What we learn about girls’ education from interventions that don’t focus on girls

David K. Evans

Fei Yuan

October 26, 2018

Abstract: Despite dramatic global gains in access to education, 130 million school-age girls remain out of school. Among those who do enter, too many fail to gain the essential skills to succeed after they complete their schooling. Many previous efforts to synthesize evidence on how to improve both access and learning for girls have focused on interventions that are principally targeted to girls, such as girls’ latrines or scholarships for girls. However, if non-targeted interventions – those that benefit both girls and boys – significantly improve girls’ education, then such a strategy may miss some of the best investments for improving educational opportunities for girls. In this review, we bring together evidence on 179 educational interventions from 54 countries and identify their impacts on girls, regardless of whether they target girls specifically. We find that for improving learning, the average impact is similar between targeted and non-targeted interventions, but the most effective interventions are non-targeted. These results suggest that improving girls’ access to school may benefit from targeted interventions, but that the best bet to improve their learning outcomes may involve simply upgrading the quality of education for all learners.

JEL Classification: I24, I21, O1

Keywords: education, girls, effect size, impact evaluation, developing countries

1 The authors would like to thank Dana Schmidt and the team at Echidna Giving, Deon Filmer, Oni Lusk-Stover, seminar audiences at the World Bank and the Brookings Institution, for suggestions. We would also like to thank Echidna Giving and the Umbrella Facility for Gender Equality (UFGE) at the World Bank for financing this work. We are grateful to Danielle Sobol, Shikhty Sunny, and Tara Siegel for excellent research assistance. The authors may be contacted at Evans (World Bank: devans2@worldbank.org) and Yuan (Harvard Graduate School of Education: fyuan@g.harvard.edu).
1. Introduction

Investing in girls’ education has been called “the world’s best investment” (Sperling and Winthrop 2015). Roughly 15 percent of the world’s population in 2016 are girls younger than 18 years old. Along with their male peers, they constitute a vibrant generation to lead the future (UNESCO 2016). Educating girls frees them to raise their aspirations and increase the potential to achieve them. It allows them to access better health, higher earnings and more control over their lives, which subsequently leads to positive impacts on their families, communities and societies (World Bank 2017b). More and better education for girls and boys contributes to economic growth, labor productivity and social mobility.

The benefits of girls’ education are particularly enormous and extend across generations. Extensive literature has demonstrated that better educated women tend to have lower fertility rates and lower child mortality rates. Data from 175 countries between 1970 and 2009 show that 51 percent of the reduction in deaths in children younger than 5 years in the past 40 years could be attributed to the increased education attainment in women of reproductive age (Gakidou et al. 2010). In addition, mothers’ education is strongly associated with children’s education attainments and achievements (Ermisch and Francesconi 2001).

Despite all the demonstrated benefits, girls face a lot of discrimination that boys do not, motivating a further emphasis on girls’ education. For example, a number of studies have demonstrated that girls’ schooling is more sensitive to both direct costs and opportunity costs (King and Winthrop 2015). When girls become teenagers, early marriage and teenage pregnancy may prevent them pursuing further schooling. Social norms in certain cultural contexts may restrict them from accessing economic opportunities as well. For example, 90 percent of women in Uttar Pradesh, India felt they needed to acquire their husband’s permission to work (World Bank 2017b). Even when women participate in the labor market, they are less likely to find a job. In the Middle East and North Africa, the female unemployment rate is twice that of their male counterparts (ILO 2017). Educating girls not only increases their control over their own lives but also empowers them to fight for gender equality within the places they live in.

Globally, the gender gap in primary and secondary schooling has significantly narrowed over the last two decades. However, global averages may mask the large variation in gender inequality at the regional and national levels. In many countries, girls are still more likely to be excluded from education, while boys have a greater chance of eventually getting into school (UNESCO 2016). Girls also complete less schooling than boys. Gender inequality in primary completion remains in nearly half of all 173 countries with available data (figure 1, upper panel) and is more prominent in poor countries: 21 out of the 25 countries with a gap larger than 5% are low- and lower middle-income countries.

Even for those girls who are already in school, many may not be able to acquire the foundational cognitive skills to help them thrive in their future lives. Data from 51 low- and middle-income countries suggest that 40 percent of women would be illiterate even if they finished primary school (Sandefur, Oye, and Pritchett 2016). The persistence in attainment and achievement gaps in education translates into lower human endowments for girls and subsequently larger asset
gaps for women relative to men. The long-run solution is to make sure that girls get into school and get the most out of their schooling.

At the same time, evidence on what works to improve the quality of education has been increasing at an unprecedented rate (Figure 2). In recent years, hundreds of impact evaluations in low- and middle-income countries have demonstrated the effectiveness – or lack thereof – of a range of interventions at improving education outcomes, for girls and boys (Evans and Popova 2016; J-PAL 2017). However, reviews that synthesize evidence tend to miss these even though they may boost girls’ education (along with boys’ education), either the separate effects for girls go unreported or they are not significantly different than those for boys. Many studies that examine the most effective ways to help girls succeed in school and subsequently enter the work force well prepared focus on interventions that target girls, for example, building girls’ latrines at schools and providing scholarships for girls (Filmer and Schady 2008; Garn et al. 2013). In fact, it may be that the interventions that are most effective for girls are also highly effective for boys. By focusing on interventions that specifically target girls, researchers and policy makers may be missing the evidence on what actually improve girls’ access and learning the most.

In this paper, we report the results of an innovative systematic review identifying the interventions most effective to improve girls’ access to education and learning outcomes within a more comprehensive evidence base that includes both girl-targeted and general education interventions. We asked three research questions: (1) Are girl-targeted interventions the most effective? (2) For non-targeted interventions, do impacts on girls tend to be larger? and (3) what are the most effective interventions for girls?

To answer these questions, we collected the primary data and test for the net effect of both general and girl-targeted interventions on girls. We went on standardizing the effects of different programs to make sure they are comparable across studies. For those studies that do not report gender differentiated impacts, we contacted the authors asking them either to run the additional analysis for us or share the data with us. Using a recent education database, we identified 105 general education studies that reported program effects for girls and 24 girl-targeted studies. With the support from authors of studies that originally did not report gender-differentiated impacts, we were able to obtain additional results of 51 studies. In total, we synthesized the effects for girls of 180 studies.

We found that girl-targeted interventions are more effective to improve girls’ access to education as a whole; and for learning outcomes, general inventions are more effective for girls and at the same time offer a larger menu of policy options. If you want to help girls learn, make schools better for all children.

2. Literature review: The missing evidence

Various researchers have undertaken systematic reviews of evidence on what works to improve girls’ education. For example, building on Herz and Sperling (2004), Sperling and Winthrop (2015) summarize eight recent reviews of education impact evaluations and provide a catalogue of information about practices and programs that improve girls’ access and learning. Unterhalter et al. (2014) conduct a systematic review of education programs that target girls. Tembon and Fort (2008) document effective polices in six developing countries to promote gender equality. J-PAL
(2017) conducts a quantitative review of randomized evaluations in 28 developing countries that tested programs designed to increase school enrollment and attendance, and compared gender differentiated effects of all reviewed studies. A number of systematic reviews focus on adolescent girl programs that may also have an impact on education outcomes (Haberland, McCarthy, and Brady 2018; Botea et al. 2017).

Traditional reviews of the most effective ways to boost girls’ education often highlight interventions that target girls. The most common approach is to use terms such as “girls” and “gender” in their search for relevant studies and identify lessons learned, such as Unterhalter et al. (2014) and Haberland, McCarthy, and Brady (2018). It is true that these girl-targeted interventions – such as offering cash transfer to girls, reducing school costs for girls, building girls’ latrines at school, reducing travel distance to school for girls and provide girls with additional skills training – are expected to be specifically effective for girls. However, this kind of review might be missing a large swath of general education interventions that are effective for both girls and boys, and are ostensibly gender neutral.

Consider two studies: Study 1 examines a girl-targeted intervention, which is effective for girls but not for boys – for example, providing inputs to help girls avoid missing school due to menstruation. Study 2 examines a general education intervention, which is much more effective for girls than intervention 1, but the effects are roughly equal for boys and girls – for example, training teachers in an innovative way. In this case the most effective intervention for improving girls’ education is the general education intervention, as Figure 3A illustrates. The fact that it also benefits boys does not take away from the fact that it is the most effective intervention for girls. However, Study 2 probably would not emphasize girls’ education or gender effects, as there are no differential effects for girls. In fact, the study may not even report gender impacts separately, since the effects are not different. So a traditional review of the best investments for girls with keywords like “girls” and “gender” could miss the best investments for girls.

Now consider two concrete interventions that demonstrate the above hypothesis (Figure 3B). A merit scholarship program for girls in Kenya increased girls’ learning by 0.19 standard deviations and had small, positive spillovers for boys (Kremer, Miguel, and Thornton 2009). Another general education program in Bangladesh involving regular parent-teacher meetings in school improved learning by 0.25 standard deviations. The impact is not significantly different between girls and boys (Islam 2016). The general intervention (parent-teacher meetings) is more effective at increasing girls’ learning than the girl-targeted merit scholarship program. But it would likely be omitted in any review of most effective interventions for girls because the study does not specifically target girls, and since the differences by gender are small, they do not come up in the title, the abstract, or the introduction of the paper. Even if the impact were significantly different between girls and boys, the apparent gender neutrality may also reduce its chance of being considered as the best investment for girls. Nevertheless, Glick (2008) has demonstrated that girls are more responsive than boys to gender neutral changes in education interventions. Therefore, general education interventions hold promise for girls’ education but have not been thoroughly reviewed.
In addition, all the systematic reviews except J-PAL (2017) examine the evidence qualitatively, usually discussing study by study, and then infer conclusion. To our best knowledge, there have been no systematic reviews that quantitatively examine the effectiveness of interventions to improve girls’ education. This systematic review complements the existing literature in two ways: first, it incorporates a wide range of general education interventions that have not been included in previous studies of what works to improve girls’ learning and access to education; second, it converts the point estimates of all included studies to effect sizes which are more comparable across studies.

3. Method

In this study, we gathered a large collection of studies that include education outcomes on either access and learning. For each of the studies, we identified whether or not they separately report impacts for boys and girls. For studies that separately report impacts for boys and girls, we extract those data, standardize the estimates, and use them to compare the impacts for boys versus girls and across programs for girls. For studies that do not separately report, we contacted the authors and asked them either to share the data or to provide the separate estimates themselves. In this section, we report on each step in detail.

**Literature search**

We began with a comprehensive database of education impact evaluations compiled for Evans and Popova (2016) and subsequently updated. The database consists of 495 studies that were cited in 10 recent systematic reviews of evidence on what works to improve learning and access in developing countries. All the reviews were published or made publicly available between 2013 and 2015 and the studies included were conducted between 1980 and 2015. Another systematic review of interventions with a special focus on access outcomes came out in 2017 (J-PAL 2017); we also tracked its references, which added four studies to the database.

To increase the coverage of studies that were published or posted online after 2015, we conducted an additional literature search between October 2017 and January 2018. We searched Google Scholar and the websites of major institutions that conduct research related to low- and middle-income countries for working papers that were published between 2015 and 2017 containing the keywords “evidence”, “education”, “access”, “learning”, “enrollment”, “dropout”, “attendance”, or “score”. We applied the same search terms to a number of economics and education journals, listed in Annex I. These two additional searches yielded another 19 new studies. In total, we reviewed 518 papers.

**Inclusion criteria**

2 The 10 reviews are: Conn (2014), Glewwe et al. (2014), Kremer et al. (2014), Krishnaratne et al. (2013), McEwan (2015), Murnane and Ganimian (2014), Asim et al. (2015), Masino and Niño-Zarazúa (2015), Glewwe and Muralidharan (2015), and Snisstveit et al. (2015). Conn, Glewwe et al, McEwan, and Masino and Niño-Zarazúa only include studies with learning outcomes. The other reviews include studies with learning outcomes and studies with access outcomes.
We included studies of education interventions (such as teacher professional development and providing textbooks), and health interventions (such as providing deworming drugs and micronutrients). We only included studies that took place in preprimary, primary, and secondary schools in low- or middle-income countries, according to the World Bank definition (World Bank 2017a). To be included studies had to be published – either as a working paper or a journal article – between 1980 and 2017 and had to report at least one of the following education outcomes: access outcomes (enrollment, dropout, or attendance) or learning outcomes (composite test score or any subject score). Skill development programs for adolescents outside of a school setting were not included.

We only included studies that used experimental or quasi-experimental design in treatment assignment. To be included, studies needed to have a valid counterfactual – in other words, what would have happened in the absence of the program? The ways that studies could construct such a counterfactual included random assignment, difference-in-differences, regression discontinuities, instrumental variables, and propensity score matching. Studies that adopted random or fixed effects for individuals in pre-post comparisons were not included.

At the same time, we restrict our analysis to studies where girls were the primary population or at least part of the intervention group. For example, Barham, Macours, and Maluccio (2013) only analyzed the program effects for boys, therefore was not included in this review.

Finally, only studies that reported sufficient data to calculate the intervention’s effect sizes for girls were included. For a few studies and/or outcomes, we could not covert the mean difference into effect size due to the lack of data on the standard error of the outcome variables (for example, Barr et al. 2012; Barrera-Osorio et al. 2011; Dumitrescu et al. 2011; Evans and Ngatia 2017; Levy and Ohls 2010; Levy et al. 2009; Lockheed et al. 2010; Loyalka et al. 2013; Malamud and Pop-Eleches 2011; Maluccio and Flores 2004; Parajuli et al. 2012; Emilie et al. 2016; Hoop De et al. 2014).

**Data Collection**

Upon reviewing the 518 identified studies, 313 studies met all the inclusion criteria. We further divided these studies into two groups: girl-targeted interventions and general interventions. Girl-targeted interventions include any intervention that is explicitly designed to boost education outcomes for girls. For example, this includes programs that provide girls with cash or in-kind transfers, reduce tuition or other school costs for girls, offer (merit) scholarship to girls, build latrines for girls in schools, reduce travel distance to schools for girls by building village schools or providing transportation, providing female teachers, or implementing girl empowerment curricula in schools. In general, if the program either specifically targets girls for benefits or explicitly states its objective as improving girls’ educational outcomes, we count it as “girl-
targeted.” In our sample, we identified 23 studies designed to increase access or learning specifically for girls. The other 290 studies were general interventions. 3

General interventions refer to programs that are gender neutral in their design. Examples include programs that offer computer-assisted learning for all students, provide school meals for all students, and distribute free school uniforms or textbooks to all students.

To collect the impacts of interventions on girls, for the 23 girl-targeted interventions, we used the results on girls directly reported in the studies. For general interventions, the average effect reported in the study covers an average across boys and girls, so we verified which studies report effects separately. 105 studies report heterogeneous intervention impacts by gender, and we incorporated those results in our review. However, that left 205 studies that do not report gender differentiated impacts in their original papers. In order to enlarge the sample of our review, we contacted the authors of these studies between January 2018 and July 2018, requesting that they either provide additional estimates of intervention effects by gender or share the data of their studies with us to perform the analysis on their behalf. Authors were given at least three months to reply to us with either new estimates or their data if they were interested.

Of the 205 studies, we received replies from the authors of 103 studies. Among them, the authors of 32 studies indicated that the data were no longer available and/or not eligible to perform the analysis by gender. Another 71 sets of authors expressed their willingness to run the additional analysis or share their data with us. By the end of July 2018, we were able to obtain new estimates of effects by gender of 51 studies.

Figure 4 demonstrates our review process. Combining girl-targeted interventions, general interventions that report impacts on girls and the new estimates we collected from authors, the final sample of this review consists of 179 studies with 274 total interventions.

Coding of effect sizes

In this paper, our unit of analysis is the estimated impact of an intervention, where a group that received an intervention is compared to another group that did not receive the intervention. For studies with multiple treatment arms, we coded the impact of each treatment arm separately (as its own intervention) and recorded the education outcomes corresponding to that intervention. For example, Berlinski et al. (2016) tested the effects of four interventions or treatment arms: (1) an active learning approach to the teaching of math, (2) an active learning approach plus an interactive white board, (3) an active learning approach plus a computer lab, and (4) an active learning approach plus one computer per student. We coded these four experiments as four separate interventions.

3 There are two general intervention studies that contain a girl-targeted intervention arm, but for the purpose of counting, but because the bulk of the benefits do not target girls, we include them in the general intervention group.
Since studies in our sample collected different outcomes using different measures, in order to compare the effectiveness of the interventions on the same scale, individual point estimates need to be standardized. In this paper, we used Cohen’s $d$ to standardize effect sizes, following McEwan (2015) and Conn (2017). Basically, Cohen’s $d$ can be estimated using the raw mean difference between a treatment group ($\bar{Y}_T$) and a control group ($\bar{Y}_C$) as well as the pooled standard deviation for the treatment and control groups combined ($S_{pooled}$) (see Equation 1).\(^4\)

\[
d = \frac{\bar{Y}_T - \bar{Y}_C}{S_{pooled}}
\]

(Equation 1)

Where the pool standard deviation was not directly reported in the studies, we calculated it using Equation 2 from Borenstein et al. (2009):

\[
S_{pooled} = \sqrt{\frac{(n_T-1)S_T^2 + (n_C-1)S_C^2}{n_T+n_C-2}}
\]

(Equation 2)

where $n_T$ and $n_C$ are the sample sizes in the treatment and control groups, and $S_T$ and $S_C$ are the standard deviations in each group.

**Qualitative variables**

We collected a set of useful variables with the aim to better analyze the most effective interventions for girls. The variables included country, region, implementation agency, location (i.e., rural, urban), intervention level (e.g., school, household, village, individual), duration of intervention (i.e., single contact or repeated contact), number of intervention components (i.e., single or multiple), the level of education at which the intervention was implemented, student age, major program components (e.g., reducing school costs, health intervention, additional teaching and learning materials, school grants), the presence of component identified by program implementers as “girl friendly,” cost data (if any), quality of outcomes (e.g., admin data, self-reported data, national tests, international tests, program designed tests). For each study, we also coded its publication type and evaluation method.

4. Results

In this section, we answer three research questions using the data collected from 180 studies. Among those studies, 85 studies measured access outcomes such as enrollment, attendance, and/or dropout; 117 studies measured learning outcomes like composite test score, math score, and/or language score. Table 1 demonstrates the descriptive statistics of our sample.

**Are girl-targeted interventions the most effective for girls?**

In terms of increasing girls’ participation in school, girl-targeted interventions are more effective than general interventions. Figure 5 demonstrates the distributions of effect sizes for

\(^4\) In this paper, wherever applicable, we collected the mean difference with controls for observable variables.
girls from both general interventions and girl-targeted interventions. On average (at the 50th percentile), the effect sizes of these two categories are very similar. The middle interventions increased girls’ enrollment or attendance by 0.05 – 0.07 standard deviations. The effect sizes of less effective interventions at the 10th and 25th percentiles are also close to each other. However, the top girl-targeted interventions at the 90th percentiles have larger effect sizes than those of general interventions. Especially at the 90th percentile, the effect size of girl-targeted interventions is 0.01 standard deviations larger than that of general interventions. Therefore, looking at the distribution, the best of girl-targeted interventions indeed are more effective to increase access to education for girls.

Notwithstanding, if we look at the individual studies of general interventions, there are programs that generated similar or larger effect sizes as those of girl-targeted interventions. The effect size of the most effective general intervention – 1.66 standard deviations – is larger than that (1.54 standard deviation) of the most effective girl-targeted intervention. At the same time, it’s worth mentioning that there are way more general interventions than specifically girl-targeted interventions. As we have seen in table 1, the number of general interventions is more than 3 times that of girl-targeted interventions. This means that in each of the effect size bins, general interventions provide a larger menu for options (Figure 6). Even for the best interventions, although on average the top girl-targeted interventions deliver larger effects, the number of interventions with big effect sizes – larger than 0.4 standard deviations – is almost the same as general interventions. Therefore, general interventions constitute an important source of ways to improve girls’ access to education.

For learning, the finding is less straightforward. Based on the distributions of effect sizes, both types of interventions have very similar impacts on girls. The middle programs at the 50th percentile are about to increase learning by 0.12 standard deviations and the top programs have quite sizable effects of about 0.45 standard deviations.

However, as discussed earlier in this section, the difference in the number of general interventions and girl-targeted interventions is significant. This is even more the case in learning outcomes: there are 16 general learning interventions compared to only 11 girl-targeted learning interventions. Due to the small number of girl-targeted interventions, the distribution of effect sizes might not demonstrate the underlying storyline accurately. In fact, we find that the large effect size of the top girl-targeted intervention (at the 90th percentile) is purely driven by two school construction interventions (Burde and Linden 2013; Kazianga et al. 2013). When schools were built in villages in rural Afghanistan and rural Burkina Faso, it dramatically improved the learning outcomes for girls. But if taking out these two interventions, the effect size of girl-targeted interventions at the 90th percentile drops to 0.2 standard deviations. Taking this into account, it is fair to say that general interventions are more effective to improve learning outcomes for girls. In addition, as access outcomes, because of the large body of studies, general interventions offer more options.

For general interventions, do impacts on girls tend to be larger?
Previous research shows that the demand for girls’ schooling tend be more responsive than boys’ to gender neutral education policies (Glick 2008). However, we do not detect much difference between girls and boys in terms of responding to interventions to increase access. The middle and the top programs at the 90\textsuperscript{th} percentile have almost identical impacts for girls and boys, which are about 0.05 and 0.25 standard deviations. At the same time, less effective programs at the 10\textsuperscript{th} percentile have similar effects for boys than girls.

Once in school, the impacts of general interventions on learning outcomes are also quite similar for girls and boys. Middle interventions increased learning for girls and boys by 0.1 standard deviations and the 75\textsuperscript{th} percentile interventions, by about 0.25 standard deviations (Figure 10). Effect for boys of top programs is 0.49 standard deviations, slightly larger than those for girls which is 0.46 standard deviations. These are sizeable effects which are equivalent of more than 2.5 years additional schooling in business-as-usual setting in developing countries (Evans and Yuan 2018).

**What are the most effective interventions for girls?**

To summarize the most effective interventions for girls, in this section we present the top 10 access and learning interventions with the largest effect sizes and try to understand their attributes. By contrast, we also compare them with the least effective 10 interventions with access and learning outcomes.

**Access**

Table 2 lists 10 studies of the most effective interventions to improve access to education for girls. These interventions greatly improved girls’ participation in school with an average effect size of 0.74 standard deviations. Four of them are girl-targeted interventions including cash transfer to dropped-out girls conditional on school attendance in Malawi (Baird et al. 2016), improving school water and sanitation system in Kenya (Garn et al. 2013), building village schools for girls in Afghanistan (Burde and Linden 2013) and providing private subsidies for girls in Pakistan (Kim, Alderman, and Orazem 1999). Five of the general interventions are related to offering cash for education in different countries (Eyal, Woolard, and Burns 2014; Maluccio, Murphy, and Regalia 2010; Edmonds and Shrestha 2014; Benhassine et al. 2015; Duflo, Dupas, and Kremer 2017) and the other one is focused on malaria prevention in the Gambia (Jukes et al. 2006).

These top interventions demonstrate that reducing school cost is the most effective way to bring girls into school, especially when enrollment and/or attendance are explicitly required for cash transfer. In addition, reducing indirect cost - commute distance to school - for girls by building village schools has also been effective to get girls into school. Nevertheless, pure cash transfer without schooling binding is less effective (Table 3) (Baird, McIntosh, and Özler 2011). Improving health conditions through either better sanitation facility or controlling malaria tend to attract more girls to school as well.
Nevertheless, there are concerns about the effectiveness of conditional cash transfer programs if only considering the most effective interventions. One of them is that the popularity of conditional cash transfers has led to an emergence of impact evaluations in this field, which might make this group of interventions overweighted in evidence base. It’s possible that conditional cash transfer interventions are the most effective and the least effective. To check the robustness of our finding, we summarize the bottom 10 interventions to increase access for girls in Table 3. There are conditional cash/in-kind transfers that did not work to bring girls into school, such as those in Burkina Faso (Kazianga et al. 2013), the Philippines (Chaudhury, Friedman, and Onishi 2013) and Uruguay (Amarante, Ferrando, and Vigorito 2013), but not all of them. In fact, compared to the most effective ones, there is more variation in the bottom programs, ranging from providing school meals to saving on education. Interestingly, we see that from the same study (Garn et al. 2013), while promoting hygiene, improving water treatment, sanitation plus water supply in Kenyan primary schools is one of the best ways to increase girls’ enrollment, promoting hygiene and improving water treatment only even reduced enrollment for girls. It is likely that girls are more responsive to sanitation conditions, however, distributing sanitary products in Nepal did not increase girls’ attendance to school.

**Learning**

For learning, the average effect size of the top interventions for girls is as large as 0.96 standard deviation. Compared to access interventions, there is more variation in learning interventions. First, only two out of 10 studies included girl-targeted interventions which is in line with the finding in the previous section that general interventions are more effective to improve girls’ learning compared to girl-targeted interventions (Table 4). One of the two girl-targeted interventions is a public private partnership initiative in schools in Pakistan providing gender differentiated subsidy that increased girls’ test score by 0.77 standard deviations. The other intervention arm in the same initiative that provided gender neutral subsidy also yielded sizeable effects but smaller than the gender-differentiated one (Barrera-Osorio et al. 2017). The other girl-targeted intervention is the Afghan village school program for girls that delivered significant impacts on girls’ access and learning outcomes (Burd and Linden 2013). Similarly, a community school program in Honduras greatly improved girls’ math score (Di Gropello and Marshall 2011).

One set of the most effective general interventions for girls in the table is structured pedagogy in early grades, which is providing teachers with clear guidance on teaching or even scripted lesson plans. These interventions have been proved to highly effective in several Sub-Saharan African countries including South Africa, Liberia and Kenya (Piper 2009; Piper, Zuilkowski, and Ong’ele 2016; Piper and Mugenda 2014; Piper and Medina 2010b). Another category of interventions that work well for girls (and boys) is teacher to the right level, either through diagnostic feedback or software as accounted in Banerjee et al. (2016) and Imbrogno (2014).

On the other hand, the least effective programs for girls’ learning are all general interventions (Table 5). Various programs had negative impacts on learning for girls compared to “business-as-usual”, but often times, the same programs did not work for boys either. For example,
technology interventions, either substituting teachers with computer or providing students with laptops did not help improve learning (Linden 2008; Sharma 2014). Although there are teacher professional development programs that work to improve student learning (Popova et al. 2018), our findings demonstrate that introducing new pedagogical methods through a short teacher training program are less likely to be effective to improve girls’ learning, and this is true no matter which education level the intervention targets (Berlinski and Busso 2017; Yoshikawa et al. 2015). In addition, school accountability interventions such as distributing school report cards to students and parents were not very effective for girls, neither in Sri Lanka nor in Liberia (Aturupane et al. 2014; Piper and Medina 2010a).

5. Discussion

Previous reviews of what works to improve girls’ education focus on girl-targeted interventions, which omits key evidence of the impact of general education interventions on girls. This review innovatively brings in a large evidence base of general interventions that report effects for girls. Based on 179 studies from 54 countries, this review finds that girls’ access to school is more responsive to changes in costs, distance and health conditions; while girls’ learning is more likely to be improved by structured pedagogy and teaching at the right level interventions. Although girl-targeted interventions are more effective to improve girls’ access to education as a whole, because of the large number of general interventions, they actually provide the same number of interventions with similar effects. For learning outcomes, general inventions are more effective for girls and at the same time offer a larger menu of policy options.

While we compared effect sizes across interventions in this review, it would be much better to bring in cost data because the most effective programs may not be the most cost-effective and not easily to scale up. However, despite a strong demand for cost data, most studies do not report them. McEwan (2015) in his review stated that 56% of studies reported no cost detail and most of the rest reported minimal info. We tried to collect data from reviewed studies as well, however, we found it hard to gather such information. In addition, even when cost data are reported, they are often not comparable due to different accounting methods used. Taking an early childhood development program in rural Mexico as an example, the cost per child estimated by World Bank researchers was $76 (Cardenas et al. 2015) but when evaluated by another group of researchers at Brookings, the cost per child almost doubled at $174-$202 (Gustafsson-Wright, Boggild-Jones, and Gardiner 2017). Ideally, we would love to collect cost data following the same set of guidelines as Dhaliwal et al. (2013) did in their cost-effective analysis.

We also tried to gather a number of program attributes with the aim to provide more information on the most effective programs. For example, the average program size of the most effective access interventions is 262 and for learning interventions, is 556. Except cash transfer programs, all others are pilot programs, and it’s the same for those less effective programs. Therefore, we cannot from this sample if pilot programs are more effective than the one that have been scaled up. Another attribute that we looked at was the level of education that the top programs targeted. In terms of access interventions, 7 out the top 10 interventions
targeted school-aged children in general, often between age 6 up to age 16, trying to get out-of-school children into school. On the other hand, 9 out the 10 learning interventions focused on the primary level and half of them were designed to improve learning in early grades, i.e. grades 1-3. And besides girl-targeted programs, only 1 general intervention out of the top 20 considered a girl-friendly component in their program design (Barrera-Osorio et al. 2017).

While this review focuses on girls’ education, we do not ignore education for boys. In fact, the global learning crisis impoverish both girls and boys in schools (WDR 2018). Our findings demonstrate that gender neutral interventions indeed hold great promise for girls’ learning. Considering the limited resources that most education systems in developing countries possess, to help girls learn, the most practical approach would be to make schools better for all children.

In addition, we should also remember that acquiring schooling and learning are not the finishing line for girls’ education. Rather, we hope girls can empower themselves through education and eventually achieve their aspirations in their lives. But currently, very few interventions have included this aspect in the program design and among all our studies. In order to generate long-lasting impacts on girls’ lives, education system has an important role to play and needs to provide additional support for girls.

**Figures and tables**

Figure 1: Gender gap in primary and lower secondary school completion
Figure 2: Evidence on what works in education has mushroomed

Access

Learning

Source: Evans and Popova (2016)

Figure 3: Interventions not targeted specifically to girls may in fact deliver bigger gains to girls

A thought experiment
Two actual examples

<table>
<thead>
<tr>
<th>Increased learning (standard deviations)</th>
<th>Girls</th>
<th>Boys</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girl-targeted intervention: Incentives to Learn, Kenya (Kremer, Miguel and Thornton 2009)</td>
<td>0.3</td>
<td>0.05</td>
<td>0.25</td>
<td>0.05</td>
</tr>
<tr>
<td>General education intervention: Parent-teacher meetings and student outcomes, Bangladesh (Islam, 2016)</td>
<td>0.25</td>
<td>0.05</td>
<td>0.3</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Impact Evaluations of Educational Interventions from a recent database (Evans and Popova 2016 and 3ie 2017 and miscellany)

[318 studies]

Without gender-differentiated impacts
[205 studies]

Authors provide new estimates OR We run the regression using shared data
[51 studies]

General interventions with gender-differentiated impacts
[105 studies]

Girl-targeted interventions
[23 studies]

Extract point estimates and calculate effect sizes

Figure 5: Effect sizes of access outcomes for girls

<table>
<thead>
<tr>
<th></th>
<th>10th</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>90th</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.07</td>
<td>0.15</td>
<td>0.27</td>
<td>[-0.2, 1.66]</td>
</tr>
<tr>
<td>Girl-targeted</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.05</td>
<td>0.16</td>
<td>0.36</td>
<td>[-0.18, 1.54]</td>
</tr>
</tbody>
</table>
Figure 6: Number of **access** outcomes by effect size

![Graph showing number of studies by effect size for General interventions and Girl-targeted interventions.]

Figure 7: Effect sizes of **learning** outcomes for girls

<table>
<thead>
<tr>
<th></th>
<th>10th</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>90th</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>-0.06</td>
<td>0.02</td>
<td>0.12</td>
<td>0.27</td>
<td>0.46</td>
<td>[-0.61, 2.56]</td>
</tr>
<tr>
<td>Girl-targeted</td>
<td>-0.03</td>
<td>0.02</td>
<td>0.13</td>
<td>0.29</td>
<td>0.45</td>
<td>[-0.45, 0.66]</td>
</tr>
</tbody>
</table>

Figure 8: Number of **learning** outcomes by effect size

![Graph showing number of studies by effect size for General interventions and Girl-targeted interventions.]


Figure 9: Effect sizes of access outcomes for girls and boys (general interventions only)

<table>
<thead>
<tr>
<th></th>
<th>10th</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>90th</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girls</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.07</td>
<td>0.15</td>
<td>0.27</td>
<td>[-0.20, 1.66]</td>
</tr>
<tr>
<td>Boys</td>
<td>-0.04</td>
<td>0.02</td>
<td>0.05</td>
<td>0.13</td>
<td>0.25</td>
<td>[-0.14, 1.11]</td>
</tr>
</tbody>
</table>

Figure 10: Effect sizes of learning outcomes for girls and boys (general interventions only)

<table>
<thead>
<tr>
<th></th>
<th>10th</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>90th</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girls</td>
<td>-0.06</td>
<td>0.02</td>
<td>0.12</td>
<td>0.27</td>
<td>0.46</td>
<td>[-0.61, 2.56]</td>
</tr>
<tr>
<td>Boys</td>
<td>-0.05</td>
<td>0.01</td>
<td>0.10</td>
<td>0.25</td>
<td>0.49</td>
<td>[-0.59, 1.98]</td>
</tr>
</tbody>
</table>
Table 1: Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>Number of studies</th>
<th>Number of interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>General studies</td>
<td>156</td>
<td>236</td>
</tr>
<tr>
<td>General studies – Access</td>
<td>65</td>
<td>88</td>
</tr>
<tr>
<td>General studies – Learning</td>
<td>106</td>
<td>178</td>
</tr>
<tr>
<td>Girl-targeted studies</td>
<td>23</td>
<td>38</td>
</tr>
<tr>
<td>Girl-targeted studies – Access</td>
<td>20</td>
<td>29</td>
</tr>
<tr>
<td>Girl-targeted studies – Learning</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>179</td>
<td>274</td>
</tr>
</tbody>
</table>

Table 2: The top 10 MOST effective interventions to improve access to education for girls

<table>
<thead>
<tr>
<th>Year</th>
<th>Year Type of publication</th>
<th>Program description</th>
<th>Country</th>
<th>Region</th>
<th>Evaluation design</th>
<th>Outcome</th>
<th>Effect size (SD)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2014 Working paper</td>
<td>Conditional cash transfer</td>
<td>South Africa</td>
<td>SSA</td>
<td>DID</td>
<td>Enrollment</td>
<td>1.657</td>
<td>Eyal &amp; Woolard 2014</td>
</tr>
<tr>
<td>3</td>
<td>2010 Journal</td>
<td>Conditional cash transfer</td>
<td>Nicaragua</td>
<td>LAC</td>
<td>DID</td>
<td>Enrollment, Yr 1</td>
<td>0.883</td>
<td>Maluccio et al. 2010</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Enrollment, Yr 2</td>
<td>0.617</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2013 Journal</td>
<td>Hygiene promotion + water treatment + sanitation + water supply</td>
<td>Kenya</td>
<td>SSA</td>
<td>RCT</td>
<td>Enrollment</td>
<td>0.634</td>
<td>Garn et al. 2013</td>
</tr>
<tr>
<td>5</td>
<td>2017 Working paper</td>
<td>Free secondary education</td>
<td>Ghana</td>
<td>SSA</td>
<td>RCT</td>
<td>Enrollment</td>
<td>0.634</td>
<td>Duflo et al. 2017</td>
</tr>
<tr>
<td>6</td>
<td>2006 Journal</td>
<td>Malaria prevention</td>
<td>Gambia</td>
<td>SSA</td>
<td>RCT</td>
<td>Enrollment – cohort w/o contamination</td>
<td>0.555</td>
<td>Jukes et al. 2006</td>
</tr>
<tr>
<td>Year</td>
<td>Publication Type</td>
<td>Program Description</td>
<td>Country</td>
<td>Region</td>
<td>Evaluation Design</td>
<td>Outcome</td>
<td>Effect Size (SD)</td>
<td>Reference</td>
</tr>
<tr>
<td>------</td>
<td>-----------------</td>
<td>---------------------</td>
<td>---------</td>
<td>--------</td>
<td>-------------------</td>
<td>---------</td>
<td>-----------------</td>
<td>-----------</td>
</tr>
<tr>
<td>7</td>
<td>Journal</td>
<td>Conditional cash transfer: school stipend</td>
<td>Nepal</td>
<td>SA</td>
<td>RCT</td>
<td>Enrollment – cohort w/ minimal contamination</td>
<td>0.457</td>
<td>Edmonds and Shrestha 2014</td>
</tr>
<tr>
<td>8</td>
<td>Journal</td>
<td>Labelled cash transfer for education</td>
<td>Morocco</td>
<td>MENA</td>
<td>RCT</td>
<td>Dropped out by the end of year 2</td>
<td>0.486 (abs. value)</td>
<td>Benhassine et al. 2015</td>
</tr>
<tr>
<td>9</td>
<td>Journal</td>
<td>Village-based school</td>
<td>Afghanistan</td>
<td>SA</td>
<td>RCT</td>
<td>Enrollment</td>
<td>0.478</td>
<td>Burde &amp; Linden 2013</td>
</tr>
<tr>
<td>10</td>
<td>Journal</td>
<td>Private school subsidies for girls</td>
<td>Pakistan</td>
<td>SA</td>
<td>RCT</td>
<td>Enrollment</td>
<td>0.441</td>
<td>Kim et al. 1999</td>
</tr>
</tbody>
</table>

Note: Interventions in red are girl-targeted while those in black are general interventions.

Table 3: The top 10 LEAST effective interventions to improve access to education for girls

<table>
<thead>
<tr>
<th>Year</th>
<th>Publication Type</th>
<th>Program Description</th>
<th>Country</th>
<th>Region</th>
<th>Evaluation Design</th>
<th>Outcome</th>
<th>Effect Size (SD)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Journal</td>
<td>School canteen</td>
<td>Burkina Faso</td>
<td>SSA</td>
<td>RCT</td>
<td>Absenteeism</td>
<td>-0.200</td>
<td>Kazianga et al. 2012</td>
</tr>
<tr>
<td>2</td>
<td>Journal</td>
<td>Conditional take-home rations for girls</td>
<td>Burkina Faso</td>
<td>SSA</td>
<td>RCT</td>
<td>Absenteeism</td>
<td>-0.182</td>
<td>Baird et al. 2011</td>
</tr>
<tr>
<td>3</td>
<td>Journal</td>
<td>Unconditional cash transfer to girls</td>
<td>Malawi</td>
<td>SSA</td>
<td>RCT</td>
<td>Attendance, Yr2</td>
<td>-0.152</td>
<td>Yi et al. 2015</td>
</tr>
<tr>
<td>4</td>
<td>Journal</td>
<td>Early financial commitment</td>
<td>China</td>
<td>EAP</td>
<td>RCT</td>
<td>Dropout</td>
<td>-0.140*</td>
<td>Garn et al. 2013</td>
</tr>
<tr>
<td>5</td>
<td>Journal</td>
<td>Hygiene promotion + water treatment</td>
<td>Kenya</td>
<td>SSA</td>
<td>RCT</td>
<td>Enrollment</td>
<td>-0.138</td>
<td>Afridi 2011</td>
</tr>
<tr>
<td>6</td>
<td>Working paper</td>
<td>Education cash saving account with parent outreach</td>
<td>Uganda</td>
<td>SA</td>
<td>RCT</td>
<td>Enrollment Y1</td>
<td>-0.110</td>
<td>Karlan and Leiden 2014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Education cash saving account</td>
<td>Uganda</td>
<td>SA</td>
<td>RCT</td>
<td>Attendance overall</td>
<td>-0.107</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Enrollment Y2</td>
<td>-0.107</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Enrollment Y1</td>
<td>-0.049</td>
<td></td>
</tr>
</tbody>
</table>
Table 4: The top 10 MOST effective interventions to improve learning for girls

<table>
<thead>
<tr>
<th>Year</th>
<th>Type of publication</th>
<th>Program description</th>
<th>Country</th>
<th>Region</th>
<th>Evaluation design</th>
<th>Outcome</th>
<th>Effect size (SD)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2009</td>
<td>Evaluation report</td>
<td>South Africa</td>
<td>SSA</td>
<td>DID</td>
<td>Letter sounding fluency</td>
<td>2.563</td>
<td>Piper 2009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Word naming fluency</td>
<td>1.840</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Reading comprehension</td>
<td>1.757</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Oral reading fluency</td>
<td>1.658</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2016</td>
<td>Journal</td>
<td>Kenya</td>
<td>SSA</td>
<td>RCT</td>
<td>Reading comprehension, Lubukusu, class 1</td>
<td>1.36</td>
<td>Piper et al. 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mother tongue instruction</td>
<td></td>
<td></td>
<td></td>
<td>Reading comprehension, Kikamba, class 2</td>
<td>1.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>*TaRL 10-day Camp</td>
<td></td>
<td></td>
<td></td>
<td>Math</td>
<td>0.870</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TaRL 10-day Camp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TaRL 20-day Camp</td>
<td></td>
<td></td>
<td></td>
<td>Language</td>
<td>0.830</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TaRL 20-day Camp</td>
<td></td>
<td></td>
<td></td>
<td>Math</td>
<td>0.730</td>
<td></td>
</tr>
</tbody>
</table>

Note: Interventions in red are girl-targeted while those in black are general interventions.
*: Adjusted negative value for comparison.
<table>
<thead>
<tr>
<th>Year</th>
<th>Type of publication</th>
<th>Program description</th>
<th>Country</th>
<th>Region</th>
<th>Evaluation design</th>
<th>Outcome</th>
<th>Effect size (SD)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Working paper</td>
<td>Computer assisted learning in school</td>
<td>India</td>
<td>SA</td>
<td>RCT</td>
<td>Test score</td>
<td>-0.613</td>
<td>Linden 2008</td>
</tr>
<tr>
<td>2</td>
<td>Journal</td>
<td>School management</td>
<td>Madagascar</td>
<td>SSA</td>
<td>RCT</td>
<td>Test score (district level intervention)</td>
<td>-0.403</td>
<td>Glewwe and Maiga 2011</td>
</tr>
<tr>
<td>3</td>
<td>Journal</td>
<td>New curriculum + OLPC</td>
<td>Costa Rica</td>
<td>LAC</td>
<td>RCT</td>
<td>Math-geometry</td>
<td>-0.378</td>
<td>Berlinski and Busso 2017</td>
</tr>
</tbody>
</table>

Note: Interventions in red are girl-targeted while those in black are general interventions.

* TaRL: Teaching at the Right Level

Table 5: The top 10 LEAST effective interventions to improve learning for girls
<table>
<thead>
<tr>
<th>#</th>
<th>Year</th>
<th>Source Type</th>
<th>Description</th>
<th>Country</th>
<th>Cluster</th>
<th>Methodology</th>
<th>Subject</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>2015</td>
<td>Journal</td>
<td>Teacher training</td>
<td>Chile</td>
<td>LAC</td>
<td>RCT</td>
<td>vocabulary</td>
<td>-0.305</td>
</tr>
<tr>
<td>5</td>
<td>2014</td>
<td>Working paper</td>
<td>OLPC</td>
<td>Nepal</td>
<td>SA</td>
<td>DID</td>
<td>English</td>
<td>-0.244</td>
</tr>
<tr>
<td>6</td>
<td>2013</td>
<td>Working paper</td>
<td>Mobile school librarian</td>
<td>India</td>
<td>SA</td>
<td>RCT</td>
<td>Language</td>
<td>-0.232</td>
</tr>
<tr>
<td>7</td>
<td>2013</td>
<td>Journal</td>
<td>Preschool voucher</td>
<td>China</td>
<td>EAP</td>
<td>RCT</td>
<td>Test score - school readiness</td>
<td>-0.223</td>
</tr>
<tr>
<td>8</td>
<td>2013</td>
<td>Journal</td>
<td>School report card</td>
<td>Sri Lanka</td>
<td>SA</td>
<td>DID</td>
<td>Science 8th grade</td>
<td>-0.221</td>
</tr>
<tr>
<td>9</td>
<td>2016</td>
<td>Working paper</td>
<td>Attendance reward</td>
<td>India</td>
<td>SA</td>
<td>RCT</td>
<td>Test score</td>
<td>-0.207</td>
</tr>
<tr>
<td>10</td>
<td>2016</td>
<td>Working paper</td>
<td>Teacher training</td>
<td>Malawi</td>
<td>SSA</td>
<td>RCT</td>
<td>Early Grade Math, 36-month follow-up</td>
<td>-0.124</td>
</tr>
</tbody>
</table>
References


Berlinski, Samuel, Matias Busso, Taryn Dinkelman, and Claudia Martinez. 2016. 'Reducing parent-school information gaps and improving education outcomes: Evidence from high frequency text messaging in Chile', *Unpublished Manuscript*.


Evans, David K, and Fei Yuan. 2018. 'Economic Returns to Interventions that Increase Learning'.


Gustafsson-Wright, Emily, Izzy Boggild-Jones, and Sophie Gardiner. 2017. 'SECT: The standardized early childhood development costing tool'.


ILO. 2017. 'World Employment and Social Outlook: Trends for Women 2017'.

Imbrogno, Jason. 2014. 'Essays on the Economics of Education'.

Islam, Asad. 2016. 'Parent-teacher meetings and student outcomes'.


Piper, Benjamin, and Korda Medina. 2010a. 'Early Grade Reading Assessment (EGRA) Plus: Liberia. Program evaluation report'.


Piper, Benjamin, Stephanie S Zuilkowski, and Salome Ong’ele. 2016. 'Implementing Mother Tongue Instruction in the Real World: Results from a Medium-Scale Randomized Controlled Trial in Kenya', *Comparative Education Review*, 60: 776-807.

Popova, Anna, David K Evans, Mary E. Breeding, and Violeta Arancibia. 2018. 'Global Landscape of Teacher Professional Development Programs: The Gap between Evidence and Practice '.

Sandefur, Justin, Mari Oye, and Lant Pritchett. 2016. 'Girls' Schooling is Good, Girls' Schooling with Learning is Better', *Background paper for the International Commission on Financing Global Education Opportunity*.


UNESCO. 2016. 'Key Messages: Girls’ Progress=Goals Progress'.

———. 2016 'Leaving no one behind: How far on the way to universal primary and secondary education?', *Policy Paper 27/Fact Sheet 37*.


World Bank. 2017a. 'World Development Indicators '.


Yoshikawa, Hirokazu, Diana Leyva, Catherine E Snow, Ernesto Treviño, M Barata, Christina Weiland, Celia J Gomez, Lorenzo Moreno, Andrea Rolla, and Nikhit D’Sa. 2015. 'Experimental impacts of a teacher professional development program in Chile on preschool classroom quality and child outcomes', *Developmental psychology*, 51: 309.
Annex I: Journals and websites searched

Journals searched:
- American Economic Review,
- American Economics Journal: Applied Economics,
- American Economics Journal: Economic Policy,
- Comparative Education Review,
- Economic Development and Cultural Change,
- Journal of Development Economics,
- Journal of Development Effectiveness,
- International Journal of Educational Development,
- Journal of Human Resources,
- Quarterly Journal of Economics,
- The Economics of Education Review,
- World Bank Economic Review,
- Journal of Public Economics,
- Journal of Human Capital.

Websites searched:
- 3ie impact evaluation database,
- RISE Programme Conference 2017,
- CASE 2017 Conference,
- the Abdul Latif Jameel Poverty Action Lab at MIT;
- Innovations for Poverty Action;
- the Inter-American Development Bank;
- the National Bureau of Economic Research;
- RTI International;
- the Rural Education Action Program at Stanford University;
- the World Bank;
- the IZA Institute of Labor Economics.