USC Dornsife

Center for Economic and Social Research

Educational Impacts and Cost-Effectiveness of Conditional Cash Transfer Programs in Developing Countries: A Meta-analysis

Juan Esteban Saavedra
Sandra Garcia

CESR WORKING PAPER SERIES

Paper Number: 2013-007 **cesr.usc.edu**

Educational Impacts and Cost-Effectiveness of Conditional Cash Transfer Programs in Developing Countries: A Meta-analysis

Juan Esteban Saavedra*

Sandra García**

January 2013

Abstract

We meta-analyze enrollment, attendance and dropout impact and cost-effectiveness estimates from forty-two CCT program evaluations in fifteen developing countries. Average impacts and cost-effectiveness estimates for all outcomes in primary and secondary schooling are statistically different from zero, with considerable heterogeneity. CCT programs are, all else constant, most impactful and cost-effective for programs that, in addition to transfers to families, also provide supply-side complements—such as infrastructure or additional teachers. Impacts are also larger in programs with infrequent payments and more stringent schooling conditions, which aligns with previous single-program evidence. Impact and cost-effectiveness estimates from randomized research designs are smaller than those from observational studies.

---

* Corresponding author. RAND Corporation, 1776 Main Street, Santa Monica, CA, 90407. Email: saavedra@rand.org

** School of Government, Los Andes University, Carrera 1 # 19-27. Bogotá, Colombia. Email: sagarcia@uniandes.edu.co

We thank seminar participants at RAND Corporation and Columbia University for comments and suggestions. We thank Felipe Alvarez, Omar Herrera, Isabella Sinisterra and especially Tatiana Velasco for research assistance. Saavedra acknowledges financial support from RAND Corporation and Universidad de los Andes. García acknowledges financial support from Los Andes University. The research reported in the paper is not the result of a for-pay consulting relationship and our employers have no financial interest in the topic of the paper, which might constitute a conflict of interest. All errors are our own.

I. Introduction

Conditional cash transfer (CCT) programs have spread rapidly over the last decade in the developing world. CCT programs provide cash transfers to poor families that are contingent on children's educational and health investments, typically school attendance and regular medical checkups, with the goal of breaking the intergenerational cycle of poverty. As of 2010, all but two countries in Latin America and over 15 countries in Asia and Africa had a CCT program as part of their social protection systems. In Latin America alone, CCT programs benefit over one hundred and ten million people (The Economist, 2010).

In most of these countries, a rigorous impact evaluation – typically a treatment/control experimental or observational setup – has accompanied CCT program implementation. In fact, the positive results on schooling and health outcomes of early impact evaluations of pioneer programs such as *Oportunidades* in Mexico and *Bolsa Escola* in Brazil helped paved the way for the rapid expansion of these programs elsewhere.

Recent qualitative review studies of CCT evaluations (Independent Evaluation Group, 2011; Fiszbein et al., 2009; Hoddinott and Bassett, 2009; Rawlings and Rubio, 2005) conclude that, on the whole, these programs have positive effects on schooling (enrollment, attendance, dropout), health (vaccinations, medical check-up) and child-nutrition outcomes. These reviews also indicate that there is substantial variation in effect sizes between countries and among different population groups within countries (for example gender, age or urban vs. rural residence).

No study to date, however, integrates quantitatively and in a systematic manner the available evidence on the educational impacts and cost-effectiveness of CCT programs or attempts to statistically understand the factors and program characteristics that mediate heterogeneity in impact and cost-effectiveness estimates. The closest available studies in scope are Manley, Gitter and Slavchevska (2011), and Leroy, Ruel and Verhofstadt (2009), which meta-analyze the impact of CCT programs on nutritional status. Our main contribution to the CCT literature is, therefore, to systematically summarize and integrate meta-analytically available evidence on CCT educational impacts and cost-effectiveness, and shed light on which factors mediate heterogeneity in these measures.

From a literature search of over 25 electronic databases conducted in the spring of 2010, we surveyed 2,931 initial references containing the words "conditional cash transfer" or "conditional cash transfers" in either title, keyword or abstract (introduction if abstract not available). After screening out duplicate references, references that did not report effect estimates on school enrollment, attendance or dropout and references that where either summary of other reports, reviews or commentaries, we narrowed down our sample to forty-two references covering CCT programs in fifteen developing countries, twenty-eight of which report effect estimates on enrollment, nineteen on attendance and nine on dropout (some references report effects in more than one of these outcomes.)

We find wide heterogeneity in educational impact and cost-effectiveness estimates from available evaluations of CCT programs in developing countries. Part of the heterogeneity can be explained by variation in contextual and program characteristics.

We find, for instance, that CCT programs in developing countries are more effective at increasing school enrollment and attendance—all else constant—in contexts with relative low levels of baseline school enrollment, and therefore, particularly effective at increasing secondary enrollment and attendance. Our results also indicate that impact and cost-effectiveness estimates from studies of programs that in addition to transfers to families also attempt to expand supply through grants, infrastructure or other resources for schools are—all else constant—significantly larger than those from studies of programs only provide transfers to families. This result is consistent with single-program evidence from the Mexico's *Oportunidades* suggesting that school enrollment impacts are larger in areas with better school infrastructure and lower pupil-teacher ratios (Berhman, Parker and Todd, 2005) and with evidence from Colombia highlighting the how resource constraints affect educational attainment (Saavedra, 2012).

We also find evidence on the association between effect sizes and other program design characteristics that is consistent with single-program evidence. For example, we find that educational effect sizes are larger, all else constant, in programs in with lower payment frequency, which is consistent with single-program evidence from Bogotá's CCT program in which payment frequency was manipulated at random (Barrera-Osorio, Bertrand, Perez-Calle and Linden, 2011).

All else constant, effect size estimates are larger in programs that impose more stringent schooling conditions. This result aligns with recent single-CCT program evidence from Brazil's *Bolsa Escola* (Bourguignon, Ferreira and Leite, 2003), Ecuador's *Bono de Desarrollo Humano* (Schady and Araujo, 2008), Malawi's CCT program (Baird, McIntosh and Özler, 2011), and Mexico's *Progresa* (de Brauw and Hoddinott, 2011; Todd and Wolpin, 2006).

The result that CCT educational impact estimates are larger when conditions are more stringent is also consistent, for example, with a broader literature of experimental evaluations of educational interventions in developing countries. These evaluations find that imposing conditions on teachers and students improves effort and school performance (Duflo, Hanna and Ryan, 2012; Bettinger, Kremer and Saavedra, 2010; Kremer, Miguel and Thornton, 2009).

We find that transfer amounts are not related, all else constant, to educational effect size estimates, which is consistent with evidence from Cambodia's *CESSP* program (Filmer and Schady, 2011) and Malawi (Baird, McIntosh and Özler, 2009). However, we find that, all else constant, cost-effectiveness estimates are larger for programs with larger transfer amounts.

Methodologically we find that—all else constant—observational evaluations yield educational impact and cost-effectiveness estimates that are larger than those from randomized research designs. This finding, in particular, is at odds with previous qualitative evidence by IEG (2011) indicating that among comparable CCT programs there are little differences between effects reported by experimental and observational evaluations.

Finally, we find some evidence indicative of publication bias and selective reporting. We find large heterogeneity in the number of effect estimates that each reference reports. With the exception of primary enrollment estimates, funnel plots for all other outcomes and corresponding linear regression (Egger) tests also suggest selective reporting. We advocate, for this reason—as Duflo, Glennerster and Kremer (2007) do for randomized evaluations in development economics—for the importance of setting clear reporting standards for CCT impact evaluations given the popularity of these programs worldwide.

II. Literature Search

We searched published and gray literature to find all available studies that report estimates of the impact of CCT programs on school enrollment, school attendance and/or school dropout. We carried out the literature search in the spring of 2010. To minimize exclusion errors we began by searching for "conditional cash transfer" or "conditional cash transfers" in reference titles, abstracts or keywords in the following electronic databases: African Healthline, CAB Direct, Database of Abstracts of Reviews of Effectiveness, EBSCO, EconLit, Effective Practice and Organization of Care Group (EPOC), Eldis, British Library for Development Studies (BLDS), EMBASE, FRANCIS, Google Scholar, Healthcare Management Information Consortium, ID21, International

Bibliography of the Social Sciences (IBSS), Internet Documents in Economics Access Service (Research Papers in Economics- IDEAS[Repec]), Inter-Science, Latin American and Caribbean Health Sciences Literature (LILACS), MEDCARIB, Medline, Pan American Health Organization (PAHO), POPLINE, ProQuest, Scielo, ScienceDirect, Social Science Research Network (SSRN), The Cochrane Central Register of Controlled Trials, Virtual Library in Health (ADOLEC), WHOLIS (World Health, Organization Library Database) and World Bank.[1]

We retrieved all references in English or Spanish language regardless of geographic focus. We limited our search to published and unpublished studies, including refereed and non-refereed journals, working papers, conference proceedings, book chapters, dissertations, government reports, non-governmental reports and other technical reports. We did not include published comments, op-eds, summaries or media briefings.

To confirm that we had not left out studies, we cross-validated the initial literature search with the reference lists of Fiszbein et al.'s (2009) CCT review book and Milazzo's (2009) annotated bibliography on CCT programs. If we found a new reference from these two sources, we included it as long as it met the language and publication type restrictions above. This initial search procedure yielded 2,931 references, of which 1,341 were duplicate references (i.e. identical references) retrieved from more than one search engine (Table 1).

We then asked two research assistants to independently flag studies that in the title, abstract (or introduction if no abstract was available) contained any of the following words: education, school, schools, enrollment, attendance, achievement, attainment, test-scores / test scores, drop-out/dropout/drop out, graduation, persistence, performance, retention, advancement, fail/ failing,

[1] The complete list of search dates and articles retrieved from each database is available upon request.

pass/passing. The two principal investigators resolved any arising discrepancies from this process. With this filter we eliminated 342 references, keeping 1,248 for additional screening.

Finally, the two principal investigators independently read the abstract, introduction, methodological sections and tables of these 1,248 remaining and only retained studies that met the following criteria:

1. *Intervention specification*: Must report CCT program effects on school enrollment, attendance or dropout. We understand CCT programs to be programs that provide monetary (i.e. not in kind) transfers to participant households in exchange of compliance with program requirements (i.e. not unconditional), which may include health visits and school enrollment/attendance.
2. *Outcome variables*: Reference must report at least one impact and its associated standard error or t-statistic on school enrollment, attendance or dropout.
3. *Geographic focus*: Study must report impacts on a CCT implemented in a developing country (i.e. studies from the United States are excluded).
4. *Research design*: Study must use a treatment-comparison research design. The comparison group can be wait-list, or no treatment. One group pre-post designs are not eligible. Non-randomized studies are eligible only if they report relevant pre-treatment characteristics of treatment and comparison groups.

Inconsistencies – the bulk of which were errors on inclusion rather than exclusion – between the two researchers were then discussed and resolved by looking at the details of the manuscripts. We retained from this filter 48 references, 6 of which were an older version of a retained reference. Our sample of analysis therefore comprises 42 references, and closely matches the sample of CCT references in IGN (2011). Appendix Table A provides details of the references in the final analysis sample. In the fall of 2012 we verified if the references found as unpublished manuscripts or working papers (in 2010) where published in journals. If so, we updated the reference in our sample.

III. Coding of References in Analysis Sample

We created a coding protocol (available upon request) to capture in a hierarchical structure (i.e. effects in references, references in programs) the following information:

Contextual and Program descriptors: Baseline enrollment; program targeting (both geographic and household targeting criteria); type of assignment to conditions (simple random assignment, random assignment after matching, stratification or blocking, nonrandom assignment); nature of the control group (whether the control group receives nothing from program or is on a waiting list); schooling conditionality (whether schooling conditionality is based on school enrollment, school attendance, grade promotion and/or other); school attendance conditionality (minimum school attendance required for schooling subsidy receipt); whether or not there is verification of school attendance; member of the household that receives the subsidy (child, mother, father or both parents); amount of schooling and health subsidies (both in US dollars and/or domestic currency); frequency of payment of both schooling and health subsidies; whether the subsidy amounts vary by gender, grade, age or other characteristic. We also collected data on whether or not the program provides supply incentives for education and health.

Reference descriptors: Type of publication (journal article, book chapter, book, working paper, thesis or doctoral dissertation, conference paper/proceeding, government report, or technical report); publication year; country and language (English or Spanish); source of data; sample sizes at baseline and follow-up of both treatment and control groups; attrition rates for both treatment and control groups; whether or not baseline data were collected before households began receiving benefits; whether or not there is balance between treatment and control groups in all reported baseline characteristics.

Effect estimates: Effect estimates for school enrollment, school attendance and school drop-out, separately for primary and secondary schooling, unless effect sizes are reported for primary and secondary overall. For each outcome, we extracted information on mean and standard deviation at baseline, effect size (value, methodology of estimation, subgroup and sample size), standard error or t-statistic of the estimated effect, and time where the effect is measured.

We coded references as follows. Two trained research assistants (A and B) independently coded 17 of the 42 references in the sample using separate paper versions of the coding protocol. During this coding stage, coders where allowed to talk to each other and PIs to resolve questions. For the remaining 25 references, the principal investigators randomized the order in which to code them and coders where not allowed to talk to each other. We then randomly assigned research assistants C and D to separately input in Excel the 42 protocols of either assistant A or B.

With two separate versions of sample descriptors and effects information, we estimated various inter-rater reliabilities (IRR) for program-, reference- and effect-level variables, defined as the percent of coincidences over the total number of variables in the set. Reassuringly, we computed IRR's of 84.8% for program-level variables, 89.5% for reference-level variables, 91.3% for enrollment effect sizes, 87.4% for attendance effect sizes and 96.9% for dropout effect sizes.

A principal investigator with the help of research assistant C, referred to the corresponding reference document for variables with detected inconsistencies and retrieved the correct response. Once we solved discrepancies one a case-by-case basis, we created a unified dataset for analysis.

IV. Sample Description

Programs in sample

Table 2 presents a summary of CCT programs in our analysis sample. Our sample contains 42 references reporting effects for 19 programs in 15 countries. Sixty-three percent of programs (12 of them) are from Latin America, 32% are from Asian (6) and one is from Africa.

Table 2 demonstrates the degree of heterogeneity in program characteristics. For example, 68% of programs condition transfer-receipt on school attendance – which is typically 80% or more of the schooling reference period, while 32% impose additional conditions on school achievement such as grade promotion or school achievement as a requirement. In most programs, school officials verify student attendance.

There is also variation in payment frequency and whether transfer amounts vary for different target groups. Fifty three percent of programs pay educational transfers on a monthly basis and over forty percent pay transfers less frequently, either bi-monthly, quarterly or bi-annually. In almost 60% of programs all children regardless of age, grade or gender are entitled to the same transfer amount. In 32% of programs, however, transfers for girls differ from boys' or transfer amounts vary by grade or age. In close to 25% of programs, only the mother is eligible to receive the payment.

For comparability across countries we constructed measures of monthly-equivalent average transfer amounts relative to PPP-adjusted GDP per capita.[2] This measure of transfer amount displays considerable variation across programs and across schooling levels. In the typical program, monthly schooling transfers for primary are 2.3% of PPP-adjusted GDP per capita, and the standard deviation

[2] Another alternative would have been to use the transfer amount as a fraction of total household expenditures in the sample. Very few references reported expenditures, which is why we opted for transfer as a fraction of PPP-adjusted GDP per capita.

is 2 percentage points. For secondary school, average transfer amount is 4.2% of PPP-adjusted-GDP per capita.

In over 70% of CCT programs, the demand-side transfer is unaccompanied by any sort of supply side intervention. In over 20% of programs in the sample, however, schools receive some form of support ranging from grants to infrastructure construction to textbook and other school inputs.

In most programs, assignment to treatment is not random and beneficiaries are usually selected using a variety of means tests. In 32% of programs, on the other hand, beneficiaries are selected randomly, most commonly after screening on the basis of geography or poverty. In close to 80% of programs the control group receives nothing, and in close to 20% controls are wait-listed.

There is also considerable variation in program costs per intended beneficiary. On average, the yearly cost per beneficiary is 80.6 US Dollars of 2011, with standard deviation is 40.3 US Dollars.

Reference Characteristics

Table 3 shows reference-level characteristics of references in our analysis sample. Forty-five percent of references are journal articles, about 30% are working papers and 17% are government or technical reports. The remaining 7% are book chapters or doctoral dissertations. Seventy six percent of references in our sample use program survey data to estimate program impacts, and the remaining use either census or household survey data or other data sources.

Sixty-seven percent of references in our analysis sample report effects on enrollment, primary, secondary or both. Forty-five percent report effects on attendance and 21% report effects on school dropout. (Some references report effects on more than one type of outcome.) We provide extensive details of each program and reference in our sample in Appendix Tables A and B, respectively.

Figure 1 shows the distribution of number of effects that each paper reports, separately by outcome and school level. For all outcomes and all levels, there is considerable heterogeneity in effect reporting, and all distributions have a long right tail. For primary enrollment, for example, conditional on reporting for the outcome, the median paper reports six effects, but the average reports ten, because four paper report 20 or more effects (different subgroups by age, grade, location or methodology). For secondary enrollment, the distribution is more symmetric conditional on reporting effects for this outcome: the median paper reports eleven effects and the average reports twelve, with four papers reporting more than twenty effects. For attendance, distributions of reported effects are fairly symmetrical, conditional on reporting. Conditional on reporting primary attendance outcomes, the median reference reports eight effects and the average nine, with two references reporting twenty-four or more effects. Conditional on reporting secondary attendance effects, median and mean number of reported effects is seven, with one reference reporting twenty-four effects. For primary dropout, conditional on reporting, the median paper reports six effects and the mean reports eight effects. One reference reports twenty-two primary dropout effects. Conditional on reporting secondary dropout effects, the median reference reports three effects, the mean reference reports five and one reference reports eighteen effects.

V. Methodological Approach to Combine and Analyze Effect Size and Cost-effectiveness Estimates

<u>Combining impact estimates</u>

All educational outcome measures we focus on are dichotomous: enrollment, attendance and dropout. Impact estimates are therefore either a post-treatment difference in conditional outcome probabilities between treatment and comparison groups, or a double difference (treatment vs. control, before vs. after).

In our approach to combine impact estimates we compute precision-weighted averages of all estimates in each reference *i* for schooling level/outcome cell *c* (for example, primary enrollment) as follows. Let T_{jci} denote the *j*'th impact estimate *j=1,2….., J* for schooling level/outcome cell *c* reported in reference *i*, v_{jci} its associated variance and $w_{jci} = 1/v_{jci}$. Then the average impact estimate for schooling level/outcome cell *c* in reference *i,* T_{ci} is:

$$T_{ci} = \frac{\sum_{j=1}^{J}(w_{jci}T_{jci})}{\sum_{j=1}^{J} w_{jci}} \qquad (1)$$

And it variance is:

$$v_{ci} = \frac{1}{\sum_{j=1}^{J}(1/v_{jci})} \qquad (2)$$

Under the null hypothesis of homogeneity (i.e. no heterogeneity) in impact estimates for schooling level/outcome cell *c* among the k references in our sample that report impact estimates for *c*, the overall mean effect size for cell *c,* $\bar{T}_c$ is therefore:

$$\bar{T}_c = \frac{\sum_{i=1}^{k}(w_{ci}T_{ci})}{\sum_{i=1}^{k} w_{ic}} \qquad (3)$$

The variance of $\bar{T}_c$ is:

$$v_c = \frac{1}{\sum_{i=1}^{k}(1/v_{ci})} \qquad (4)$$

The homogeneity test for impact estimates of schooling level/outcome cell *c* is given by:

$$Q_c = \sum_{i=1}^{k} w_{ci}\,(T_{ci} - \bar{T}_c)^2 \qquad (5)$$

which under the null hypothesis of homogeneity has a chi-square distribution with $k - 1$ degrees of freedom. An unbiased estimate of the variance in true impacts for schooling level/outcome cell *c* is then obtained by computing:

$$\widehat{\tau_{\theta c}^2} = [Q_c - (k-1)]/b_c \qquad (6)$$

where,

$$b_c = \sum_{i=1}^{k} w_{ci} - \left[\sum_{i=1}^{k} w_{ci}^2 / \sum_{i=1}^{k} w_{ci}\right]$$

We estimate (3) and (4) using Method of Moments estimators.

Computing cost-effectiveness estimates

In addition to program impact estimates, we also construct cost-effectiveness estimates for programs with available cost data. We obtained program cost data from Grosch, del Ninno, Tesliuc and Ouerghi's (2008) Table B.5 for Bangladesh's Female Stipend Program, Brazil's Bolsa Escola and Bolsa Escola/Familia, Colombia's Familia's en Acción, Costa Rica's Superémonos, Ecuador's Bono de Desarrollo Humano, Hondura's PRAF II, Indonesia's JPS Scholarship Program, Jamaica's PATH, Mexico's Progresa/Oportunidades, Nicaragua's Red de Protección Social, and Turkey's Social Risk Mitigation Project.[3] For each of these programs, Grosch et al. (2008) report total expenditures—including the cost of the transfers—in a year expressed in nominal US dollars and the number of beneficiaries in that year or the closest year available.

To convert to comparable monetary figures for total yearly program costs, we use the Bureau of Labor Statistic's Consumer Price Index.[4] We used 2011 as the year of analysis, so we converted all cost figures to US dollars of 2011. We then obtained estimates of program costs per year per intended beneficiary dividing total yearly costs in 2011 US dollars by total program beneficiaries in that year or closest available.

With the available cost data, we model a program's cost-effectiveness following Dhaliwal, Duflo, Glennerster and Tulloch's (2011) methodology of

[3] The cost data in Grosh et al. (2008) was compiled from various sources and it is unlikely that these sources collected program costs in a uniformly comparable fashion, for instance, using the ingredients method as advocated, for example, by Dhaliwal, Duflo, Glennerster and Tulloch (2011). These data, however, are the most comparable cost data available for CCT programs.
[4] Available at ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt, retrieved December 19, 2012.

dividing the marginal change in outcomes by the marginal change in costs as a result of the program:

$$Cost\ Effectiveness = \frac{Outcome\ with\ Program - Outcome\ without\ Program}{Cost\ with\ Program - Cost\ without\ Program}$$

We assume that the marginal change in cost is cost per year per intended beneficiary. This is implicitly making three assumptions. The first is that the cost without the program—the "cost comparator" case—includes the costs of teachers and infrastructure, for example, which would also be incurred in the absence of the program and thus cancel out. Second, that the relevant program duration is one year, which is consistent with the time horizon of most impact estimates. Third, that because total program costs as reported by Grosch et al. (2008) include transfer costs, the cost of the transfer is a cost of the program.

In the forest plots of cost-effectiveness estimates we report in the results section, the numerator of the cost effectiveness formula is the precision-weighted impact in percentage points. We use formula (2) above and apply the delta method to obtain confidence intervals on each reference-level cost-effectiveness estimate.

For the meta-regression results, in which we pool all outcomes, we compute the numerator of the cost-effectiveness formula by standardizing T_{ci} with respect to its mean and standard deviation in the sample such that effect sizes for all schooling level/outcome cells *c* are expressed in standard deviation units. Dropout estimates as reported are problematic because, unlike estimates for enrollment and attendance, the signs are reversed, so that more negative is better. So instead, we standardize 1-dropout.

If a reference reports impacts for more than one schooling level/outcome cell *c*, we divide all the impacts by the yearly program cost per intended beneficiary. We therefore do not attempt to "allocate" costs to different outcomes, which is not possible with the available cost data. This approach to include multiple outcomes in the cost-effectiveness analysis therefore takes the perspective of an implementing government interested in judging the cost-effectiveness of an education CCT program against all schooling outcomes, not just enrollment, for instance.

Analyzing effect size and cost-effectiveness estimates

To explore how contextual and program characteristics explain variability in effect size and cost-effectiveness estimates, we pursue the following approach that combines all reference-level estimates in one meta-regression model. By pooling primary and secondary enrollment, primary and secondary attendance, and (sign-reversed) primary and secondary dropout estimates in one model, this approach allows us to maximize statistical power.

Separately for each dependent variable of interest (standardized effect size, cost-effectiveness) we estimate the following hierarchical model:

$$Y_{ic} = \alpha_c + X\beta + \mu_i + \varepsilon_i \qquad (7)$$

where α_c are schooling level/outcome cell indicators, μ_i is a random effect and ε_i is sampling error and Y_{ic} is either the precision weighted standardized effect size or cost-effectiveness estimate for reference i in schooling level/outcome cell *c*. In the vector X we include context and program characteristics such as baseline enrollment and whether the program is in Latin-America; whether benefits are randomly assigned, whether the program complements cash transfers with any form of supply-side intervention such as infrastructure or additional teachers, payment frequency (monthly vs. less frequently), whether the program imposes conditions on achievement beyond the standard school attendance conditions, transfer recipient (mother vs. other) and minimum school attendance required for schooling transfer receipt expressed as percent of total attendance days in the reference period. In model (7) we adjust standard errors for hierarchical dependence of effect estimates (i.e. clustering) at the reference-level using the methods of Hedges, Tipton and Johnson (2010).

Assessing publication bias and selective reporting

We employ two techniques to assess the extent to which publication bias and selective reporting are issues of potential concern in the CCT evaluation literature: funnel plots and Egger linear regression tests. The first is funnel plots in which we plot each impact estimate against the sample size used to calculate it. The intuition behind this test is straightforward. When sample sizes are small, there is likely a lot of variation in estimated effects around the overall (random effects) average effect size. As sample sizes increase, estimates on both sides of the overall effect will gradually converge to the overall effect, rendering a funnel-shaped plot of effect estimates. In the absence of publication bias and selective reporting, the funnel plot should look symmetrical and the number of effects should be evenly distributed around the overall effect (Sutton, 2009). The suppression of some effects that is associated with publication bias and selective reporting results in the plot being asymmetrical, with patchy spots of "missing effects."

Egger linear regression tests are a statistical formalization of the intuition behind funnel plots. In Egger tests, we regress standardized effect sizes against the reciprocal of the standard errors and a constant term. The constant provides a measure of asymmetry and thus we can test the null hypothesis of no asymmetry using a standard t-test on the constant term. None of these tests are "magic bullets," however. The funnel plot might be asymmetric if, for example, smaller studies take place under less rigorous conditions. Similarly, Egger's regression analysis has inflated type I errors in meta-analyses with dichotomous outcome variables. Moreover, asymmetry might be due to heterogeneity in effect sizes. For these reasons, we take the results from these tests as suggestive, not conclusive (Sutton, 2009.)

VI. Results

School Enrollment Impacts

Figures 2a and 2b show the forest plots (distribution) of impact estimates from all studies reporting enrollment impacts on primary and secondary school, respectively. Forest plot figures report the reference-level precision-weighted impact and 95% confidence interval that we compute using equations (1) and (2) above.

We highlight three aspects of Figure 2a. First, the overall random-effects average primary enrollment impact is 5.2 percentage points, with a 95% confidence interval between 3.7 and 6.7 percentage points. Relative to the mean baseline primary enrollment of 84%, the average impact represents a 6.2 percent enrollment increase. Second, with the exception of one reference reporting impacts from the *SRMP* CCT program in Turkey, all reference-level average impacts are positive and most are statistically distinguishable from zero. Third, there is ample variation in estimated impacts across studies.[5]

Reference-level impact estimates for Nicaragua's *Red de Protección Social* are the largest as a whole, ranging from close to 8 to 29 percentage points, and statistically positive. For Colombia's *Familias en Acción* and Brazil's *Bolsa Escola*, reference-level effects are, on the other hand, consistently small and generally statistically positive.

Figure 2b displays the forest plot of precision-weighted secondary enrollment impact estimates that we estimate using equations (1) and (2) above. The average secondary enrollment impact is similar in percentage points to that of primary enrollment – 6 percentage points – although as a fraction of baseline enrollment it is notably larger. Baseline secondary enrollment is on average 59%, implying

[5] The chi-square test-statistic for the null hypothesis of homogeneity in primary enrollment impact estimates is 727 (p-value 0.000). We obtain similar conclusions when we estimate the average effect size estimate using all impact estimates in all references. The overall estimate is 5.2 percentage points with a standard error of 0.02. The chi-square test statistic for the homogeneity test is 1,761 (p-value 0.000).

that the average secondary enrollment impact estimate represents a 10 percent secondary enrollment increase. The secondary enrollment plot displays considerable effect-size variation, with evaluations of programs like Cambodia's JFPR Scholarship and CESSP programs reporting average secondary enrollment impacts of close to twenty percentage points.[6]

The finding that CCT programs on average are more effective at increasing secondary than at increasing primary enrollment resonates with previous CCT review findings in Fiszbein et al. (2009). Note, however, that this finding might simply reflect the fact that CCT programs are more effective at increasing enrollment in contexts in which baseline enrollment is low, which is usually the case for secondary schooling in developing countries.

School Enrollment Cost-Effectiveness Estimates

Figures 3a and 3b show the distribution of precision-weighted cost-effectiveness estimates. For these figures, we construct reference-level cost-effectiveness estimates dividing the precision-weighted enrollment impact in percentage points obtained from equation (1) above by the yearly cost per intended beneficiary of the program in US dollars of 2011, when the latter is available. As advocated by Dhaliwal, Duflo, Glennerster and Tulloch (2011), we account for uncertainty in our cost-effectiveness estimates. Specifically, we construct the 95% confidence interval on each cost-effectiveness estimate using the delta method and the variance of the precision-weighted impact estimate obtained from equation (2) above.

We only show cost-effectiveness estimates' forest plots for enrollment outcomes for illustrative purposes, although in reality education CCT have impacts on multiple education outcomes including attendance and dropout as

[6] The chi-square test-statistic for the null hypothesis of homogeneity in secondary enrollment impact estimates is 1,302 (p-value 0.000). We obtain a similar overall secondary enrollment estimate (5 percentage points) when we use all estimates from all references, and similarly reject the null hypothesis of homogeneity in impact estimates (chi-square statistic=2,409, p-value 0.000).

well. In our meta-regression analysis we do include cost-effectiveness estimates for all outcomes, which implicitly takes the perspective of an implementing government interested in judging the cost-effectiveness of an education CCT program against all schooling outcomes, not just enrollment.

Figure 3a shows precision-weighted cost-effectiveness estimates for primary enrollment. The overall cost-effectiveness mean, computed from equation (3) above, is 0.06 percentage points per 2011 US dollar per intended beneficiary. As is the case with precision-weighted impact estimates, there is considerable heterogeneity in cost-effectiveness estimates across references and programs.[7] With respect to primary enrollment—and without allocating costs to multiple outcomes—Brazil's *Bolsa Escola/Bolsa Familia* and Colombia's *Familias en Acción* are less cost-effective than the average program. On the other hand, Indonesia's *JPS Scholarship and Grant Program* and Nicaragua's *Red de Protección Social* are more cost-effective that the average program.

Figure 3b shows precision-weighted cost-effectiveness estimates for secondary enrollment. The overall cost-effectiveness mean, computed from equation (3) above, is also 0.06 percentage points per 2011 US dollar per intended beneficiary. As is the case with precision-weighted impact estimates, there is considerable heterogeneity in secondary enrollment cost-effectiveness estimates across references and programs.[8] Without allocating costs to multiple outcomes, Bangladesh's *Female Stipend Program,* Colombia's *Familias en Acción* and Indonesia's *JPS Scholarship and Grant Program* are the most cost-effective at improving secondary enrollment, while Turkey's *SRMP* and Mexico's *Oportunidades* are the least cost-effective.

School Attendance Impact Estimates

[7] The chi-square test-statistic for the null hypothesis of homogeneity in primary enrollment cost-effectiveness estimates is 826 (p-value 0.000).
[8] The chi-square test-statistic for the null hypothesis of homogeneity in secondary enrollment cost-effectiveness estimates is 1,519 (p-value 0.000).

Figure 4a displays the precision-weighted primary attendance impact distribution. Fewer references report primary attendance effects relative to those reporting primary enrollment. The average random-effects primary attendance effect is 2.5 percentage points, which off of a baseline attendance of 80% represents a three percent attendance effect – and is statistically significantly different from zero.

A clear outlier in the distribution of primary attendance impact estimates is Nicaragua's *Red de Protección Social*, with reported average attendance effect of thirteen percentage points. For this program, as we noted earlier, primary enrollment effects are also notoriously large. With the exception of Uruguay's *Ingreso Ciudadano*, all primary attendance reference-level effects are positive and the majority statistically different from zero. We strongly reject the null hypothesis of estimate homogeneity (chi-square statistic=113, p-value 0.000).

Figure 4b displays the distribution of precision-weighted secondary attendance impact estimates in percentage points obtained from equation (1) above. The CCT average secondary attendance impact estimate is 7.7 percentage points and statistically significant. This impact represents a 12% increase in attendance relative to the average baseline secondary attendance level of 68%.

There is considerable heterogeneity in secondary attendance impact estimates across programs.[9] At the most impactful extreme stands Cambodia's *CESSP* with average secondary attendance impact estimates of twenty to thirty percentage points. (This program's evaluation also reports notoriously high secondary enrollment effects.) At the other extreme we find Malawi's *CCT for Schooling Program* with average secondary attendance impact estimates that although positive and statistically significant are small – close to half of a percentage point.

<u>School Dropout Average Effect Sizes</u>

[9] We reject the null hypothesis of homogeneity in impact estimates, with a homogeneity test chi-square statistic of 4,470 and associated p-value of 0.000.

Compared to enrollment and attendance, few CCT evaluation references report dropout effects: nine for primary and six for secondary. This relatively low number of references reflects on the uncertainty with which we calculate average impacts. Unlike enrollment and attendance, only two papers report baseline dropout rates so we opted for not reporting an average to avoid potential issues of sample selection in converting effect sizes to relative magnitudes.

Figure 5a shows the distribution of precision-weighted primary dropout impact estimates calculated from equation (1) and (2) above. The overall average dropout impact estimate for primary is negative one percentage point and statistically different from zero, although the 95% confidence interval is relatively wide. Evaluations for Nicaragua's *Red de Protección Social* and Brazil's *Bolsa Escola* report the largest dropout reduction impact estimates, while the evaluation of Ecuador's *Bono de Desarrollo Humano* suggests, if any, increases in dropout as a consequence of program participation. We reject the null hypothesis of homogeneity in primary dropout impact estimates (chi-square statistic=3,603, p-value 0.000).

The average secondary dropout effect of negative four percentage points is three times larger—in percentage points—than that for primary dropout (Figure 5b). Although we estimate the average secondary dropout effect with a high degree of uncertainty due to the fact that only six references in our sample report estimates for this outcome, we still reject the null hypothesis that CCT programs do not, on average, reduce secondary. All reported secondary dropout impact estimates are negative and statistically different from zero and those from Brazil's *Bolsa Escola* and Mexico's *Progresa* stand out as the largest effects in secondary dropout reduction, close to eight percentage points.[10]

<u>Program impacts meta-regression results</u>

[10] We reject the null hypothesis of homogeneity in secondary dropout estimates (chi-square statistic=1,238, p-value 0.000).

Table 4 reports meta-regression estimation results of equation (7). Column (1) presents estimates in which the dependent variable is the standardized effect size estimate (standardized to have mean zero and standard deviation one within each schooling level/outcome cell *c*). Column (1) includes precision-weighted estimates for all cells *c*, with signs of dropout estimates appropriately reversed.

We highlight five main findings from meta-regression results in Column (1). First, controlling for program design characteristics and geographic location, CCT programs from Latin America and those put in place in contexts with relatively low baseline school enrollment levels are significantly more effective at improving educational outcomes. Such contexts might include settings with a large rural share of the population or secondary schooling, which is typically low in developing countries.

Our second main finding is that educational impacts are stronger in programs that complement cash transfers to families with supply-side interventions such as school infrastructure, additional teachers, grants or textbooks. Holding all else constant, effect sizes are one standard deviation statistically higher in CCT programs with a complementary supply side intervention relative to those without one.

The positive association between educational effect size estimates and supply-side complementary interventions is consistent, for instance, with evidence from Mexico's *Oportunidades* program suggesting that school enrollment impacts are larger in areas with better school infrastructure and lower pupil-teacher ratios (Behrman, Parker and Todd, 2005). This result is also consistent with results in Saavedra (2012) demonstrating, more generally, that resource constraints in developing countries negatively affect educational attainment.

Our third main finding is that educational effect sizes are stronger in programs in with lower payment frequency. Holding all else constant, relative to programs that pay families monthly, effect sizes in programs that pay families less frequently (for example bimonthly) are 0.49 standard deviations higher, and the difference is statistically significant.

The meta-regression finding that payment frequency is negatively associated with educational effect sizes is consistent with single-program evidence from Bogotá's CCT program in which payment frequency was manipulated at random (Barrera-Osorio, Bertrand, Perez-Calle and Linden, 2011). The authors of the Bogotá study argue that fully or partially delaying transfers increases re-enrollment because doing so might help families relax savings constraints.

From a theoretical standpoint, savings constraints might arise if families have limited attention with respect to lumpy expenditures (Karlan, McConnell, Mullainathan and Zinman, 2011) or face self-control problems (Ashraf, Karlan and Yin, 2006). Our meta-regression results—in conjunction with results from the Bogotá CCT evaluation—suggest that among target populations of CCT programs savings constraints exist and that programs that delay fully or partially delay transfers can be therefore be more effective.

Our fourth main finding is that, all else constant, effect size estimates are larger in programs with more stringent conditions. Effect sizes are statistically significantly larger when the minimum school attendance required for the schooling transfer receipt is higher—going, for example, from 80% to 90% of days in the reference period. At the same time, effect sizes are larger when transfer continuation is conditional on achievement, although this particular conditional correlation is imprecisely estimated.

These results are consistent with recent single-program evidence stressing the role that conditionality plays in program design and subsequent impacts. For example, de Brauw and Hoddinott (2011) for the case of Mexico's *Progresa*, and Schady and Araujo (2008) for the case of Ecuador's *Bono de Desarrollo Humano* program, argue that the belief that there was a school enrollment requirement attached to the transfers explains the positive effect of the program on schooling outcomes.

Similarly, taking a structural modeling approach, Bourguignon, Ferreira and Leite (2003) for the case of Brazil's *Bolsa Escola* and Todd and Wolpin (2006) for the case of Mexico's *Progresa* argue that program impacts would be considerably lower were the school enrollment conditionality to be removed. In a randomized research design in Malawi, Baird, McIntosh and Özler (2011) manipulate the conditionality requirement and find that educational outcomes—including school dropout reductions—were significantly better in the conditional transfer treatment relative to the unconditional transfer treatment.

Our result that CCT educational impacts are larger when conditions are more stringent is also consistent, for example, with a broader literature of experimental evaluations of educational interventions in developing countries. These evaluations find, for example, that, imposing conditions on teachers reduces teacher absenteeism and improves student performance (Duflo, Hanna and Ryan, 2012), and that imposing conditions on students motivates them to exert more effort in school (Bettinger, Kremer and Saavedra, 2010; Kremer, Miguel and Thornton, 2009).

Our fifth main finding is that, all else constant, transfer amounts are not statistically correlated to effect sizes. This finding is consistent with single-program evidence from Cambodia and Malawi. For Cambodia's *CESSP* program, for instance Filmer and Schady (2011) find that, relatively to children in observationally similar households receiving a smaller transfer, children in households receiving a larger transfer showed similar enrollment improvements. Similarly Baird, McIntosh and Özler (2009), randomly manipulate transfer amounts in Malawi's CCT program and find that educational outcomes including school dropout are insensitive to transfer amounts.

Column (2) of Table 4 presents a robustness check to the results in Column (1) by dropping outlier impact estimates reported by Filmer and Schady (2011) for Cambodia's *JFPR* program as well as those reported by Dammert (2009) for Nicaragua's *Red de Protección Social*. Meta-regression results in Column (2) confirm our main findings. The magnitudes of the conditional correlations are similar to those in Column (1) and all previously estimated statistically significant correlations remain so after the exclusion of these outlier impact estimates.

<u>Cost-effectiveness meta-regression results</u>

In Columns (3) and (4) of Table 4 we explore moderators of cost-effectiveness estimates, defined as the standardized effect size divided by the yearly program cost per intended beneficiary in US dollars of 2011. Recall that if a reference reports impacts for more than one schooling level/outcome cell *c*, we divide all the impacts by the yearly program cost per intended beneficiary. We therefore do not attempt to "allocate" costs to different outcomes, and divide all impacts for a given schooling level/outcome cell by the same per-intended beneficiary cost.

Because we do not have cost data for all programs for which we have impact estimates, our first step of the cost-effectiveness meta-regression analysis is to investigate for potential selection into having cost data based on observed contextual and program characteristics. We carry out this selection analysis in a simple OLS regression framework in which the dependent variable is a dichotomous indicator that takes the value of one for schooling level/outcome cells for programs with cost data and zero for those without.

Column (3) of Table 4 presents the result for selection into having cost data. These estimates indicate that, all else constant, we are more likely to have cost data for programs in settings with higher baseline enrollment levels, such as those in richer, more urbanized contexts. Similarly, holding other contextual and program attributes constant, we are more likely to have cost data for programs that, in addition to the transfer also provide a supply-side complementary intervention. These selection estimates therefore suggest that the sample in which we carry out our cost-effectiveness meta-regression analysis might not be entirely representative of the sample of programs for the impact meta-regression results. With this caveat in mind, we report cost-effectiveness meta-regression results in Column (4) of Table 4.

We highlight three main findings from our cost-effectiveness meta-regression results. The first is that, as is the case with impact meta-regression results, program cost-effectiveness estimates, all else constant, are larger when programs provide a supply-side complementary intervention. This result suggests that the increased impacts as a result of the complementary supply side intervention more than outweigh the additional costs.

Our second main finding from the cost-effectiveness meta-regression results is that, holding all else constant, programs are more cost-effective when the average monthly subsidy is larger. This finding suggests that although in absolute terms transfer amounts are not significantly correlated with effect size estimates, once the marginal cost per program beneficiary is accounted for, programs with larger transfer amounts are associated with larger impacts per dollar spent.

Our third main finding is that cost-effectiveness estimates are lower in programs that randomly assign beneficiaries to conditions. Part of this result is explained by the fact that, all else constant, program impacts are considerably lower in magnitude—0.4 to 0.5 standard deviations—in programs with random assignment, although this correlation is imprecisely estimated, as Columns (1) and (2) indicate. This result therefore suggests that programs evaluated with observational research designs are, all else constant, likely overstating CCT program cost-effectiveness.

Publication bias and selective reporting

We have already reported some suggestive evidence of publication bias and selective reporting in CCT impact evaluation reports. For instance, we noted the wide degree of heterogeneity in the number of effects that references report: median number of reported effects ranges from six to eleven across schooling level/outcome cells and some references report more than twenty effects.

In this section we report graphical and linear regression results from additional publication bias and selective reporting tests. We use two tests: funnel plots and linear regression Egger-type tests. Figures 6 through 8 display funnel plots separately for each outcome. Table 5 reports Egger tests for each outcome and level separately. Effects for primary enrollment do converge to the overall random effects average effect size, but the density of effects is not symmetric around the overall mean (Figure 6a). Column 1 of Table 5 confirms this asymmetry: we strongly reject the null hypothesis that the constant is zero. Effects for secondary enrollment are also converge to the overall mean as sample size increases, but the funnel plot is considerably more symmetric than that for primary enrollment (Figure 6b). Results in column 2 of Table 5 support the symmetry conclusion for secondary enrollment effects, as we cannot reject the null hypothesis that the constant is different from zero.

Effects for primary attendance converge to the overall mean as sample size increases (Figure 7a). The funnel plot is visibly asymmetric, with a large patch of missing effects to the left of the overall mean. The funnel plot for secondary attendance effect sizes is also visibly asymmetric (Figure 7b). Statistical analysis in columns 3 and 4 of Table 5 reject the hypothesis of funnel plot symmetry for both of these outcomes.

Figures 8a and 8b display funnel plots for primary and secondary dropout effects. Effects for both levels tend to converge to the overall effect size as sample size increases, but they are both visibly asymmetrical, with patches of missing positive effects (for instance, smaller reductions in dropout than the overall effect size). Results in columns 5 and 6 of Table 5 confirm the visual inspection of the funnel plots and for the case of primary dropout reject the null hypothesis of funnel plot symmetry. For secondary dropout, the magnitude of the constant is large (in standard deviation units) but the test is underpowered due to the small number of effects. Overall we conclude that for most outcomes – perhaps with the exception of secondary enrollment – there is suggestive evidence in support of publication bias and/or selective reporting. The heterogeneity in the number of effects that each paper reports provides additional support to this conjecture.

VII. Conclusion

CCT programs in developing countries are more impactful in contexts with relative low levels of baseline school enrollment, and therefore, particularly effective at improving secondary schooling outcomes that include enrollment, attendance and dropout.

On the whole, our meta-regression results are consistent with single-program evidence results that exploit variation in program design features. For example our results indicate that CCT Programs that in addition to cash transfers to families also attempt to expand supply through grants, infrastructure or other resources for schools are both significantly more impactful and more cost-effective than those programs only provide transfers to families, all else constant. This result is consistent with evidence from Mexico's *Oportunidades* program.

Our meta-analysis also suggests that lower payment frequency and more stringent conditions for transfer receipt are, all else constant, associated with larger impact estimates. These results are consistent with evidence from Bogotá's CCT program, and with evidence from Brazil's *Bolsa Escola*, Ecuador's *Bono de Desarrollo Humano* and Mexico's *Progresa* CCT programs, respectively.

Although our finding that transfer amounts are not associated with larger impact estimates is consistent with previous single-program evidence from Cambodia's *CESSP* and Malawi's CCT Program, our cost-effectiveness meta-regression results indicate that, all else constant, larger transfer amounts are associated with more cost-effective educational interventions. This finding, which is consistent with evidence from structural econometric models for Brazil's *Bolsa Escola* and Mexico's *Progresa*, in conjunction with our finding that programs with supply-side complementary interventions are also more cost-effective for improving educational outcomes highlight how cost-minimization—as opposed to cost-effectiveness maximization—might not be the most relevant objective when designing social assistance programs.

Our meta-regression results also indicate that, all else constant, evaluations with an observational research design report, on average, impact and cost-effectiveness estimates that are larger than those from evaluations that take advantage of random assignment. This finding, in particular, is at odds with previous qualitative evidence by IEG (2011) indicating that among comparable CCT programs there are little differences between effects reported by experimental and observational evaluations.

Finally, we find some evidence indicative of publication bias and selective reporting. We find large heterogeneity in the number of effect estimates that each reference reports. With the exception of primary enrollment estimates, funnel plots for all other outcomes and corresponding linear regression (Egger) tests also suggest selective estimate reporting. From an impact evaluation policy perspective we therefore advocate for setting clear reporting standards for CCT impact evaluations given the popularity of these programs around the world.

References

Ahmed, A. U., D. Gilligan, A. Kudat, R. Colasan, H. Tatlidil, and B. Ozbilgin. 2006. "Interim Impact Evaluation of the Conditional Cash Transfers Program in Turkey: A Quantitative Assesment." International Food Policy Research Institute.

Attanasio, O., C. Meghir, and A. Santiago. 2012. "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA." *Review of Economic Studies* 79(1): 37-66.

Attanasio, O., E. Fitzsimons, A. Gomez, M. I. Gutiérrez, C. Meghir, and A. Mesnard. 2010. "Children's Schooling and Work in the Presence of a Conditional Cash Transfer Program in Rural Colombia." *Economic Development and Cultural Change* 58 (2): 181-210.

Attanasio, O. and L.C. Gómez. 2004. "Evaluación del Impacto del Programa Familias en Acción - Subsidios Condicionados de la Red de Apoyo Social." Bogotá D.C.: National Planning Department.

Attanasio, O., M. Syed, M. Vera-Hernandez. 2004. "Early Evaluation of a New Nutrition and Education Programme in Colombia." Institute for Fiscal Studies Briefing Note 44.

Ashraf, N., D. Karlan, and W. Yin. 2006. "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines." *The Quarterly Journal of Economics*, 121(2): 635-672.

Baird, S., C. McIntosh, and B. Özler. 2009. "Designing Cost-Effective Cash Transfer Programs to Boost Schooling Among Young Women in Sub-Saharan Africa." World Bank Policy Research Working Paper 5090.

Baird, S., C. McIntosh, and B. Özler. 2011. "Cash or Condition? Evidence from a Cash Transfer Experiment." *Quarterly Journal of Economics* 126(4): 1709-1753.

Barrera-Osorio, F., M. Bertrand, L.Linden, and F. Perez-Calle.(2011). "Improving the Design of Conditional Transfer Programs: Evidence from a Randomized Education Experiment in Colombia." *American Economic Journal: Applied Economics* 3(2): 167-195.

Behrman, J. R., J. Gallardo-García, S. W. Parker, P.E. Todd and V.Vélez-Grajales. 2012. "Are conditional cash transfers effective in urban areas? Evidence from Mexico" *Education Economics* 20(3): 233-259

Behrman, J. R., S. W. Parker, and P. E. Todd. 2004. "Medium-Term Effects of the Oportunidades Program Package, Including Nutrition, on Education of rural children Age 0-8 in 1997." Instituto Nacional de Salud Publica (INSP) Technical Document 9.

Behrman, J. R., P. Sengupta, and P. E. Todd. 2005. "Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment in

Rural Mexico." *Economic Development and Cultural Change* 54 (1): 237-275.
Bettinger, E., M. Kremer, and J. E. Saavedra. 2010. "Are Educational Vouchers Only Redistributive?" *The Economic Journal* 120 (546): F204-F228.
Borraz, F. and N. González. 2009. "Impact of the Uruguayan Conditional Cash Transfer Program." *Cuadernos de Economía* 46 (November): 243-271.
Bourguignon, F., F. Ferreira, and P. G. Leite. 2003. "Conditional Cash Transfers, Schooling, and Child Labor: Micro-Simulating Brazil"s Bolsa Escola Program," *The World Bank Economic Review,* 17: 229-254.
Cameron, L. 2009. "Can a Public Scholarship Program Successfully Reduce School Drop-outs in a Time of Economic Crisis? Evidence from Indonesia." *Economics of Education Review* 28 (3): 308-317.
Cardoso, E., and A. P. Souza. 2004. "The Impact of Cash Transfers on Child Labor and School Attendance in Brazil." Department of Economics, Vanderbilt University Working Paper 04-W07.
Chaudhury, N. and D. Parajuli. 2010. "Conditional Cash Transfers and Female Schooling: The Impact of the Female School Stipend Programme on Public School Enrolments in Punjab, Pakistan." *Applied Economics* 42(28-30): 3565-3583.
Coady, D. and S. W. Parker. 2004. "Cost-Effectiveness Analysis of Demand- and Supply-Side Education Interventions: The Case of PROGRESA in Mexico." *Review of Development Economics* 8(3): 440-451.
Cooper, H., L. V. Hedges, and J. C. Valentine, editors. 2009. *Handbook of Research Synthesis and Meta-analysis.* New York, NY: Russell Sage.
Dammert, A. C. 2009. "Heterogeneous Impacts of Conditional Cash Transfers: Evidence from Nicaragua." *Economic Development and Cultural Change* 58 (1): 53-83.
Davis, B., S. Handa, M. Ruiz-Arranz, M. Stampini, and P. Winters. 2002. "Conditionality and the Impact of Programme Design on Household Welfare: Comparing Two Diverse Cash Transfer Programmes in Rural Mexico". Unpublished manuscript.
de Brauw, A., and J. Hoddinott. 2011. "Must Conditional Cash Transfer Programs be Conditioned to be Effective? The Impact of Conditioning Transfers on School Enrollment in Mexico," *Journal of Development Economics,* 96(2): 359-370.
De Janvry, A., F. Finan and E. Sadoulet. 2006. "Evaluating Brazil's Bolsa Escola Program: Impact on Schooling and Municipal Roles." Berkeley: University of California at Berkeley.

De Souza, P. Z. 2005. “An Impact Evaluation of the Conditional Cash Transfers to Education under PRAF: An experimental Approach.” Rio de Janeiro: Fundacao Getulio Vargas.
Dhaliwal, I., E. Duflo, R. Glennerster and C. Tulloch. 2011. “Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Application to Education. Mimeo, MIT.
Duflo, E., R. Glennerster, and M. Kremer. 2007. “Using Randomization in Development Economics Research: A Toolkit.” CEPR Discussion Paper 6059.
Duflo, E., R. Hanna, and S. P. Ryan. (forthcoming). “Incentives Work: Getting Teachers to Come to School.” *American Economic Review* 102(4): 1241-1278.
Duryea, S. and A. Morrison. 2004. “The Effect of Conditional Transfers on School Performance and Child Labor: Evidence from an Ex-post Impact Evaluation in Costa Rica.” Inter-American Development Bank Working Paper 505.
Filmer, D. and N. Schady. 2008. Getting Girls into School: Evidence from a Scholarship Program in Cambodia. *Economic Development and Cultural Change* 56 (3): 581-617.
Filmer, D. and N. Schady. 2011. “Does More Cash in Conditional Cash Transfer Programs Always Lead to Larger Impacts on School Attendance?.” *Journal of Development Economics* 96(1): 150-157.
Filmer, D. and N. Schady. 2009. “School Enrollment, Selection and Test Scores.” World Bank Policy Research Working Paper 4998.
Fiszbein, A., N. Schady, F. Ferreira, M. Grosh, N. Kelleher, P. Olinto, and E. Skoufias. 2009. *Conditional Cash Transfers: Reducing Present and Future Poverty*. Washington D.C.: World Bank.
Ford, D. B. 2007. “Household Schooling Decisions and Conditional Cash Transfers in Rural Nicaragua.” Washington, D.C.: Georgetown University
Gitter, S. R., and B. L. Barham. 2009. “Conditional Cash Transfers, Shocks, and School Enrolment in Nicaragua.” *The Journal of Development Studies* 45 (10): 1747-1767.
Glewwe, P. and A. L. Kassouf. 2012. “The impact of the Bolsa Escola/Familia Conditional Cash Transfer Program on Enrollment, Dropout Rates and Grade Promotion in Brazil.” *Journal of Development Economics* 97(2): 505-517.
Glewwe, P. and P. Olinto. 2004. “Evaluating the Impact of Conditional Cash Transfers on Schooling: An experimental Analysis of Honduras' PRAF Program.” Unpublished manuscript.
Grosch, M., C. del Ninno, E. Tesliuc and A. Ouergui. 2008. *For Protection and Promotion: The Design and Implementation of Social Safety Nets.* Washington D.C.: World Bank.

Hasan, A. 2010. “Gender-targeted Conditional Cash Transfers: Enrollment, Spillover Effects and Instructional Quality.” World Bank Policy Research Working Paper 5257.

Hedges, L. V., E. Tipton, and M. C. Johnson. 2010. “Robust Variance Estimation in Meta-regression with Dependent Effect Size Estimates.” *Research Synthesis Methods* 1(1): 39-65.

Hoddinott, J. and L. Bassett. 2009. “Conditional Cash Transfer Programs and Nutrition in Latin America: Assessment of Impacts and Strategies for Improvement.” United Nations Food and Agriculture Organization Working Paper # 9, April.

Independent Evaluation Group (IEG). 2011. *Evidence and Lessons Learned from Impact Evaluations on Social Safety Nets*. Washington D.C.: World Bank.

Karlan, D., M. McConnell, S. Mullainathan and J. Zinman, J. 2011. “Getting to the Top of Mind: How Reminders Increase Saving.” Mimeo, Yale University.

Khandker, S. R., M. M. Pitt, and N. Fuwa. 2003. “Subsidy to Promote Girls’ Secondary Education: The Female Stipend Program in Bangladesh.” Unpublished manuscript.

Kremer, M., E. Miguel and R. Thornton. 2009. “Incentives to Learn.” *Review of Economics and Statistics* 91 (3): 437-456.

Levy, D. and J. Ohls. 2007. *Evaluation of Jamaica’s PATH Program: Final Report*. Washington D.C.: Mathematica Policy Research.

Leroy, J. L., M. Ruel and E. Verhofstadt. 2009. “The Impact of Conditional Cash Transfer Programmes on Child Nutrition: A Review of Evidence Using a Programme Theory Framework.” *Journal of Development Effectiveness* 1(2): 103-129.

Manley, J., S. Gitter and V. Slavchevska. 2011. “How Effective are Cash Transfer Programs at Improving Nutritional Status?”. Towson University Department of Economic Working Paper Series. Working Paper No. 2010-18

Maluccio, J., A. Murphy, and F. Regalia. 2010. "Does supply matter? Initial Schooling Conditions and the Effectiveness of Conditional Cash Transfers for Grade Progression in Nicaragua." *The Journal of Development Effectiveness* 2(1): 87-116.

Maluccio, J. A., and R. Flores. (2005). “Impact Evaluation of a Conditional Cash Transfer Program.” International Food Policy Research Institute Research Report 141.

Milazzo, A. (2009). *Conditional Cash Transfers: An Annotated Bibliography*. Retrieved from http://siteresources.worldbank.org/SAFETYNETSANDTRANSFERS/Resources/281945-1131738167860/CCT_Biblio_6Feb2009.pdf

National Planning Department. 2006. "Programa Familias en Acción: Condiciones Iniciales de los Beneficiarios e Impactos Preliminares." Bogota, D.C.: National Planning Department.
Oosterbeek, H., J. Ponce, and N. Schady. 2008. "The Impact of Cash Transfers on School Enrollment: Evidence from Ecuador." World Bank Policy Research Working Paper 4645.
Parker, S., P. E. Todd, and K. I. Wolpin. 2006. "Within-family Treatment Effect Estimators: The Impact of Oportunidades on Schooling in Mexico." Unpublished manuscript.
Ponce, J. 2006. "The impact of a conditional cash transfer program on school enrolment: the 'Bono de Desarrollo Humano'of Ecuador". Facultad latinoamerica de Ciencia Sociales- Sede Ecuador. Working Paper 06/302.
Rawlings, L. B., and G. M. Rubio. 2005. "Evaluating the Impact of Conditional Cash Transfer Programs." *The World Bank Research Observer* 20 (1): 29-55.
Raymond, M., and E. Sadoulet. 2003. "Educational Grants Closing the Gap in Schooling Attainment between Poor and Non-poor." Unpublished manuscript.
Saavedra, J. E. (2012). "Resource Constraints and Educational Attainment in Developing Countries: Colombia 1945-2005." *Journal of Development Economics*, 99: 80-91.
Schady, N. and M. C. Araujo. 2008. "Cash Transfers, Conditions, and School Enrollment in Ecuador." *Economia* 8 (2): 43-77.
Schultz, T. P. 2004. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program." *Journal of Development Economics* 74: 199–250.
Skoufias, E. and S.Parker. 2009. "The Impact of PROGRESA on Child Labor and Schooling". In P. F. Orazem, G. Sedlacek, Z. Tzannatos (Eds.) , *Child Labor and Education in Latin America: An Economic Perspective* (pp. 167-185). Houndmills, U.K. and New York: Palgrave Macmillan.
Sparrow, R. 2007. "Protecting Education for the Poor in Times of Crisis: An Evaluation of a Scholarship Programme in Indonesia." *Oxford Bulletin of Economics and Statistics* 69 (1): 99-122.
Sutton, A. J. (2009). "Publication Bias." In *Handbook of Research Synthesis and Meta-analysis,* edited by H. Cooper, L.V. Hedges, and J. C. Valentine. New York, NY: Russell Sage.
The Economist. 2010. "Societies on the move." The Economist, U.S. edition September 11, 2010.
Todd, P. E., J. Gallardo-Garcia, J. R. Behram, and S. W. Parker. (2005). "Impacto de Oportunidades sobre la Educación de Niños y Jóvenes de Áreas Urbanas Después de un Año de Participación en el Programa." In *Evaluación Externa de Impacto del Programa Oportunidades*, edited by B. Hernández-Prado and

M. Hernández-Ávlia, 165-227. Mexico D.F.: Instituto Nacional de Salud Pública.
Todd, P. E., and K. I. Wolpin. 2006. "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, 96(5): 1384–1417.
World Bank. (2012). World Development Indicators. http://data.worldbank.org/data-catalog/world-development-indicators,

Table 1. Reference screening procedure to obtain analysis sample

Phase 1	
Total references	2,931
Duplicates	1,341
No education-related words in abstract or title	342
Total eligible references phase 1	1,248
Phase 2	
Articles that did not meet inclusion criteria	
Intervention specification (unconditional transfer, scholarships, in-kind transfers)	24
Outcomes variables not related to education	146
Research design does not meet requirements	15
Other topic or type of document (policy briefs, comments, descriptive reports, reviews, etc.)	1,015
Total ineligible references	1,200
Phase 3	
Old version of an eligible paper	6
Total eligible references	**42**

Notes: See text for additional details of search procedure, and inclusion/exclusion criteria.

Table 2. Characteristics of CCT Programs in analysis sample

	Freq	%	N	Min	Max
Total number of programs	19	100			
Region [a]					
Latin America	12	63.2			
Asia	6	31.6			
Africa	1	5.3			
Education conditionality requirements					
School attendance	13	68.4			
Grade promotion or achievement	6	31.6			
Minimum school attendance for transfer receipt [b] (mean, SD)	84.1	.06	17	75	95
Verification of school attendance					
Yes	16	84.2			
No	2	10.3			
No information reported	1	5.3			
Payment frequency					
Monthly	10	52.6			
Bimonthly	5	26.3			
Other	4	21.1			
Monthly average subsidy amount as a % of PPP- adjusted GDP per capita (mean, SD)					
Primary	2.3	2.0	13	0.4	6.9
Secondary	4.2	4.3	17	0.8	17.3
School subsidy amount varies by					
Gender	3	15.8			
Grade or age	3	15.8			
None	11	57.9			
Other [c]	2	10.5			
Only mother eligible to receive transfer					
Yes	5	26.3			
No	11	57.9			
No information reported	3	15.8			
Supply incentives for education					
Yes	4	21.1			
No	14	73.7			
No information	1	5.3			
Type of assignment to conditions					
Random	6	31.5			
Non-random	13	68.4			
Nature of the control group					
Receives nothing from program	15	79.0			
Wait list, delayed entry	4	21.0			
Yearly cost per intended beneficiary in 2011 US Dollars (mean, SD)	80.6	40.3	13	16.7	143.2

[a] Programs that have changed their name are counted as different program because in some cases these changes were accompanied with changes in the program (these cases are: Bolsa Familia and Bolsa Escola in Brazil; Progresa and Oportunidades in Mexico; and CESSP and JFPR Scholarship Program in Cambodia.
[b] Percentage of time in school (month, every two months or school year).
[c] Dropout risk (CESSP program in Cambodia) and random (CCT for Schooling program in Malawi).

Table 3. Characteristics of references in analysis sample

Total number of references	42	
Publication type		
Journal article	19	45.2
Working paper	13	31.0
Government/technical reports	7	16.7
Book chapter	1	2.4
Thesis or doctoral dissertation	2	4.8
Source of data		
Program survey	32	76.2
National household survey	3	7.1
Census data	4	9.5
Other	3	7.1
Reports effects on		
Enrollment	28	66.7
Attendance	19	45.2
Dropout	9	21.4

See notes to Table 1 for reference screening procedure and Appendix Tables A and B for reference details.

.

Table 4. Meta-regression of effect size and cost-effectiveness estimates' moderators

	Dependent variable is:			
	Standardized Effect Size	Standardized Effect Size	Cost Data Available (1=yes)	Standardized Effect Size Divided by Yearly Project Cost per Intended Beneficiary
	(1)	(2)	(3)	(4)
Contextual Characteristics				
Baseline school enrollment	-4.537	-4.068	2.461	-0.009
	(1.524)**	(1.384)**	(0.665)**	(0.019)
	[0.007]	[0.008]	[0.001]	[0.646]
Latin America (1=yes)	0.753	0.583	-0.202	-
	(0.376)+	(0.328)+	(0.383)	
	[0.056]	[0.089]	[0.600]	
Average monthly subsidy as percent of per-capita GDP (PPP)	0.045	0.026	-0.005	0.001
	(0.052)	(0.048)	(0.028)	(0.002)*
	[0.403]	[0.594]	[0.868]	[0.018]
Program Characteristics				
Random assignment to conditions	-0.503	-0.375	-0.267	-0.019
	(0.422)	(0.368)	(0.233)	(0.006)**
	[0.245]	[0.319]	[0.258]	[0.003]
Supply-side complement (1=yes)	1.009	0.819	0.415	0.021

	(0.345)**	(0.290)**	(0.195)*	(0.006)**
	[0.007]	[0.010]	[0.039]	[0.003]
Payment frequency (1=monthly, 0 less frequently)	-0.486	-0.370	-0.121	-0.0005
	(0.260)+	(0.201)+	(0.111)	(0.002)
	[0.073]	[0.080]	[0.283]	[0.834]
Transfer continuation conditional on achievement (1=yes)	0.380	0.314	-0.457	-0.004
	(0.426)	(0.362)	(0.392)	(0.004)
	[0.381]	[0.396]	[0.252]	[0.356]
Only mother eligible to receive transfer (1=yes)	-0.070	0.007	-0.167	-0.001
	(0.320)	(0.281)	(0.093)+	(0.004)
	[0.830]	[0.982]	[0.079]	[0.796]
Minimum school attendance required for schooling transfer receipt	6.189	5.137	1.060	0.046
	(1.675)**	(1.418)**	(1.514)	(0.060)
	[0.001]	[0.002]	[0.488]	[0.451]
Number of Level 1 (Effects) Observations	74	70	74	60
Number of Level 2 (References) Observations	39	37	39	31

Notes: In all columns standard errors are adjusted for hierarchical dependence (clustering) of estimates at the reference level in parentheses and corresponding p-values in brackets. Estimates in column (1) are from variance-weighted Method of Moments estimation of regression equation (7) in text and use the full sample. Estimates in column (2) are from variance-weighted Method of Moments estimation of regression equation (7) in text and exclude estimates from Cambodia's JFPR Program from Filmer and Schady (2008) and from Nicaragua's Red de Protección Social Program from Dammert (2009). Estimates in column (3) are from an OLS regression and use the full sample. Estimates in column (4) are from variance-weighted Method of Moments estimation of regression equation (7) in text using the full sample. In Column (4) the Latin America indicator is dropped because it is perfectly collinear with transfer continuation conditional on achievement. Baseline net school enrollment is from the World Development Indicators data

source for the year the program began in a given country or the closest year available if data is not available for the year the program began. For primary school outcomes baseline enrollment is net primary enrollment. For secondary outcomes baseline enrollment is net secondary enrollment. All columns include schooling level/outcome cell indicators in addition to the reported coefficients.
** significant at 1% level, * significant at 5% level, + significant at 10% level.

Table 6. Egger's linear regression tests for publication bias and selective reporting

	Primary Enrollment (1)	Secondary Enrollment (2)	Primary Attendance (3)	Secondary Attendance (4)	Primary Dropout (5)	Secondary Dropout (6)
Constant (Asymmetry)	1.67	0.24	0.96	4.45	-3.27	-2.00
Standard Error	(0.29)	(0.23)	(0.39)	(0.45)	(0.85)	(1.26)
p-value	0.00	0.29	0.01	0.00	0.00	0.12
Number of Estimates	187	258	86	131	72	31

Notes: Each column reports estimates from a different regression in which the effect size divided by its standard error is regressed against the standard error and a constant term. In each column, we use all the effect estimates reported in all references reporting estimates for a given outcome-schooling level combination.

Figure 1. Distribution of effects reported in each reference in sample, by outcome and level

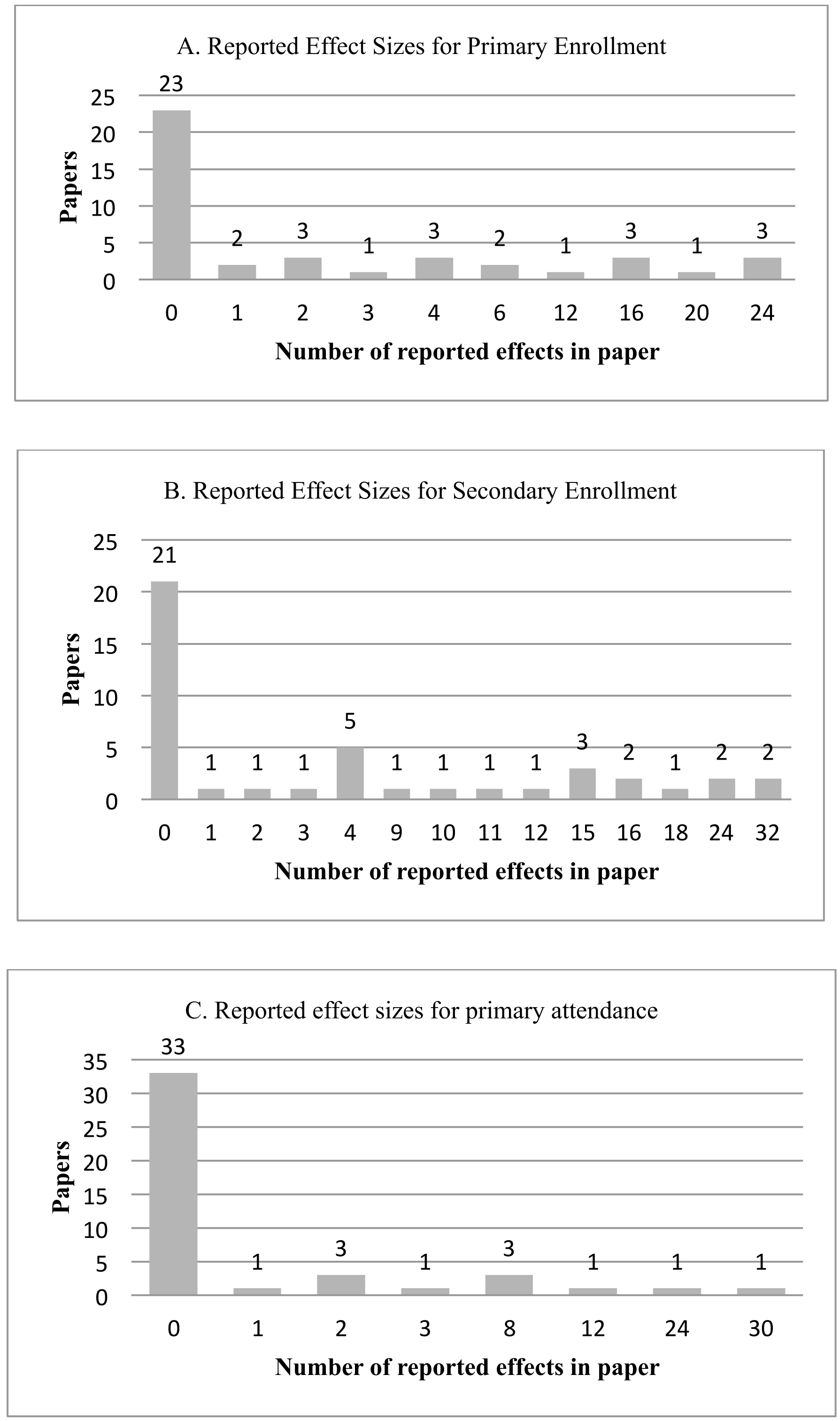
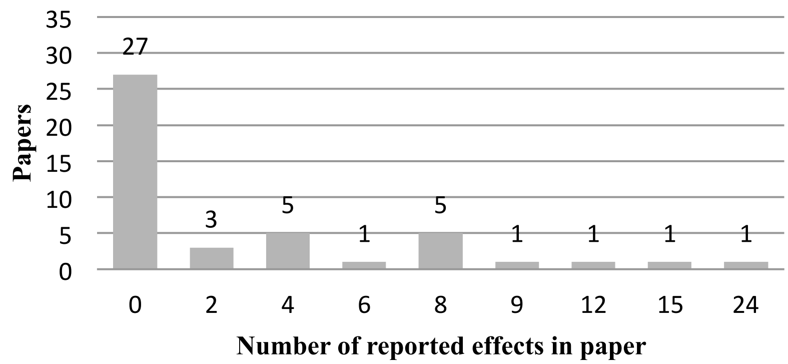
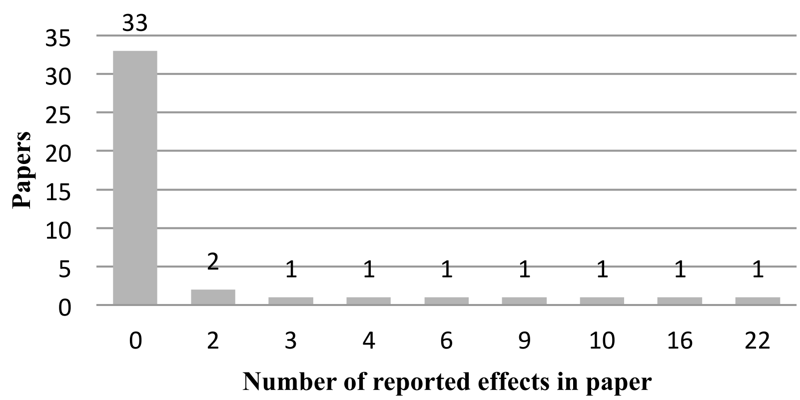


Figure 1. (cont.)

D. Reported effect sizes for secondary attendance



E. Reported effect sizes for primary dropout



F. Reported effect sizes for secondary dropout

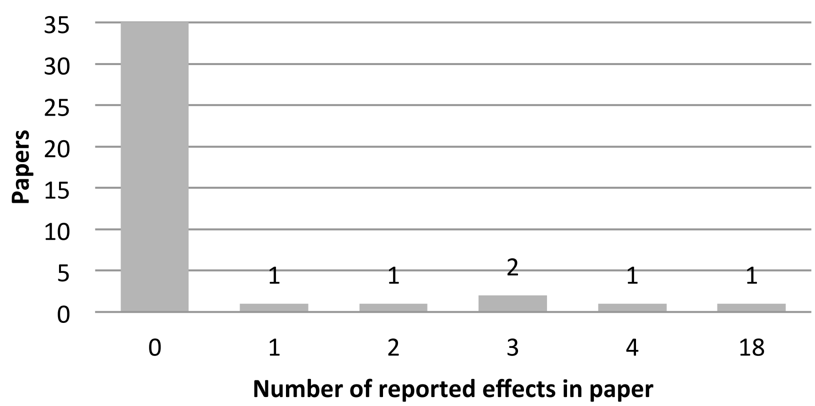
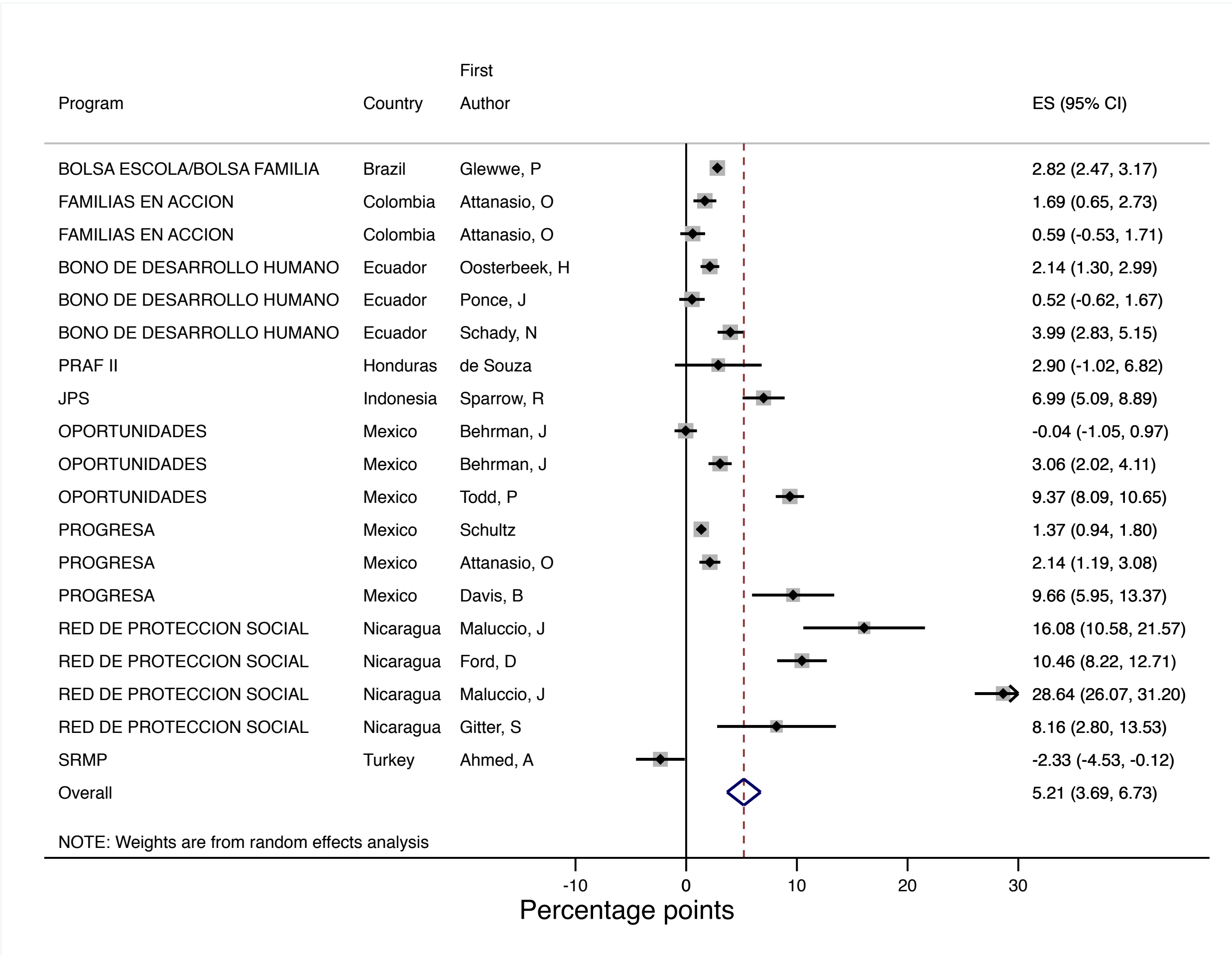
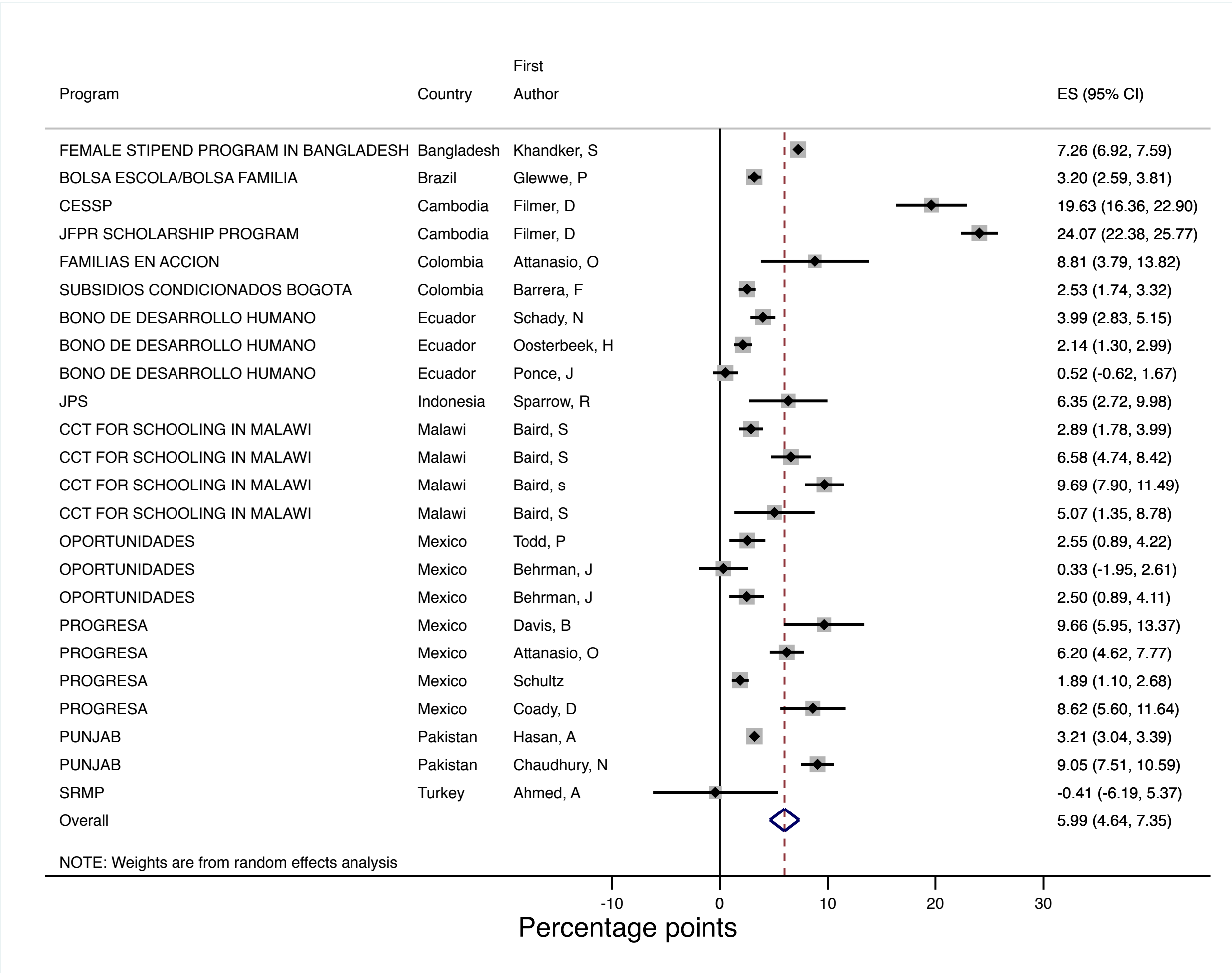


Figure 2a. Forest plot of precision-weighted impact estimates for primary enrollment



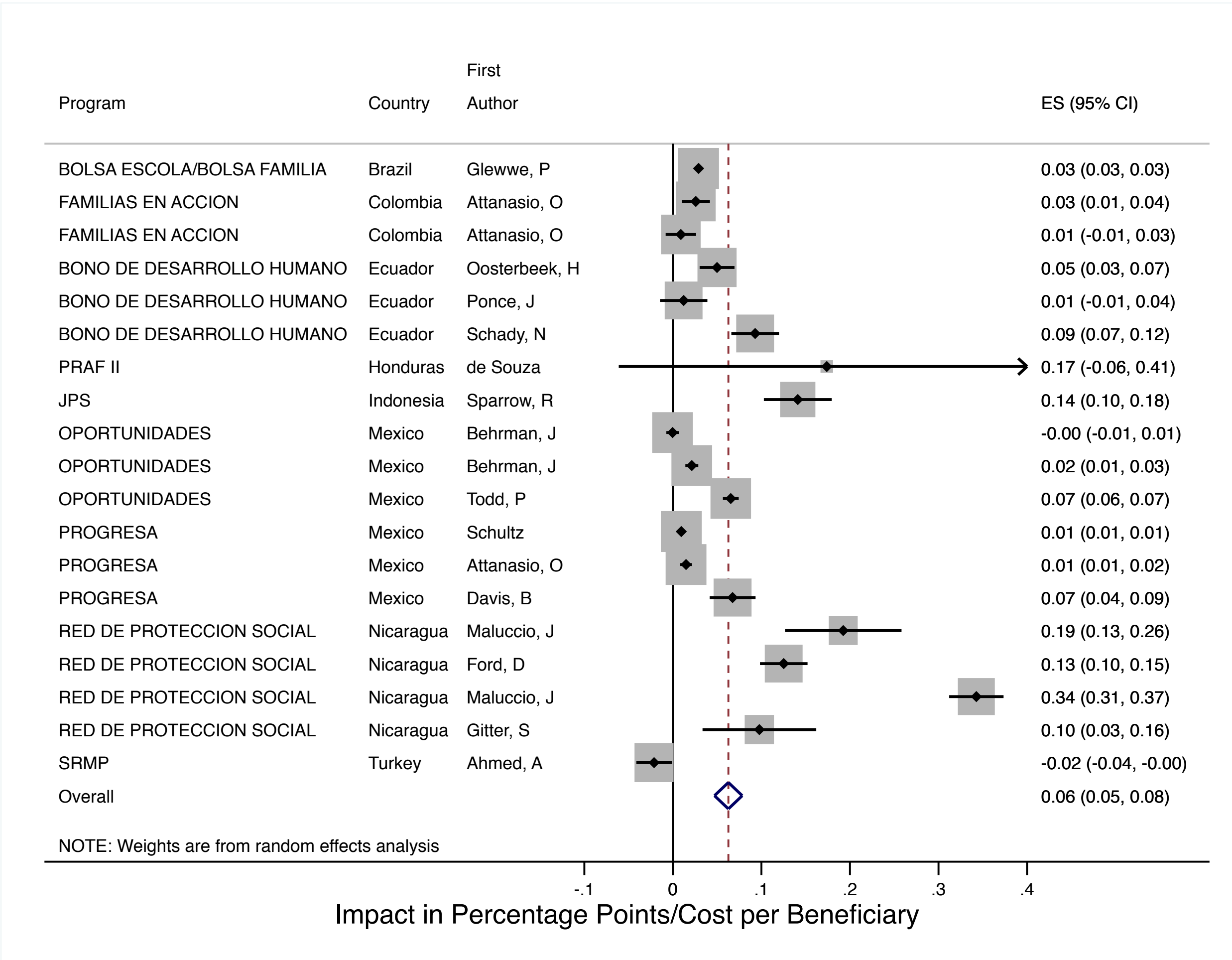
Notes: For each reference we compute the precision-weighted impact estimate and its variance using formulas (1) and (2) in the text. The overall impact estimate is from a random-effects method of moments model. The chi-square test statistic for the null hypothesis of homogeneity in primary enrollment impact estimates in the random effects model is 727 (p-value 0.000). Mean baseline primary enrollment from the World Development Indicators data source for the year the program began or closest available is 84%.

Figure 2b. Forest plot of precision-weighted impact estimates for secondary enrollment



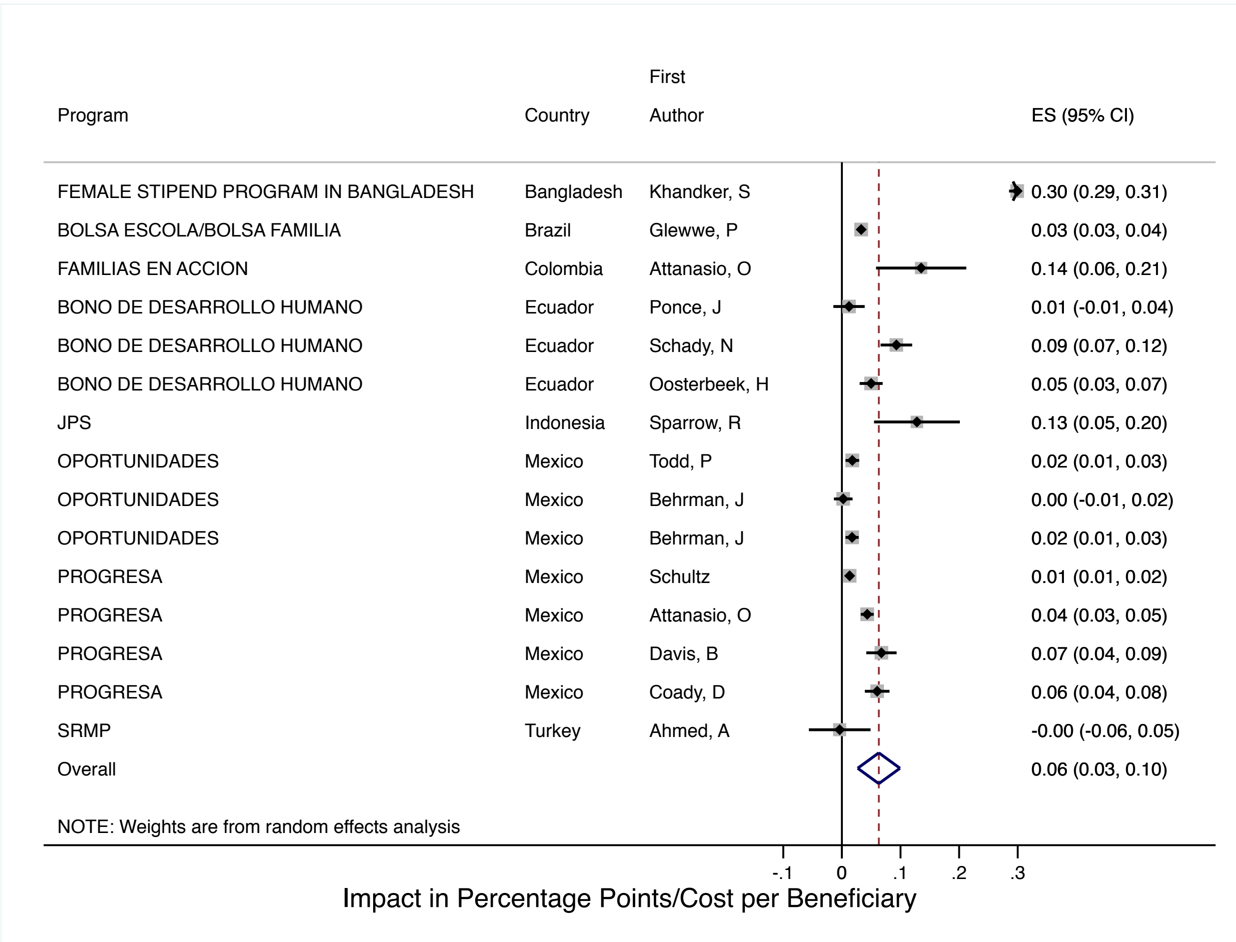
Notes: For each reference we compute the precision-weighted impact estimate and its variance using formulas (1) and (2) in the text. The overall impact estimate is from a random-effects method of moments model. The chi-square test statistic for the null hypothesis of homogeneity in secondary enrollment impact estimates in the random effects model is 1,300 (p-value 0.000). Mean baseline secondary enrollment from the World Development Indicators data source for the year the program began or closest available is 59%.

Figure 3a. Forest plot of precision-weighted cost-effectiveness estimates for primary enrollment



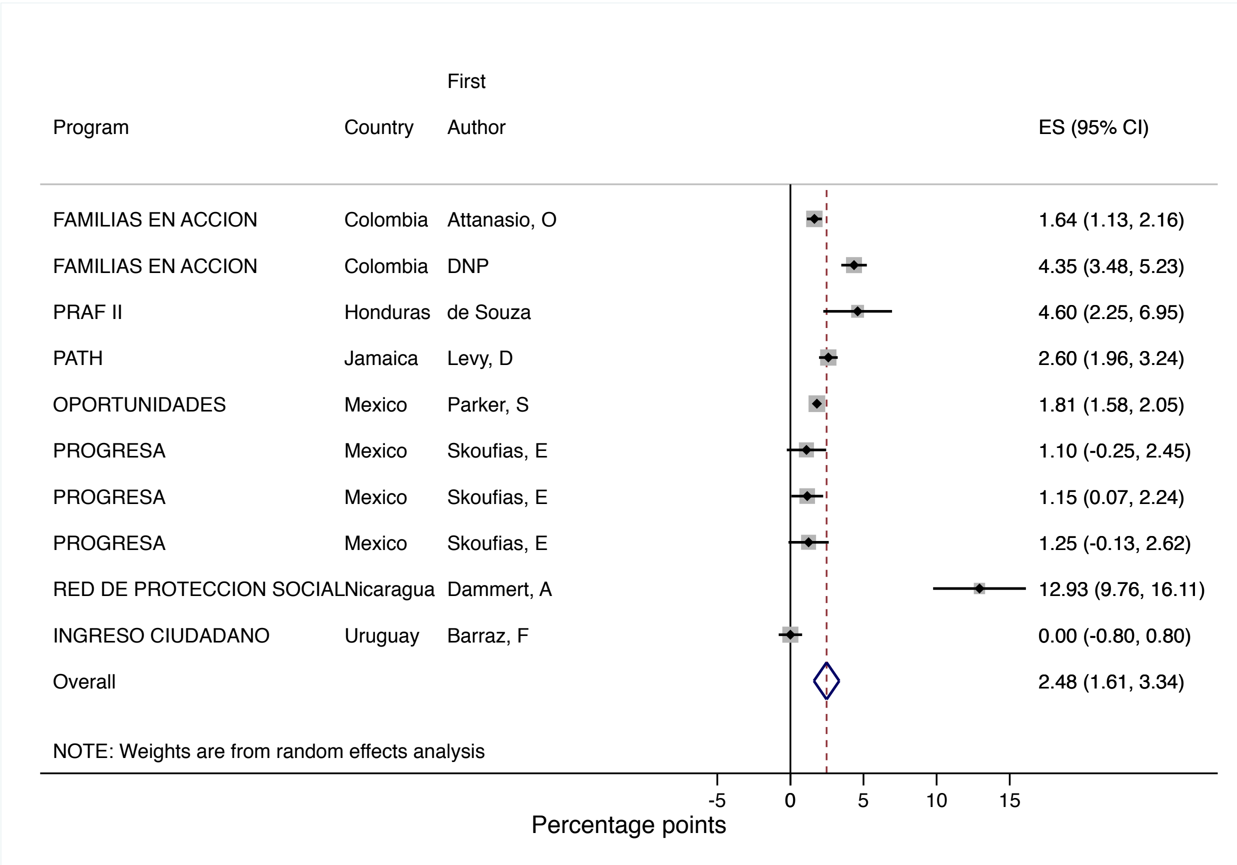
Notes: For each reference we compute the precision-weighted cost effectiveness estimate dividing the impact estimate by the yearly program cost per intended beneficiary in US dollars of 2011. We compute the standard error of the cost-effectiveness estimate using the delta method. The overall cost-effectiveness estimate is from a random-effects method of moments model. The chi-square test statistic for the null hypothesis of homogeneity in cost-effectiveness estimates in the random effects model is 826 (p-value 0.000).

Figure 3b. Forest plot of precision-weighted cost-effectiveness estimates for secondary enrollment



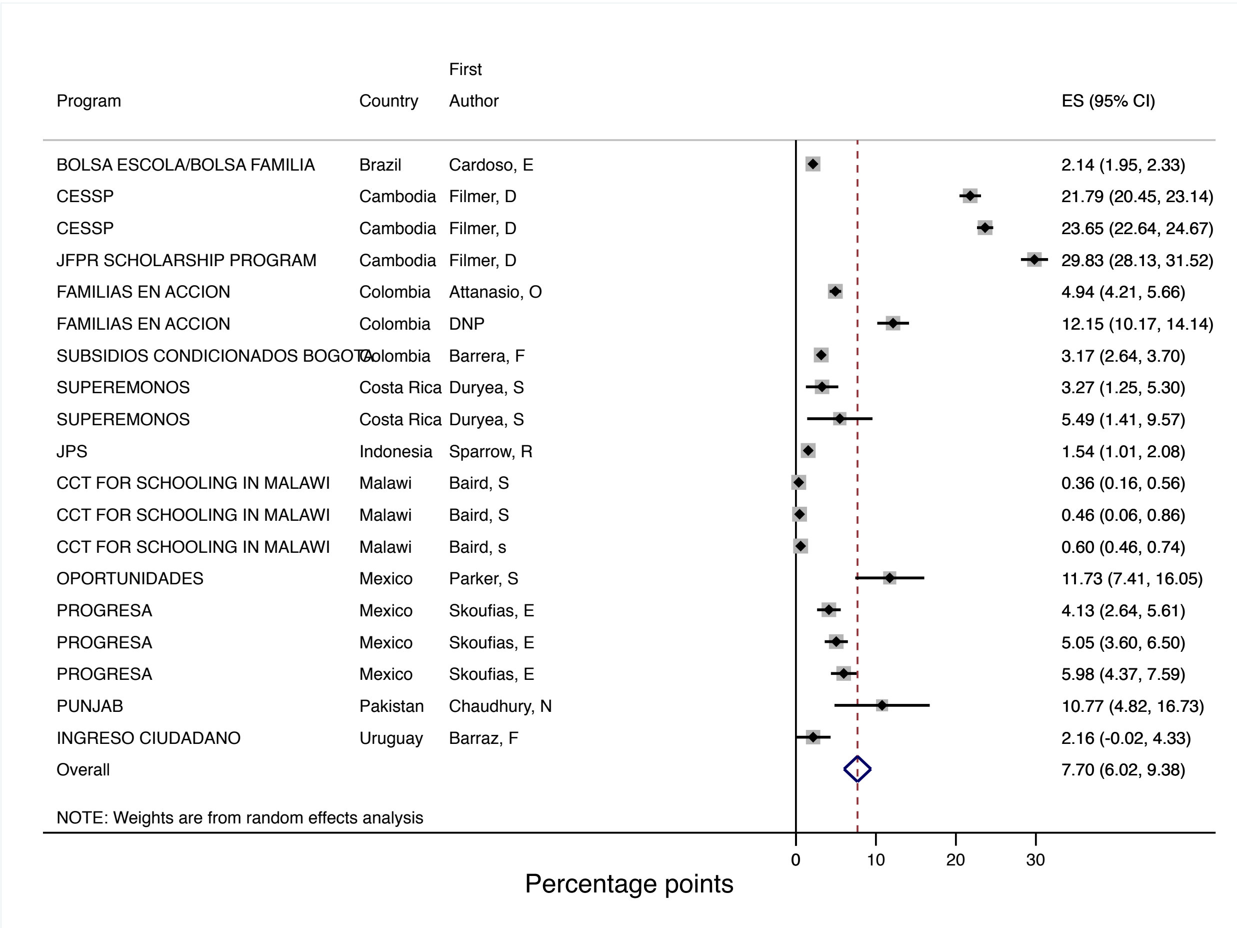
Notes: For each reference we compute the precision-weighted cost effectiveness estimate dividing the impact estimate by the yearly program cost per intended beneficiary in US dollars of 2011. We compute the standard error of the cost-effectiveness estimate using the delta method. The overall cost-effectiveness estimate is from a random-effects method of moments model. The chi-square test statistic for the null hypothesis of homogeneity in cost-effectiveness estimates in the random effects model is 1,519 (p-value 0.000).

Figure 4a. Forest plot of precision-weighted impact estimates for primary attendance



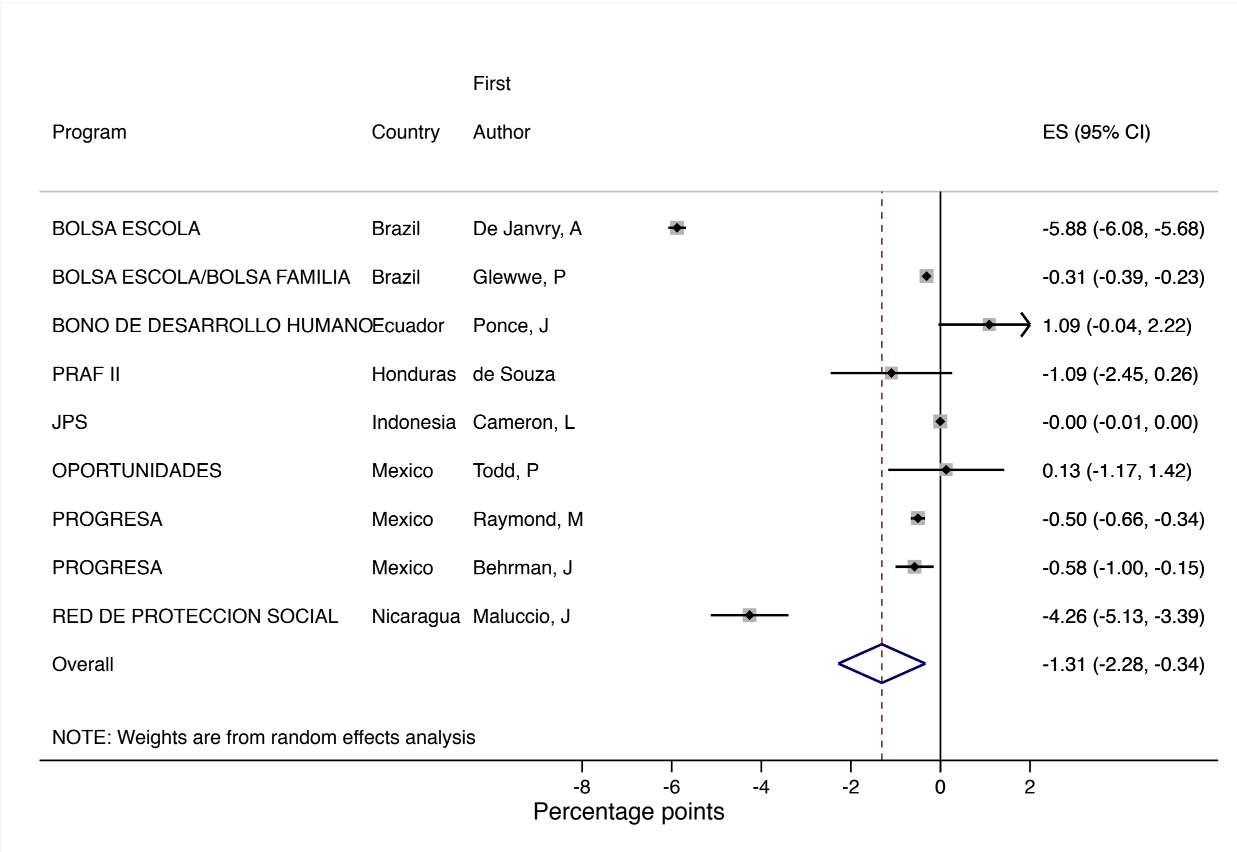
Notes: For each reference we compute the precision-weighted impact estimate and its variance using formulas (1) and (2) in the text. The overall impact estimate is from a random-effects method of moments model. Skoufias and Parker (2001) (First Author Skoufias for Progresa) reports effects for three post-treatment periods and we compute one effect size per measurement period. The chi-square test statistic for the null hypothesis of homogeneity in primary attendance effect size estimates in the random effects model is 113 (p-value 0.000). Mean baseline primary attendance computed from studies in the sample reporting it is 80%.

Figure 4b. Forest plot of precision-weighted impact estimates for secondary attendance



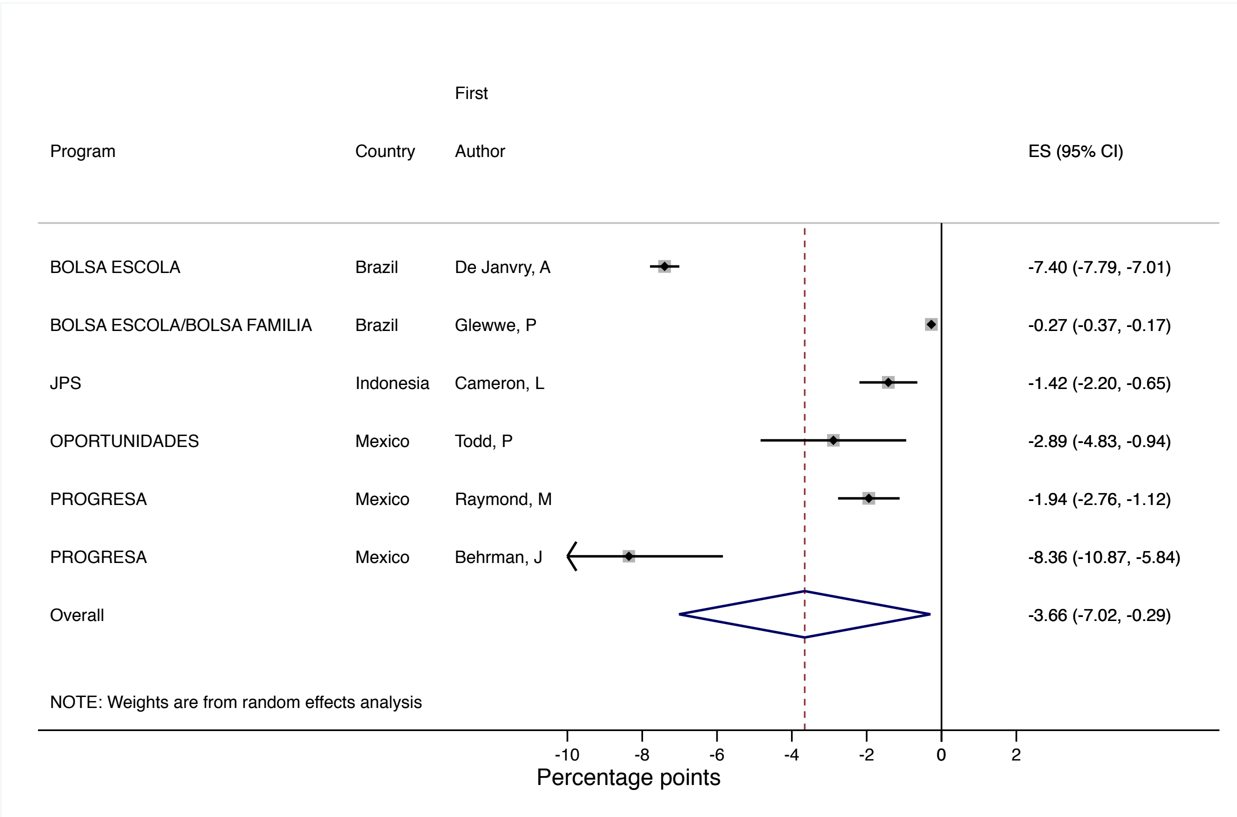
Notes: For each reference we compute the precision-weighted impact estimate and its variance using formulas (1) and (2) in the text. The overall impact estimate is from a random-effects method of moments model. Duryea and Morrison (2004) reports effects for two post-treatment periods and we compute one effect size per measurement period. Skoufias and Parker (2001) (First Author Skoufias for Progresa) reports effects for three post-treatment periods and we compute one effect size per measurement period. The chi-square test statistic for the null hypothesis of homogeneity in secondary attendance effect size estimates in the random effects model is 4,470 (p-value 0.000).Mean baseline secondary attendance computed from studies in the sample reporting it is 68%.

Figure 5a. Forest plot of precision-weighted impact estimates for primary dropout



Notes: For each reference we compute the precision-weighted impact estimate and its variance using formulas (1) and (2) in the text. The overall impact estimate is from a random-effects method of moments model. The chi-square test statistic for the null hypothesis of homogeneity in primary dropout effect size estimates in the random effects model is 3603 (p-value 0.000).

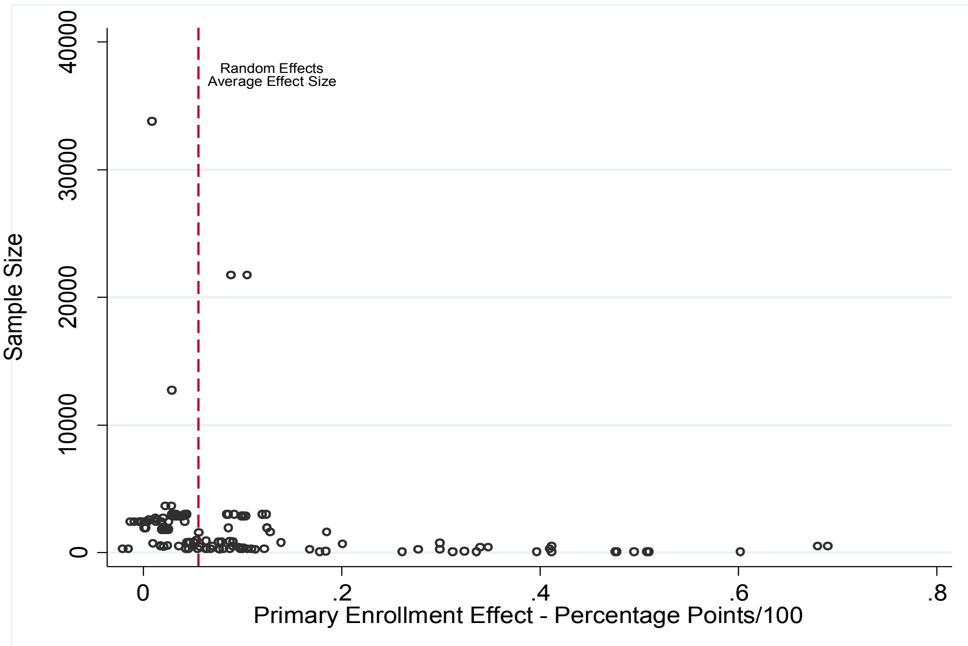
Figure 5b. Forest plot of precision-weighted impact estimates for secondary dropout



Notes: For each reference we compute the precision-weighted impact estimate and its variance using formulas (1) and (2) in the text. The overall impact estimate is from a random-effects method of moments model. The chi-square test statistic for the null hypothesis of homogeneity in secondary dropout effect size estimates in the random effects model is 1238 (p-value 0.000).

Figure 6. Funnel plot of sample size on reported enrollment impact estimate (all estimates)

a. Primary enrollment



b. Secondary enrollment

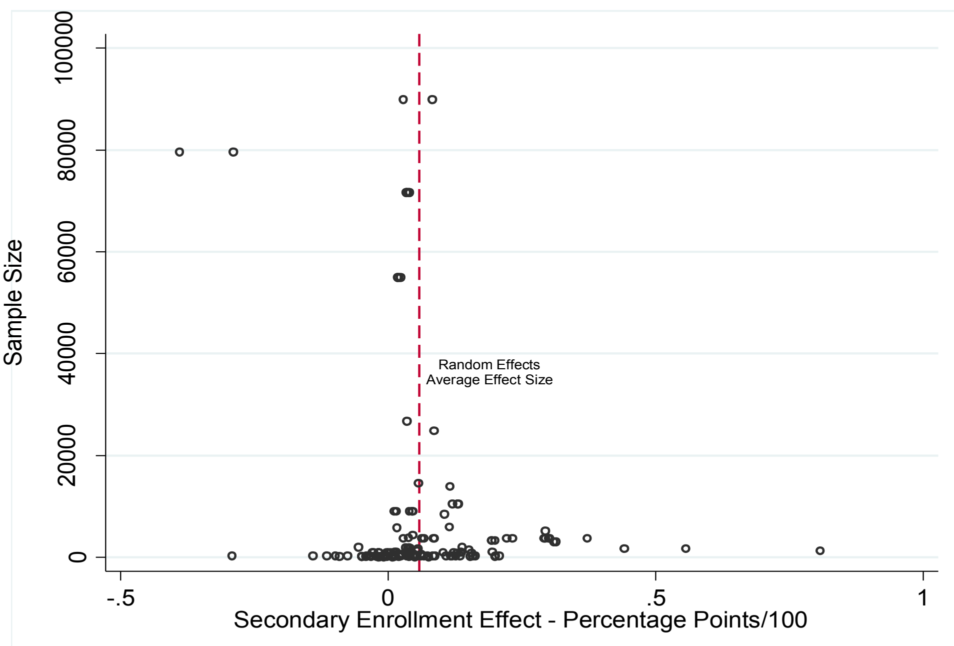
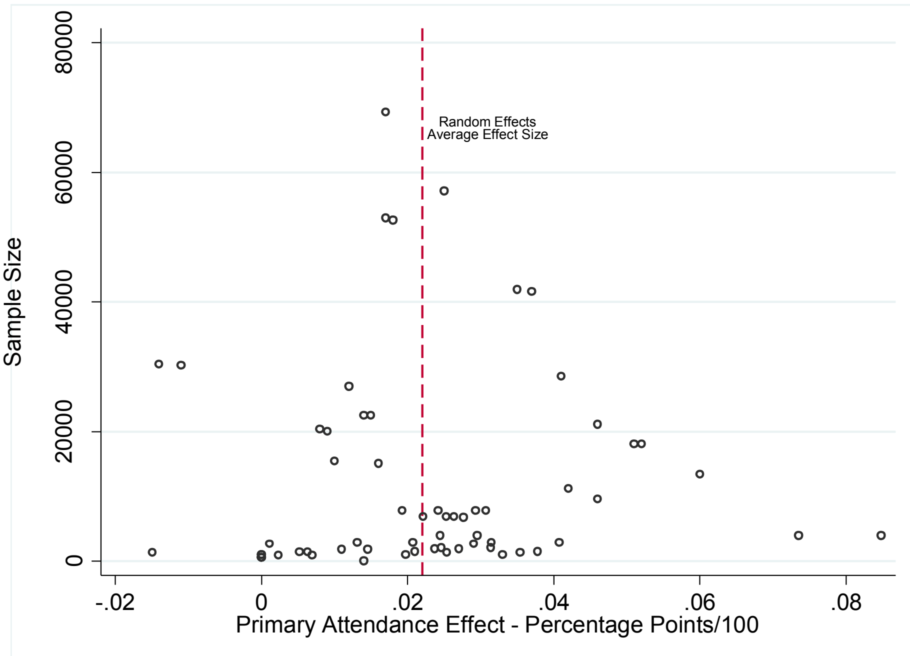


Figure 7. Funnel plot of sample size on reported attendance impact estimate (all estimates)

a. Primary attendance



b. Secondary attendance

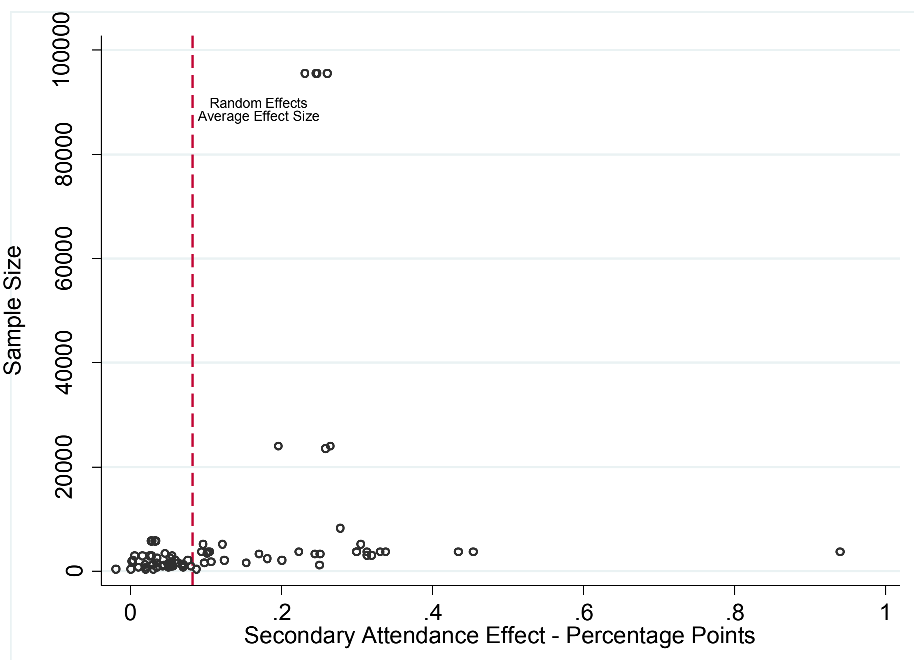
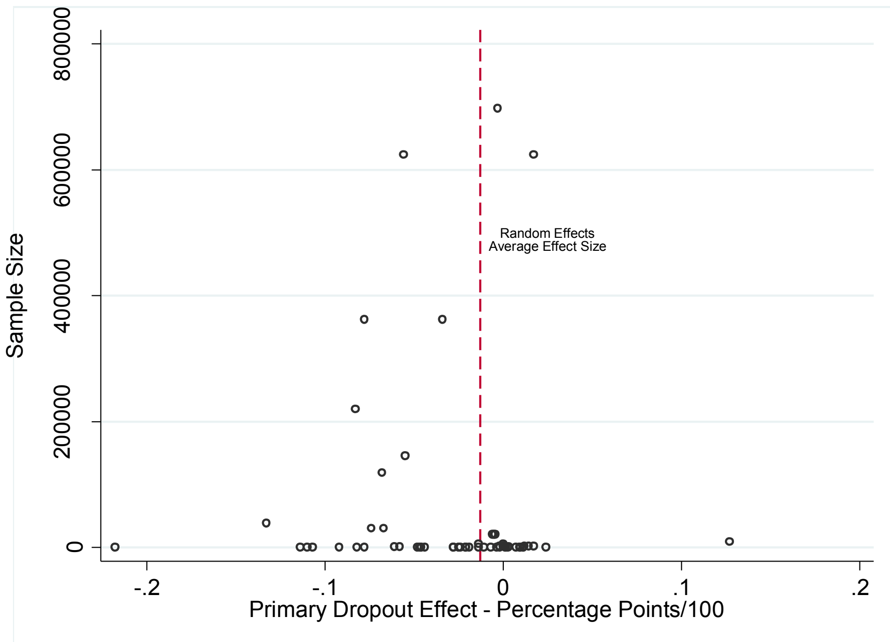
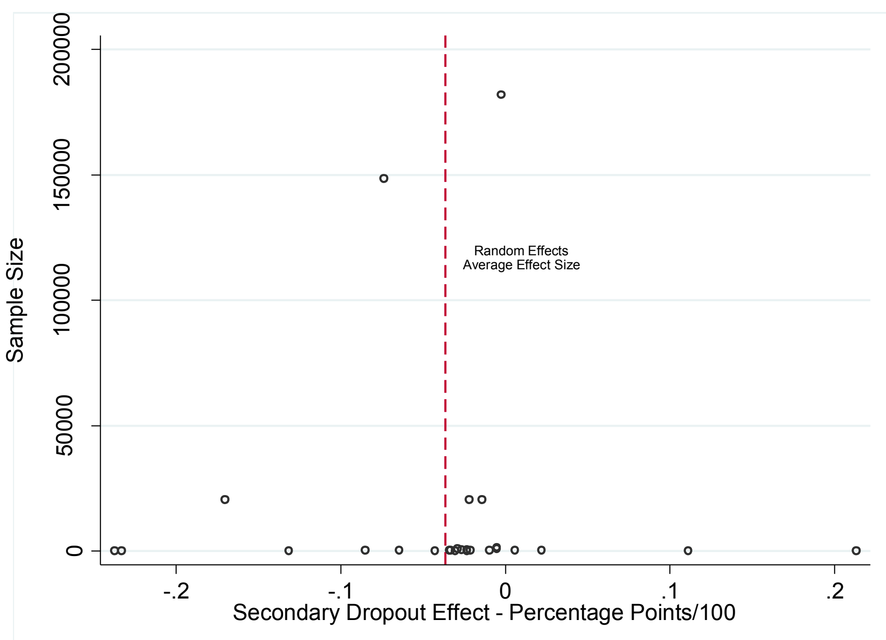


Figure 8. Funnel plot of sample size on reported dropout impact estimate (all estimates)

a. Primary dropout



b. Secondary dropout



Appendix Table A (NOT FOR PUBLICATION). Characteristics of references in final analysis sample

Country	Program name	First author	Year	Publication type	Source of data	Sample size[a]	Reports effects on Enrollment	Attendance	Dropout
Bangladesh	Female Stipend Program	Khandker, S.	2003	Working paper	Household survey and school data	89,861	Yes	No	No
Brazil	Bolsa Escola	De Janvry, A.	2006	Working paper	Administrative data	624,077	No	No	Yes
Brazil	Bolsa Escola/Bolsa Familia	Glewwe, P.	2012	Journal article	Census data	699,255	Yes	No	Yes
Brazil	PETI/Bolsa Escola/Renda Minima	Cardoso, E.	2004	Working paper	Census data	428,740	No	Yes	No
Cambodia	CESSP	Filmer, D.	2011	Journal article	Program survey	95,493	No	Yes	No
Cambodia	CESSP	Filmer, D.	2009	Working paper	Program survey	3,225	Yes	Yes	No
Cambodia	JFPR Scholarship Program	Filmer, D.	2008	Journal article	Program survey	5,138	Yes	Yes	No
Colombia	Familias en Acción	Attanasio, O.	2010	Journal article	Program survey	3,648	Yes	No	No
Colombia	Familias en Acción	Attanasio, O.	2004	Technical Report	Program survey	3,935	No	Yes	No
Colombia	Familias en Acción	Attanasio, O.	2004	Government report	Program survey	2,691	Yes	No	No
Colombia	Familias en Acción	National Planning Department	2006	Government report	Program survey	3,935	No	Yes	No
Colombia	Subsidios Condicionados a la Asistencia Escolar en Bogotá	Barrera, F.	2011	Journal article	Program survey	8,980	Yes	Yes	No
Costa Rica	Superémonos	Duryea, S.	2004	Working paper	Program survey	1,109	No	Yes	No
Ecuador	Bono de Desarrollo Humano	Oosterbeek, H.	2008	Working paper	Program survey	3,004	Yes	No	No
Ecuador	Bono de Desarrollo Humano	Ponce, J.	2006	Working paper	Program survey	2,384	Yes	No	Yes
Ecuador	Bono de Desarrollo Humano	Schady, N.	2008	Journal article	Program survey	2,875	Yes	No	No
Honduras	PRAF II	De Souza	2005	Doctoral dissertation	Program survey	12,741	Yes	Yes	Yes
Indonesia	JPS	Cameron, L.	2009	Journal article	National household survey	5,358	No	No	Yes
Indonesia	JPS	Sparrow, R.	2007	Journal article	National household survey	120,022	Yes	Yes	No
Jamaica	PATH	Levy, D.	2007	Technical report	Program survey	7,751	No	Yes	No

Country	Program name	First author	Year	Publication type	Source of data	Sample size[a]	Reports effects on		
							Enrollment	Attendance	Dropout
Malawi	CCT for Schooling	Baird, S.	2009	Working paper	Program Survey	5,914	Yes	Yes	No
Malawi	CCT for Schooling	Baird, S.	2011	Journal article	Program survey	2,023	Yes	Yes	No
Mexico	Oportunidades	Behrman, J.	2012	Journal article	Program survey	1,796	Yes	No	No
Mexico	Oportunidades	Behrman, J.	2004	Technical report	Program survey	1,013	Yes	No	No
Mexico	Oportunidades	Parker, S.	2006	Working paper	Program survey	69,261	No	Yes	No
Mexico	Oportunidades	Todd, P.	2005	Technical report	Program survey	1,994	Yes	No	Yes
Mexico	Progresa	Attanasio, O.	2012	Journal article	Program survey	N/A	Yes	No	No
Mexico	Progresa	Behrman, J.	2005	Journal article	Program survey	75,000	No	No	Yes
Mexico	Progresa	Coady, D.	2004	Journal article	Program survey	N/A	Yes	No	No
Mexico	Progresa	Davis, B.	2002	Working paper	Program survey	21,709	Yes	No	No
Mexico	Progresa	Raymond, M.	2003	Working paper	Program survey	20,541	No	No	Yes
Mexico	Progresa	Schultz, P.	2004	Journal article	Program survey	33,795	Yes	No	No
Mexico	Progresa	Skoufias, E.	2009	Book chapter	Program survey	27,845	No	Yes	No
Nicaragua	Red de Protección Social	Dammert, A.	2009	Journal article	Program survey	1,745	No	Yes	No
Nicaragua	Red de Protección Social	Ford, D.	2007	Doctoral dissertation	Program survey	1,946	Yes	No	No
Nicaragua	Red de Protección Social	Gitter, S.	2009	Journal article	Program survey	1,561	Yes	No	No
Nicaragua	Red de Protección Social	Maluccio, J.	2010	Journal article	Program survey	1,227	Yes	No	Yes
Nicaragua	Red de Protección Social	Maluccio, J.	2005	Technical report	Program survey	1,594	Yes	Yes	No
Pakistan	PUNJAB	Chaudury, N.	2010	Journal article	Census data	5,164	Yes	Yes	No
Pakistan	PUNJAB	Hasan, A.	2010	Working paper	Census data	71,620	Yes	No	No
Turkey	SRMP	Ahmed, A.	2006	Working paper	Program survey	2,905	Yes	No	No
Uruguay	Ingreso Ciudadano	Barraz, F.	2009	Journal article	National household survey	1,011	No	Yes	No

[a]Maximum sample size to compute effect sizes or sample size reported in the text (if no sample size reported in effect sizes results).

Appendix Table B (NOT FOR PUBLICATION). Programs characteristics

Country	Program name	Year program started	Conditionality	Minimum attendance rate (%)	Conditions verification	Transfer amount[a]		Payment frequency	Only mother receives the pay-ment	Subsidy varies by	Supply compo-nent	Rando m Assign-ment
						Primary	Secondary					
Bangladesh	Female Stipend Program	1994	Attendance, academic proficiency and remain unmarried	75	Yes	Not applicable	1.42	Monthly	No	Grade	Yes	No
Brazil	Bolsa Escola	2001	Attendance	85	Yes	0.77	0.77	Monthly	N/A	None	No	No
Brazil	Bolsa Escola/Bolsa Familia	1995	Enrollment and attendance	85	Yes	1.05	1.05	Monthly	Yes	None	N/A	No
Cambodia	CESSP	2005	Enrollment, attendance and grade promotion	95	Yes	Not applicable	8.95	3 times per year	No	Dropout risk	No	No
Cambodia	JFPR Scholarship Program	2004	Enrollment, attendance and grade promotion	95	Yes	Not applicable	10.01	3 times per year	No	None	No	No
Colombia	Familias en Acción	2001	Enrollment and attendance	80	Yes	1.10	2.21	Bimonthly	Yes	Age	No	No
Colombia	Subsidios Condicionados a Asistencia Escolar en Bogotá	2005	Attendance, grade promotion, graduation and enrollment in higher education institution	80	Yes	Not applicable	2.46	Bimonthly plus lump-sum at the end of school year or upon graduation[b]	No	None	No	Yes
Costa Rica	Superémonos	2001	Enrollment and attendance	N/A	Yes	4.47	4.47	Monthly	N/A	None	No	No
Ecuador	Bono de Desarrollo Humano	2004	Enrollment and attendance	90	No	3.08	3.08	Monthly	No	None	No	Yes
Honduras	PRAF II	1998	Enrollment and attendance	85	No	2.06	Not applicable	Monthly	No	None	Yes	Yes

Country	Program name	Year program started	Conditionality	Minimum attendance rate (%)	Conditions verification	Transfer amount[a]		Payment frequency	Only mother receives the pay-ment	Subsidy varies by	Supply compo-nent	Rando m Assign-ment
						Primary	Secondary					
Indonesia	JPS	1998	Enrollment and passing grades	85	N/A	0.39	0.98	3 times per year	No	Grade	No	No
Jamaica	PATH	2001	Attendance	85	Yes	1.11	1.11	Bimonthly	No	None	No	No
Malawi	CCT for schooling	2007	Enrollment and attendance	75	Yes	Not applicable	17.3	Monthly	No	Randomly	No	Yes
Mexico	Oportunidades	2002	Attendance	85	Yes	1.21	3.92	Bimonthly	Yes	Gender and grade	No	No
Mexico	Progresa	1997	Attendance	85	Yes	1.05	2.49	Monthly	Yes	Gender and grade	Yes	Yes
Nicaragua	Red de Protección Social	2000	Enrollment and attendance	85	Yes	5.23	Not applicable	Bimonthly	No	None	Yes	Yes
Punjab	Pakistan	2004	Attendance	80	Yes	Not applicable	2.28	Monthly	No	None	No	No
Turkey	SRMP	2004	Attendance and not repeating a grade more than once	80	Yes	1.56	2.62	Bimonthly	Yes	Gender and grade	No	No
Uruguay	Ingreso Ciudadano	2005	Enrollment and attendance	N/A	Yes	6.94	6.94	Monthly	N/A	None	No	No

[a] As percentage of PPP-adjusted GDP/capita.

[b] This program was part of an experiment that included 3 different treatments that varied in the timing of subsidy delivery: (1) a subsidy with bimonthly payments conditioned on attendance, (2) subsidy with bimonthly payments conditioned on attendance and a lump sum at the end of the school year conditioned on school enrollment the following year, and (3) a subsidy with bimonthly payments conditioned on attendance and a lump sum upon graduation and enrollment in a higher education institution.

[c] The program included two transfers: one to the household and another one to the student (girl).