

Education For All?

The Gender Literacy Gap and the Distributional Effects of World Bank Assistance

Brandon Merrell

Jackson Institute for Global Affairs
Yale University

July 15, 2021

Word Count: 11,898

Abstract

Does foreign development assistance improve social equity, or do governments use aid to reinforce existing hierarchies? I address this question by combining data from more than 6,000,000 survey respondents in 110 low-income countries to develop new measures of female and male literacy dating back to the 1930s. Using these data, I show that although countries exhibit improved literacy after joining the World Bank's International Development Association (IDA) or receiving IDA development funds, these improvements are almost entirely restricted to males. As a result, overall gender literacy gaps *increase* substantially following IDA membership. I use several additional tests to argue these effects are driven by autocratic governments' selective channeling of development aid as opposed to alternative mechanisms. In addition to supplying improved measures of historical education, my findings directly contribute to growing literatures on the ability of development assistance to reach designated targets and the impact of international shocks on the domestic provision of public goods. In particular, even well-intended development initiatives may yield perverse distributional effects due to underlying political, cultural, and economic conditions in participating countries.

I thank Anjali Thomas, Zoe Nemerever, Abigail Vaughn, Kevin Calderwood, David Lake, and audience members during a talk at the Georgia Institute of Technology for their comments and suggestions. For the most recent version, please contact the author at brandon.merrell@yale.edu.

*“If we educate a boy, we educate one person.
If we educate a girl, we educate... a whole nation.”*
—World Bank President James Wolfensohn (1995)

Introduction

Education is a valuable conduit for economic progress,¹ is strongly associated with a broad range of political development and public health indicators,² and is increasingly recognized as a human right. Unfortunately, the right to education is also widely violated. Government funding and civilian access to education vary substantially across and within countries.³ In particular, studies of rural and low-income communities document significant and persistent gender gaps in which women obtain less schooling and suffer higher illiteracy rates than men.⁴

Because illiteracy restricts economic productivity and stymies efforts to achieve gender equity, international relations researchers and development practitioners devote considerable energy to understanding the effects of foreign assistance programs that aim to improve education access and achievement. At the micro level, international donors and development organizations contribute to, among other things, student retention and feeding programs, teacher training and recruitment, school construction, curriculum development, and the supply of instructional materials. The World Bank alone provides more than four billion dollars of education-related financial assistance each year.⁵ Nevertheless, despite optimism regarding such assistance, individual program evaluations as well as larger, cross-national comparisons present a mixed record of results, particularly on the question of whether large-scale development initiatives translate into actual learning.⁶

The scarcity of conclusive evidence regarding development financing is unsurprising for two distinct but related reasons. First, substantial variation exists in how recipient countries respond to

¹ Lucas, 1988; Becker et al., 1990; Rebelo, 1991; Mankiw et al., 1992, Psacharopoulos, 1994; Barro, 2001; Gennaioli et al., 2013; Wantchekon et al., 2015.

² Gregorio and Lee, 2002; Breierova and Duflo, 2004; Cutler et al., 2006; Trudell, 2009; Woodberry, 2012.

³ Stromquist, 2006; Ansell, 2008.

⁴ Behrman and Knowles, 1999; Lewis and Lockheed, 2006; Björkman-Nyqvist, 2013.

⁵ Broad, macro-level development initiatives that do not directly target the education sector may nevertheless generate spillover effects from economic growth, political development, and international exposure writ large.

⁶ Even micro-level analyses of textbook provision (Glewwe et al., 2002), expansions in teaching staff (Glewwe and Kremer, 2006), and enrollment initiatives find mixed evidence of improvements in test scores (Banerjee et al., 2007). See also Hanushek, 1986.

windfalls in aid. If autocratic and democratic states differ in the degree to which they efficiently harness foreign assistance to educate their citizens, pooled estimates will paint an incomplete portrait of the overall effects of development efforts. Second, improvements in country-wide education outcomes are exceedingly difficult to track systematically—let alone to attribute to specific programs—due to the dubious quality of historical education data from developing countries. Popular datasets provide sparse or inconsistent coverage even in recent decades, rely on information that cannot be directly compared across countries or time periods, and often report proxy measures that correlate poorly with researchers’ preferred quantities of interest. Collectively, these challenges create substantial barriers for scholars who hope to explain cross-country shifts in education outcomes or identify the net effects of foreign assistance.

This paper addresses each of the preceding obstacles and, in the process, provides troubling evidence that windfalls in education funding can reify existing educational cleavages and widen prevailing literacy gaps between males and females. I begin by introducing a theory of how governments direct newly-available resources within the education sector. My argument, which is grounded in political economy and development research, acknowledges the robust empirical relationship between educational attainment and political engagement as well as the rich literature on government misuse of foreign aid.⁷ In short, I argue autocratic leaders fear that extending education to new social groups or diverse segments of society could provoke social unrest or instability that would threaten the leader’s tenure in office. As a result, these leaders prefer to channel education improvements within blocs whose members already enjoy greater access to education compared to outsiders. Democratic leaders, in contrast, are broadly accountable to a wider array of actors and are therefore more inclined to provide education across social lines or to groups whose members have historically lacked access. Thus, although windfalls in education funding might improve nationwide education rates and boost overall education spending under both types of regimes, the distributional effects will differ substantially, with foreign assistance reinforcing preexisting social cleavages within autocracies.

⁷ Nie et al., 1996; Dee, 2004; Hillygus, 2005; Milligan et al., 2004; Stromquist, 2006; Sondheimer and Green, 2010; Berinsky and Lenz, 2011; Beath et al., 2013; Wantchekon et al., 2015).

Because testing this argument requires access to long-term, historical measures of educational attainment that allow direct comparisons across countries, social groups, and time periods, my first contribution is constructing a new dataset suited to this task. I combine evidence from more than 6,000,000 face-to-face interviews conducted as part of roughly 400 Demographic and Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS) in developing countries. Importantly, each interview includes an assessment of the respondent's literacy by the enumerator.⁸ I use these data to introduce and validate new annual estimates of the Lifetime Literacy Rate (LLR) for females and males born in more than 100 developing countries between 1935 and 2005.⁹ The resulting data provide a useful means of examining shifts in literacy that occurred within and across different countries over much of the twentieth century.

I use the new measures to assess the impact of World Bank funding on education outcomes. Using generalized difference-in-differences estimators, I provide evidence that gaining access to World Bank (IDA) assistance on average causes a country's gender literacy gap to widen rather than shrink. Consistent with my theory of autocratic targeting, however, I find that such expansions are primarily evident within autocratic countries, whereas literacy gaps remain relatively stable under more democratic regimes. Estimated effect sizes are substantively large: on average, the gulf between male and female literacy rates grows by more than 5% in the years after a country gains access to IDA assistance. Evidence of parallel pre-trends, inclusion of unit-specific linear trends along with country and year fixed effects, several placebo checks, and results using counterfactual estimators support a causal interpretation. The results are also substantively robust to estimation on alternative subsamples; the inclusion of additional regressors intended to control for potential economic, political, and international confounders; and the substitution of alternative outcome, treatment, and control measures.

Finally, I offer several supplemental tests that further support the autocratic targeting mechanism in contrast to alternative processes. I first assess whether shifts in gender literacy gaps stem

⁸ This paper addresses the functional/skills-oriented literacy paradigm as opposed to interpretations of literacy-as-cultural-fluency.

⁹ LLR, which I describe below, is designed to capture the proportion of individuals born in a particular country-year who acquire literacy during their lifetimes.

not from government actions but rather from within-household biases such as gender queuing or male prioritization. To test this, I compare literacy rates among second-born women, assessing whether such individuals obtain distinct benefits depending on whether their elder sibling is male or female. If parents prioritize educating male children, we should expect girls with older brothers to obtain literacy at lower rates than girls with older sisters. Nevertheless, I find no evidence that son-preference systematically varies in conjunction with access to IDA funding, limiting the possibility that this explanation drives my main results. Second, I explore whether differences in expected economic returns to education might lead women to pursue schooling at lower rates than their male peers. To this end, I compare improvements in female literacy between ethnic groups known to follow “bride price” traditions and groups that do not. Although other researchers have shown financial transfers from male grooms to the families of female brides encourage female education by raising its expected returns for prospective brides, I find no evidence that these behaviors influence my estimates.

Two final sets of supplemental results directly support my theory that financial windfalls can reify existing educational hierarchies. In one test, I examine an alternative cleavage along which governments might target education efforts: the rural-urban divide. I find that IDA assistance widens the literacy gap in favor of urban communities in countries where governing regimes draw support from urban areas, but this result is reversed in countries where a regime’s political supporters are geographically diverse or are located primarily in rural environments. Finally, I show that gender literacy gaps *shrink* when autocrats lose access to IDA assistance, consistent with what we would expect if governments can no longer channel excess funding toward preferred groups.

My findings directly contribute to growing literatures on the ability of development assistance to reach designated targets and the impact of international financing on the domestic provision of public goods. In particular, I show that even large and well-intended development initiatives can yield perverse distributional effects due to underlying political, cultural, and economic conditions in participating communities.

A Theory of Autocratic Targeting

Governments receiving foreign assistance allocate aid in ways that will maximize their chances of retaining power.¹⁰ At one extreme, governments can use newly available resources to provide public goods that will benefit the broadest possible set of citizens; on the other hand, incumbent governments may choose to target their efforts toward specific groups whose support the leaders deem particularly important. The political environment and selection process under which a regime operates will influence its choice of policy.¹¹ In general, governments pursue strategies in accordance with their regime type, with autocratic countries and weak democracies more likely to target narrow constituencies, adopt clientelist policies, or provide patronage.¹²

Preferential channeling of development aid to specific groups should be especially pronounced in the education sector, where major investments can reshape the political landscape. The positive relationship between education and political participation is among the most robust in the social sciences.¹³ In developing countries, education programs are associated with subsequent democratization, and recent development initiatives have been shown to increase participation (particularly by women) in local governance, community life, and village decision-making.¹⁴ Autocratic leaders must therefore weigh the economic benefits of educating citizens against the possibility that such citizens would demand additional public goods, seek democratization, or gain the capacity to organize efficiently—actions that could threaten the state’s political stability and therefore the incumbent government’s hold on power.¹⁵ Leaders’ strategic incentives to avoid extending education are especially strong in the context of foreign assistance windfalls. Although extending new public goods could temporarily appease citizens and reduce revolutionary threats, citizens may react poorly to policy reversals or interruptions that suggest the leader is uncommitted to providing such goods

¹⁰ Ahmed, 2012; Bueno de Mesquita and Smith, 2009; Faye and Niehaus, 2012; Jablonski, 2014; Kono and Montinola, 2009; Cruz and Schneider, 2017.

¹¹ Bueno de Mesquita et al., 2005; Lake and Baum, 2001; Stasavage, 2005.

¹² Isham et al., 1997; Dollar and Levin, 2005; Wright and Winters, 2010; Hicken, 2011.

¹³ Hillygus, 2005.

¹⁴ Lankina and Getachew, 2012; Woodberry, 2012; Beath et al., 2013; Wantchekon et al., 2015.

¹⁵ Lipset, 1960; Verba and Almond, 1963; Smith, 2008; Miller, 2012; Sanborn and Thyne, 2014. See also Haggard and Kaufman, 2012.

permanently.¹⁶ Unfortunately for autocrats and citizens alike, durable financial commitments to public service provision are difficult to provide in contexts where aid flows are inconsistent, leaders face short time horizons, and foreign assistance may not be sustained over time.¹⁷ Collectively, these challenges provide autocrats with incentives to avoid broad investments in human capital.¹⁸

Nevertheless, an emerging body of empirical research finds that autocratic countries often behave similarly to democracies in the domain of education provision. Ansell and Lindvall (2013) demonstrate that both democratic and autocratic governments have increased their influence over primary education, with substantial centralization occurring under fascist, authoritarian regimes as well as liberal, democratic governments. Likewise, Gift and Wibbels (2014), López-Cariboni and Cao (2019), and Paglayan (2021) each show that autocracies sometimes devote significant effort toward education, while Ross (2006) finds no substantial differences between democracies and non-democracies on education outcomes. As Paglayan (2021, p. 18) summarizes, “[t]he most important puzzle” to emerge from recent empirical work “concerns the high levels of primary education provision observed under non-democracies.” Equally striking is the absence of evidence showing that expanding access to education provokes democratization or destabilization in autocracies.¹⁹ Why are autocratic regimes using development assistance to improve public education, and how are they doing so in ways that minimize the political consequences?

I argue that autocratic governments navigate this dilemma by selectively targeting education programming to preferred groups. In doing so, governments simultaneously satisfy several criteria. First, governments sidestep corruption monitoring methods that track only whether funding receipts translate into public spending—in this case, resources are indeed being spent.²⁰ Likewise, because members of targeted groups benefit from the investments, aggregate levels of education will continue to rise when measured at the national level, enabling governments to reassure donors and

¹⁶ Merrell and Abrahams, 2020.

¹⁷ Wright and Winters, 2010; Altincekic and Bearce, 2014.

¹⁸ Bourguignon and Verdier, 2000; Acemoglu and Robinson, 2006; López-Cariboni and Cao, 2019.

¹⁹ Acemoglu et al., 2005.

²⁰ Altincekic and Bearce (2014) find autocrats are more likely to increase public-sector spending following foreign aid receipt than are democracies, suggesting that autocrats do not siphon assistance away from the education sector.

provide concrete evidence that their education investments are effective.²¹ Third, increases in aggregate education should translate into improvements in overall economic productivity, even if those figures fall short of what the country might have achieved if funding was allocated inclusively. Most importantly, the government avoids the risks of political destabilization associated with expansive increases in education among groups that previously lacked access.

Gender provides a visible and likely cleavage along which governments can selectively target new education investments.²² Gender literacy gaps are common in developing countries, creating an easily identifiable group whose members have been denied equitable access to education.²³ Moreover, improvements in women's education and labor force participation are strongly associated not only with increases in political engagement, but also with subsequent extensions of suffrage, the development of successful women's movements, and even democratization—potential outcomes that motivate autocratic leaders to think cautiously about expanding female access to education.²⁴ Finally, preferential targeting or prioritization of male education is feasible in practice, and could be achieved by, for example, funneling resources toward male schools in countries where single-sex education remains relatively common, by channeling resources toward secondary schools in regions where girls typically exit formal education at earlier stages than boys, or by other investments that boost female enrollment without improving female learning outcomes.²⁵

World Bank education assistance during the twentieth century serves as an ideal testing ground for evidence of preferential targeting. Despite the Bank's general reputation for strong oversight, limited aid fungibility, and "donor control,"²⁶ throughout much of its history, the Bank's educa-

²¹ Bisbee et al. (2019) show governments respond to performance-based incentives with respect to Millennium Challenge goals.

²² In the robustness section, I show that governments also target aid geographically toward areas where supporters reside. See also Ansell (2008) for discussion of how autocrats might prioritize tertiary education spending for elites and World Bank (2017) for brief discussion of potential ethnic biases in the allocation of education aid.

²³ Gray et al., 2006.

²⁴ Gray et al., 2006; Iversen and Rosenbluth, 2006; Stromquist, 2006; Beath et al., 2013, Wyndow et al., 2013.

²⁵ For example, hiring prioritization of male teachers could create differences in learning outcomes across student gender groups. Emphasizing female enrollment without attendant learning is also consistent with theories that autocrats provide primary schooling not because they aim to raise citizens' functional education but rather in hopes of forging a consistent national identity or inspiring regime loyalty (Darden and Grzymala-Busse, 2006; Cantoni et al., 2017; Testa, 2018; Paglayan, 2021).

²⁶ Milner et al., 2016 and Brazys et al., 2017, although see also Kilby, 2009 and Birchler et al., 2016.

tion lending allowed for substantial influence by recipients. In the 1960s, World Bank President George Woods mandated that education assistance go only toward capital projects (e.g., buildings and equipment) developed jointly by bank staff and officials from host countries, thereby creating opportunities for state employees to channel efforts toward preferred communities.²⁷ Moreover, until recent periods the Bank paid little attention to potential distributional biases related to its education lending. In its early decades, for example, the Bank heavily favored secondary school and even university-level projects over primary education, reasoning that the former categories would more quickly stimulate economic growth by supplying highly-skilled workers.²⁸ By emphasizing the training of engineers, architects, and technicians in countries where women were disproportionately excluded from such professions, the Bank encouraged countries to focus on males' educational interests at the expense of their female peers. Even after the Bank expanded its investments in primary education during the 1980s, it promoted policies such as privatization and schooling fees that were likely to generate or reify within-country divergences in education access and outcomes.²⁹ Given these policies, I offer the following general hypotheses:

H1: Countries that receive World Bank (IDA) assistance should subsequently exhibit greater improvements in male education than female education.

H2: The difference between male educational improvements and female educational improvements described in H1 should be more pronounced in autocracies than in democracies.

Unfortunately, these types of within-country education gaps are difficult to discern in available data that suffer from limited coverage, low inter-temporal and inter-country comparability, questionable quality, and an overemphasis on education provision rather than education outcomes. The next section reviews these limitations, then introduces and validates new measures of male and female literacy that facilitate appropriate comparisons.

²⁷ Dorn and Ghodsee, 2012.

²⁸ During the 1960s, approximately 96% of the Bank's education lending went to secondary or higher education; even by the late 1970s primary education comprised only 14 percent of all education-related lending (Dorn and Ghodsee, 2012). See also Jones (1997), Heyneman (2003), and Wickens and Sandlin (2007).

²⁹ Heyneman, 2003; Wickens and Sandlin, 2007. The Bank began to advocate the abolishment of school fees in the early 2000s, at the end of the time period covered by this paper (Mundy and Menashy, 2014).

Measuring Literacy

Available Education Data

In low-income countries, the adult literacy rate is a natural benchmark measurement for education. Viewed as an outcome, rapid increases in literacy can reflect investments in education and allow researchers to assess the efficacy of new schooling policies or development programs. On the input side, shifts in literacy can affect other processes of interest, including public health, economic growth, and political development.³⁰ Given these dual roles, researchers seek reliable measures of historical literacy that provide broad coverage across developing countries and time periods. Unfortunately, as I detail in this section, available data suffer from two significant limitations: low comparability and substantial sparsity. Moreover, although the problems related to literacy have motivated researchers to adopt alternative education measures—such as estimates of school enrollment and attainment—even these proxies fail to fully resolve underlying data quality concerns. In short, fundamental measurement problems pose significant obstacles to researchers seeking to conduct cross-country or within-country analyses of changes in educational outcomes.

The most commonly used source of aggregated historical literacy data is UNESCO, which has produced empirical reports on literacy since the mid-1950s.³¹ UNESCO's data, however, are not consistently collected by the organization itself. Instead, many of the estimates are constructed using results published in government censuses or reports released by national education authorities.³² Some of the figures are estimated indirectly from aggregate data on school enrollment or attendance. Elsewhere, literacy rates are calculated from censuses in which respondents are asked to report whether they or their family members are literate, raising concerns about the reliability of individuals' self-descriptions or assertions about their relatives' educational histories.³³ UNESCO itself warns that its data involve “definitions and methods of data collection that differ across coun-

³⁰ Literacy is also frequently incorporated in larger development indices. The Human Development Index, for example, entails equally weighted measures of life expectancy at birth, literacy, and real income per capita (Srinivasan, 1994).

³¹ Behrman and Rosenzweig, 1994, Wagner, 2011. UNESCO data are often obtained indirectly via the World Bank.

³² Behrman and Rosenzweig, 1994, Srinivasan, 1994.

³³ Stromquist, 2006, Wagner, 2011.

tries” and should therefore “be used cautiously.” Even setting aside the variation in source material, sparsity remains a major concern. Despite the development of representative sampling methods in the early twentieth century, frequent and large-scale data collection projects remained impractical for many countries long thereafter. As a result, historical estimates of national literacy are available only at irregular intervals and with significant gaps in coverage. For example, across a set of more than 100 countries, UNESCO provides estimates of adult female literacy for only 5% of all country-years between 1970 and 2000, with no estimates at all prior to this date (see Figure 1). As Behrman and Rosenzweig (1994, p. 153) summarize, “data simply do not exist for confident estimates of literacy” in many countries prior to the 1990s.³⁴

Given these challenges, researchers have adopted alternative education measures. Several authors have assembled or constructed national estimates of educational attainment (i.e., years-of-schooling), then collated those estimates into longitudinal panel datasets. Although these efforts improve upon available literacy measures in several ways, they fail to fully resolve underlying problems regarding data sparsity and incomparability. For example, although datasets may not appear to contain gaps upon first perusal, in many cases the educational attainment figures they provide are estimates based on interpolation, projection, or even the substitution of enrollment data in place of attainment itself.³⁵ Not only do these processes rely on substantial and sometimes untested assumptions for validity, the periods of data missingness are likely systematically related to other social and development indicators of interest.³⁶ Even when raw data on educational attainment is available and reliable, inter-country and inter-temporal comparability concerns are potentially magnified when compared to literacy because characteristics including school quality, school year

³⁴ Even in OECD countries, representative estimates gathered via direct evaluations of literacy are available only beginning in the late 1980s.

³⁵ Kyriacou (1991) estimates missing schooling data using a regression of educational stocks on lagged flows, a process that requires the relationship remain stable over time (De la Fuente and Doménech, 2006). In Barro and Lee (1996), census data provide information for less than half of the cells on educational attainment, with the remainder filled using information on enrollment rates as a substitute (Knowles et al., 2002). Updated data by the same authors utilizes “forward-and-backward extrapolation of the census/survey observations... to fill in missing observations” (Barro and Lee, 2013). See also discussion in Behrman and Rosenzweig (1994).

³⁶ Behrman, 1996.

length, and curriculum content vary across countries and over time.³⁷ It is perhaps no surprise that different datasets yield diverging estimates of national education trends and produce contradictory or implausible results when used as inputs in other analyses—even in estimates of something as fundamental as the relationship between education and economic growth.³⁸

Finally, education researchers and development organizations increasingly recognize the necessity of studying concrete education outcomes as opposed to proxies. Historical data on student enrollment, attendance, and even completed years of schooling are more readily available than direct evaluations of individuals' skills, but researchers who use those proxy measures implicitly assume time spent in school translates into actual learning.³⁹ Unfortunately, a growing body of evidence demonstrates this is often not the case. For example, although Barro and Lee's (2013) data on years-of-schooling is widely used as a proxy for human capital, Paglayan (2021) finds a correlation of only .07 between the measure and national performances on PISA tests. Given these discrepancies, reliable measures of genuine learning would be valuable. As Hanushek and Woessmann (2012) show, such measures perform significantly better than years-of-schooling at predicting economic growth—perhaps because, as Angrist et al. (2019, p. 2) note, divergences between years-of-schooling and learning are “particularly acute in developing countries.” The World Bank, UNESCO, and development researchers have alike acknowledged the existence of a “learning crisis” and urged the collection and analysis of data on concrete learning outcomes.⁴⁰ To address the need for measures of this type, the following section introduces new annualized literacy estimates drawn from in-person evaluations conducted equivalently in more than 100 developing countries between 1935-2005. The data should facilitate renewed analysis not only of how policy changes influence education but also how improvements in literacy affect other social or economic outcomes of interest.

³⁷ Angrist et al. (2019), for example, calculate learning-adjusted years of schooling for countries in the post-2000 period. They find that although children in South Africa spend on average roughly eight years in school, this translates to only 4.59 years of learning-adjusted years. See also Behrman and Rosenzweig, 1994 and Srinivasan, 1994.

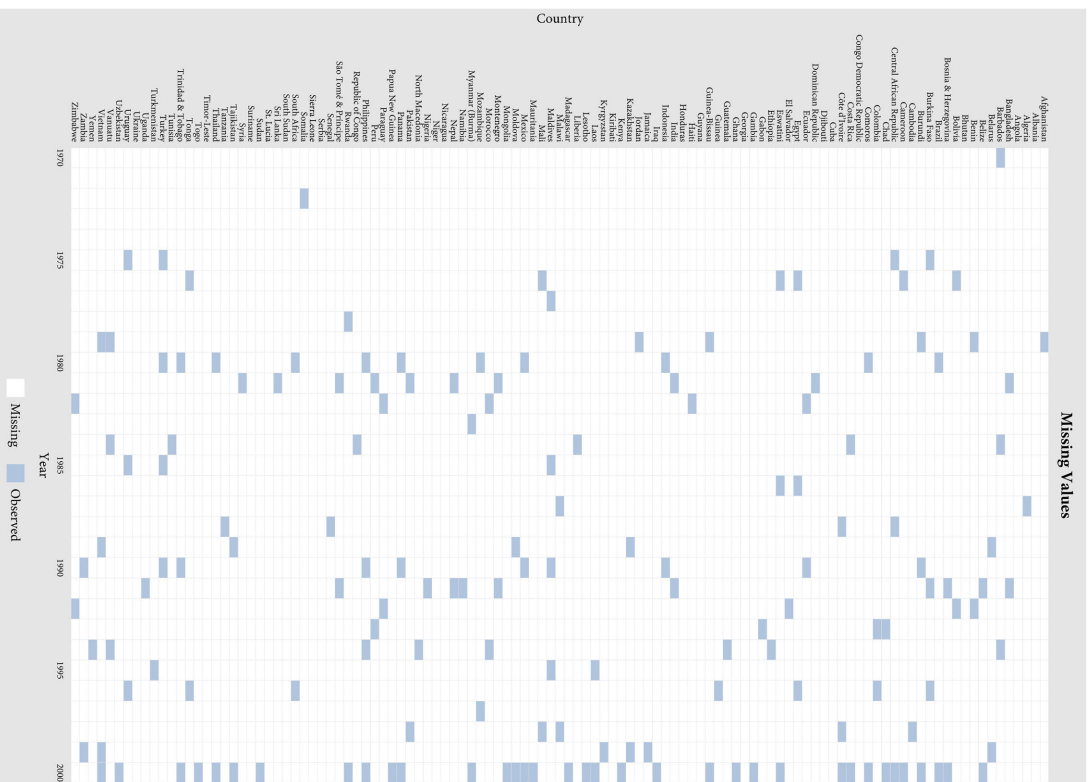
³⁸ De la Fuente and Doménech, 2006, Krueger and Lindahl, 2001, Cohen and Soto, 2007, Behrman, 1996.

³⁹ De la Fuente and Doménech, 2006, Angrist et al., 2019, World Bank, 2017.

⁴⁰ World Bank, 2017, Angrist et al., 2019.

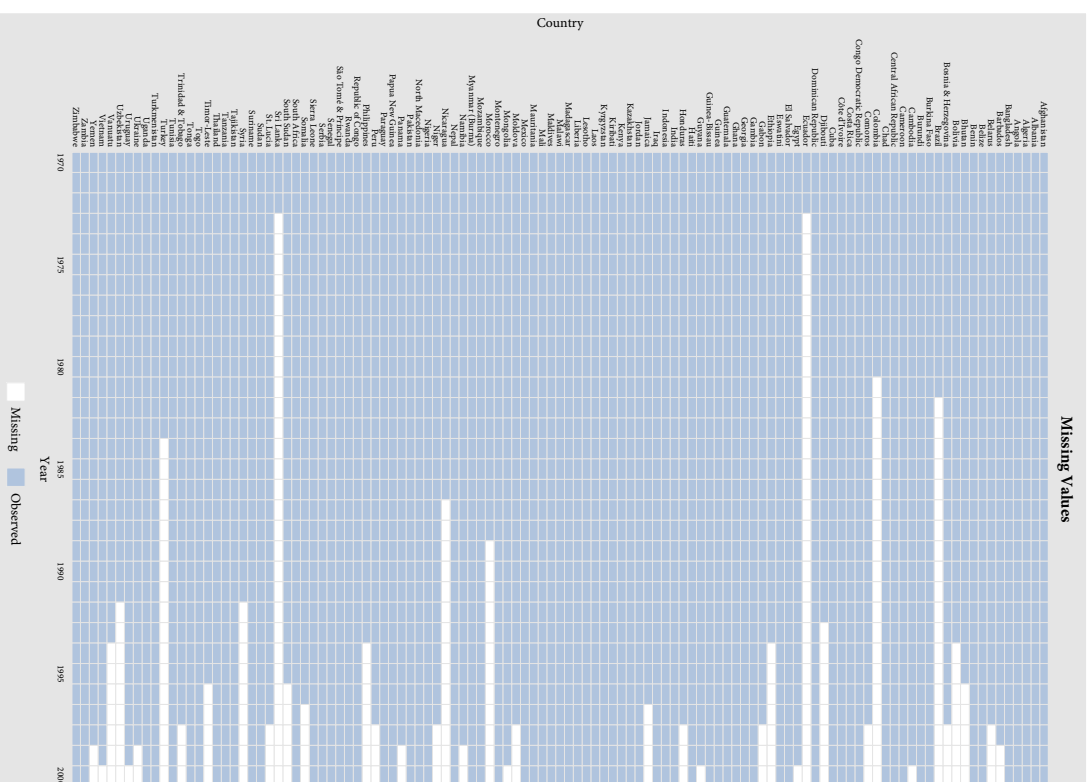
Figure 1: Data Coverage and Missingness Comparison (1970-2000)

UNESCO Estimates of Adult Female Literacy



Notes: Data from UNESCO Institute for Statistics (uis.unesco.org), last updated in September 2020. Estimates of adult female literacy are available for shaded cells and unavailable for white cells.

Female Lifetime Literacy Estimates (This Paper)



Notes: Calculated by the author using DHS and MICS data. Estimates are available for shaded cells and unavailable for white cells. For a complete coverage map from 1935-2005, see Appendix A2.

New Estimates of Lifetime Literacy

I estimate changes in literacy using 405 DHS and MICS studies fielded in 111 countries (see lists in Appendix A1). DHS and MICS are large, nationally representative, face-to-face surveys widely used in development research. Surveys are conducted by local partners and are designed by USAID, UNICEF, and participating governments with the explicit goal of facilitating cross-country comparisons of public health, gender equity, and other development indicators. Both programs provide extensive documentation about data collection methods. The pooled sample includes roughly 6,300,000 respondents, with females representing a disproportionately large share due to interest on behalf of DHS/MICS in collecting data on maternal health. Participation in most surveys is restricted to individuals older than 15 and below the age of either 49 or 59.

The large majority of surveys, including all DHS, include a direct evaluation of literacy in which respondents are asked to read a simple, one-sentence prompt in their language of choice. Enumerators indicate whether respondents are able to read the entire prompt, only a portion of the prompt, or none of the prompt. MICS include equivalent direct tests of respondent literacy, but in some of the earliest surveys these tests were waived for respondents who reported educational histories that exceeded a threshold survey designers considered an appropriate literacy benchmark for that country—typically the equivalent of either completing or enrolling in secondary school. Although I include surveys utilizing this exemption process in the main text, I demonstrate in subsequent sections that my findings are robust to the exclusion of such data.

For each survey, I assign respondents to country-birth-cohorts based on their country of origin and year of birth.⁴¹ Roughly 4-5% of respondents do not provide an explicit birth year. For these, I calculate a birth year by comparing the respondent's self-reported age and the interview date. To account for the possibility that international migration could impede accurate assignment to country birth-cohorts, I exclude the small proportion of respondents who indicate they lived “abroad” during childhood. I then calculate the proportion of remaining female and male respondents within each birth-cohort whom enumerators graded as “fully literate,” omitting individuals

⁴¹ For work implementing a similar approach to study Female Genital Mutilation/Cutting, see Engelsma et al., 2020.

whose tests were not completed or conducted for technical reasons, e.g., respondents who were blind or visually impaired or for whom no prompt was available in the preferred language. I refer to the resulting proportion as the cohort's estimated *Lifetime Literacy Rate*, (LLR) because it estimates the proportion of surveyed individuals born in each cohort who ultimately attained literacy.

Figure 2 depicts female LLR estimates for Rwanda, with estimates plotted separately for each survey. Visual inspection suggests the results are similar across surveys, with estimates for equivalent birth-cohorts largely consistent.⁴² The overlapping surveys also highlight equivalent overall trends: LLRs steadily increased among cohorts of Rwandan girls born from the early 1940s through the late 1970s, dipped temporarily across a series of cohorts whose education was potentially compromised by civil violence, then finally increased again toward the end of the millennium. Importantly, although the disruptions in education are apparent in the figure and are consistent with first-hand accounts,⁴³ they are not observable in UNESCO's data, which includes only two estimates of adult female literacy during this window: ~27% in 1978 and ~60% in 2000. Likewise, neither the Barro and Lee (2013) years-of-schooling data nor measures of human capital from the Penn World Tables reflect these disruptions to Rwandan education.

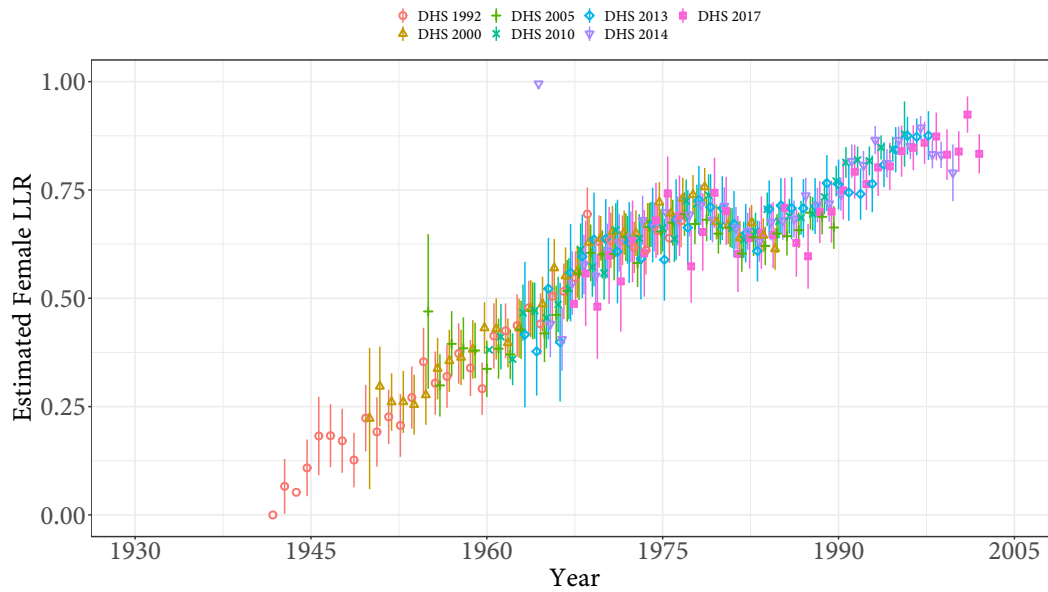
Because the survey-specific LLR estimates are similar, I pool overlapping survey cohorts to create a single estimate of LLR for each gender group and country-year, as depicted for Rwandan Females in Figure 3.⁴⁴ Figure 4 maps all countries for which I estimate LLR, with shading denoting the earliest available dates. Coverage is widespread, including many countries throughout Latin America and Africa; parts of Eastern Europe; and Central and Southeast Asia.

⁴² To more precisely assess consistency across surveys, I identify all country-birth-cohorts that appear in multiple surveys and find a median absolute difference between overlapping estimates of approximately 4.5%. Fewer than 1% of Fisher exact tests on overlapping estimates indicate a significant association between estimates and survey waves. If literacy predicts life expectancy, sample attrition might also lead to inflated LLR estimates among older birth cohorts whose illiterate members are less likely to survive until surveying occurs. However, I find no evidence that LLR estimates are systematically higher for equivalent cohorts in more recently-fielded surveys.

⁴³ See, for example, Akresh and De Walque (2008).

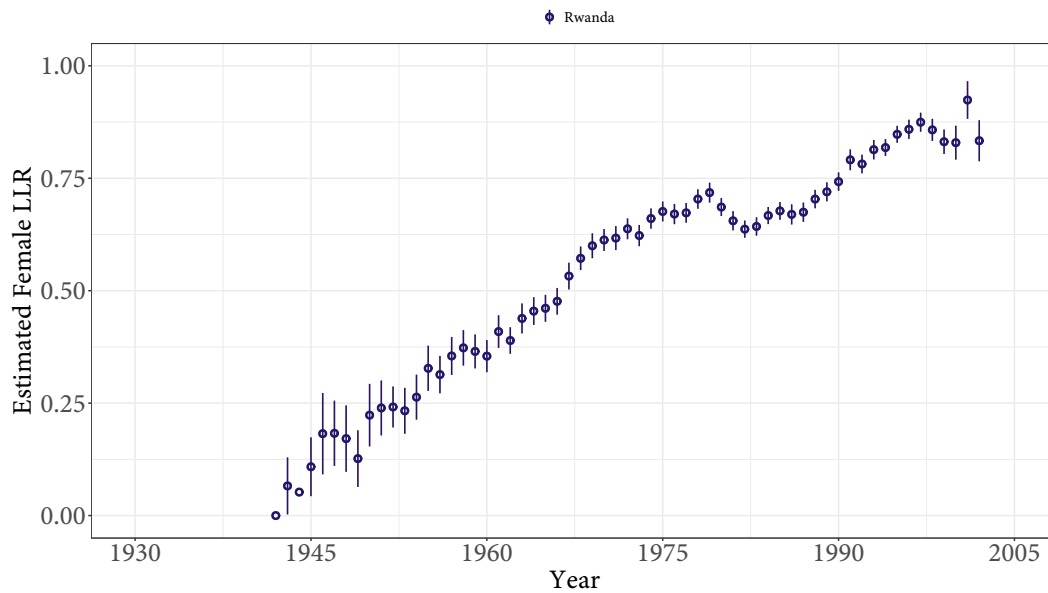
⁴⁴ The weights included in DHS surveys are inappropriate for use when pooling data because sampling probabilities shift across survey waves in conjunction with population demographics. I therefore follow DHS recommendations by creating denormalized weights when pooling observations across surveys. I obtain the requisite population estimates from World Bank data. For further discussion of this procedure, contact the author, consult the replication code, or see the explanation by Ruilin Ren in "Note on DHS standard weight de-normalization."

Figure 2: Female Lifetime Literacy Estimates for Rwanda (Distinct Survey Estimates)



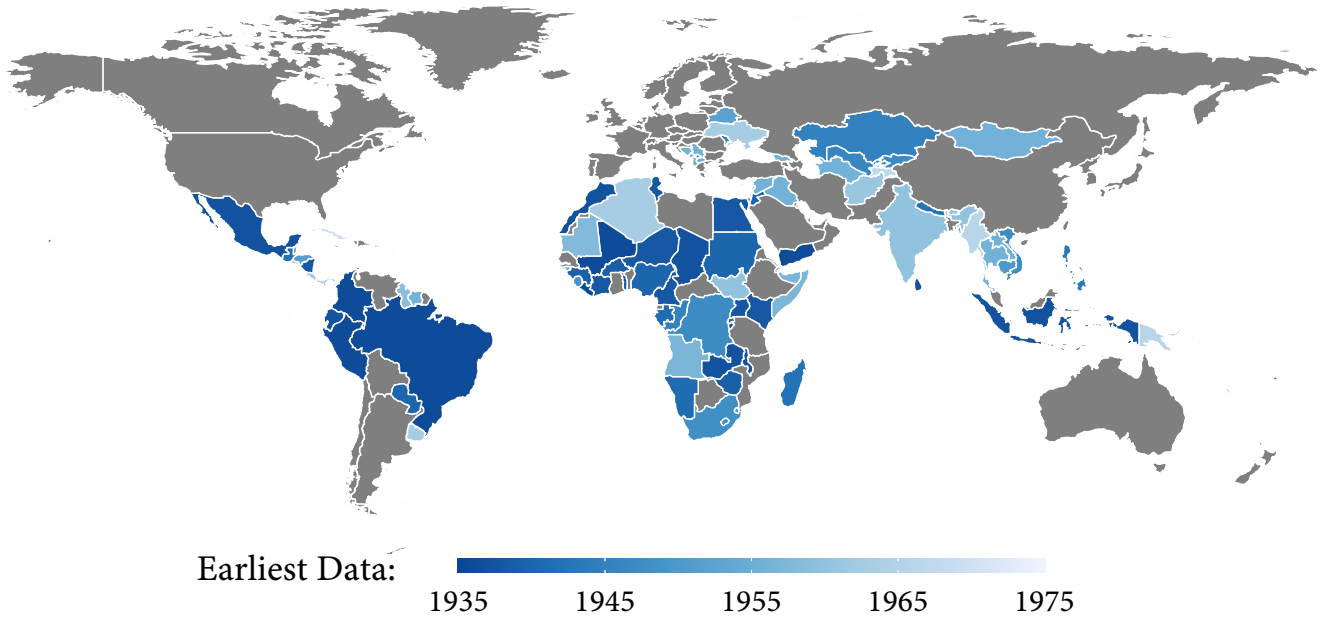
Notes: Calculated by the author using data from the 1992, 2000, 2005, 2010, 2013, 2014, and 2017 Rwandan DHS. Points depict the estimated proportion of women in each survey-year cohort who acquired literacy. Vertical bars depict 95% confidence intervals.

Figure 3: Female Lifetime Literacy Estimates for Rwanda (Pooled Country Estimates)



Notes: Calculated by the author using data from the 1992, 2000, 2005, 2010, 2013, 2014, and 2017 Rwandan DHS. Points depict the estimated proportion of women in each birth-year cohort who acquired literacy. Vertical bars depict 95% confidence intervals.

Figure 4: Map of Female Lifetime Literacy Estimates



Notes: LLR calculated by the author using data from DHS and MICS.

Validating the Lifetime Literacy Measures

Although patterns like those observed for Rwanda provide some face validity to the LLR measures, I offer two additional assessments of data quality. First, Table 1 presents correlation matrices for male and female Lifetime Literacy Rates and several alternative measures of literacy and human capital accumulation. The first, UNESCO's measure of *adult literacy* was described above. The second measure, *youth literacy*, is also assembled by UNESCO and entails an estimated literacy rate for residents of a country between the ages of 15 and 24. I obtain the final two measures from Barro and Lee, 2013 (henceforth B&L): one depicts the percentage of individuals in a country who have obtained *any schooling*, while the latter estimates the average *years of schooling* attained by a country's residents.⁴⁵ UNESCO's data are available sporadically in accordance with the missingness chart in Figure 1, while the B&L estimates are computed for years evenly divisible by five. Each of these measures are estimated separately for males and females, as illustrated in Table 1.

⁴⁵ For ease of interpretation, I calculate *any schooling* by taking the complement of B&L's *no schooling* measure.

Female LLR is moderately correlated with each of the alternative education measures at rates ranging from $r = 0.79$ (*years of schooling*) to $r = 0.90$ (*youth female literacy*). Correlations with Male LLR are slightly weaker, ranging from $r = 0.77$ and $r = 0.75$ with regard to with UNESCO's *youth* and *adult literacy* estimates to lows of $r = 0.70$ and $r = 0.66$ with *any schooling* and the *years of schooling*, respectively. Overall, the reasonably strong correlation between LLR and the UNESCO literacy data indicates that the new measures should prove useful, particularly due to their expanded coverage and annual periodicity. In contrast, the more modest correlations between LLR and *years of schooling* could raise questions about whether the two sets of data measure consistent outcomes. Indeed, because LLR estimates the proportion of individuals who achieve a binary outcome (literacy), one might worry it contains less useful information than years-of-schooling.

Table 1: Lifetime Literacy Correlation Matrices

<u>Female Estimates</u>				
	This Paper Lifetime Lit.	UNESCO Adult Lit.	UNESCO Youth Lit.	B&L 2013 Any School.
Female Lifetime Literacy (This Paper)				
Adult Female Literacy % (UNESCO)	0.86			
Youth Female Literacy % (UNESCO)	0.90	0.97		
% Any Schooling (Barro & Lee 2013)	0.86	0.91	0.87	
Years of Schooling (Barro & Lee 2013)	0.79	0.80	0.73	0.94

<u>Male Estimates</u>				
	This Paper Lifetime Lit.	UNESCO Adult Lit.	UNESCO Youth Lit.	B&L 2013 Any School.
Male Lifetime Literacy (This Paper)				
Adult Female Literacy % (UNESCO)	0.75			
Youth Female Literacy % (UNESCO)	0.77	0.96		
% Any Schooling (Barro & Lee 2013)	0.70	0.85	0.78	
Years of Schooling (Barro & Lee 2013)	0.66	0.72	0.65	0.94

Female and Male LLR calculated by the author using DHS and MICS surveys. Schooling data from Barro and Lee (2013). Literacy estimates from UNESCO via the World Bank Indicators.

To evaluate the information content of the new LLR measures, I use OLS to estimate a rudimentary model of economic growth. In doing so, I follow the approach of Hanushek and Woessmann (2012) and Angrist et al. (2019), each of whom sought to assess the performance of student-learning measures compared to years-of-schooling data in predicting economic growth. Compared to the previous authors, the LLR data offer substantially improved periodicity and coverage across coun-

tries and/or time periods. For example, Hanushek and Woessmann (2012) estimate a relationship between educational performance for a cross-section of only 50 countries, roughly half of which were OECD members. Angrist et al. (2019), by comparison, evaluate data for a broader set of 117 countries but with a restricted time period of 2000-2009. The new LLR estimates therefore create an opportunity to assess whether measures of concrete learning outcomes can usefully predict economic growth in a broader set of developing countries in earlier decades.

Because the B&L estimates are available only at five-year intervals, my unit of analysis is a five-year country-period (e.g., Rwanda 1980-1984). Time periods begin in 1970 and extend through 2009. The outcome is annual per capita GDP growth averaged across the period, calculated using World Bank data. In all regressions, I control for per capita GDP at the start of the period.⁴⁶ For models 1-8 I follow Hanushek and Woessmann (2012), Angrist et al. (2019), and others in the economic growth literature by excluding countries experiencing civil war, suffering inflation crises, or whose rents from natural resources comprise more than 25 percent of GDP, because growth in such states may respond substantially to factors other than human capital.⁴⁷ These exclusions do not substantially change the results; models (9) and (10) show results with the full sample.

For each of the education regressors (LLR, years of schooling, and human capital), I use the estimated value in the first year of the time period. Years of schooling are taken from the B&L data discussed earlier. I also obtain an additional human capital measure from the Penn World Tables, v10.⁴⁸ Because LLR estimates the proportion of individuals born in a year who will eventually obtain literacy—as opposed to the proportion of residents who are already literate—in models (7) and (8) I substitute a measure of LLR with a fifteen-year lag meant to capture the approximate delay between a respondent's birth and her attainment of literacy. Thus, for example, the estimated LLR in 1955 is used to predict the average growth a country experienced from 1970 through 1974. Standard errors across all models are clustered by country.

⁴⁶ I use the log of initial GDP per capita in models (5) and (6) as favored by the neoclassical growth model.

⁴⁷ Mankiw et al. (1992).

⁴⁸ Human capital is calculated using a combination of the Barro and Lee (2013) data, schooling measures constructed by Cohen and Leker (2014), and UNESCO enrollment data.

Table 2 summarizes female estimation results.⁴⁹ Estimates for males are similar and are available in Appendix A4. Female Lifetime Literacy Rates strongly predict economic growth. In general, an increase of female LLR by one-standard deviation predicts an associated increase in growth of 1-2%; a reasonably large amount relative to the sample mean of $\sim 1.6\%$. Moreover, although this exercise is simple and is intended merely as a validation exercise, LLR exhibits greater predictive performance than the alternative measures of schooling years and human capital, neither of which are significantly associated with growth when included independently.⁵⁰ In summary, LLR, as a measure of actual learning, captures valuable information that should prove of interest to researchers.

Table 2: Validity Check of Female L.L.R. as a Predictor of GDP Growth

	GDP Per Capita Growth (Five-Year Average)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lifetime Literacy Rate	0.976*** (0.251)			1.967*** (0.384)	1.126*** (0.257)	2.041*** (0.388)			0.706*** (0.189)	1.293*** (0.264)
L.L.R. (15-Year Lag)							0.781*** (0.271)	2.006*** (0.422)		
Schooling Years (B&L)		-0.040 (0.330)		-0.372 (1.150)		-0.278 (1.154)		-0.340 (1.288)		-0.766 (0.751)
Human Capital (PWT)			-0.166 (0.355)	-1.209 (1.246)		-1.259 (1.259)		-1.345 (1.304)		-0.186 (0.665)
Period Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Initial GDP Per Capita	✓	✓	✓	✓			✓	✓	✓	✓
Log of Initial GDP PC					✓	✓				
No Excluded Countries									✓	✓
Observations	296	296	296	296	296	296	313	313	526	526
Adjusted R^2	0.107	0.065	0.066	0.152	0.112	0.154	0.101	0.142	0.134	0.156
F	7.555	0.008	0.124	7.282	9.683	7.617	4.153	6.017	9.645	9.444

Linear models estimated using OLS. Standard errors are listed in parentheses and are clustered by country. GDP Per Capita Growth is average annualized growth over a five-year period as per the World Bank Indicators, e.g., (1970-1974). All models include a control for observed GDP per capita at the start of the period. The following regressors are also measured in the first year of each five-year period: (1) Lifetime Literacy Rate, which refers to a country's Female LLR; (2) Schooling Years, which refers to the estimated average years of schooling among residents (Barro & Lee 2013); (3) Human Capital, which is taken from the Penn World Tables, v10. In Models (7) and (8), Lifetime Literacy Rate is lagged by fifteen years relative to the outcome, e.g., LLR in 1950 is used to predict average growth from 1965-1969. Models (1-8) exclude observations from countries suffering inflation crises (Reinhart and Rogoff, 2009), for whom natural resource rents comprise more than 25% of GDP (World Bank Indicators), and that are experiencing civil conflict (MEPV).

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$ DV mean (GDP Growth): 1.635.

⁴⁹ All models include period fixed effects to account for global shocks that might influence economic growth or stagnation across all states within a period, but the significance and direction of LLR are robust to omitting these regressors.

⁵⁰ Because years of schooling measures are correlated with learning outcomes but exert no independent effect on growth other than via student learning, in supplemental work I argue years of schooling can function as a useful instrument for the effect of literacy on growth if conditioned on government education sector spending as a proportion of GDP.

World Bank Assistance and the Gender Literacy Gap

Estimation Strategy

I estimate the influence of World Bank (IDA) membership on gender LLR gaps using two-way fixed effects (TWFE). A generalization of difference-in-differences (DiD), the TWFE estimator allows me to control for a large set of unit- and time-invariant confounders. Country fixed effects account for differences in literacy levels associated with time-invariant factors such as distinct cultural traditions, political institutions, or colonial legacies. Similarly, year fixed effects account for the potential influence of annual shocks that affect all countries in a given year, such as worldwide economic downturns or increases in global attention to female education.

A key identifying assumption of TWFE is an absence of unobserved time-varying confounders. If time-varying, unaccounted for factors affect literacy and correlate with World Bank membership, those factors might threaten my ability to discern the effects of membership itself.⁵¹ This requirement is colloquialized in the literature as an assumption of parallel trends. In other words, the average counterfactual outcomes for treated units would have followed a parallel trajectory to that of the control units. Although I cannot test the assumption directly, I implement several common practices that should increase confidence that it is satisfied. As I discuss in the following section, I observe no divergence in average pre-treatment trends between countries that join the IDA and those that do not. Moreover, I include country-specific linear trends for all results in the main text—a conservative requirement intended to account for the possibility that levels of preexisting gender literacy trended distinctly in each country.⁵² Finally, although the baseline model includes no additional regressors, subsequent models incorporate time-varying controls to account for the influence of alternative factors that might affect both World Bank membership and education outcomes.

⁵¹ A second requirement of generalized DiD estimators is for treatment effects to remain constant across units and time periods (Athey and Imbens, 2018; Goodman-Bacon, 2018; Callaway and Sant’Anna, 2020; Sun and Abraham, 2020; Imai and Kim, 2020, etc.). To reduce concerns that results stem from violations of this assumption, I replicate the analysis using counterfactual estimators that allow treatment effects to be arbitrarily heterogeneous across units and periods (Liu et al., 2020). Because results are substantively consistent across estimators, I focus on the TWFE results but offer further discussion in the robustness section.

⁵² Mora and Reggio, 2012. I show in the robustness section that results do not rely on these trends.

I therefore estimate the following equation, in which $Y_{i,t}$ depicts the *female literacy deficit* (Male LLR - Female LLR) in country i for year t ; the terms γ , θ , and δ represent country fixed effects, birth-year fixed effects, and country-specific linear trends, respectively; $X_{i,t}$ denotes a vector of unit- and time-varying control characteristics; and $\epsilon_{i,t}$ represents the disturbance term. Finally, β , the coefficient of interest, is a binary indicator with a value of zero for years in which a country has not become a member of the World Bank's International Development Association (IDA) and a value of one in all years thereafter.⁵³ The IDA provides loans and grants to the poorest developing countries, and membership is necessary to obtain funding. Conditional on these regressors and assumptions, the coefficient of interest, β , provides an unbiased estimate of the average change in the female literacy deficit countries experience upon becoming eligible for IDA assistance:

$$Y_{i,t} = \gamma_i + \theta_t + \delta_{trend\ i} + \beta Member_{i,t} + \kappa' X_{i,t} + \epsilon_{i,t}$$

Baseline Results

Table 3 summarizes my estimates for the overall effect of IDA membership on female literacy deficits. As described above, all results include country fixed effects, year fixed effects, and country-specific linear trends. Standard errors are clustered separately by country and birth year to allow for correlation either among observations from different countries within the same year or, alternatively, from the same country across successive years. Column 1 provides baseline results that exclude additional regressors. Columns 2-5 introduce sets of additional time-varying covariates designed to guard against the possibility that alternative causal processes impede inference. I provide information on the control variables in Appendix A3 and the notes below Table 3, but the different groups include core measures of population and urbanization; variables related to political regime type, government accountability, and the presence of civil or international conflict; and measures of economic growth, membership in international organizations, and a country's receipts of both foreign direct investment and official development assistance.⁵⁴

⁵³ Information on the membership dates for each country are from the World Bank's website. I find evidence of consistent results if I modify the treatment variable to also account for IDA graduation and reentry. See results in Appendices A6 and A7, along with the discussion in the robustness section.

⁵⁴ Concessional IDA loans comprise a subset of ODA flows. For discussion and interpretation, please contact the author.

Across all models, I find that gender literacy gaps increase by a significant amount when countries gain access to IDA assistance. The strength of the relationship does not change substantially across models, and estimated effect sizes range between 4-6%.⁵⁵ Thus, although women born in countries without IDA assistance are on average already less likely than their male peers to obtain literacy, the estimated size of this deficit increases by roughly 5% among women born following IDA accession. While this result does not imply that women born in the wake of IDA membership are less likely to become literate in *absolute* terms, it does suggest that IDA assistance disproportionately benefits males rather than females.⁵⁶

Table 3: Baseline Estimates

	Estimated Effect on Female Literacy Deficit				
	(1)	(2)	(3)	(4)	(5)
W.B. (IDA) Membership	0.047** (0.020)	0.044** (0.020)	0.054*** (0.015)	0.061** (0.023)	0.064*** (0.019)
<u>Controls:</u>					
Population & Urban.		✓			✓
Government & Conflict			✓		✓
Economic Growth & Exposure				✓	✓
Observations	1488	1488	1488	1488	1488
Adjusted R^2	0.778	0.778	0.779	0.780	0.781

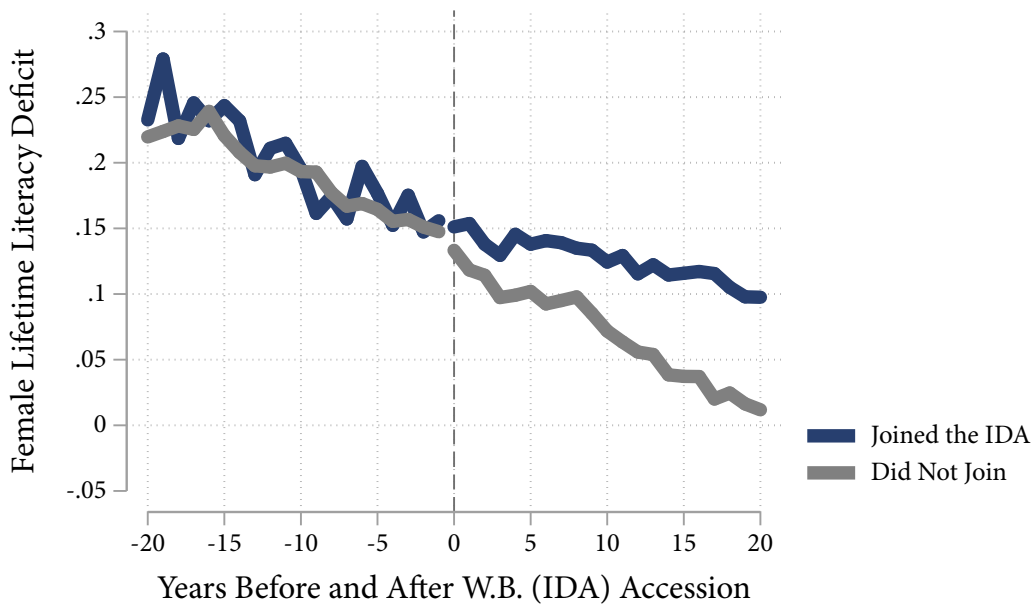
All models include country fixed effects, year fixed effects, and country-specific linear trends. Standard errors are listed in parentheses and are clustered separately by country and year. Controls for Models (2) & (5): Population, Population Growth (annual %), Urban Population (% of total). Controls for Models (3) & (5): VDem Polyarchy, Natural Resource Rents (% of GDP), Presence/Intensity of Civil and International Conflict, Occurrence of a Successful Coup. Controls for Models (4) & (5): Per Capita GDP (current USD), Per Capita GDP Growth (annual %), WTO Membership, UNESCO Membership, Net ODA Received (current USD), Net FDI Inflows (current USD).

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$ DV mean (Female Literacy Deficit): .0809.

⁵⁵ The set of observations in Table 3 is standardized across models to facilitate consistent comparisons. For estimates using alternative samples, see Appendix Table A7.1.

⁵⁶ Appendix Tables A5.1/A5.2 summarize estimates of the absolute effects of IDA membership on Male LLR and Female LLR, respectively, and provide mild evidence that Male LLR increases after a country acquires IDA membership, whereas Female LLR remains roughly constant.

Figure 5: Pre- and Post-Treatment Trends



Notes: Trends for IDA member countries shown blue and for non-members in gray. For each member country I compute the average female literacy deficit for the comparison group, which consists of all countries that were non-IDA members in a given year. The gray line is the average literacy deficit across all comparison groups. Female Literacy Deficits are calculated by the author by subtracting the estimated Female LLR from Male LLR. Data on timing of IDA accession is from the World Bank.

Figure 5 demonstrates that female literacy deficits follow similar downward trends for future IDA members and non-members in the pre-membership period but diverge substantially thereafter.⁵⁷ The figure plots, in blue, the average female literacy deficit among countries that gained access to IDA assistance during the period covered by the data. In contrast, the gray line displays the average female literacy deficit across all non-member comparison groups. Although future members exhibit greater within-trend movement (i.e., greater amplitude of peaks and troughs) relative to the comparison group, overall the two lines trend similarly prior to membership.⁵⁸ The time-varying controls in Models 2-5 of Table 3 should reduce concerns that alternative factors are respon-

⁵⁷ This figure plots trends among all observations in the dataset that include information on a country's level of Electoral Democracy (VDem Polyarchy)—an important variable in subsequent analysis. For a figure including all countries, even those with missing data, see Appendix Figure AF6.2.

⁵⁸ To facilitate additional comparison, Appendix Figure AF6.1 reproduces Figure 5 with fitted pre- and post-membership regression lines for IDA members and an overall regression line fitted for non-members.

sible for this change in trends. To further assess the possibility that the shift in behavior coincides with IDA membership timing merely by happenstance, I conduct 1000 Monte Carlo simulations, repeatedly assigning each country a placebo IDA membership year drawn randomly within the range of its observed LLR data. I find a significant relationship between these placebo dates and shifts in a country's Female Literacy Deficit in only 5% of simulations.

Evidence of Heterogeneous Effects by Regime Type

The theory predicts that education windfalls should yield different effects in autocracies compared to democracies. Regimes in the former category should be wary of extending education to groups that previously lacked access and should therefore attempt to direct aid toward individuals similar to those who are currently educated. In contrast, more democratic regimes are accountable to broader constituencies and should be more willing to extend education benefits broadly and to distribute assistance to new communities. I test this prediction by estimating a new set of models, each of which includes an interaction between World Bank (IDA) membership and regime type. I measure the latter characteristic using the Electoral Democracy Index (Polyarchy) from the Varieties of Democracy dataset, version 11. The measure is coded continuously from 0 to 1, with higher values depicting regimes with broad suffrage; active political and civil society organizations; and clean, competitive elections that directly influence the head of state.⁵⁹

Table 4 summarizes the estimates from the interaction models, which are otherwise equivalent to those from the previous section. Across all results, I find evidence consistent with my theory: the estimated effect of gaining access to World Bank assistance differs substantially between autocracies and democracies. This pattern is robust to the introduction of additional covariates. On average, extreme autocracies with an electoral democracy scores approaching zero experience increases in the female literacy deficit of roughly 7-8% upon gaining access to IDA assistance. In contrast, the estimates predict that a fully democratic country would experience a mild decrease in its female literacy deficit. The sample, however, is inherently limited to developing countries whose electoral

⁵⁹ Lindberg et al., 2014.

democracy scores during the time period largely fell toward the lower end of this range.⁶⁰ Figure 6 provides a graphical illustration of the relationship from Table 4, Model 5. In supplemental tests, I substitute alternative measures of democracy and civic empowerment, including a dichotomous indicator of autocracy as per the Polity2 index, an assessment of whether a country constitutes an egalitarian democracy, and measures of female political participation/representation. These estimates are available in Appendix Table A8.2.

Finally, to account for the possibility that IDA funding influences subsequent democratization, I estimate models interacting IDA membership with the country's last recorded electoral democracy score prior to the date of accession.⁶¹ Results from these estimates are substantively consistent with those presented below and are available in Appendix Table A8.1.

Table 4: Interaction Effect (Membership & Electoral Democracy)

	Estimated Effect on Female Literacy Deficit				
	(1)	(2)	(3)	(4)	(5)
W.B. Membership	0.070*** (0.013)	0.068*** (0.013)	0.070*** (0.013)	0.081*** (0.016)	0.079*** (0.016)
Electoral Democracy	0.017 (0.015)	0.017 (0.025)	0.015 (0.016)	0.018 (0.026)	0.018 (0.033)
Interaction Term	-0.098*** (0.031)	-0.101*** (0.033)	-0.097*** (0.031)	-0.089** (0.034)	-0.092** (0.036)
<u>Controls:</u>					
Population & Urban.		✓			✓
Government & Conflict			✓		✓
Economic Growth & Exposure				✓	✓
Observations	1488	1488	1488	1488	1488
Adjusted R^2	0.780	0.780	0.779	0.781	0.781

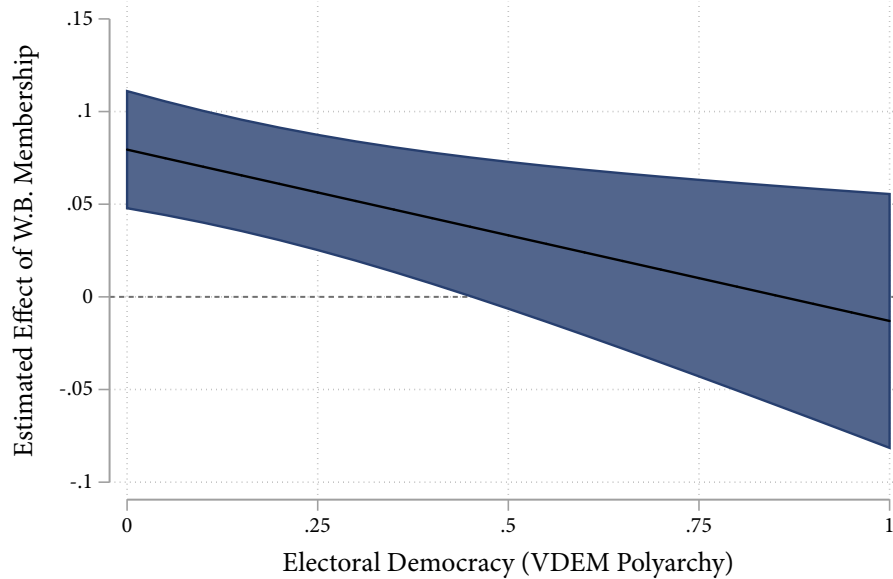
All models include country fixed effects, year fixed effects, and country-specific linear trends. Standard errors are listed in parentheses and are clustered separately by country and year. Controls for Models (2) & (5): Population, Population Growth (annual %), Urban Population (% of total). Controls for Models (3) & (5): Natural Resource Rents (% of GDP), Presence/Intensity of Civil and International Conflict, Occurrence of a Successful Coup. Controls for Models (4) & (5): Per Capita GDP (current USD), Per Capita GDP Growth (annual %), WTO Membership, UNESCO Membership, Net ODA Received (current USD), Net FDI Inflows (current USD).

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$ DV mean (Female Literacy Deficit): .0809.

⁶⁰ India and Suriname's scores were the highest in the sample, approaching 0.8 in the late 1990s and early 2000s.

⁶¹ To include this measure with country fixed effects, I allow electoral democracy scores to vary until IDA accession.

Figure 6: Conditional Marginal Effect of W.B. Membership & Regime Type



Notes: Estimated marginal effect of IDA membership on the female literacy deficit, conditional on a country's electoral democracy score. The estimated results are equivalent to those summarized in Table 4, model 5. Measures of the female literacy deficit are calculated by the author using data from DHS and MICS. Data on IDA membership are from the World Bank. Data on electoral democracy from the Varieties of Democracy dataset, v11.

Robustness and Exploration of Additional Mechanisms

Alternative Samples, Measurement Checks, and Estimators

As described earlier, both the baseline results and the estimates from the interaction models are robust to estimation on alternative subsets of the available data. Several of the alternative samples are designed to assess the influence of potential measurement error in LLR. For example, to assess whether minor differences in the implementation of literacy tests between the DHS and MICS could influence the results, I estimate models that include only data from the DHS.⁶² Likewise, to reduce the possibility that results are subject to influence from outliers in the LLR estimates, I estimate models excluding observations with values that fall below various thresholds of precision.⁶³ Third,

⁶² I do not estimate models using only observations from MICS surveys, as those comprise less than 15% of the data.

⁶³ I exclude observations with fewer than 50 respondents or the standard error of the LLR proportion exceeds 5%. Likewise, I assess the influence of outlier countries by repeatedly re-estimating Table 3, Model 5 and Table 4, Model

in case early LLR ceilings prevent me from observing improvements in female literacy, I estimate results while omitting countries with initial Female LLR values above 95%. Fourth, to reduce the possibility that literacy gaps merely reflect differences in the typical ages at which males or females acquire education, I estimate models that exclude respondents under the age of 18 at the time of surveying and, separately, respondents who indicated they had participated in adult literacy programs.⁶⁴ Fifth, several countries in the sample gained independence during the observed period. To reduce the possibility that the political or social changes that accompany independence drive the main results, I estimate a model in which I omit countries whose timing of independence closely coincides with IDA accession. Finally, although unit-specific linear trends are generally considered a conservative requirement, some authors have noted that they can complicate the interpretation of treatment effects when data cover long periods.⁶⁵ I therefore estimate models in which such trends are excluded. Estimates are largely consistent across each of the alternative samples discussed above; results are available in Appendix A6.⁶⁶

I also assess the relationship between IDA membership and female literacy deficits using alternative estimators. Liu et al. (2020) describe a class of estimators that take the observations under dichotomous treatment as missing data in order to estimate their counterfactuals. Although they are a new addition to the literature, these estimators provide three advantages relative to the TWFE results I described above. First, they relax the constant treatment effect assumption and correct the biases induced by treatment effect heterogeneity that have provoked substantial recent discussion.⁶⁷ Second, by extension they enable researchers to estimate and plot dynamic treatment effects with-

5, iteratively excluding different countries. I find no evidence that results rely on the inclusion of any specific country. Across all estimates, variables of interest retain consistent signs and remain significant.

⁶⁴ These steps address the possibility that women may, for example, acquire education more gradually than males during childhood or, alternatively, that males enjoy greater access to literacy programming in adulthood.

⁶⁵ See discussion in Wolfers (2006).

⁶⁶ To account for possible variation in measurement error across survey waves, I replicate the main results using LLR measures estimated at the survey-year level, substituting survey fixed effects in place of country fixed effects. Results are available in Appendix Table A6.5. In these results, the interaction term falls slightly shy of traditional significance, potentially due to reduced within-unit variation in electoral democracy scores because temporal coverage within individual surveys is reduced relative to country-level data. An alternative interaction between IDA membership and a country's electoral democracy score prior to membership remains significant.

⁶⁷ Goodman-Bacon, 2018; De Chaisemartin and D'Haultfoeuille, 2020.

out assuming treatment effect homogeneity. Researchers can therefore conduct systematic tests to gauge the validity of the no-time-varying-confounder assumption. These include tests of whether the means of pre-treatment estimates differ from zero as well as placebo tests that estimate effects for several pre-treatment periods. Finally, counterfactual estimators can more easily accommodate complex effects such as treatment reversal and reentry.⁶⁸ From this category, I use matrix completion (MC) and interactive fixed effects counterfactual (IFEct) estimators to assess the effects of IDA membership.⁶⁹ I find consistent evidence that IDA membership widens the gender literacy gap in autocracies, whereas in more democratic states membership is associated with either no change or a decline in the gender literacy gap.⁷⁰ Results are available in Appendix A12, are substantively consistent across the IFEct and MC estimators, and are robust to the use of pre-membership democracy measures as well as a treatment condition that allows for IDA graduation and reentry.

Assessing Male Prioritization and Differential Returns to Education

Two alternative mechanisms merit particular attention. Both possibilities assert that divergent LLR outcomes for males and females stem from distinct educational preferences at the family or individual level rather than from selective targeting by governments. Although a priori there may be little reason to believe that such preferences correlate with both IDA membership timing and regime type, these explanations nevertheless bear mentioning. Estimation results from the procedures I discuss below are located in Appendix A11.

First, parental preferences for male children are well-established in many parts of the world. Although excess mortality among female infants remains the most notorious example, males also enjoy within-family prioritization in health care, superior nutrition, lower rates of child labor, and, of particular relevance to this study, access to education.⁷¹ Even if new education-sector investments

⁶⁸ Several IDA members “graduated” from access to assistance by exceeding an income-per-capita threshold during the data coverage, although some subsequently reentered eligibility.

⁶⁹ For MC estimators, see Athey and Imbens (2018).

⁷⁰ Because MC and IFEct estimators do not allow the inclusion of interactive covariates, I split the sample at the median Electoral Democracy value, then assess whether the estimated treatment effects are statistically distinct.

⁷¹ Sen, 1992; Rose, 1999; Behrman, 1988; Edmonds and Pavcnik, 2005; Alderman and King, 1998; Kumar, 2013.

provide resources and opportunities that on paper are equally available across genders, parents who follow gender-queuing or son-preference patterns might nonetheless choose to educate their male children first.⁷² In general, differences in gender equity norms between countries are tied to cultural traditions, historical legacies, and other fixed or slow-moving processes that are largely absorbed in the estimations by unit fixed effects. Assessing these behaviors therefore requires that I find within-country variation among groups of respondents who differ on individual characteristics. To do so, I construct LLR estimates for female DHS respondents who possess only one older sibling, then compare LLR estimates between those whose older sibling is male versus those whose sibling is female. In principle, the two groups should be similar to one another but for the difference in their sibling's sex. Although this test is necessarily crude and estimates of LLR may be underpowered due to the limited sub-sample, I find no evidence that a "sibling literacy gap" varies in conjunction with either IDA membership or a country's level of electoral democracy, substantially reducing the possibility that these dynamics drive my results.⁷³

Second, human capital theorists argue that families and individuals pursue different levels of education depending on how levels of attainment translate into expected earnings.⁷⁴ Thus, even if IDA assistance provides new and equal access to education for all groups, males will be more likely to participate if the expected returns to male education are higher than those for females.⁷⁵ Unfortunately, although persistent gender wage gaps are widely evident, systematic data comparing male and female earnings at various education levels is difficult to obtain for many of the countries in question.⁷⁶ Moreover, as with the previous example, any time-invariant characteristics would be absorbed during estimation by unit fixed effects. I therefore turn once again to within-country variation between groups. Recent work in economics finds that groups who follow bride price and bride wealth traditions are more responsive to development projects aimed at boosting female ed-

⁷² Dollar and Gatti, 1999; Sundaram and Vanneman, 2008; Björkman-Nyqvist, 2013. To obtain loans, some countries agreed to raise or establish fees for primary education, placing additional pressure on parents.

⁷³ The estimated relationship with IDA membership is insignificant with a coefficient opposite of the predicted direction.

⁷⁴ Becker, 1981; Rosenzweig and Schultz, 1982; Kumar, 2013; Bussemakers et al., 2017.

⁷⁵ Women might likewise forgo additional education if they face high opportunity costs for schooling (Becker, 1985).

⁷⁶ Oaxaca, 1973; Weichselbaumer and Winter-Ebmer, 2005.

ucation.⁷⁷ Ashraf et al. (2020), for example, show that parents who follow these customs are more likely to educate their daughters when new opportunities emerge—likely because educated women can command larger bride price payments. To assess whether similar effects might correlate with IDA assistance and regime type, I record the ethnicity of each respondent in my sample, then match respondents to the roughly 1300 societies listed in Murdock’s (1967) *Ethnographic Atlas*, which provides information on the customary direction (if any) of financial transfers at marriage. Finally, I estimate whether women from groups known to practice bride price traditions obtain larger improvements in LLR following IDA membership than do women from alternative groups. As with the previous test, these results should be interpreted cautiously due the relatively small sample and the difficulty of obtaining high quality data on marriage institutions. Nevertheless, I find no evidence supporting this mechanism or suggesting that it distorts my overall findings.

Evidence of Urban Targeting and Effect Reversal at IDA Graduation

Finally, I provide two supplemental tests that further support my overall theory that governments channel windfalls in education funding to preferred communities. First, respondents to DHS surveys provide information on their childhood place of residence. Answers range from a country’s capital city to other urban areas or even rural and unincorporated communities. Likewise, the VDem dataset includes a measure of the location from which a country’s federal government draws its primary political support: from the capital, from other cities, from rural communities, or the regime’s supporters are not geographically consolidated within the country. If my theory is correct that governments direct newly available education aid toward preferred communities, I should find evidence that urban communities enjoy more rapid improvements in literacy than rural communities when the head of state draws support from urban areas, along with opposite results when the regime draws support from rural areas.

To test this, I construct new LLR estimates for individuals who describe their childhood place of residence as either the nation’s capital or another urban setting, along with LLR estimates for indi-

⁷⁷ These customs involve large financial transfers—sometimes in excess of a year’s income in money, livestock, or commodities—paid by the groom or his family to the bride’s parents upon marriage.

viduals raised in more rural communities. I calculate an Urban Literacy Gap by taking the difference of these measures (Urban LLR - Rural LLR). As with previous models, I then use TWFE to estimate the effect of an interaction between *IDA membership* and the government's *source of political support* on the Urban Literacy Gap. Estimates for both males and females are available in Appendix A11, models 3-6. I find strong evidence of the predicted pattern among males: regardless of whether a regime's degree of urban support is measured dichotomously or continuously, IDA membership is associated with an increase in the Urban Literacy Gap within countries whose regimes draw primarily from urban supporters, but results are neutral or mildly negative where regimes draw primarily from rural groups. Among females, coefficients are in the expected directions but the relationship is only significant when urban support is measured continuously.

My final test relates to IDA graduation. When the per capita GDP of an IDA member state exceeds a maximum threshold, that state is no longer eligible for development assistance. Twenty-five countries for which I provide LLR estimates experienced IDA graduation during the period covered by the data, although several of these subsequently experienced a reversal of this process.⁷⁸ If autocratic governments are indeed manipulating the use of foreign education funding while they are eligible for such assistance, I should expect to observe an end or reversal of the effect once IDA access is switched off. To assess this behavior, I use TWFE to estimate the effect of both IDA graduation and a country's electoral democracy score on the female literacy gap.⁷⁹ Estimates are provided in Appendix Table A11.2. Although the results are relatively weak, I find that IDA graduation is consistently associated with a decline in a country's Female Literacy Deficit, even when a full set of controls are included. Likewise, in accordance with the theory I find evidence in some models that the reversal of behavior is particularly strong among autocratic states. Collectively, the evidence suggests that women begin to close the literacy gap once autocratic governments lose the ability to manipulate external funds or direct excess educational resources toward males.

⁷⁸ I estimate results of exit and reentry with the counterfactual estimators in Appendix A12.

⁷⁹ To create an appropriate comparison group, I subset the sample, removing observations from countries that have never joined the IDA, observations prior to IDA accession, and observations following re-entry to the IDA.

Discussion and Conclusion

The World Bank plays a critical role in development lending. During the twentieth century, its assistance extended lifespans, lifted communities from poverty, and improved education outcomes for millions of individuals across the globe. Nevertheless, as this paper highlights, the distributional effects of World Bank lending are uneven. Using new measures of lifetime literacy that address many of the quality and comparability problems that are common in other datasets, I find strong evidence of distributional discrepancies in education outcomes. Among recipients of IDA assistance, educational improvements accrue predominantly among male citizens, leading to expansions of gender literacy gaps, particularly in autocratic countries. The identification strategy and supplemental tests suggest these outcomes are causally tied to IDA membership and are not plausibly attributable to other potential mechanisms.⁸⁰

My findings offer valuable insight into how development funding can reify preexisting social hierarchies. Although autocratic leaders seek to obtain the benefits of economic growth associated with an educated citizenry, they also worry that expansions of education will undermine social stability by facilitating collective action or encouraging political participation among groups that were previously excluded. As a result, foreign aid providers, international financial institutions, and development organizations should tailor their lending strategies accordingly, acknowledging the likelihood that leaders will channel funds and development assistance toward preferred communities and the possibility that aggregate measures of improvement can easily mask hidden inequities. Even when development funding and technical knowledge abound, disadvantaged communities are unlikely to realize large educational improvements if political roadblocks limit access and obscure accurate assessment. This is not to say aid providers should treat the dismantlement of preexisting hierarchies as their primary objective; rather, when extending assistance they should gauge the potential consequences with an eye toward the Hippocratic oath: first, do no harm.

Because aid recipients will weigh the incentives donors establish against their own political

⁸⁰ In separate work, I show receiving larger *amounts* of IDA assistance also increases gender literacy gaps. Because this requires a substantially different research design and lengthy discussion, I omit analysis in the current paper but can provide results upon request.

exigencies, development organizations should continue working to define contracts that induce recipients to pursue desired actions. An important step in this process is the establishment of reliable metrics and monitoring programs that assess not only countrywide improvements in outcomes, but also within-country variation across likely cleavage groups. As the data in this paper demonstrate, detailed measures of concrete learning outcomes—even when measured crudely, via dichotomous indicators of individual literacy—can expose substantial government manipulation and targeting of development assistance. Funders should encourage participation in standardized testing and systematic learning assessments that, albeit imperfect, will facilitate improved analysis of performance across social groups. When analyzed appropriately, such data should foster improved monitoring of education outcomes, encouraging more efficient service provision and implementation in line with funders' desires.

Although this study provides evidence of overall trends and suggests a number of potential channels—including male hiring initiatives, prioritization of secondary education, and preferential funding of male-only schools—pinpointing the operative policy within each country will only be feasible through careful, case-by-case analysis. An important task for future research is discerning the precise actions through which individual governments selectively target education aid as well as the specific groups they seek to prioritize.

Finally, the data and method presented in this paper should stimulate additional research not only on the educational consequences of specific policy interventions but also on the relationship between human capital accumulation and economic growth or other outcomes of interest. By highlighting a process through which researchers can construct reliable measures of historical conditions using contemporary survey data, this paper may facilitate the study of other important issues. Asking survey respondents to report information about historic conditions or events offers a fruitful means of gathering data when reliable evidence from time periods of interest is otherwise difficult to obtain. Just as researchers use individual interviews to construct historical accounts of earlier events, so too can we use representative surveys to develop a clearer picture of history where none might otherwise exist.

References

- Acemoglu, Daron, Simon Johnson, James Robinson, and Yared Pierre (2005). From education to democracy? *American Economic Review* 95.2, 44–49.
- Acemoglu, Daron and James Robinson (2006). Economic backwardness in political perspective. *American Political Science Review* 100.1, 115–131.
- Ahmed, Faisal (2012). The perils of unearned foreign income: Aid, remittances, and government survival. *American Political Science Review*, 146–165.
- Akresh, Richard and Damien De Walque (2008). *Armed conflict and schooling: Evidence from the 1994 Rwandan genocide*. The World Bank.
- Alderman, Harold and Elizabeth King (1998). Gender differences in parental investment in education. *Structural Change and Economic Dynamics* 9.4, 453–468.
- Altincekic, Ceren and David Bearce (2014). Why there should be no political foreign aid curse. *World Development* 64, 18–32.
- Angrist, Noam, Simeon Djankov, Pinelopi Koujianou Goldberg, and Harry Anthony Patrinos (2019). *Measuring human capital*. The World Bank.
- Ansell, Ben (2008). Traders, teachers, and tyrants: democracy, globalization, and public investment in education. *International Organization*, 289–322.
- Ansell, Ben and Johannes Lindvall (2013). The Political Origins of Primary Education Systems: Ideology, Institutions, and Interdenominational Conflict in an Era of Nation-Building. *American Political Science Review*, 505–522.
- Ashraf, Nava, Natalie Bau, Nathan Nunn, and Alessandra Voena (2020). Bride price and female education. *Journal of Political Economy* 128.2, 591–641.
- Athey, Susan and Guido Imbens (2018). *Design-based analysis in difference-in-differences settings with staggered adoption*. Tech. rep. National Bureau of Economic Research.
- Banerjee, Abhijit, Shawn Cole, Esther Duflo, and Leigh Linden (2007). Remedying education: Evidence from two randomized experiments in India. *Quarterly Journal of Economics* 122.3, 1235–1264.
- Barro, Robert (2001). Human capital and growth. *American Economic Review* 91.2, 12–17.
- Barro, Robert and Jong Wha Lee (1996). International measures of schooling years and schooling quality. *American Economic Review* 86.2, 218–223.
- (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics* 104, 184–198.
- Beath, Andrew, Fotini Christia, and Ruben Enikolopov (2013). Empowering women through development aid: Evidence from a field experiment in Afghanistan. *American Political Science Review*, 540–557.
- Becker, Gary (1981). Altruism in the Family and Selfishness in the Market Place. *Economica* 48.189, 1–15.
- (1985). Human capital, effort, and the sexual division of labor. *Journal of Labor Economics* 3.1.2, S33–S58.
- Becker, Gary, Kevin Murphy, and Robert Tamura (1990). Human capital, fertility, and economic growth. *Journal of Political Economy* 98.5, Part 2, S12–S37.
- Behrman, Jere (1988). Intrahousehold allocation of nutrients in rural India: Are boys favored? Do parents exhibit inequality aversion? *Oxford Economic Papers* 40.1, 32–54.

- Behrman, Jere (1996). Measuring the effectiveness of schooling policies in developing countries: Revisiting issues of methodology. *Economics of Education Review* 15.4, 345–364.
- Behrman, Jere and James Knowles (1999). Household income and child schooling in Vietnam. *World Bank Economic Review* 13.2, 211–256.
- Behrman, Jere and Mark Rosenzweig (1994). Caveat emptor: cross-country data on education and the labor force. *Journal of Development Economics* 44.1, 147–171.
- Berinsky, Adam and Gabriel Lenz (2011). Education and political participation: Exploring the causal link. *Political Behavior* 33.3, 357–373.
- Birchler, Cassandra, Sophia Limpach, and Katharina Michaelowa (2016). Aid modalities matter: The impact of different world bank and imf programs on democratization in developing countries. *International Studies Quarterly* 60.3, 427–439.
- Bisbee, James, James Hollyer, Peter Rosendorff, and James Vreeland (2019). The Millenium Development Goals and Education: Accountability and Substitution in Global Assessment. *International Organization*, 547–578.
- Björkman-Nyqvist, Martina (2013). Income shocks and gender gaps in education: Evidence from Uganda. *Journal of Development Economics* 105, 237–253.
- Bourguignon, François and Thierry Verdier (2000). Oligarchy, democracy, inequality and growth. *Journal of Development Economics* 62.2, 285–313.
- Brazys, Samuel, Johan A Elkind, and Gina Kelly (2017). Bad neighbors? How co-located Chinese and World Bank development projects impact local corruption in Tanzania. *The Review of International Organizations* 12.2, 227–253.
- Breierova, Lucia and Esther Duflo (2004). *The impact of education on fertility and child mortality: Do fathers really matter less than mothers?* Tech. rep. National Bureau of Economic Research.
- Bueno de Mesquita, Bruce and Alastair Smith (2009). A political economy of aid. *International Organization*, 309–340.
- Bueno de Mesquita, Bruce, Alastair Smith, Randolph Siverson, and James Morrow (2005). *The logic of political survival*. MIT press.
- Bussemakers, Carlijn, Kars van Oosterhout, Gerbert Kraaykamp, and Niels Spierings (2017). Women’s worldwide education–employment connection: a multilevel analysis of the moderating impact of economic, political, and cultural contexts. *World Development* 99, 28–41.
- Callaway, Brantly and Pedro Sant’Anna (2020). Difference-in-differences with multiple time periods. *Journal of Econometrics*.
- Cantoni, Davide, Yuyu Chen, David Yang, Noam Yuchtman, and Jane Zhang (2017). Curriculum and ideology. *Journal of Political Economy* 125.2, 338–392.
- Cohen, Daniel and Laura Leker (2014). Health and education: Another look with the proper data.
- Cohen, Daniel and Marcelo Soto (2007). Growth and human capital: good data, good results. *Journal of Economic Growth* 12.1, 51–76.
- Cruz, Cesi and Christina Schneider (2017). Foreign aid and undeserved credit claiming. *American Journal of Political Science* 61.2, 396–408.
- Cutler, David, Angus Deaton, and Adriana Lleras-Muney (2006). The determinants of mortality. *Journal of Economic Perspectives* 20.3, 97–120.
- Darden, Keith and Anna Grzymala-Busse (2006). The great divide: Literacy, nationalism, and the communist collapse. *World Politics* 59.1, 83–115.
- De Chaisemartin, Clément and Xavier D’Haultfoeuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review* 110.9, 2964–96.

- De la Fuente, Angel and Rafael Doménech (2006). Human capital in growth regressions: how much difference does data quality make? *Journal of the European Economic Association* 4.1, 1–36.
- Dee, Thomas (2004). Are there civic returns to education? *Journal of Public Economics* 88.9-10, 1697–1720.
- Dollar, David and Roberta Gatti (1999). *Gender inequality, income, and growth: are good times good for women?* Vol. 1. Development Research Group, The World Bank Washington, DC.
- Dollar, David and Victoria Levin (2005). *Sowing and reaping: institutional quality and project outcomes in developing countries*. The World Bank.
- Dorn, Charles and Kristen Ghodsee (2012). The Cold War politicization of literacy: communism, UNESCO, and the World Bank. *Diplomatic History* 36.2, 373–398.
- Edmonds, Eric and Nina Pavcnik (2005). Child labor in the global economy. *Journal of Economic Perspectives* 19.1, 199–220.
- Engelsma, Brian, Gerry Mackie, and Brandon Merrell (2020). Unprogrammed abandonment of female genital mutilation/cutting. *World Development* 129, 104845.
- Faye, Michael and Paul Niehaus (2012). Political aid cycles. *American Economic Review* 102.7, 3516–30.
- Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer (2013). Human capital and regional development. *Quarterly Journal of Economics* 128.1, 105–164.
- Gift, T and E Wibbels (2014). Reading, writing, and the regrettable status of education research in comparative politics. *Annual Review of Political Science* 17, 291–312.
- Glewwe, Paul and Michael Kremer (2006). Schools, teachers, and education outcomes in developing countries. *Handbook of the Economics of Education* 2, 945–1017.
- Glewwe, Paul, Michael Kremer, Sylvie Moulin, et al. (2002). *Textbooks and test scores: Evidence from a prospective evaluation in Kenya*. Tech. rep. Citeseer.
- Goodman-Bacon, Andrew (2018). *Difference-in-differences with variation in treatment timing*. Tech. rep. National Bureau of Economic Research.
- Gray, Mark, Miki Kittilson, and Wayne Sandholtz (2006). Women and globalization: A study of 180 countries, 1975-2000. *International Organization*, 293–333.
- Gregorio, Jose De and Jong-Wha Lee (2002). Education and income inequality: new evidence from cross-country data. *Review of Income and Wealth* 48.3, 395–416.
- Haggard, Stephan and Robert Kaufman (2012). Inequality and regime change: Democratic transitions and the stability of democratic rule. *American Political Science Review*, 495–516.
- Hanushek, Eric (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature* 24.3, 1141–1177.
- Hanushek, Eric and Ludger Woessmann (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth* 17.4, 267–321.
- Heyneman, Stephen (2003). The history and problems in the making of education policy at the World Bank 1960–2000. *International Journal of Educational Development* 23.3, 315–337.
- Hicken, Allen (2011). Clientelism. *Annual Review of Political Science* 14, 289–310.
- Hillygus, Sunshine (2005). The missing link: Exploring the relationship between higher education and political engagement. *Political Behavior* 27.1, 25–47.
- Imai, Kosuke and In Song Kim (2020). On the use of two-way fixed effects regression models for causal inference with panel data. *Political Analysis*, 1–11.
- Isham, Jonathan, Daniel Kaufmann, and Lant Pritchett (1997). Civil liberties, democracy, and the performance of government projects. *World Bank Economic Review* 11.2, 219–242.

- Iversen, Torben and Frances Rosenbluth (2006). The political economy of gender: Explaining cross-national variation in the gender division of labor and the gender voting gap. *American Journal of Political Science* 50.1, 1–19.
- Jablonski, Ryan (2014). How aid targets votes: the impact of electoral incentives on foreign aid distribution. *World Politics* 66, 293.
- Jones, Phillip (1997). The World Bank and the literacy question: Orthodoxy, heresy and ideology. *International Review of Education* 43.4, 367–375.
- Kilby, Christopher (2009). The political economy of conditionality: An empirical analysis of World Bank loan disbursements. *Journal of Development Economics* 89.1, 51–61.
- Knowles, Stephen, Paula Lorgelly, and Dorian Owen (2002). Are educational gender gaps a brake on economic development? Some cross-country empirical evidence. *Oxford Economic Papers* 54.1, 118–149.
- Kono, Daniel and Gabriella Montinola (2009). Does foreign aid support autocrats, democrats, or both? *Journal of Politics* 71.2, 704–718.
- Krueger, Alan and Mikael Lindahl (2001). Education for growth: Why and for whom? *Journal of Economic Literature* 39.4, 1101–1136.
- Kumar, Alok (2013). Preference based vs. market based discrimination: Implications for gender differentials in child labor and schooling. *Journal of Development Economics* 105, 64–68.
- Kyriacou, George (1991). *Level and growth effects of human capital: a cross-country study of the convergence hypothesis*. Tech. rep.
- Lake, David and Matthew Baum (2001). The invisible hand of democracy: political control and the provision of public services. *Comparative Political Studies* 34.6, 587–621.
- Lankina, Tomila and Lullit Getachew (2012). Mission or empire, word or sword? The human capital legacy in postcolonial democratic development. *American Journal of Political Science* 56.2, 465–483.
- Lewis, Maureen and Marlaine Lockheed (2006). *Inexcusable absence: why 60 million girls still aren't in school and what to do about it*. Ctr for Global Development.
- Lindberg, Staffan, Michael Coppedge, John Gerring, and Jan Teorell (2014). V-Dem: A new way to measure democracy. *Journal of Democracy* 25.3, 159–169.
- Lipset, Seymour Martin (1960). *The social bases of politics*. Baltimore: The Johns Hopkins University Press.
- Liu, Licheng, Ye Wang, and Yiqing Xu (2020). A practical guide to counterfactual estimators for causal inference with time-series cross-sectional data. SSRN-3555463.
- López-Cariboni, Santiago and Xun Cao (2019). When do authoritarian rulers educate: Trade competition and human capital investment in Non-Democracies. *Review of International Organizations* 14.3, 367–405.
- Lucas, Robert (1988). On the mechanics of economic development. *Journal of Monetary Economics* 22.1, 3–42.
- Mankiw, Gregory, David Romer, and David Weil (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107.2, 407–437.
- Merrell, Brandon and Alexei Abrahams (2020). Monopolies of Violence: Civil Insurgency and the Accountability Deficit. *Unpublished Manuscript*.
- Miller, Michael (2012). Economic development, violent leader removal, and democratization. *American Journal of Political Science* 56.4, 1002–1020.

- Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos (2004). Does education improve citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics* 88.9-10, 1667–1695.
- Milner, Helen, Daniel Nielson, and Michael Findley (2016). Citizen preferences and public goods: Comparing preferences for foreign aid and government programs in Uganda. *The Review of International Organizations* 11.2, 219–245.
- Mora, Ricardo and Iliana Reggio (2012). Treatment effect identification using alternative parallel assumptions.
- Mundy, Karen and Francine Menashy (2014). The World Bank and private provision of schooling: A look through the lens of sociological theories of organizational hypocrisy. *Comparative Education Review* 58.3, 401–427.
- Murdock, George (1967). Ethnographic atlas: a summary. *Ethnology* 6.2, 109–236.
- Nie, Norman, Jane Junn, and Kenneth Stehlik-Barry (1996). *Education and democratic citizenship in America*. University of Chicago Press.
- Oaxaca, Ronald (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 693–709.
- Paglayan, Agustina (2021). The non-democratic roots of mass education: evidence from 200 years. *American Political Science Review* 115.1, 179–198.
- Psacharopoulos, George (1994). Returns to investment in education: A global update. *World Development* 22.9, 1325–1343.
- Rebelo, Sergio (1991). Long-run policy analysis and long-run growth. *Journal of Political Economy* 99.3, 500–521.
- Reinhart, Carmen and Kenneth Rogoff (2009). *This time is different: Eight centuries of financial folly*. Princeton University Press.
- Rose, Elaine (1999). Consumption smoothing and excess female mortality in rural India. *Review of Economics and Statistics* 81.1, 41–49.
- Rosenzweig, Mark and Paul Schultz (1982). Market opportunities, genetic endowments, and intrafamily resource distribution: Child survival in rural India. *American Economic Review* 72.4, 803–815.
- Ross, M (2006). Is democracy good for the poor? *American Journal of Political Science* 50.4, 860–874.
- Sanborn, Howard and Clayton Thyne (2014). Learning democracy: Education and the fall of authoritarian regimes. *British Journal of Political Science*, 773–797.
- Sen, Amartya (1992). Missing women. *BMJ: British Medical Journal* 304.6827, 587.
- Smith, Alastair (2008). The perils of unearned income. *Journal of Politics* 70.3, 780–793.
- Sondheimer, Rachel and Donald Green (2010). Using experiments to estimate the effects of education on voter turnout. *American Journal of Political Science* 54.1, 174–189.
- Srinivasan, Thirukodikaval (1994). Human development: a new paradigm or reinvention of the wheel? *American Economic Review* 84.2, 238–243.
- Stasavage, David (2005). Democracy and education spending in Africa. *American Journal of Political Science* 49.2, 343–358.
- Stromquist, Nelly (2006). Women's rights to adult education as a means to citizenship. *International Journal of Educational Development* 26.2, 140–152.
- Sun, Liyang and Sarah Abraham (2020). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*.

- Sundaram, Aparna and Reeve Vanneman (2008). Gender differentials in literacy in India: The intriguing relationship with women's labor force participation. *World Development* 36.1, 128–143.
- Testa, Patrick (2018). Education and propaganda: Tradeoffs to public education provision in non-democracies. *Journal of Public Economics* 160, 66–81.
- Trudell, Barbara (2009). Local-language literacy and sustainable development in Africa. *International Journal of Educational Development* 29.1, 73–79.
- Verba, Sidney and Gabriel Almond (1963). *The civic culture: Political attitudes and democracy in five nations*. Princeton University Press.
- Wagner, Daniel (2011). What happened to literacy? Historical and conceptual perspectives on literacy in UNESCO. *International Journal of Educational Development* 31.3, 319–323.
- Wantchekon, Leonard, Marko Klašnja, and Natalija Novta (2015). Education and human capital externalities: evidence from colonial Benin. *Quarterly Journal of Economics* 130.2, 703–757.
- Weichselbaumer, Doris and Rudolf Winter-Ebmer (2005). A meta-analysis of the international gender wage gap. *Journal of Economic Surveys* 19.3, 479–511.
- Wickens, Corrine and Jennifer Sandlin (2007). Literacy for what? Literacy for whom? The politics of literacy education and neocolonialism in UNESCO- and World Bank-sponsored literacy programs. *Adult Education Quarterly* 57.4, 275–292.
- Wolfers, Justin (2006). Did unilateral divorce laws raise divorce rates? A reconciliation and new results. *American Economic Review* 96.5, 1802–1820.
- Woodberry, Robert (2012). The missionary roots of liberal democracy. *American Political Science Review*, 244–274.
- World Bank (2017). *World development report 2018: Learning to realize education's promise*. The World Bank.
- Wright, Joseph and Matthew Winters (2010). The politics of effective foreign aid. *Annual Review of Political Science* 13.
- Wyndow, Paula, Jianghong Li, and Eugen Mattes (2013). Female empowerment as a core driver of democratic development: A dynamic panel model from 1980 to 2005. *World Development* 52, 34–54.

Online Appendix

Table of Contents

A1: Lists of DHS and MICS Files	A2
A2: Missingness Maps of Lifetime Literacy Estimates	A6
A3: Lifetime Literacy Rate Summary Statistics	A8
A4: Male LLR and Economic Growth	A10
A5: Estimated Effects of IDA Membership on Absolute Levels of Lifetime Literacy	A11
A6: Robustness — Alternative Pre/Post Trends	A12
A7: Robustness — Alternative Samples	A13
A8: Robustness — Alternative Regime Type Measures	A14
A9: Robustness — Survey-Year Observations & Survey Fixed Effects	A16
A10: Robustness — Omit IDA Graduates	A17
A11: Supplemental Tests — Sibling Gap, Bride-Price Gap, Urban Gap & Graduation	A19
A12: Matrix Completion and Interactive Fixed Effects Counterfactual Estimates	A21

Appendix omitted to meet IPES maximum file size requirements
A version with the Appendix included is [linked here](#)