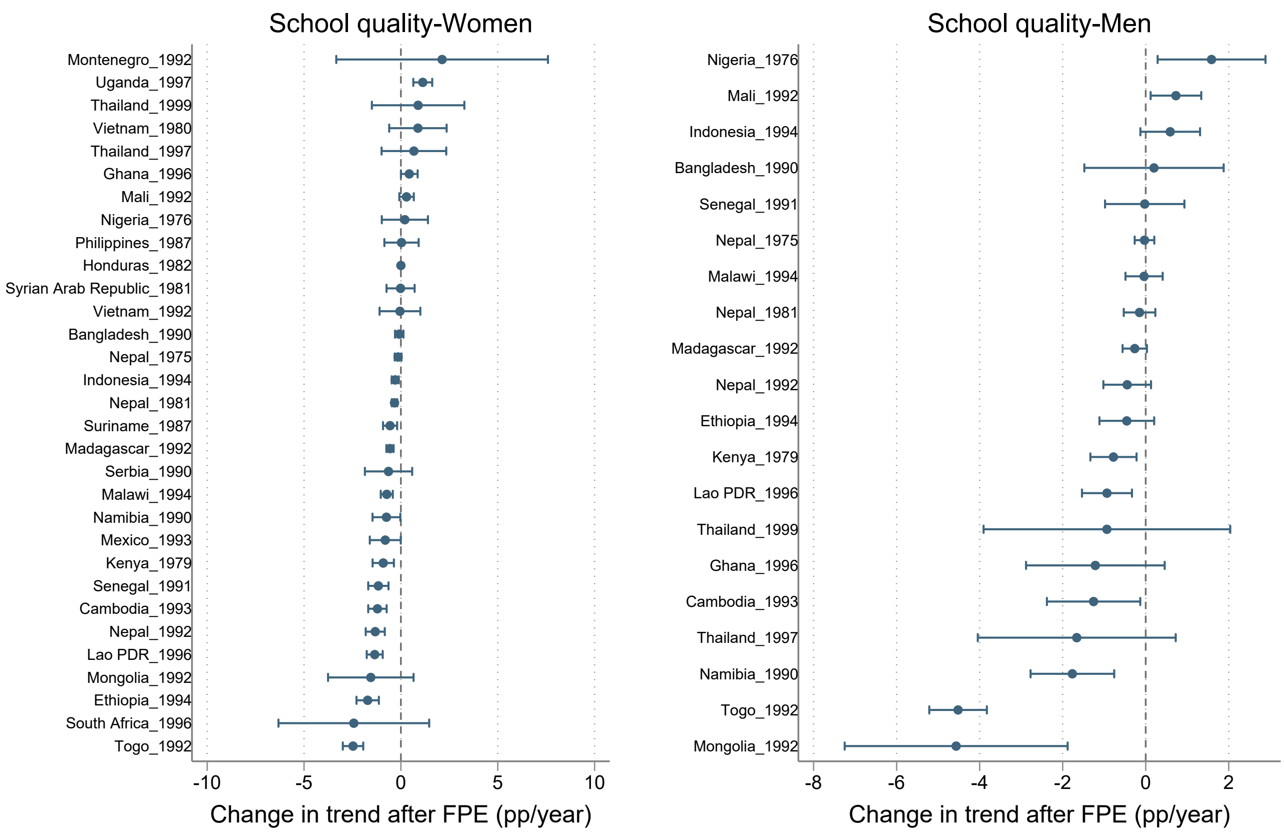


Table 1: Coverage

	Surveys	Countries	Countries with > 1 survey	Pop. share	Sample
Women					
East Asia and Pacific	27	10	8	.91	498559
Europe and Central Asia	8	4	2	.03	39137
Latin America and Caribbean	37	16	10	.41	520634
Middle East and North Africa	18	10	4	.75	389544
South Asia	22	7	5	.99	1246963
Sub-Saharan Africa	167	40	36	.97	1942746
Total	279	87	65	.81	4637583
Men					
East Asia and Pacific	16	9	5	.77	120530
Europe and Central Asia	2	1	1	.01	7505
Latin America and Caribbean	15	7	3	.1	96277
Middle East and North Africa	0	0	0	0	0
South Asia	12	6	4	.99	213121
Sub-Saharan Africa	107	35	32	.86	519795
Total	152	58	45	.67	957228

Population share excludes high-income countries except Panama and Trinidad and Tobago. Excludes China

Table 2: Literacy, share with at least five years of schooling and school quality

	Women				
	1950s	1960s	1970s	1980s	1990s
Literacy					
East Asia and Pacific	77.9	81.3	89.2	92.1	94.1
	(0.03)	(0.02)	(0.01)	(0.01)	(0.02)
Latin America and Caribbean	72.1	85.7	90.1	94.0	96.5
	(0.10)	(0.04)	(0.03)	(0.02)	(0.01)
Middle East and North Africa	36.5	50.6	64.9	73.4	83.8
	(0.08)	(0.05)	(0.03)	(0.03)	(0.03)
South Asia	33.6	36.0	46.5	62.3	75.4
	(0.02)	(0.01)	(0.01)	(0.00)	(0.01)
Sub Saharan Africa	29.8	42.4	48.1	53.4	62.7
	(0.07)	(0.03)	(0.02)	(0.01)	(0.02)
Share with $S \geq 5$					
East Asia and Pacific	58.7	66.9	83.0	87.9	92.2
	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)
Latin America and Caribbean	56.3	76.5	81.9	89.1	94.6
	(0.12)	(0.05)	(0.04)	(0.03)	(0.02)
Middle East and North Africa	34.0	48.1	62.2	70.5	83.3
	(0.08)	(0.05)	(0.03)	(0.03)	(0.03)
South Asia	30.2	34.1	45.5	62.3	76.2
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Sub Saharan Africa	26.5	41.5	48.8	54.6	66.0
	(0.07)	(0.03)	(0.02)	(0.01)	(0.02)
School quality					
East Asia and Pacific	86.5	86.9	89.7	89.1	89.1
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Latin America and Caribbean	84.4	89.4	90.7	90.5	86.2
	(0.04)	(0.02)	(0.01)	(0.01)	(0.01)
Middle East and North Africa	62.2	69.7	67.8	66.4	59.4
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)
South Asia	91.9	87.1	76.5	65.7	56.1
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sub Saharan Africa	71.2	67.8	61.8	54.1	43.6
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Countries					
East Asia and Pacific	7	10	10	10	9
Latin America and Caribbean	9	16	16	16	14
Middle East and North Africa	5	10	10	10	7
South Asia	4	7	7	7	7
Sub Saharan Africa	35	40	40	40	38

Table 3: Literacy, share with at least five years of schooling and school quality

	Men				
	1950s	1960s	1970s	1980s	1990s
Literacy					
East Asia and Pacific	86.5	88.2	91.9	93.4	93.9
	(0.05)	(0.03)	(0.02)	(0.02)	(0.04)
South Asia	69.9	64.8	72.4	78.4	82.2
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)
Sub Saharan Africa	60.6	64.9	66.3	70.3	76.4
	(0.12)	(0.06)	(0.03)	(0.03)	(0.03)
Share with $S \geq 5$					
East Asia and Pacific	74.1	78.0	86.1	89.4	90.4
	(0.07)	(0.04)	(0.03)	(0.03)	(0.05)
South Asia	63.8	59.1	69.0	76.7	81.8
	(0.02)	(0.02)	(0.01)	(0.02)	(0.03)
Sub Saharan Africa	52.8	59.5	63.4	68.6	76.6
	(0.13)	(0.06)	(0.03)	(0.03)	(0.03)
School quality					
East Asia and Pacific	90.2	90.4	90.1	87.6	88.4
	(0.00)	(0.09)	(0.08)	(0.06)	(0.05)
South Asia	95.6	89.7	79.2	68.9	55.6
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sub Saharan Africa	83.4	76.9	70.3	62.0	52.4
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Countries					
East Asia and Pacific	2	9	9	9	8
South Asia	3	6	6	6	6
Sub Saharan Africa	25	35	35	35	35

Table 4: Gender gap in school quality

	1960s	1970s	1980s	1990s
Sub-Saharan Africa				
Gender gap	-10.57	-10.46	-9.93	-8.79
	(0.00)	(0.00)	(0.00)	(0.00)
	35	35	35	35
East Asia & Pacific				
Gender gap	-.91	.57	1.41	-1.12
	(0.05)	(0.04)	(0.03)	(0.03)
	9	9	9	8
South Asia				
Gender gap	-2.47	-2.76	-3.3	.25
	(0.00)	(0.00)	(0.00)	(0.00)
	6	6	6	6

Gender gap=female school quality- male school quality. Results weighted by population of the birth cohort

Table 5: School quality and economic growth

VARIABLES	(1) 1990	(2) 1990	(3) 2000	(4) 2000	(5) 2010	(6) 2010
School quality	0.0245***	0.0157**	0.0301***	0.0220**	0.0383**	0.0289**
	(0.00786)	(0.00741)	(0.0108)	(0.00926)	(0.0149)	(0.0129)
Years of schooling	0.00474***		0.00430***		0.00330	
	(0.00111)		(0.00130)		(0.00228)	
Share accessing grade 5		0.0346***		0.0312**		0.0294*
		(0.0106)		(0.0124)		(0.0165)
Log of initial GDP pc (WB)	-0.0121***	-0.0108***	-0.0123***	-0.0106***	-0.00945*	-0.00936**
	(0.00343)	(0.00325)	(0.00353)	(0.00357)	(0.00493)	(0.00414)
Constant	0.0582***	0.0670***	0.0604***	0.0672***	0.0416	0.0520**
	(0.0206)	(0.0199)	(0.0194)	(0.0200)	(0.0272)	(0.0245)
Observations	62	74	65	81	66	83
R-squared	0.311	0.234	0.274	0.196	0.128	0.117

Note: The dependent variable is real economic growth per capita in constant USD from the initial year to 2019, as reported by the World Bank, World Development Indicators. Education in single years for adults is taken from Barro and Lee (2018). School quality and share accessing grade five are the averages at initial year for the population aged 20 to 40 (or the highest available age) weighted by cohort population. Robust standard errors are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 6: School quality and enrollment

	(1)	(2)	(3)	(4)
Women				
Birth Year	-0.465***	-0.292***	-0.317***	-0.357***
	(0.0242)	(0.0588)	(0.0473)	(0.0459)
Share with schooling		-10.67***	-11.83***	-0.418
		(3.427)	(3.972)	(4.305)
Share with secondary schooling		-8.423**	-6.104*	-11.92***
		(3.928)	(3.521)	(4.030)
Repetition			11.53***	7.611***
			(1.418)	(1.314)
Height w/ schooling				0.966***
				(0.287)
N	3231	3231	2186	1401
R^2	0.788	0.790	0.875	0.915

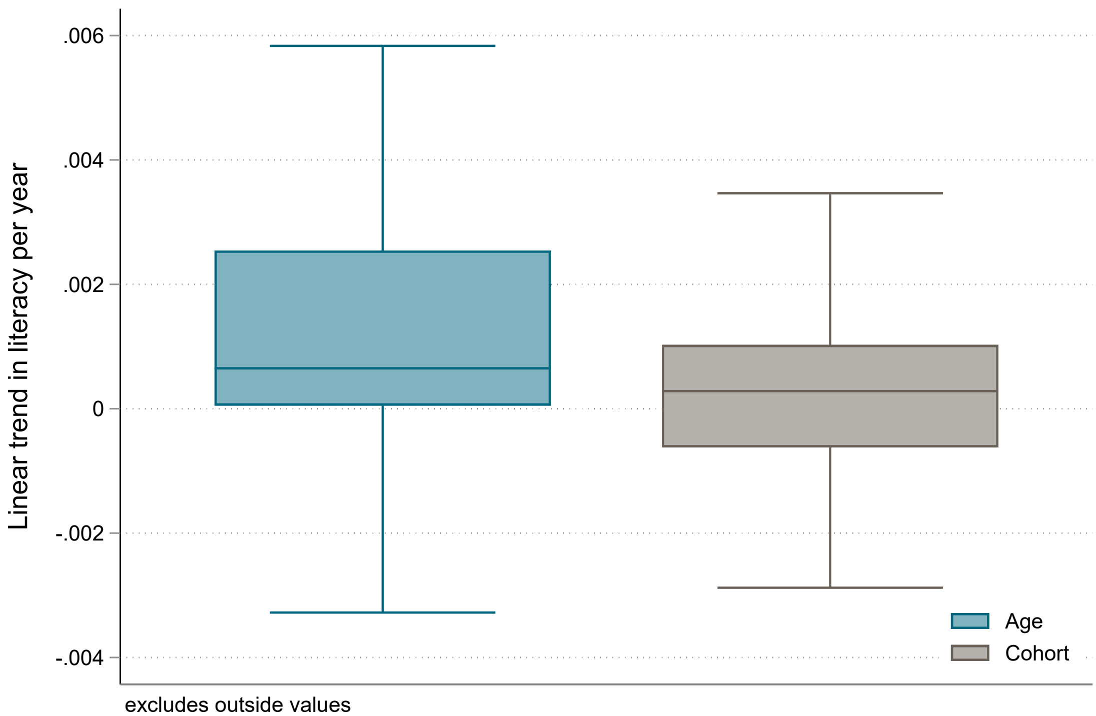
	(1)	(2)	(3)	(4)
Men				
Birth Year	-0.603***	-0.574***	-0.647***	
	(0.0283)	(0.0471)	(0.0482)	
Share with schooling		7.748	5.879	
		(6.249)	(6.690)	
Share with secondary schooling		-7.933*		
		(4.548)		
Repetition			6.385***	
			(1.416)	
N	2124	2124	1510	
R^2	0.727	0.727	0.768	

Note: Each column presents a separate regression, estimated separately for men and women, of predicted literacy at grade 5 on the covariates shown in the left column. The unit of observation is a country-birth cohort. All estimates control for country fixed effects.
* for p < 0.05, ** for p < 0.01, and *** for p < 0.001.

A Appendix

A.1 APC decomposition of literacy for individuals with no schooling

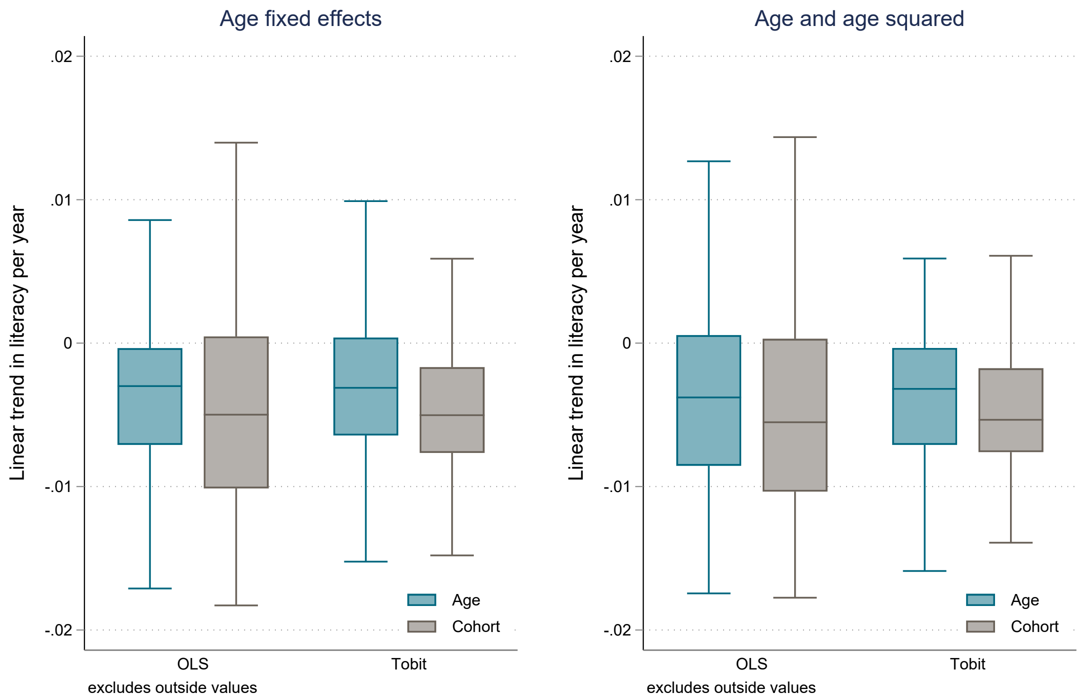
Figure 9: Age and cohort effects of literacy for individuals with no schooling. Deaton-Paxson normalization.



A.2 Age and cohort trends estimated with OLS and Tobit estimators

Results are summarized by looking at the linear trends of the age and cohorts as seen in and show OLS and Tobit estimates for equations 2 and 4. Period effects trends are not shown as they are zero by construction.

Figure 10: Comparison of age and cohort trends for OLS and Tobit estimates with age fixed effects and age and age squared



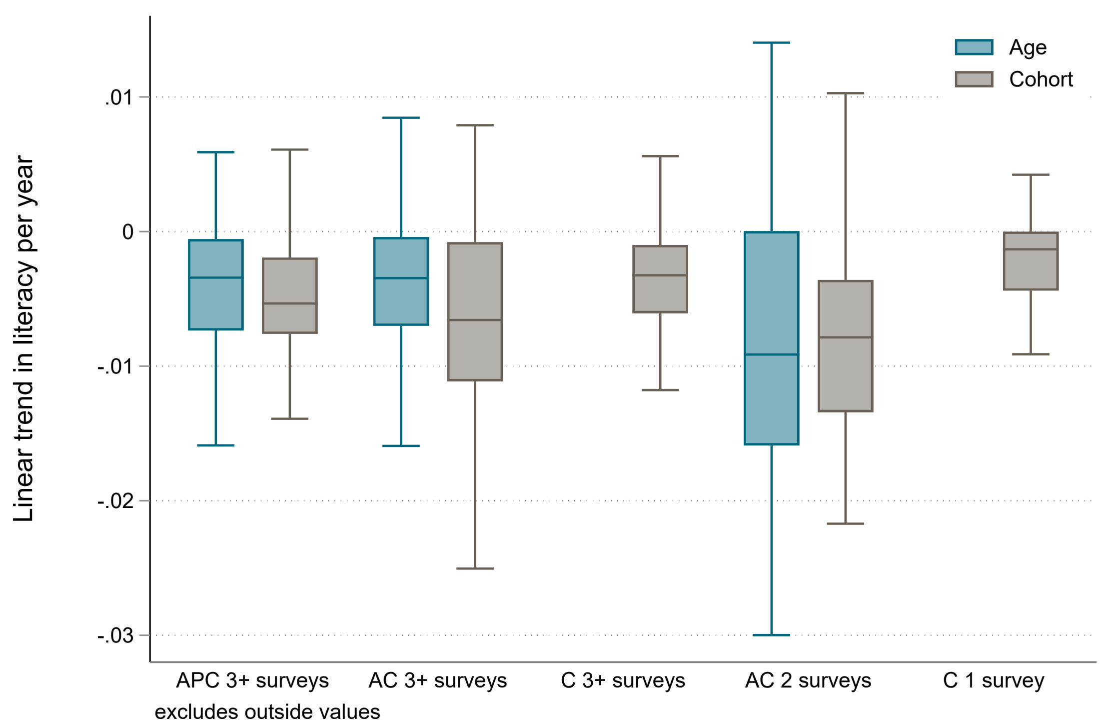
The different specifications show similar results. Age and cohort effects tend to be small and negative with about three quarters of estimated age and cohort trends being negative. Thus, both age and cohort effects tend to show that in the majority of cases individuals tend to lose literacy as they age and that more recent cohorts have lower rate of literacy after five years of schooling. In 12 cases, OLS and Tobit estimates using age fixed effects have estimation issues with parameters being omitted because of multicollinearity. Except for these 12 cases, age and cohort trends are almost identical when controlling for age fixed effects or age and age squared with a correlation of 0.99 for cohort effects and 0.9 for age effects for OLS and Tobit estimates. In order to keep a maximum number of cases, we therefore use the specification with age and age squared.

Compared to OLS estimates, Tobit estimates tend to have trends closer to 0 with a median age effect trend of -0.3 percentage points per year, or 9 percentage points over 30 years, while the median trend for age effect is -0.4 percentage points per year with OLS estimates. Similarly, median trends of cohort effects are -0.5 and -0.6 percentage points per year with

Tobit and OLS estimates respectively. The main difference lies in the trimming of large negative cohort effects. With Tobit estimates, there are 10 cases with cohort effects larger than -1 percentage point per year while there are 20 cases with OLS estimates. Interestingly, maximum values for positive cohort or age effects are similar with OLS or Tobit estimates. The trimming of large cohort effects makes results more plausible. For instance, for Ethiopian women OLS estimates predict a very large drop in school quality but this is due to the fact that expected literacy at grade five of 20 year old is 1.3 for the first cohorts and then decrease to about 0.7 for the last cohorts, whereas Tobit estimates predict a more realistic drop from 1 to 0.7. For this reason, we use Tobit estimates in our main discussion.

A.3 Comparison of APC, AC and C models

Figure 11: Comparison of APC, AC and C models by number of available surveys



For cases with at least three available surveys, dropping the normalized period effects (AC model) gives very similar results to the APC model estimated with DP normalization. Median trends in cohort and age effects are almost identical, which is not surprising given the

DP normalization assumes no trend in period effects. For comparison, age and cohort effects of AC estimates for country/sex cases with two surveys are shown in figure 11. They tend to show similar median trends than country/cases with more more than two surveys available but a wider distribution of effects. Finally, estimating a C model (with no controls for age or period) on country/sex cases with at least three available surveys shows that, as expected, the median cohort effects tend to be biased upwards (median cohort effects of -0.3 pp instead of -0.5 pp with the full APC decomposition), which is consistent with age effects being negative. Thus, for country/sex cases with only one available surveys we cannot rule out that the small negative median cohort trend (-0.1 pp) might be due to the absence of controls for age effects.

A.4 Country results

Table 7: Cohort-based estimates of trends: Women

	Birth Cohort		Literacy Rate			Share with $S \geq 5$			Literacy @ $S = 5$		
	First	Last	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***
Afghanistan+	1963	1995	6.7	17.5	**.5**	6	16.4	**.5**	87.2	55.4	**-1.4**
			(.4)	(.3)		(.4)	(.3)		(0)	(0)	
Algeria+	1965	1999	66.5	95	**1.2**	64.6	94.6	**1.3**	87.1	77	**-.4**
			(.4)	(.2)		(.5)	(.2)		(0)	(0)	
Angola	1958	1995	45.8	60.1	**.5**	35.6	60.1	**.9**	92.3	55.8	**-1.4**
			(1.8)	(.5)		(1.9)	(.5)		(0)	(0)	
Bangladesh	1958	1999	29.5	83.8	**1.8**	28.5	84.3	**1.8**	87.3	73.8	**-.4**
			(.3)	(.2)		(.3)	(.2)		(0)	(0)	
Belize	1958	1995	86.8	95.1	**.3**	83.2	92.7	**.4**	36.7	79.7	**1.6**
			(1.2)	(.4)		(1.3)	(.5)		(4.2)	(1.4)	
Benin	1953	1997	14.5	38.1	**.7**	14.5	39.4	**.7**	63.1	43.3	**-.6**
			(.8)	(.5)		(.9)	(.5)		(0)	(0)	
Bhutan*	1962	1990	14.4	49.9	**2**	8.4	41.4	**1.8**	72.2	70.1	**-.1**
			(.6)	(.6)		(.5)	(.7)		(0)	(0)	
Bolivia+	1955	1988	77.5	96	**.8**	51	85.7	**1.5**	100	100	**0**

	Survey Year		Literacy Rate			Share with $S \geq 5$			Literacy @ $S = 5$		
	First	Last	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***
			(.6)	(.2)		(.7)	(.4)		(0)	(0)	
Bosnia*	1958	1986	97.6	99.4	**.1**	88.5	98	**.5**	93.1	90	**-.2**
			(.4)	(.1)		(1)	(.3)		(1.9)	(2.4)	
Burkina Faso	1955	1997	6.4	27.3	**.7**	7	27.8	**.6**	46.2	58.6	**.4**
			(.4)	(.7)		(.4)	(.7)		(0)	(0)	
Burundi	1957	1996	33.4	63.8	**1**	11.7	45	**1.1**	88.3	96	**.3**
			(1)	(.5)		(.8)	(.5)		(0)	(0)	
CAR+	1958	1986	22.9	31.6	**.5**	24.1	36.2	**.7**	90.7	47.5	**-2.4**
			(.9)	(.7)		(1)	(.7)		(0)	(0)	
Cambodia	1952	1994	47.4	73.8	**.8**	16.3	68.1	**1.6**	97.6	77.3	**-.6**
			(.6)	(.4)		(.6)	(.5)		(0)	(0)	
Cameroon	1956	1998	53.5	69.6	**.5**	56.3	74.3	**.6**	71.8	49.9	**-.7**
			(.8)	(.5)		(.9)	(.5)		(0)	(0)	
Chad	1956	1994	7.8	21.4	**.5**	6.2	25.4	**.7**	82.4	28.8	**-1.9**
			(.6)	(.5)		(.6)	(.5)		(0)	(0)	
Comoros*	1964	1992	34.8	65	**1.7**	34.8	66.8	**1.8**	46.6	43.7	**-.2**
			(1.7)	(1.1)		(1.8)	(1.1)		(0)	(0)	
Congo, Dem Rep+	1959	1993	44.7	58.1	**.6**	45.2	64.7	**.8**	72.8	37.4	**-1.5**
			(1)	(.5)		(1.1)	(.5)		(0)	(0)	
Congo, Rep+	1957	1991	69	77.1	**.3**	72.7	85.1	**.5**	33.3	28.8	**-.2**
			(1.3)	(.6)		(1.3)	(.5)		(0)	(0)	
Costa Rica+	1963	1998	95.4	99.4	**.2**	89.4	97.6	**.3**	91.9	93.6	**.1**
			(.6)	(.1)		(.8)	(.3)		(1.6)	(1.2)	
Cote d'Ivoire	1958	1996	23.4	41.2	**.6**	24.2	41.3	**.6**	81.9	60.8	**-.8**
			(.8)	(.6)		(.8)	(.7)		(0)	(0)	
Djibouti*	1958	1986	17.5	44.9	**1.5**	21.1	52.4	**1.7**	34.5	36.9	**.1**
			(1.3)	(1)		(1.6)	(1)		(2.9)	(2.3)	
Dominican Rep.	1954	1994	79.8	95.2	**.5**	65.7	95	**1**	95.6	73.1	**-.7**

	Survey Year		Literacy Rate			Share with $S \geq 5$			Literacy @ $S = 5$		
	First	Last	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***
			(.5)	(.2)		(.6)	(.2)		(0)	(0)	
Egypt, Arab Rep	1952	1994	36.8	80.3	**1.4**	39.6	80.1	**1.3**	35.1	40.1	**.2**
			(.5)	(.4)		(.5)	(.4)		(0)	(0)	
El Salvador*	1966	1994	83.2	95.3	**.7**	59.3	87.2	**1.5**	96.4	94.8	**-.1**
			(.7)	(.3)		(1)	(.5)		(0)	(0)	
Eswatini	1958	1994	81.8	95	**.5**	72.4	92.7	**.8**	83.9	85.9	**.1**
			(1)	(.3)		(1.3)	(.4)		(0)	(0)	
Ethiopia	1952	1996	7.6	44.9	**1.1**	4	41.6	**1.1**	97.5	68.2	**-.9**
			(.4)	(.5)		(.3)	(.6)		(0)	(0)	
Gabon*	1964	1992	82.3	90	**.4**	84.2	91.6	**.4**	64.5	60.9	**-.2**
			(.9)	(.5)		(.9)	(.5)		(0)	(0)	
Ghana	1955	1999	46.9	76	**.9**	53	82.6	**.9**	5.1	13.2	**.2**
			(1.1)	(.5)		(1.1)	(.5)		(0)	(0)	
Guatemala*	1967	1994	64.3	84.5	**1.2**	39.6	67.3	**1.6**	95.5	95.3	**0**
			(.6)	(.4)		(.7)	(.5)		(0)	(0)	
Guinea	1957	1998	8.7	31.9	**.7**	9.6	37.2	**.9**	26	19.4	**-.2**
			(.6)	(.6)		(.7)	(.6)		(0)	(0)	
Guinea-Bissau	1958	1998	17.4	51.9	**1.2**	13.8	51.8	**1.3**	88.8	58.8	**-1**
			(1)	(.7)		(1)	(.7)		(0)	(0)	
Guyana	1961	1994	91.2	96.3	**.2**	83.3	95	**.5**	95.7	66.4	**-1.3**
			(.6)	(.3)		(.9)	(.4)		(0)	(0)	
Haiti	1952	1996	28.4	82.5	**1.6**	20.7	81.8	**1.8**	87.5	72.6	**-.4**
			(.8)	(.4)		(.7)	(.4)		(0)	(0)	
Honduras+	1958	1991	82.4	94.3	**.5**	54.1	81.1	**1.2**	100	100	**0**
			(.5)	(.2)		(.7)	(.4)		(0)	(0)	
India+	1958	1995	38	75.8	**1.4**	36.6	77.3	**1.5**	89.8	51.4	**-1.4**
			(.3)	(.1)		(.3)	(.1)		(0)	(0)	
Indonesia	1954	1997	68	97.7	**.9**	58.5	96.6	**1.2**	83.9	91.7	**.2**

	Survey Year		Literacy Rate			Share with $S \geq 5$			Literacy @ $S = 5$		
	First	Last	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***
			(.4)	(.1)		(.4)	(.1)		(0)	(0)	
Iraq	1958	1998	56.9	75.9	**.6**	52.8	73	**.7**	87.1	77.6	**-.3**
			(.6)	(.3)		(.6)	(.4)		(0)	(0)	
Jordan*	1969	1997	94.5	96.8	**.1**	94.7	96.6	**.1**	64.2	67.6	**.2**
			(.3)	(.3)		(.3)	(.3)		(0)	(0)	
Kenya	1955	1995	56.3	85.5	**1**	61.3	89.8	**1**	66.5	61.4	**-.2**
			(1.1)	(.3)		(1.2)	(.3)		(0)	(0)	
Lao PDR+	1958	1997	55.2	73.3	**.6**	38.5	73.5	**1.2**	99.4	79	**-.7**
			(1.2)	(.5)		(1.2)	(.5)		(0)	(0)	
Lesotho	1956	1994	87.6	95.2	**.3**	80.7	96.2	**.6**	98.2	83	**-.5**
			(.7)	(.3)		(1)	(.3)		(0)	(0)	
Liberia	1959	1996	24.7	56.2	**1.2**	26.7	61.7	**1.3**	48.8	37.9	**-.4**
			(.8)	(.8)		(.8)	(.9)		(0)	(0)	
Madagascar	1956	1998	65.7	68.3	**.1**	38.5	54.5	**.5**	99.5	87.6	**-.4**
			(.8)	(.4)		(.9)	(.5)		(0)	(0)	
Malawi	1952	1997	35	71.9	**1.1**	25.8	71.5	**1.3**	82.5	64.6	**-.5**
			(.8)	(.5)		(.7)	(.5)		(0)	(0)	
Maldives*	1968	1996	90.8	99.6	**.5**	72.8	99.1	**1.5**	89.8	92.7	**.2**
			(.5)	(.1)		(1.1)	(.2)		(0)	(0)	
Mali	1953	1998	9.4	28.2	**.5**	9.2	29.2	**.6**	57.1	33.9	**-.7**
			(.5)	(.5)		(.5)	(.5)		(0)	(0)	
Mauritania	1959	1995	36.2	56.3	**.8**	18.7	48.8	**1.2**	74.9	68.4	**-.2**
			(.8)	(.5)		(.6)	(.6)		(0)	(0)	
Mexico*	1967	1995	94	98.3	**.2**	88.5	97.8	**.5**	91.1	87.6	**-.2**
			(.4)	(.2)		(.6)	(.2)		(0)	(0)	
Mongolia	1958	1998	95.2	98.2	**.1**	90.2	96	**.2**	84	80.7	**-.1**
			(.5)	(.2)		(.8)	(.3)		(0)	(0)	
Montenegro+	1957	1993	97.9	99.2	**0**	96	99	**.1**	70.9	53.7	**-.7**

	Survey Year		Literacy Rate			Share with $S \geq 5$			Literacy @ $S = 5$		
	First	Last	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***
			(.6)	(.2)		(.9)	(.3)		(3.7)	(4.3)	
Morocco*	1955	1983	29.4	52.9	**1.3**	19.2	40.9	**1.2**	94	94.6	**0**
			(.7)	(.7)		(.7)	(.7)		(0)	(0)	
Mozambique	1955	1998	20.9	51.1	**.9**	12.2	61.4	**1.5**	99.8	41.9	**-1.8**
			(.7)	(.8)		(.5)	(.8)		(0)	(0)	
Myanmar*	1968	1995	77.5	85.2	**.5**	52.3	74.4	**1.3**	92.2	93.9	**.1**
			(.6)	(.5)		(.8)	(.7)		(0)	(0)	
Namibia	1952	1993	74.7	93.5	**.6**	63.5	93.3	**1**	51	64.6	**.4**
			(1.1)	(.3)		(1.3)	(.4)		(0)	(0)	
Nepal	1952	1998	17.9	80.7	**1.7**	5.4	76.5	**2**	99.5	82.8	**-.5**
			(.7)	(.3)		(.4)	(.4)		(0)	(0)	
Nicaragua*	1953	1981	74	85.1	**.6**	47.6	69.5	**1.2**	98.6	97.3	**-.1**
			(.9)	(.5)		(1)	(.7)		(0)	(0)	
Niger+	1958	1992	5.8	11.8	**.3**	6.3	14.6	**.3**	54.3	36	**-.8**
			(.5)	(.5)		(.5)	(.5)		(0)	(0)	
Nigeria	1955	1998	33.9	57	**.7**	40.9	62.4	**.7**	52.3	21.2	**-.9**
			(.8)	(.3)		(.8)	(.3)		(0)	(0)	
North Macedonia*	1957	1985	93.3	96.3	**.2**	84.2	93.5	**.5**	90.5	91.3	**0**
			(.6)	(.3)		(1)	(.4)		(0)	(0)	
Pakistan	1958	1997	26	52.3	**.9**	22.4	48.9	**.9**	68.9	85.9	**.6**
			(.7)	(.7)		(.7)	(.7)		(0)	(0)	
Panama*	1965	1993	95.1	96.9	**.1**	91.6	95.9	**.2**	96	91	**-.3**
			(.4)	(.3)		(.6)	(.4)		(0)	(0)	
Pap. N. Guinea*	1969	1996	48.5	60.8	**.7**	56.2	65.7	**.6**	54.8	51.8	**-.2**
			(.8)	(.6)		(.9)	(.7)		(0)	(0)	
Paraguay*	1968	1996	90.4	96.8	**.4**	83.5	94.9	**.6**	88.2	88.1	**0**
			(.7)	(.3)		(1)	(.4)		(0)	(0)	
Peru	1952	1992	82.3	96	**.5**	71.7	93.7	**.7**	69.4	84.9	**.5**

	Survey Year		Literacy Rate			Share with $S \geq 5$			Literacy @ $S = 5$		
	First	Last	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***
			(.4)	(.1)		(.5)	(.1)		(0)	(0)	
Philippines+	1955	1988	92.2	96.7	**.2**	88.4	95.7	**.3**	92.4	83.8	**-.4**
			(.4)	(.2)		(.5)	(.3)		(0)	(0)	
Rwanda	1952	1997	40.2	81.9	**1.2**	23	61.5	**1.1**	97.4	97	**0**
			(.9)	(.5)		(.8)	(.7)		(0)	(0)	
Sao Tome	1960	1999	65.1	88.7	**.8**	40.6	83.8	**1.5**	96.1	79	**-.6**
			(1.7)	(.8)		(1.9)	(1)		(0)	(0)	
Senegal	1957	1998	21.5	47.8	**.8**	21.4	49.6	**.9**	75	55.9	**-.6**
			(.5)	(.4)		(.5)	(.4)		(0)	(0)	
Serbia	1957	1999	98.6	99.1	**0**	96.6	98.9	**.1**	91.8	82.5	**-.3**
			(.2)	(.2)		(.4)	(.3)		(2.1)	(4.8)	
Sierra Leone	1957	1999	14.6	53.2	**1.2**	16.6	59.4	**1.3**	40.8	16.3	**-.8**
			(.8)	(.4)		(.9)	(.4)		(0)	(0)	
Somalia*	1958	1986	15.8	23	**.4**	9.2	10.6	**.1**	85.9	89.2	**.2**
			(1.3)	(.8)		(1)	(.6)		(0)	(0)	
South Africa*	1968	1996	91.5	97.3	**.3**	89.2	98.9	**.5**	71.4	47.9	**-1.3**
			(.6)	(.3)		(.7)	(.2)		(0)	(0)	
Sudan+	1962	1994	37.7	54.4	**.8**	34.4	54.8	**.9**	78.7	60	**-.8**
			(.8)	(.5)		(.9)	(.5)		(0)	(0)	
Suriname	1958	1998	88	95.7	**.3**	83.6	94.9	**.4**	99	88	**-.4**
			(.7)	(.4)		(.8)	(.4)		(0)	(0)	
Syria*	1958	1986	63.3	88.1	**1.4**	60	87.3	**1.5**	74.9	63.7	**-.6**
			(.7)	(.3)		(.8)	(.4)		(0)	(0)	
Tanzania	1956	1997	53.8	76.6	**.7**	49	79.4	**1**	77.4	65	**-.4**
			(1)	(.5)		(1)	(.5)		(0)	(0)	
Thailand	1958	1999	87.7	96.4	**.3**	44.1	96.9	**1.7**	100	74.8	**-.8**
			(.3)	(.2)		(.4)	(.2)		(1.8)	(1.9)	
The Gambia	1958	1998	11.6	57.4	**1.5**	13.3	62.5	**1.6**	39.7	16.1	**-.8**

	Survey Year		Literacy Rate			Share with $S \geq 5$			Literacy @ $S = 5$		
	First	Last	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***
			(.8)	(.6)		(.8)	(.6)		(0)	(0)	
Timor-Leste+	1961	1996	36	79.5	**1.7**	31.8	79	**1.9**	96.9	64.8	**-1.3**
			(.8)	(.5)		(.8)	(.6)		(0)	(0)	
Togo	1958	1997	29.5	57.2	**1**	32	64.4	**1.1**	74.9	42.6	**-1.1**
			(1.1)	(.6)		(1.2)	(.6)		(0)	(0)	
Trini. and Toba.+	1958	1991	96.2	98.7	**.1**	95.8	98.8	**.1**	99.8	90.2	**-.4**
			(.4)	(.2)		(.5)	(.2)		(0)	(0)	
Tunisia+	1964	1998	66.5	95.1	**1.2**	65.1	95.9	**1.3**	98.8	67.4	**-1.3**
			(.8)	(.3)		(.8)	(.3)		(0)	(0)	
Uganda	1952	1998	41.4	68.3	**.7**	32	80.7	**1.4**	85.1	45.1	**-1.1**
			(1.4)	(.5)		(1.3)	(.4)		(0)	(0)	
Vietnam	1958	1994	90.1	93.5	**.1**	76.7	88.7	**.5**	72	94.1	**.8**
			(.4)	(.5)		(.7)	(.6)		(0)	(0)	
West Bank*	1966	1994	91.3	99.1	**.4**	93.3	99.5	**.3**	58.9	26.9	**-1.8**
			(.5)	(.1)		(.5)	(.1)		(0)	(0)	
Yemen, Rep*	1965	1993	20	52.4	**1.8**	16.1	49	**1.8**	68.3	70	**.1**
			(.7)	(.5)		(.7)	(.5)		(0)	(0)	
Zambia	1954	1998	56.4	70.8	**.4**	58.4	85.4	**.8**	61.2	22.6	**-1.1**
			(1.2)	(.5)		(1.2)	(.4)		(0)	(0)	

Table 8: Cohort-based estimates of trends: Men

	Birth Cohort		Literacy Rate			Share with $S \geq 5$			Literacy @ $S = 5$		
	First	Last	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***

	Survey Year		Literacy Rate			Share with $S \geq 5$			Literacy @ $S = 5$		
	First	Last	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***
Afghanistan*	1967	1995	36.9	48.1	**.6**	31.4	46.9	**.9**	55.8	51.7	**-.2**
			(.9)	(.9)		(.9)	(.9)		(0)	(0)	
Albania+	1961	1997	96.6	96.9	**0**	98.3	96.9	**-.1**	48.7	75.1	**1**
			(.5)	(.4)		(.3)	(.4)		(5.2)	(2.7)	
Angola*	1967	1995	81.3	83.2	**.1**	69.1	81.7	**.7**	81	66.4	**-.8**
			(1.2)	(.8)		(1.6)	(.9)		(0)	(0)	
Bangladesh+	1959	1991	50.1	61	**.5**	41.7	53.6	**.5**	97.2	82.2	**-.7**
			(1.1)	(1.7)		(1.2)	(1.9)		(0)	(0)	
Belize*	1967	1995	86.4	94.1	**.4**	79.6	93.3	**.8**	81.6	84.1	**.1**
			(1.2)	(.6)		(1.6)	(.7)		(0)	(0)	
Benin	1953	1997	37.4	58.4	**.6**	34.6	58.9	**.7**	79.7	51.5	**-.8**
			(1.7)	(.8)		(1.8)	(.9)		(0)	(0)	
Bolivia+	1955	1988	95.5	99.1	**.2**	75.1	94.7	**.9**	93.8	94.5	**0**
			(.5)	(.2)		(1.1)	(.4)		(1.4)	(1)	
Burkina Faso+	1955	1990	18.5	39.5	**.8**	14.9	37.6	**.9**	25.8	54.1	**1.1**
			(1.3)	(1)		(1.3)	(1)		(0)	(0)	
Burundi+	1962	1996	65.7	78.2	**.5**	29.1	60.9	**1.3**	99.8	97	**-.1**
			(1.5)	(.7)		(1.5)	(.9)		(0)	(0)	
Cambodia+	1962	1994	69.9	77.8	**.4**	56.2	76.7	**.9**	73.6	61.6	**-.5**
			(.8)	(.7)		(1)	(.8)		(0)	(0)	
Cameroon	1956	1998	72	84.6	**.4**	71.3	87.1	**.5**	66.4	59.7	**-.2**
			(1.3)	(.6)		(1.4)	(.5)		(0)	(0)	
Chad+	1956	1994	25.6	55.8	**1.1**	22.3	57.8	**1.3**	81.6	35.2	**-1.7**
			(2.7)	(1.3)		(2.8)	(1.4)		(0)	(0)	
Congo, Dem Rep+	1959	1993	83	85.9	**.1**	82.5	87	**.2**	91.9	61.4	**-1.3**
			(1)	(.6)		(1.2)	(.6)		(0)	(0)	
Congo, Rep+	1957	1991	92.5	86.9	**-.2**	93.1	90.4	**-.1**	58	46.6	**-.5**
			(1)	(.8)		(1)	(.7)		(0)	(0)	

	Survey Year		Literacy Rate			Share with $S \geq 5$			Literacy @ $S = 5$		
	First	Last	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***
Cote d'Ivoire+	1964	1996	52.9	62.1	**.4**	49.1	60.9	**.5**	95.3	60.6	**-1.6**
			(1.4)	(1)		(1.4)	(1)		(0)	(0)	
Dominican Rep.	1954	1993	78	88.9	**.4**	68.5	88	**.7**	97.8	74.2	**-.8**
			(1.2)	(.4)		(1.5)	(.4)		(0)	(0)	
Eswatini	1958	1994	79.7	92.1	**.5**	72.7	87.8	**.6**	68.3	78.9	**.4**
			(1.3)	(.7)		(1.6)	(.9)		(0)	(0)	
Ethiopia	1952	1996	39.5	69.7	**.9**	18.6	59.7	**1.2**	97.2	82.8	**-.4**
			(1.6)	(.6)		(1.2)	(.7)		(0)	(0)	
Gabon*	1964	1992	87.4	90.5	**.2**	86.9	90.8	**.2**	65.2	55.6	**-.5**
			(.9)	(.7)		(1)	(.8)		(0)	(0)	
Ghana	1955	1997	68.7	86.2	**.5**	70.9	89.5	**.6**	33.9	23.9	**-.3**
			(1.3)	(.6)		(1.3)	(.6)		(0)	(0)	
Guatemala*	1967	1994	80.2	91.1	**.6**	52	77.6	**1.5**	96.3	95.1	**-.1**
			(.8)	(.5)		(1.2)	(.8)		(0)	(0)	
Guinea	1957	1998	30.6	63.9	**1.1**	30.1	66.1	**1.2**	54.5	45	**-.3**
			(1.8)	(1.3)		(1.8)	(1.4)		(0)	(0)	
Guinea-Bissau+	1966	1998	57	77.7	**.9**	46.4	74	**1.3**	99.9	86.5	**-.6**
			(1.6)	(1)		(1.7)	(1.1)		(0)	(0)	
Guyana*	1961	1989	86.3	91.8	**.3**	81	92	**.6**	62.8	46.3	**-.9**
			(1.1)	(.9)		(1.3)	(.9)		(0)	(0)	
Haiti	1952	1996	45	85.2	**1.2**	33.9	84.4	**1.5**	92.1	73	**-.6**
			(1.6)	(.5)		(1.6)	(.5)		(0)	(0)	
Honduras*	1964	1991	85.5	91.6	**.4**	58.8	79.1	**1.2**	97.4	93.3	**-.2**
			(1)	(.6)		(1.4)	(.9)		(.5)	(.6)	
India+	1958	1995	66.7	88	**.8**	61.2	87.9	**1**	91.5	56.2	**-1.3**
			(.4)	(.2)		(.4)	(.2)		(0)	(0)	
Indonesia	1954	1997	84.5	96.8	**.4**	71.1	93.9	**.7**	91.8	90.8	**0**
			(.5)	(.6)		(.7)	(.9)		(0)	(0)	

	Survey Year		Literacy Rate			Share with $S \geq 5$			Literacy @ $S = 5$		
	First	Last	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***
Kenya	1955	1994	81.8	89.7	**.3**	79.8	93.4	**.5**	89.8	60.5	**-1**
			(1.5)	(.4)		(1.7)	(.4)		(0)	(0)	
Lao PDR*	1969	1997	77	83.8	**.4**	70.8	82.2	**.6**	83.3	75.9	**-.4**
			(.7)	(.6)		(.9)	(.6)		(0)	(0)	
Lesotho	1956	1994	64.1	81.2	**.6**	49.3	77.7	**1**	92.8	75.7	**-.6**
			(2)	(.9)		(2.1)	(1.1)		(0)	(0)	
Liberia+	1959	1993	68.3	76	**.3**	68.8	77.9	**.4**	97.6	66.4	**-1.3**
			(1.2)	(1)		(1.2)	(1)		(0)	(0)	
Madagascar	1956	1998	73.1	73.8	**0**	43.6	62.3	**.6**	96.3	90.1	**-.2**
			(1.3)	(.8)		(1.6)	(1)		(0)	(0)	
Malawi	1952	1995	72.1	79.9	**.2**	55.2	79	**.7**	93.5	67.8	**-.8**
			(1.7)	(.5)		(1.9)	(.6)		(0)	(0)	
Maldives*	1969	1996	90.5	98.9	**.5**	84.1	99.1	**.9**	84.5	69.8	**-.9**
			(.8)	(.2)		(1.2)	(.2)		(2)	(5.5)	
Mali	1953	1998	26.1	50.8	**.7**	22.7	50.8	**.8**	56.5	38.4	**-.5**
			(1.4)	(1)		(1.5)	(1)		(0)	(0)	
Mauritania*	1967	1995	54.8	70.9	**.9**	38.3	57.6	**1.1**	82.3	76.3	**-.3**
			(1.6)	(1.1)		(1.7)	(1.3)		(0)	(0)	
Mongolia+	1965	1998	92.7	96	**.1**	85.9	90.7	**.2**	87.4	64.3	**-1**
			(.8)	(.4)		(1.2)	(.8)		(0)	(0)	
Mozambique+	1955	1991	63.3	70.3	**.3**	44.5	67.8	**.9**	99.9	76.4	**-.9**
			(2.1)	(1.2)		(2.2)	(1.3)		(0)	(0)	
Myanmar*	1968	1995	84.2	86.7	**.1**	60.5	76.1	**.9**	93.4	85.1	**-.5**
			(.9)	(.8)		(1.4)	(1.1)		(0)	(0)	
Namibia	1952	1993	74	88.3	**.5**	59.8	88	**.9**	88.7	67.2	**-.7**
			(1.6)	(.7)		(2.1)	(.8)		(0)	(0)	
Nepal	1952	1996	61.1	91.8	**.9**	35.6	85.6	**1.5**	98.8	91.1	**-.2**
			(1.7)	(.6)		(1.7)	(.9)		(0)	(0)	

	Survey Year		Literacy Rate			Share with $S \geq 5$			Literacy @ $S = 5$		
	First	Last	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***
Niger+	1958	1992	20.5	37	**.7**	15.2	36.9	**.9**	41	45.7	**.2**
			(1.2)	(1.4)		(1.2)	(1.5)		(0)	(0)	
Nigeria	1955	1998	62.6	76.9	**.4**	66.4	80.2	**.4**	71.6	31.2	**-1.2**
			(1.7)	(.5)		(1.8)	(.4)		(0)	(0)	
Pakistan+	1964	1997	57.6	64.8	**.3**	56.7	62.6	**.3**	74.2	68.9	**-.2**
			(1.2)	(1.6)		(1.3)	(1.8)		(0)	(0)	
Pap. N. Guinea*	1969	1996	66.7	75.7	**.5**	69.7	78.9	**.5**	64.5	58.1	**-.4**
			(1.1)	(.8)		(1.2)	(.9)		(0)	(0)	
Philippines*	1955	1983	90.5	94.6	**.2**	83	91.5	**.5**	84.8	83.8	**-.1**
			(.8)	(.5)		(1.3)	(.8)		(0)	(0)	
Rwanda	1952	1994	65.5	80.6	**.5**	39.8	57.3	**.5**	98.4	96.5	**-.1**
			(1.7)	(.6)		(1.8)	(.9)		(0)	(0)	
Sao Tome	1960	1999	90.1	89.6	**0**	67.3	83.9	**.6**	98.3	78.8	**-.7**
			(1.3)	(.9)		(2.2)	(1.3)		(0)	(0)	
Senegal	1957	1998	40.9	60.4	**.6**	34.8	59	**.8**	76.8	61.4	**-.5**
			(1.6)	(.7)		(1.6)	(.7)		(0)	(0)	
Sierra Leone	1960	1999	36	75.1	**1.3**	39.2	78.7	**1.4**	42.1	26.7	**-.5**
			(1.4)	(.6)		(1.5)	(.6)		(0)	(0)	
Tanzania	1956	1995	82.8	80.8	**-.1**	77.5	82.3	**.2**	76	64.5	**-.4**
			(1.5)	(1)		(1.9)	(1)		(0)	(0)	
Thailand+	1967	1999	93.7	95.4	**.1**	91.2	95.9	**.2**	100	84	**-.7**
			(.2)	(.3)		(.3)	(.2)		(4.2)	(2.4)	
The Gambia+	1965	1998	54.4	72	**.8**	51.3	72.8	**.9**	65.8	30.3	**-1.5**
			(1.6)	(.9)		(1.7)	(1)		(0)	(0)	
Timor-Leste+	1961	1996	61.9	81.7	**.8**	53.1	77.5	**1**	94.4	71	**-.9**
			(1.5)	(1)		(1.6)	(1.1)		(0)	(0)	
Togo+	1966	1997	67.5	79.7	**.6**	67.7	83.3	**.7**	100	64.5	**-1.7**
			(1.4)	(1)		(1.5)	(.9)		(0)	(0)	

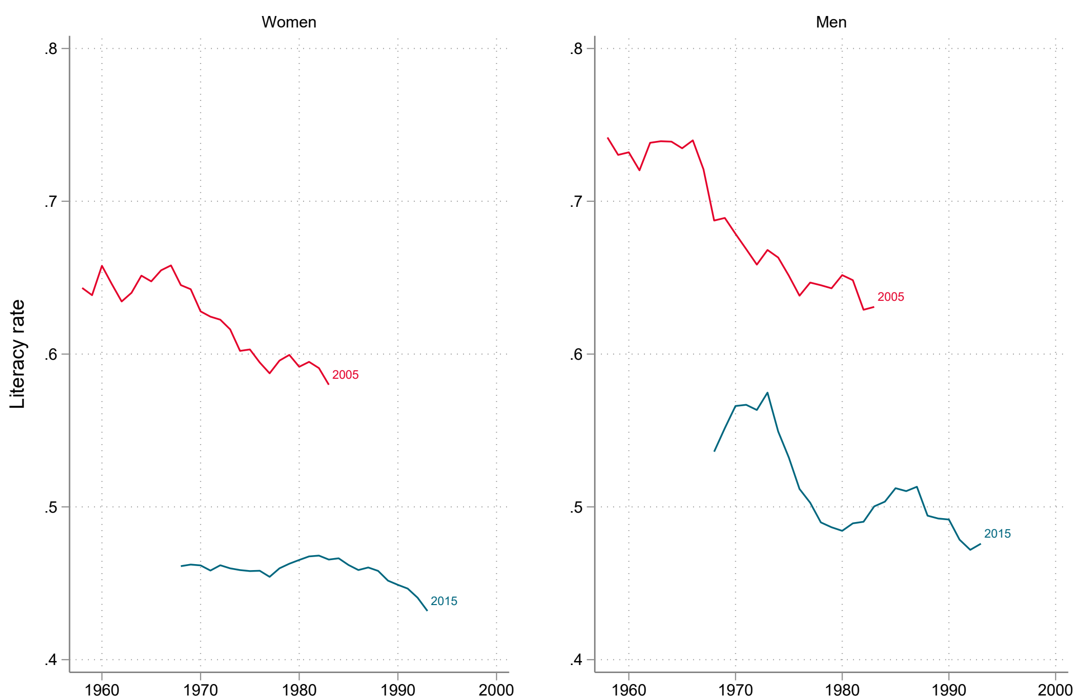
	Survey Year		Literacy Rate			Share with $S \geq 5$			Literacy @ $S = 5$		
	First	Last	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***	First	Last	**Δ% *p.a.***
Uganda	1952	1996	72.4	73.8	**0**	59.7	84.7	**.7**	93.2	46.6	**-1.4**
			(2.2)	(.9)		(2.6)	(.8)		(0)	(0)	
Zambia	1954	1998	81.6	83.7	**.1**	81.3	91.8	**.3**	74.5	36.1	**-1.1**
			(1.4)	(.5)		(1.6)	(.4)		(0)	(0)	

A.5 The case of India

In this section, we explore further the case of India. India makes up 33 percent of the country population of the female sample but Indian estimates are only based on two surveys. We find a large drop in school quality in India but this might be due to the fact that we are unable to estimate age and period/survey effects simultaneously.

To understand why the trend in school quality is negative in India, it is useful to look at the raw data. Figure 12 shows the literacy rate of women and men with four to five years of primary of schooling by years of birth for the surveys 2005/06 and 2015/16. For women, literacy rates are seem to be flat within each survey but there is a drop of literacy of more than 10 percentage points between data collected in 2005/06 and data collected in 2015/16 for women born the same year. The picture is similar for men although literacy rates are also declining within each survey.

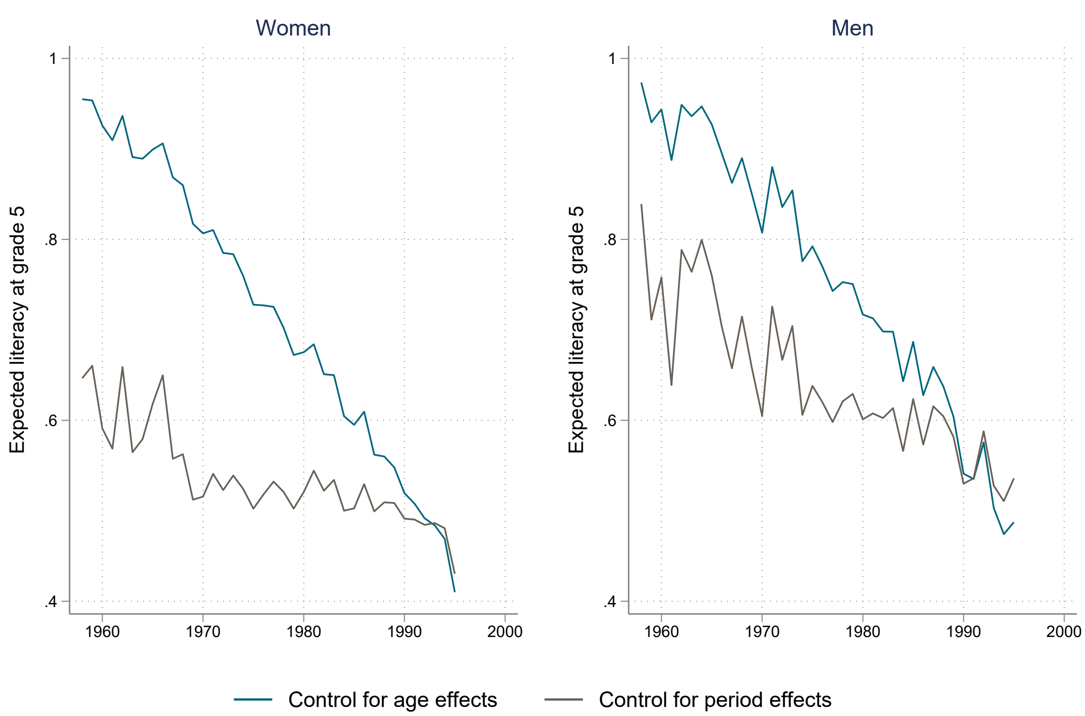
Figure 12: Literacy rates for women and men with 4 to 5 years of education in India by survey



As we control for cohort fixed effects and age and age squared, the gap in literacy for women or men born the same year but interviewed at different dates is entirely attributed to age effects. Estimates age effects are relatively large with women expected to lose 25 percentage points of literacy between age 20 and 40 and men 20 percentage points. Since predictions of literacy at grade 5 are estimated for individuals at age 20, the model predicts that a steep downward trend in school quality between the first and the last cohorts. However, with only two available surveys it is hard to know if the systematic difference in literacy rates we observe between individuals interviewed in different surveys is due to age effects, as our model assumes, or period/survey effects.

To test the sensitivity of our results to this modelling assumption, we estimated the same model without age effects but controlling for period/survey effects and predicted the literacy at grade five by year of birth. Results can been seen in figure 13 for women and men. In both cases, the trend in school quality is steeper when estimated with a model controlling for age effects. In this case, it is predicted that Indian women and men born in the 1960's had an expected literacy at grade five greater than 0.9 but that it declined to less than 0.5 for the cohort born in the 1990's. However, when the model only includes a survey effect, trends in school quality are flatter as the expected literacy at grade five of first cohorts is only 0.6 for women and around 0.7 for men.

Figure 13: Trends in school quality with different model specifications for women and men.



The two models presented in figure 13 can be thought as two extreme modelling assumptions: either attributing all systematic differences between the two surveys to age effects or attributing it fully to period/survey effects. In both cases, school quality is declining but the magnitude of this decrease is really different, especially for women and without additional data, it is not possible to know what modelling assumption is the most appropriate. However, there are reasons to think that our original modelling assumption controlling for age effects best fits the data.

In the model with survey effects, the coefficient associated with survey effect is 13 percentage points for women and 9.7 for men. That is a very large effect; for comparison the largest absolute survey effects for country/sex cases with at least three surveys is 22 percentage points and there are only seven survey effects out of 313 that are larger than 13 (and 15 greater than 9.7). What could explain such a large period/survey effect? First, we do not believe that there can be large period effect for adult literacy in a relatively short period of time. We find unlikely that the Indian context would have changed so drastically in 10 years that millions of adults would have become illiterate. A change in participation in literacy program could be a factor but only 3.3 percent of women said they ever took part to a literacy program in the 2005/06 survey (the question was not asked in 2015/16), which is too low to impact adult literacy rates.

Therefore, the difference should come from survey effects, such as differences in sampling or data collection methods. Both surveys are designed to be nationally representative and sampling design is identical. However, the sampling frame has changed between the two surveys as there has been a new census in 2011. Could this change explain the systematic difference between the two surveys? We think it is unlikely. First, similar differences between the two surveys are found within urban and rural areas. Moreover, analysis by Indian state shows that a negative school quality trend and negative age effects are found in most states (see figure 14), showing that the difference in literacy rates between the two surveys is not driven by composition effects. We can also see in figure 15 that the average number of years of schooling for men and women is identical across the two surveys for individuals born the same year, suggesting that there has not been systematic differences in sampling between the two surveys.

Figure 14: Trends in school quality by Indian state for women

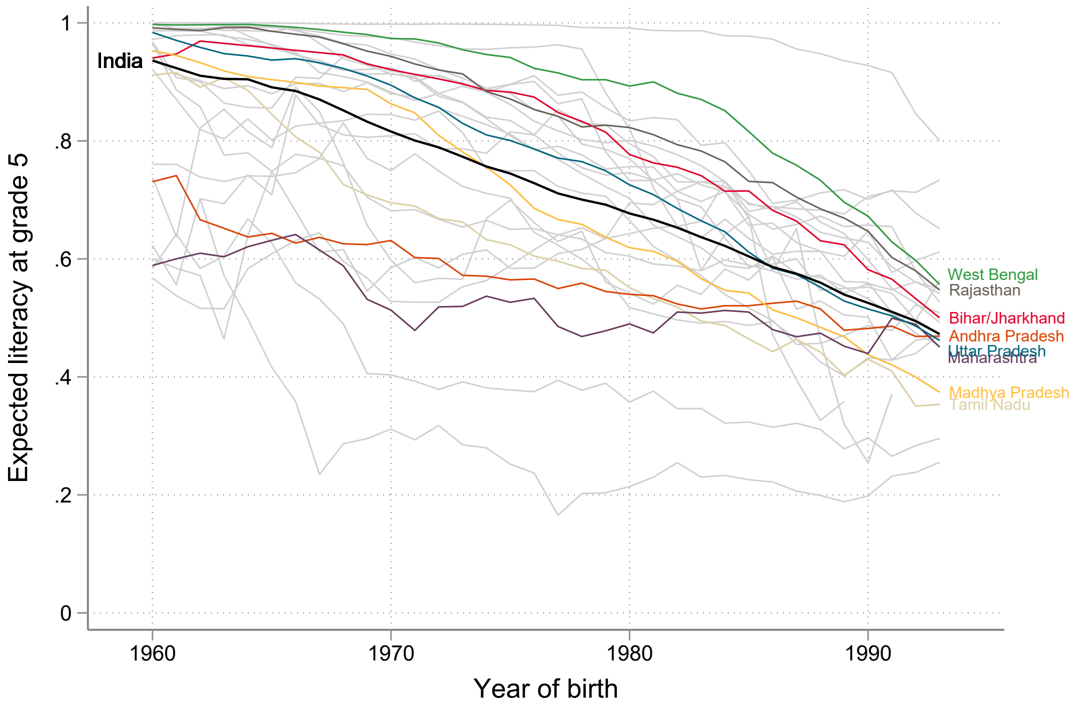
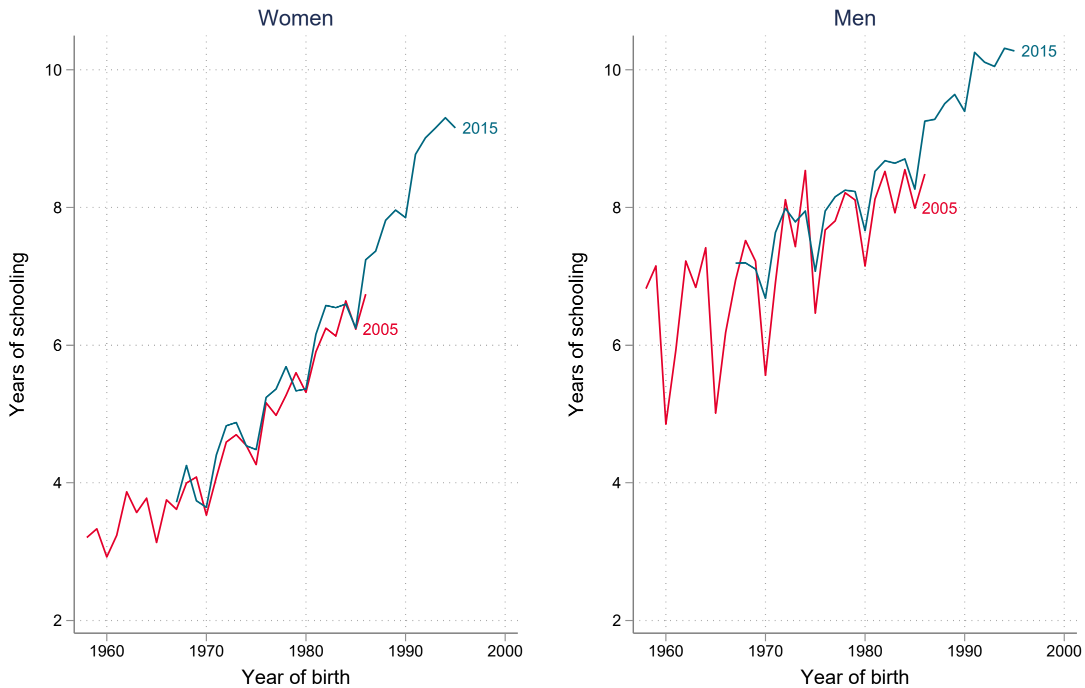


Figure 15: Years of schooling by survey for women and men.



Another possible source of survey effect could be a change in the way literacy is measured. In both surveys, individuals have to read a sentence and are graded on a 0-1-2 scale. The way interviewers grade literacy could have changed but results remain the same if people who are graded as 1 are classified as literate or illiterate. Also, differences in literacy rates are not due to differences in response rates that are high (>99 percent) and identical between the two surveys.

Estimated age effects for India are large but, as can be seen in figure 16, they are not outliers. Actually, they are close to estimated age effects for Bangladesh that have been estimated with seven surveys and also close to age effects for Nepalese women estimated with six surveys. Thus, it seems likely that there are indeed large age effects in India and that our estimates of a large drop in school quality are genuine.

Figure 16: Age effects for women (red) and men (blue)