Impacts of Conditional Cash Transfer Programs on Educational Outcomes in Developing Countries: A Meta-analysis

Juan Esteban Saavedra*

Sandra García**

To be presented at the Association for Public Policy Analysis and Management Annual Research Conference, Baltimore, MD November 8-10

* 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.

Abstract

We meta-analyze enrollment, attendance and dropout effect estimates from forty-two CCT program evaluations in fifteen developing countries. Average effect size estimates for all outcomes in primary and secondary schooling are statistically different from zero, with considerable heterogeneity. CCT programs are most effective—all else constant—at improving enrollment and attendance when baseline enrollment is relatively low. Estimates are statistically larger for programs that in addition to transfers to families also provide supply-side complements—such as infrastructure or additional teachers— and for programs that condition transfer continuation on achievement. We find evidence in support of publication bias and selective reporting.

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; Rawlings and Rubio, 2005) conclude that, on the whole, these programs have positive effects on schooling (enrollment, attendance, dropout) and health (vaccinations, medical check-up) 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 effects of CCT programs on schooling outcomes nor attempts to statistically understand the factors and program characteristics that mediate heterogeneity in program effect estimates. The closest available study in scope is Manley, Gitter and Slavchevska (2011), which metaanalyzes 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 effects on educational outcomes, and shed light on which factors mediate heterogeneity in treatment effects.

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 estimates on school enrollment, attendance and dropout 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 effect size 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 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, in press).

We do not find evidence that characteristics such as relative transfer amounts—as percent of per capita (PPP) GDP—or payment frequency are associated with larger enrollment and attendance effect estimates. Some evidence, however suggest that conditioning transfer continuation on school achievement is associated with larger effect estimates, controlling for other contextual, program and study characteristics. This latter finding is consistent with recent evidence highlighting the importance of educational incentives (see for example Duflo, Hanna and Ryan, forthcoming; Bettinger, Kremer and Saavedra, 2010; Kremer, Miguel and Thornton, 2009) and with single CCT-program evidence from Bogotá suggesting larger enrollment effects from a program variant that conditions payment on high-school graduation (Barrera-Osorio, Bertrand, Perez-Calle and Linden, 2009).

Methodologically we find evidence that suggests that—all else constant—observational evaluations yield school enrollment and attendance estimates that are larger than those from randomized evaluations. This finding, in particular, is at odds with previous qualitative evidence by IGN (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 search 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

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

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 2,921 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, 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.

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 two 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 30% of programs, however, transfers for girls differ from boys' or transfer amounts vary by grade or age.

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 is 2 percentage points. For secondary school, average transfer amount is 4% 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 30% 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.

Reference Characteristics

Table 3 shows reference-level characteristics of references in our analysis sample. Over fifty percent of references are working papers, less than 25% are journal articles and the remaining 25% are either government or technical reports

[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.

and unpublished manuscripts/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 Sizes

Universe of generalization

The most important decision in choosing the method for statistical inference in meta-analysis is the universe to which the study aims to generalize (Cooper et al., 2009). In our study we seek to make inferences beyond those CCT programs in our sample to gain knowledge about potential CCT impacts in different populations and under potentially different conditions. As such, CCT programs in our hypothetical universe of study might differ from those in our sample along three dimensions: i) study characteristics; ii) true effect size parameter, and iii) effect-size estimates due to sampling variation.

For these reasons, given that we do not hold fixed study characteristics potentially related to effect sizes, the appropriate method for statistical inference in our case is a random-effects model (Cooper, Hedges and Valentine, 2009). Formally, let T_i denote the estimated effect size of study i, with population effect size θ and *i=1,.....,k*. Under a random-effects model, T_i has a random component u_i in addition to sampling variation, e_i:

$$T_i = \theta + u_i + e_i \qquad (1)$$

The variability of T_i – v_i^* – stems from variability arising from sampling variation v_i and variation of u_i – τ_θ^2. (In a fixed effects model $\tau_\theta^2 = 0$.)

$$v_i^* = v_i + \tau_\theta^2 \qquad (2)$$

Effect size estimates

All educational outcome measures we focus on are dichotomous: enrollment, attendance and dropout. Estimates of T_i in our study are therefore either the post-treatment difference in the corresponding probability between treatment and comparison groups, $T_i = p_{it} - p_{ic}$, or a double difference (treatment v. control, before v. after). Treatment-control contrasts are a natural measure in our context, and the way in which almost all references report program effects.[3]

Combining estimates of effect sizes

We pursue two approaches to combine effects sizes. We apply both approaches to estimate average effect sizes separately for each outcome and schooling level. In the first approach, *separately for each outcome and schooling level*, we combine all of a reference effects in one reference-level average effect under a fixed-effects assumption.[4] Specifically, for each outcome and schooling level, let T_{ij} denote the j'th effect estimate $j=1,2.....,J$ of study i, v_{ij} its associated variance and $w_{ij} = 1/v_{ij}$. Then the average study-level effect estimate T_i is:

$$T_i = \frac{\sum_{j=1}^{J}(w_{ij}T_{ij})}{\sum_{j=1}^{J} w_{ij}} \qquad (3)$$

And it variance is:

$$v_i = \frac{1}{\sum_{j=1}^{J}(1/v_{ij})} \qquad (4)$$

Under a fixed-effects model, for the k studies in our sample, the overall mean effect size $\bar{T}$ is therefore:

$$\bar{T} = \frac{\sum_{i=1}^{k}(w_i T_i)}{\sum_{i=1}^{k} w_i} \qquad (5)$$

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

$$v = \frac{1}{\sum_{i=1}^{k}(1/v_i)} \qquad (6)$$

---

[3] The only exceptions are Davis et al. (2002) which reports probit coefficients that we convert to (approximate) marginal effects by dividing the probit coefficient by 2.5 (Wooldridge, 2005), and Raymond and Sadoulet (2003) that report hazard ratios for dropout that we convert to percentage points using baseline hazards for each grade.

[4] When a reference reports effects for different follow-up periods – one year and two years after baseline data collection, for instance – we we compute separate average effect sizes for each measurement period. This occurred in the case of Duryea and Morrison (2004) who report primary attendance effects for two follow-up periods and for Skoufias and Parker (2001) who report primary and secondary attendance effects for three follow-up periods.

The homogeneity test to test whether σ_θ^2, the variation of θ_i around its population mean is zero is:

$$Q = \sum_{i=1}^{k} w_i \, (T_i - \bar{T})^2 \qquad (7)$$

which under the null hypothesis of fixed-effects (i.e. H_0: τ_θ^2=0) has a chi-square distribution with $k - 1$ degrees of freedom. An unbiased estimate of τ_θ^2 is then obtained by computing:

$$\widehat{\tau_\theta^2} = [Q - (k - 1)]/c \qquad (8)$$

where,

$$c = \sum_{i=1}^{k} w_i - \left[\sum_{i=1}^{k} w_i^2 / \sum_{i=1}^{k} w_i\right]$$

We then calculate the random-effects mean of T_i, $\bar{T}_*$ and its variance v_* by computing $v_i^* = v_i + \hat{\tau}_\theta^2$ and $w_i^* = 1/v_i^*$ and using them instead of v_i and w_i in equations (5) and (6). We do this procedure separately for each outcome (enrollment, attendance, dropout) and each schooling level (primary and secondary). We estimate (5) and (6) using Method of Moments estimators.

In the second approach to estimate overall average effect sizes, separately for each outcome and school level, we take all estimates from all papers and combine them directly in a random effects model. The second approach is useful for two reasons. The second approach allows us to test the sensitivity of the average effect size estimate to the fixed effects assumption. Although the fixed-effects assumption is arguably justified in our context by the fact that all effects in a given study correspond to the same underlying population and thus share the same institutional characteristics, it is important to test how it affects our estimates.

The second approach of directly combining all effect sizes in a random effects model also allows us to test for selective reporting using funnel plots and Egger linear regression tests that we explain in detail in the "analyzing effect sizes" section of the methodology, below.

Analyzing effect size estimates

To explore how program and study characteristics explain variability in effect size estimates, we pursue the following meta-regression approach that combines enrollment and attendance estimates for all schooling levels in one model. By pooling all enrollment and attendance estimates in one model, this approach allows us to maximize statistical power, and the number of degrees of freedom. Following with notation, denote by T_{ic} the per-study fixed-effects average effect size estimate of study i for outcome school-level group c (primary enrollment, primary attendance, secondary enrollment, secondary attendance). We estimate the following mixed-effects model:

$$T_{ic} = \alpha_c + X\beta + \mu_i + \varepsilon_i \qquad (9)$$

where α_c are outcome-by-school-level group fixed effects, μ_i are random effects and ε_i is sampling error. In the vector X we include context, program and study 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 complements such as infrastructure or additional teachers, payment frequency (monthly vs. less frequently) and whether the program imposes conditions on achievement beyond the standard school attendance conditions; and the number of reported estimates. Because in model (9) we might use multiple estimates per study (for example, one for primary enrollment and one for secondary enrollment), in all specifications of model (9) we adjust standard errors for hierarchical dependence of effect estimates (i.e. clustering) at the study-level using the methods of Hedges, Tipton and Johnson (2010) for random effects meta-analysis models.

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

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

Figures 2 and 3 show the forest plots (distribution) of average effect sizes from all studies reporting enrollment effects on primary and secondary school, respectively. In all forest plot figures we report the average effect size per study, combining all estimates into one using a fixed-effects model.

We highlight three aspects of Figure 2. First, the overall random-effects average primary enrollment effect size is 5.1 percentage points, with a 95% confidence interval between 3.7 and 6.6 percentage points. Relative to the mean baseline primary enrollment of 93%, the average effect size represents a 5.5 percent enrollment increase. Second, with the exception of one reference reporting effects from the *SRMP* CCT program in Turkey, all reference-level average effects are positive and most are statistically distinguishable from zero. Third, there is ample variation in estimated effects across studies.[5]

Reference-level effect sizes for Nicaragua's *Red de Protección Social* are an exception, however, 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 5 displays the forest plot of secondary enrollment effect estimates, with one effect per study that we estimate under a fixed-effects model. The average secondary enrollment effect 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 50%, implying that the average secondary enrollment effect estimate represents a 12 percent secondary enrollment increase. The secondary enrollment plot displays considerable effect-size variation, with evaluations of programs like Cambodia's

[5] The chi-square test-statistic for the null hypothesis of homogeneity in primary enrollment effect size estimates is 735 (p-value 0.000). We obtain similar conclusions when we estimate the average effect size estimate using all references in all studies. 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 1761 (p-value 0.000).

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 Attendance Average Effect Sizes

Figure 4 displays the primary attendance effect size distribution (one effect per paper). 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. The overall primary-attendance estimate using all estimates from all studies is 2.2 percentage points, also statistically significant.

A clear outlier 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=113.4, p-value 0.000).

Figure 5 displays the secondary attendance forest plot with one effect per reference computed using a fixed effects model. The CCT average secondary

---

[6] The chi-square test-statistic for the null hypothesis of homogeneity in secondary enrollment effect size estimates is 1302 (p-value 0.000). We obtain a similar overall secondary enrollment estimate (5 percentage points) when we use all estimates from all studies, and similarly reject the null hypothesis of homogeneity in effect estimates (chi-square statistic=2409, p-value 0.000).

attendance effect is 8.1 percentage points (8 percentage points when computed with all estimates from all studies) and statistically significant. This effect represents a 12% increase in attendance relative to the average baseline secondary attendance level of 68%.

There is considerable heterogeneity in secondary attendance effects across programs.[7] At one extreme stands Cambodia's CESSP with average secondary attendance effect sizes 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 program with average secondary attendance effect sizes that although positive and statistically significant are small – close to half of a percentage point.

School Dropout Average Effect Sizes

Compared to enrollment and attendance, few CCT 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 effects.

Figure 6 shows primary dropout's forest plot. The overall average dropout effect size for primary is negative one percentage point and statistically different from zero, although the 95% confidence interval is relatively wide. Nicaragua's *Red de Protección Social* and Brazil's *Bolsa Escola* have the largest effects on dropout reduction, while the evaluation of Ecuador's *Bono de Desarrollo Humano* suggests, if any, increases in dropout as a consequence of program participation.

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. Average effect sizes are similar across different model specifications, as the overall estimate using all

[7] We reject the null hypothesis of homegeneity in effect estimates, with a homogeneity test chi-square statistic of 4050 and associated p-value of 0.000.

estimates is -2 percentage points. We reject the null hypothesis of homogeneity in primary dropout effect size estimates (chi-square statistic=3603, 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 7). Although we estimate the average secondary dropout effect with a high degree of uncertainty due to the fact that only six studies in our sample report effect estimates for this outcome, we still reject the null hypothesis that CCT programs do not affect secondary dropout outcomes. All reported secondary dropout effects 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. Estimates for secondary dropout for *Progresa*, however, differ drastically across references.[8]

Meta-regression results

Table 4 reports meta-regression estimation results of estimation equation (9). Because we are pooling estimates for various outcomes and schooling level combinations, all specifications, in addition to the reported coefficients, include outcome-by-schooling level fixed effects. In all specifications we adjust standard errors for hierarchical dependence (i.e. clustering) of effect size estimates at the study-level. We present various model specifications, beginning with a model that only includes contextual characteristics: baseline school enrollment, whether the program is in a Latin American country, and the average monthly subsidy as a percent of per capita (PPP) GDP.

In columns (3) to (7) of Table 4 we add—one at a time—program characteristics that include whether benefits are assigned at random, whether the program provides a supply-side complement such a infrastructure grants or books, payment frequency, whether transfer continuation is conditional on school

[8] We reject the null hypothesis of homogeneity in secondary dropout estimates (chi-square statistic=1238, p-value 0.000).

achievement and the year the program began. In column (8) we also control for the number of reported effect estimates in the study.

We highlight five main findings from Table 4. First, controlling for program design characteristics and geographic location, CCT programs in contexts with relatively low baseline school enrollment levels are significantly more effective at improving school enrollment and attendance outcomes. Such contexts might include settings with a large rural share of the population or secondary schooling, which is typically low in developing countries. As results in Table 4 indicate, this finding is consistent across all specifications.

Our second main finding is that enrollment and attendance estimates from studies of programs that complement cash transfers to families with supply-side interventions such as school infrastructure, additional teachers, grants or textbooks are statistically significantly larger—about four percentage points—than of programs that only provide cash transfers to families. This result is robust to controlling for contextual characteristics and other program attributes. The positive association between 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).

Our third main finding is that program design characteristics such as payment frequency and relative subsidy amounts are not systematically related to enrollment and attendance effect size estimates. The lack of association between payment frequency and program estimates contests single-program evidence from Bogotá's program in which payment frequency was manipulated at random (Barrera-Osorio, Bertrand, Perez-Calle and Linden, 2009). The authors of the Bogotá study argue that fully or partially delaying transfers increases reenrollment because doing so might help families relax savings constraints. We do not find such association across programs and contexts.

At the same time, our finding that average monthly subsidy is not related to the size of enrollment and attendance estimates is consistent with evidence from Cambodia suggesting decreasing returns to transfer amount (Filmer and Schady, 2009a). Controlling for other contextual and program design characteristics, this finding suggests that more generous transfers need not better compensate for the opportunity cost of sending children to school.

Fourth, controlling for contextual, other program design characteristics and study characteristics, effect size estimates from studies of programs in which transfer continuation is conditional on school achievement are statistically larger—about six percentage points—than those from programs that condition only on school attendance. This result is consistent, for example, with recent evidence from randomized controlled trials in developing countries highlighting the importance of educational incentives (see for example Duflo, Hanna and Ryan, forthcoming; Bettinger, Kremer and Saavedra, 2010; Kremer, Miguel and Thornton, 2009).

Fifth, estimates from CCT evaluations that use random assignment are statistically smaller—between two and three percentage points—than estimates from non-randomized evaluations, all else constant. It is not clear, ex-ante, the sign of the bias from observational evaluations of CCT programs relative to those that employ randomization. On the one hand, given the targeted nature of most CCT programs, program participants in observational studies are likely negatively selected. On the other hand, CCT programs such as *Familias en Acción* in Colombia, disburse transfers through the banking system, and for that reason, targeted rural municipalities are required to have a bank. Relative to families in comparison municipalities, participant families are therefore positively selected along observable and likely unobservable characteristics. Our finding indicates that across programs and contexts, participants might be positively selected relative to comparison groups in observational evaluations.

In Table 5 we show robustness test results that eliminate from the sample outlier estimates from the JFPR Program in Cambodia from Filmer and Schady (2009a) and from Red de Protección Social in Nicaragua from Dammert (2009). Consistent with our first three main findings using the full sample, CCT programs estimates are: i) significantly larger, all else constant, in contexts with low baseline school enrollment levels, ii) significantly larger when transfers are accompanied by a supply-side complement, iii) no different depending on payment frequency or relative subsidy amount.

While the coefficient on random assignment still suggests that relative to randomized evaluations, observational ones tend to yield larger estimates—about three percentage points—the finding is no longer significant as the two outlier studies where based on observational data. Similarly, the coefficient on whether transfer continuation is conditional on school achievement suggests that such conditionality might be associated with larger enrollment and attendance estimates, the correlation is not significant at conventional levels once we remove the outlier studies, with p-values between 0.12 and 0.16.

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, effect sizes for both secondary enrollment and secondary attendance are significantly larger in published references than in unpublished ones. Similarly, 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 outcomes and levels 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 8 through 10 display funnel plots separately for each outcome. Table 8 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 8a). Column 1 of Table 8 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 8b). Results in column 2 of Table 8 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 9a). 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 9b). Statistical analysis in columns 3 and 4 of Table 8 reject the hypothesis of funnel plot symmetry for both of these outcomes.

Figures 10a and 10b 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 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 effective in contexts with relative low levels of baseline school enrollment, and therefore, particularly effective at increasing secondary enrollment and attendance. Programs that in addition to cash transfers to families also attempt to expand supply through grants, infrastructure or other resources for schools are significantly larger than those from studies of programs only provide transfers to families, all else constant.

We do not find evidence that characteristics such as transfer amounts or payment frequency are associated with larger effect estimates, although some evidence suggest that conditioning transfer continuation on school achievement is associated with larger effect estimates, all else constant.

Observational evaluations report larger estimates, on average, than evaluations that take advantage of random assignment. This finding, in particular, is at odds with previous qualitative evidence by IGN (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, Akhter U., Daniel Gilligan, Ayse Kudat, Refik Colasan, Huseyin Tatlidil, and Bulent Ozbilgin. 2006. "Interim Impact Evaluation of the Conditional Cash Transfers Program in Turkey: A Quantitative Assesment." International Food Policy Research Institute.

Attanasio, Orazio, Costas Meghir, and Ana Santiago. 2005. "Education Choices in Mexico: Using a Structural Model and a Randomised Experiment to Evaluate Progresa." Institute for Fiscal Studies Working Paper EWP05/01.

Attanasio, Orazio, Emla Fitzsimons, Ana Gomez, Martha I. Gutiérrez, Costas Meghir, and Alice 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, Orazio, and Luis 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, Orazio, Murtaza Syed, Marcos Vera-Hernandez. 2004. "Early Evaluation of a New Nutrition and Education Programme in Colombia." Institute for Fiscal Studies Briefing Note 44.

Baird, Sarah, Craig McIntosh, and Berk 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, Sarah, Craig McIntosh, and Berk Özler. 2010. "Cash or Condition? Evidence from a Randomized Cash Transfer Program." World Bank Policy Research Working Paper 5259.

Barrera-Osorio, Felipe, Marianne Bertrand, Leigh L. Linden, and Francisco Perez-Calle. 2009. "Improving the Design of Conditional Transfer Programs: Evidence from a Randomized Education Experiment in Colombia." Unpublished manuscript.

Barrera-Osorio, Felipe, Marianne Bertrand, Leigh L. Linden, and Francisco Perez-Calle. 2008. "Conditional Cash Transfers in Education: Design Features, Peer and Sibling Effects: Evidence from a Randomized Experiment in Colombia." World Bank Policy Research Working Paper 4580.

Behrman, Jere R., Jorge Gallardo-Garcia, Susan W. Parker, Petra E. Todd, and Viviana Vélez-Grajales. 2005. "How Conditional Cash Transfers Impact Schooling and Working Behaviors of Children and Youth in Urban Mexico." Unplished manuscript.

Behrman, Jere R., Susan W. Parker, and Petra 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, Jere R., Pilyali Sengupta, and Petra 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, Eric, Michael Kremer, and Juan E. Saavedra. 2010. "Are Educational Vouchers Only Redistributive?" *The Economic Journal* 120 (546): F204-F228.

Borraz, Fernando, and Nicolás González. 2009. "Impact of the Uruguayan Conditional Cash Transfer Program." *Cuadernos de Economía* 46 (November): 243-271.

Cameron, Lisa. 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, Eliana, and André 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, Nazmul, and Dilip Parajuli. 2006. "Conditional Cash Transfers and Female Schooling: The Impact of the Female School Stipend Program on Public School Enrollments in Punjab, Pakistan." World Bank Policy Research Working Paper 4102.

Coady, David and Susan W. Parker. 2002. "A Cost-effectiveness Analysis of Demand- and Supply-side Education Interventions: The Case of Progresa in Mexico." International Food Policy Research Institute Discussion Paper 127.

Cooper, Harris, Larry V. Hedges, and Jeffrey C. Valentine, editors. 2009. *Handbook of Research Synthesis and Meta-analysis.* New York, NY: Russell Sage.

Dammert, Ana C. 2009. "Heterogeneous Impacts of Conditional Cash Transfers: Evidence from Nicaragua." *Economic Development and Cultural Change* 58 (1): 53-83.

Davis, Benjamin, Sudhanshu Handa, Marta Ruiz-Arranz, Marco Stampini, and Paul Winters. 2002. "Conditionality and the Impact of Programme Design on Household Welfare: Comparing Two Diverse Cash Transfer Programmes in Rural Mexico". Unpublished manuscript.

De Janvry, Alain, Federico Finan and Elisabeth Sadoulet. 2006. "Evaluating Brazil's Bolsa Escola Program: Impact on Schooling and Municipal Roles." Berkeley: University of California at Berkeley.

De Souza, Priscila Zeraik. 2005. "An Impact Evaluation of the Conditional Cash Transfers to Education under PRAF: An experimental Approach." Rio de Janeiro: Fundacao Getulio Vargas.

National Planning Department. 2006. *Programa Familias en Acción: Condiciones Iniciales de los Beneficiarios e Impactos Preliminares*. Bogota, D.C.: National Planning Department.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Using Randomization in Development Economics Research: A Toolkit." CEPR Discussion Paper 6059.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan. (forthcoming). "Incentives Work: Getting Teachers to Come to School." *American Economic Review.*

Duryea, Suzanne. and Andrew 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, Deon, and Norbert Schady. 2008. Getting Girls into School: Evidence from a Scholarship Program in Cambodia. *Economic Development and Cultural Change* 56 (3): 581-617.

Filmer, Deon and Norbert Schady. 2009a. "Are There Diminishing Returns to Transfer Size in Conditional Cash Transfers?" World Bank Policy Research Working Paper 4999.

Filmer, Deon and Norbert Schady. 2009b. "School Enrollment, Selection and Test Scores." World Bank Policy Research Working Paper 4998.

Fiszbein, Ariel, Norbert Schady, N, Francisco H.G. Ferreira, Margaret Grosh, Nial Kelleher, Pedro Olinto, and Emmanuel Skoufias. 2009. *Conditional Cash Transfers: Reducing Present and Future Poverty*. Washington D.C.: World Bank.

Ford, Deanna B. 2007. "Household Schooling Decisions and Conditional Cash Transfers in Rural Nicaragua." Washington, D.C.: Georgetown University

Gitter, Seth R., and Bradford L. Barham. 2009. "Conditional Cash Transfers, Shocks, and School Enrolment in Nicaragua." *The Journal of Development Studies* 45 (10): 1747-1767.

Glewwe, Paul, and Ana L. Kassouf. 2008. "The Impact of the Bolsa Escola/Familia Conditional Cash Transfer Program on Enrollment, Grade Promotion and Drop Out Rates in Brazil." Unpublished manuscript.

Glewwe, Paul, and Pedro Olinto. 2004. "Evaluating the Impact of Conditional Cash Transfers on Schooling: An experimental Analysis of Honduras' PRAF Program." Unpublished manuscript.

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

Hedges, Larry V., Elizabeth Tipton, and Mathew C. Johnson. 2010. "Robust Variance Estimation in Meta-regression with Dependent Effect Size Estimates." *Research Synthesis Methods*, 1 (1): 39-65.

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

Khandker, Shahidur R., Mark M. Pitt, and Nobuhiko Fuwa. 2003. "Subsidy to Promote Girls' Secondary Education: The Female Stipend Program in Bangladesh." Unpublished manuscript.

Kremer, Michael, Edward Miguel and Rebecca Thornton. 2009. "Incentives to Learn." *Review of Economics and Statistics*, 91 (3): 437-456.

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

Maluccio, John A., Murphy, Alexis, and Ferdinando Regalia. 2009. "Does Supply Matter? Initial Supply Conditions and the Effectiveness of Conditional Cash Transfers for Grade Progression in Nicaragua." Middlebury College Economics Discussion Paper 0908.

Maluccio, John A., and Rafael Flores. (2005). "Impact Evaluation of a Conditional Cash Transfer Program." International Food Policy Research Institute Research Report 141.

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

Oosterbeek, Hessel, Juan Ponce, and Norbert Schady. 2008. "The Impact of Cash Transfers on School Enrollment: Evidence from Ecuador." World Bank Policy Research Working Paper 4645.

Parker, Susan, Petra E. Todd, and Kenneth I. Wolpin. 2006. "Within-family Treatment Effect Estimators: The Impact of Oportunidades on Schooling in Mexico." Unpublished manuscript.

Ponce, Juan. 2006. "The Impact of Conditional Cash Transfer Programs on Achievement Test Scores: An Impact Evaluation of the "Bono de Desarrollo Humano" of Ecuador." Facultad Latinoamericana de Ciencias Sociales Sede Ecuadro Working Paper 06/302.

Rawlings, Laura B., and Gloria M. Rubio. 2005. "Evaluating the Impact of Conditional Cash Transfer Programs." *The World Bank Research Observer* 20 (1): 29-55.

Raymond, Mélanie, and Elisabeth Sadoulet. 2003. "Educational Grants Closing the Gap in Schooling Attainment between Poor and Non-poor." Unpublished manuscript.

Saavedra, Juan E. (in press). "Resource Constraints and Educational Attainment in Developing Countries: Colombia 1945-2005" *Journal of Development Economics*.

Schady, Norbert and Maria C. Araujo. 2008. "Cash Transfers, Conditions, and School Enrollment in Ecuador." *Economia* 8 (2), 43-77.

Schultz, T. Paul. 2004. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program." *Journal of Development Economics* 74: 199–250.

Skoufias, Emmanuel, and Susan W. Parker. 2001. "Conditional Cash Transfers and their Impact on Child Work and Schooling: Evidence from the Progresa Program in Mexico." International Food Policy Research Institute Discussion Paper 123.

Sparrow, Robert. 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, Alexander J. (2009). "Publication Bias." In *Handbook of Research Synthesis and Meta-analysis,* edited by Harris Cooper, Larry V. Hedges, and Jeffrey C. Valentine. New York, NY: Russell Sage.

The Economist. 2010. "Societies on the move." The Economist, U.S. edition September 11, 2010.

Todd, Petra E., Jorge Gallardo-Garcia, Jere R. Behram, and Susan 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 Bernardo Hernández-Prado and Mauricio Hernández-Ávlia, 165-227. Mexico D.F.: Instituto Nacional de Salud Pública.

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
Total number of programs		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 subsidy receipt [b] (mean, SD)	82.5	.04	14
Verification of school attendance			
Yes	9	47.4	
No	2	10.5	
No information reported	8	42.1	
Payment frequency			
Monthly	10	52.6	
Bimonthly	4	21.1	
Other	4	21.1	
No information reported	1	5.3	
Monthly average subsidy amount as a % of PPP- adjusted GDP per capita (mean, SD)			
Primary	2.3	2.0	13
Secondary	4.2	4.3	17
School subsidy amount varies by			
Gender	3	15.8	
Grade or age	3	15.8	
None	11	57.9	
Other [c]	2	10.5	
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	

[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	10	23.8
Working paper	22	52.4
Government/technical reports	7	16.7
Unpublished	3	7.1
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 results of enrollment and attendance effect size estimate moderators

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Contextual Characteristics								
Baseline school enrollment	-0.357	-0.438	-0.444	-0.374	-0.375	-0.476	-0.471	-0.476
	(0.108)	(0.116)	(0.121)	(0.109)	(0.088)	(0.099)	(0.100)	(0.100)
	[0.002]	[0.001]	[0.001]	[0.002]	[0.000]	[0.000]	[0.000]	[0.000]
Latin America (1=yes)	0.030	0.031	0.043	0.038	0.034	0.115	0.116	0.118
	(0.021)	(0.022)	(0.023)	(0.022)	(0.019)	(0.036)	(0.037)	(0.037)
	[0.170]	[0.155]	[0.074]	[0.101]	[0.078]	[0.004]	[0.004]	[0.004]
Average monthly subsidy as percent of per-capita GDP (PPP)		-0.004	-0.003	0.0002	-0.0003	0.002	0.001	0.002
		(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
		[0.333]	[0.458]	[0.946]	[0.931]	[0.467]	[0.611]	[0.526]
Program Characteristics								
Random assignment to conditions			-0.019	-0.049	-0.018	-0.038	-0.043	-0.044
			(0.017)	(0.024)	(0.020)	(0.023)	(0.024)	(0.025)
			[0.282]	[0.055]	[0.360]	[0.118]	[0.080]	[0.086]
Supply-side complement (1=yes)				0.050	0.052	0.047	0.057	0.059
				(0.024)	(0.021)	(0.023)	(0.025)	(0.027)
				[0.044]	[0.017]	[0.048]	[0.029]	[0.041]
Payment frequency (1=monthly, 0 less frequently)					-0.058	-0.023	-0.021	-0.017
					(0.017)	(0.020)	(0.020)	(0.021)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					[0.002]	[0.246]	[0.312]	[0.435]
Transfer continuation conditional on achievement (1=yes)						0.088	0.089	0.093
						(0.037)	(0.038)	(0.039)
						[0.026]	[0.027]	[0.026]
Year program began							0.002	0.001
							(0.002)	(0.003)
							[0.531]	[0.718]
Reference Characteristics								
Number of reported effects								0.001
								(0.001)
								[0.423]
Number of Observations	64	64	64	64	64	64	64	64

Notes: Standard errors adjusted for hierarchical dependence (clustering) of estimates at the study level in parentheses and corresponding p-values in brackets. The dependent variable in each specification is the fixed-effects average effect size estimate for a study and an outcome-by-school level combination (primary enrollment, primary attendance, secondary enrollment, secondary attendance). 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 outcome-by-level fixed effects in addition to the reported coefficients.

Table 5. Robustness check: Meta-regression results of enrollment and attendance effect size estimate moderators excluding outlier studies

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Contextual Characteristics								
Baseline net school enrollment	-0.223	-0.301	-0.304	-0.263	-0.278	-0.351	-0.344	-0.338
	(0.085)	(0.094)	(0.099)	(0.099)	(0.086)	(0.110)	(0.109)	(0.116)
	[0.014]	[0.003]	[0.005]	[0.013]	[0.003]	[0.004]	[0.004]	[0.008]
Latin America (1=yes)	0.015	0.017	0.022	0.020	0.020	0.074	0.074	0.072
	(0.017)	(0.016)	(0.018)	(0.019)	(0.017)	(0.039)	(0.040)	(0.042)
	[0.376]	[0.32]	[0.236]	[0.297]	[0.242]	[0.070]	[0.074]	0.101
Average monthly subsidy as percent of per-capita GDP (PPP)		-0.004	-0.003	-0.0012	-0.0013	0.0005	0.0000	-0.0003
		(0.002)	(0.002)	(0.002)	(0.002)	0.002	(0.002)	(0.003)
		[0.140]	[0.183]	[0.613]	[0.591]	[0.834]	[0.999]	0.915
Program Characteristics								
Random assignment to conditions			-0.010	-0.029	-0.010	-0.025	-0.031	-0.030
			(0.013)	(0.017)	(0.014)	(0.020)	(0.022)	(0.023)
			[0.465]	[0.098]	[0.479]	[0.218]	[0.170]	[0.193]
Supply-side complement (1=yes)				0.033	0.037	0.034	0.045	0.045
				(0.017)	(0.015)	(0.017)	(0.021)	(0.022)
				[0.059]	[0.022]	[0.056]	[0.044]	[0.050]
Payment frequency (1=monthly, 0 less frequently)					-0.041	-0.020	-0.017	-0.019

					(0.014)	(0.015)	(0.016)	(0.017)
					[0.006]	[0.201]	[0.297]	[0.292]
Transfer continuation conditional on achievement (1=yes)						0.060	0.060	0.058
						(0.037)	(0.038)	(0.040)
						[0.121]	[0.122]	[0.162]
Year program began							0.002	0.002
							(0.003)	(0.003)
							[0.493]	[0.454]
Reference Characteristics								
Number of reported effects								-0.0003
								(0.001)
								[0.685]
Number of Observations	61	61	61	61	61	61	61	61

Notes: Standard errors adjusted for hierarchical dependence (clustering) of estimates at the study level in parentheses and corresponding p-values in brackets. The dependent variable in each specification is the fixed-effects average effect size estimate for a study and an outcome-by-school level combination (primary enrollment, primary attendance, secondary enrollment, secondary attendance). 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 outcome-by-level fixed effects in addition to the reported coefficients. Results in Table 5 exclude estimates from Cambodia's JFPR Program from Filmer and Schady (2008) and from Nicaragua's Red de Protección Social Program from Dammert (2009).

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

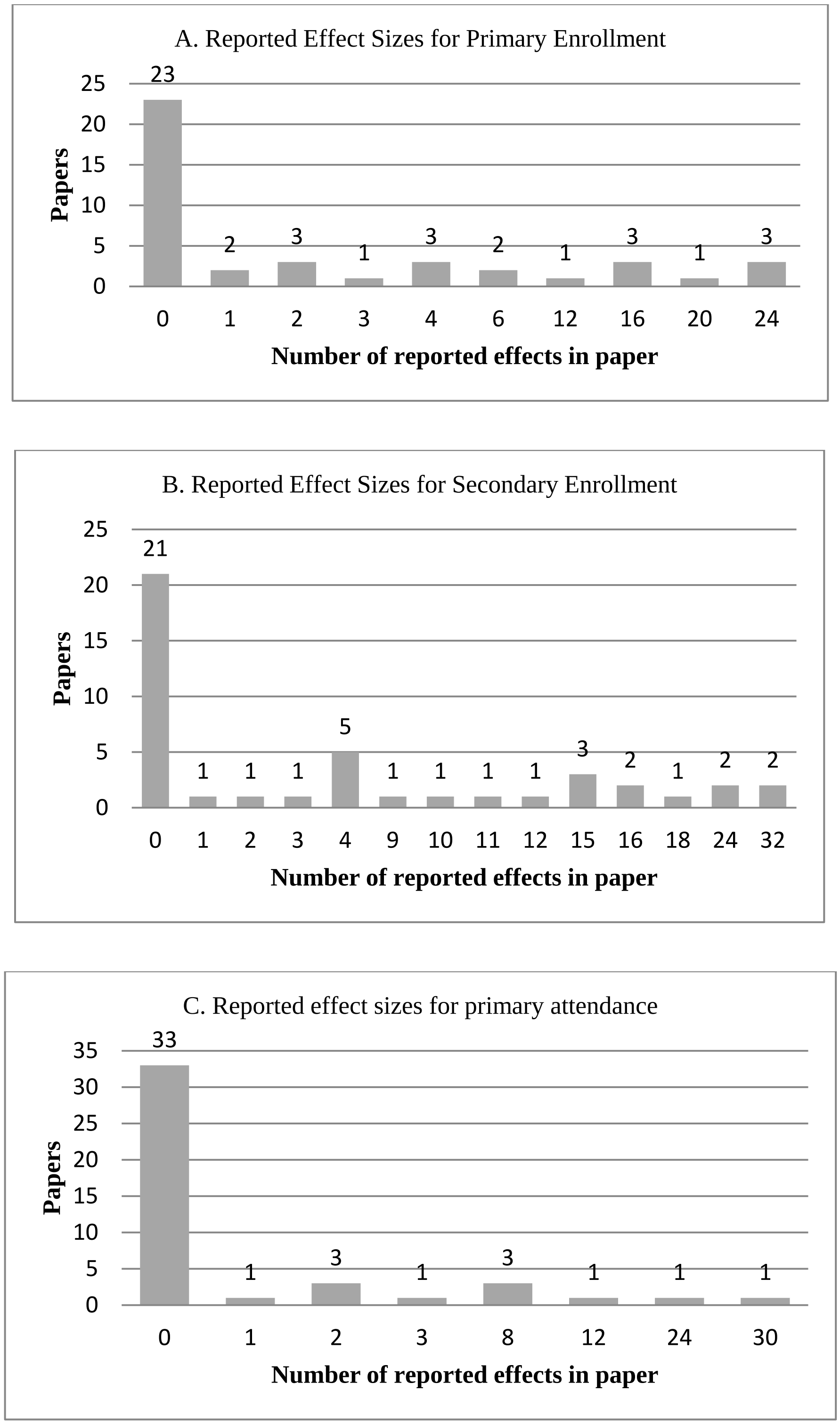
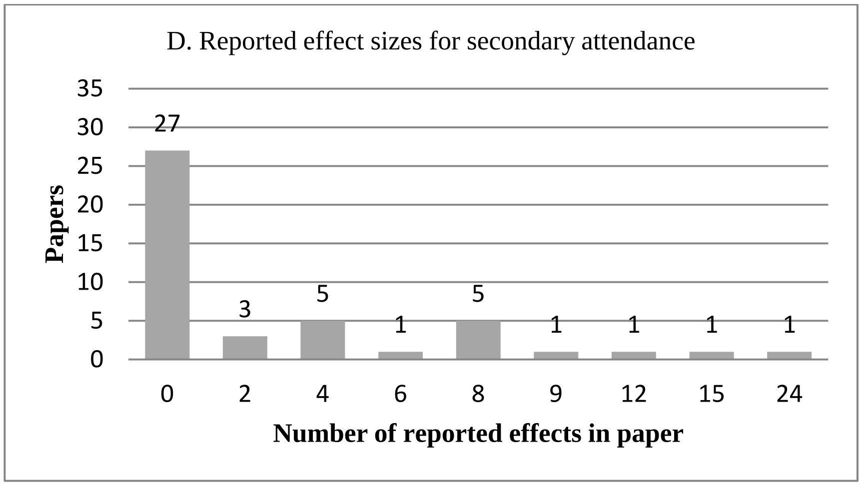
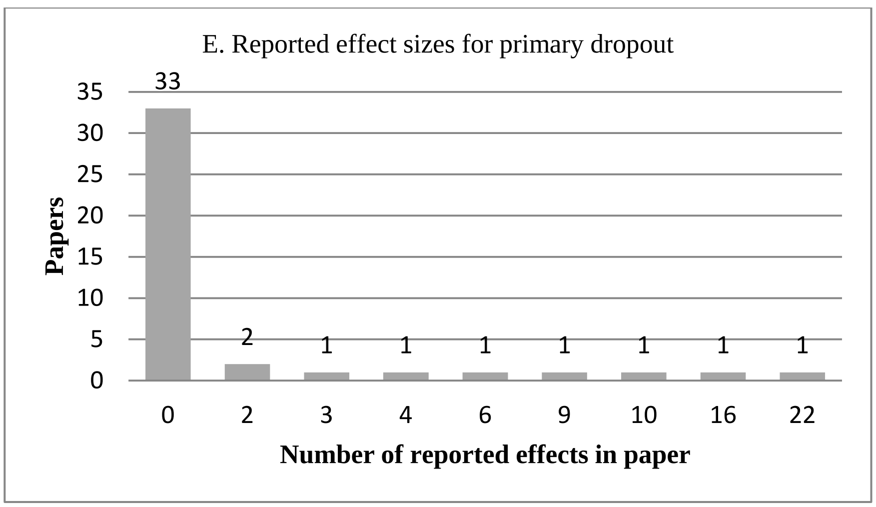


Figure 1. (cont.)





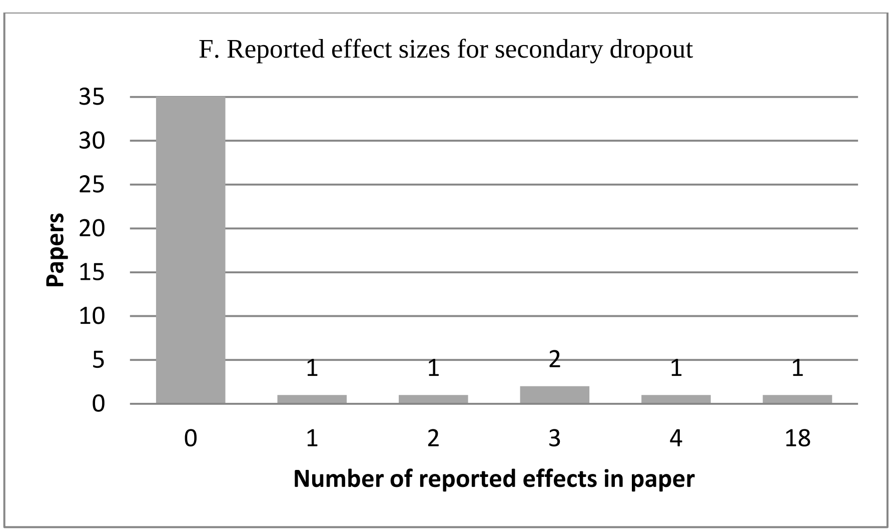
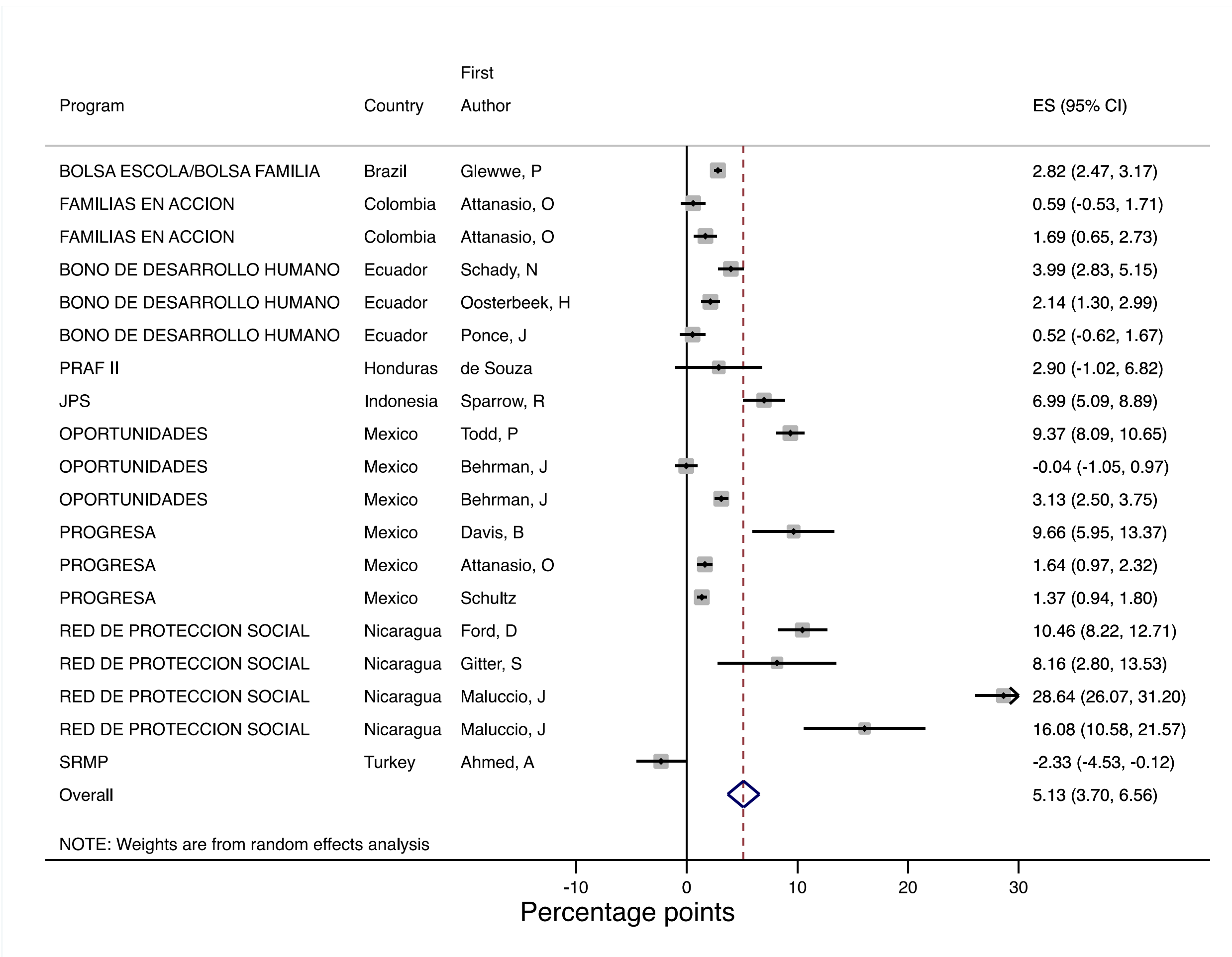
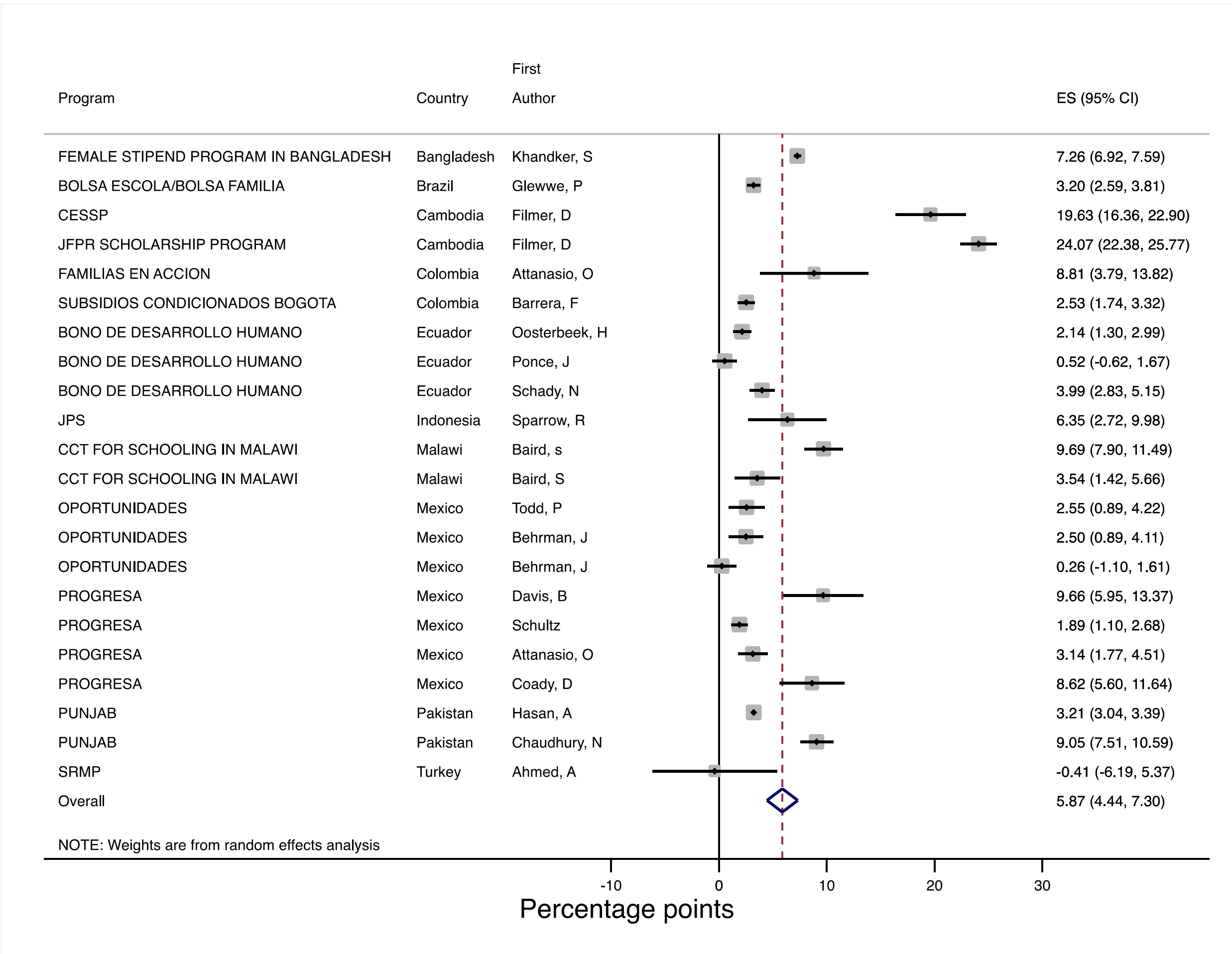


Figure 2. Forest plot of effect sizes on primary enrollment (with one fixed effects average estimate per paper)



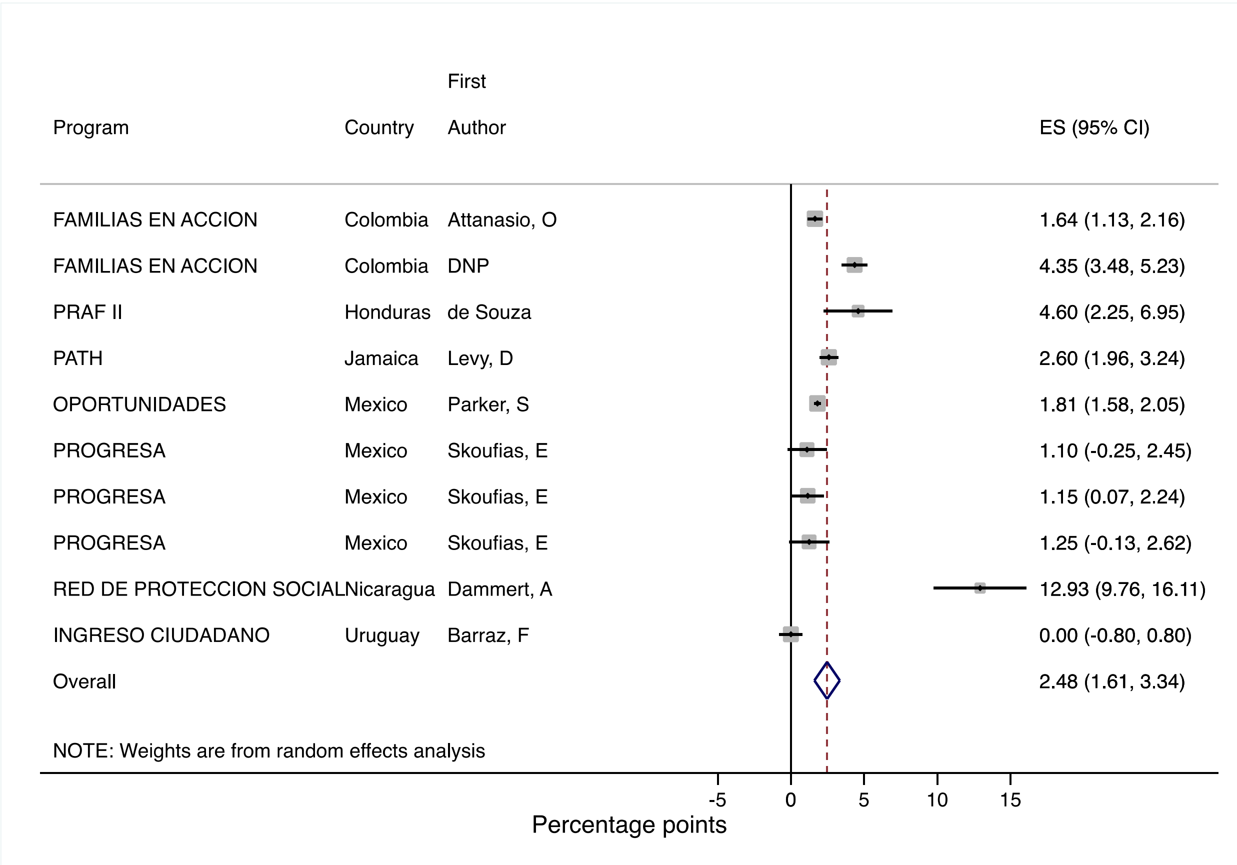
Notes: For each study we compute one average effect size using a fixed effects model to combine all estimates in the study. The overall average effect estimate is from a random effects methods of moments model. The chi-square test statistic for the null hypothesis of homogeneity in primary enrollment effect size estimates in the random effects model is 735 (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 93%.

Figure 3. Forest plot of effect sizes on secondary enrollment (with one fixed effects average estimate per paper)



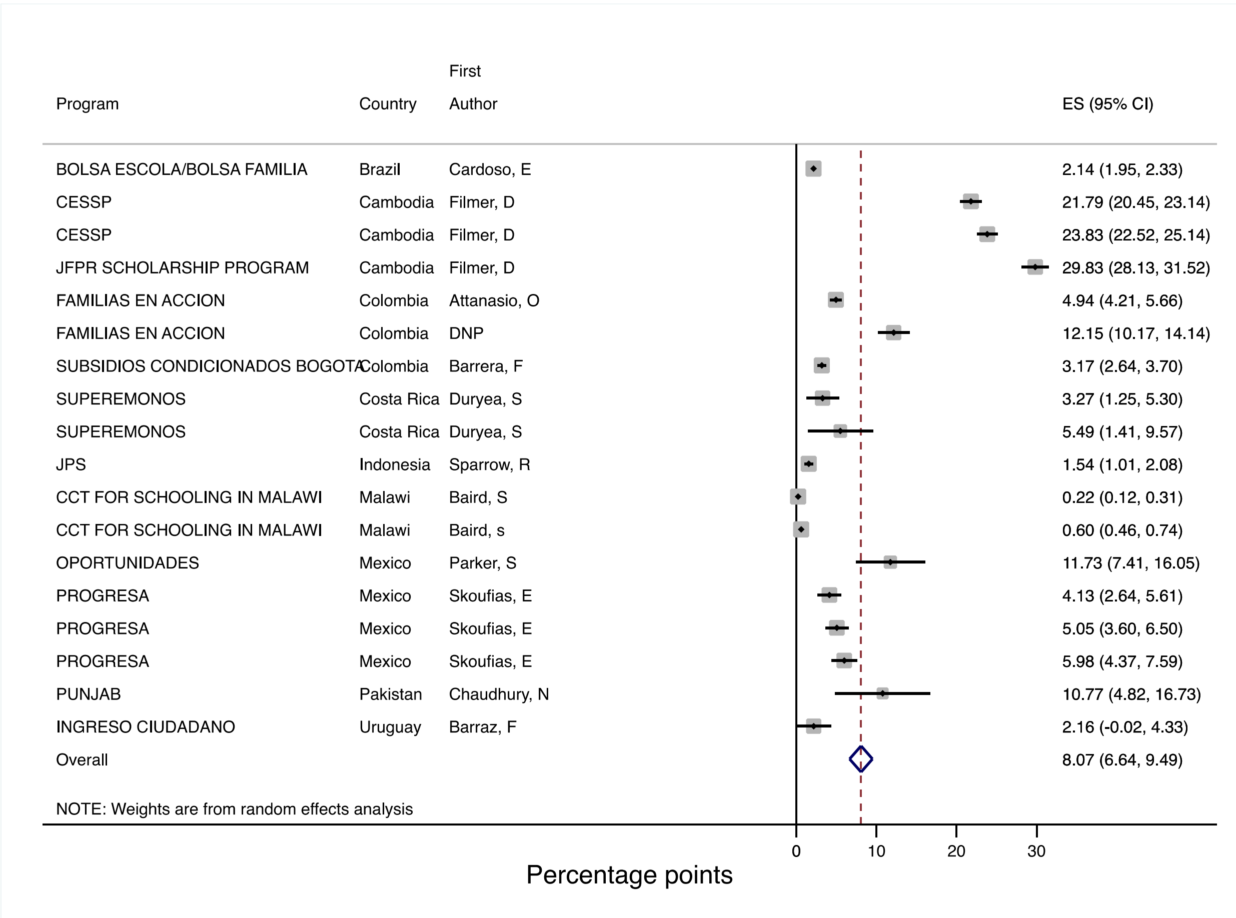
Notes: For each study we compute one average effect size using a fixed effects model to combine all estimates in the study. The overall average effect estimate is from a random effects methods of moments model. The chi-square test statistic for the null hypothesis of homogeneity in secondary enrollment effect size estimates in the random effects model is 1302 (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 50%.

Figure 4. Forest plot of effect sizes on primary attendance (with one fixed effects average estimate per paper)



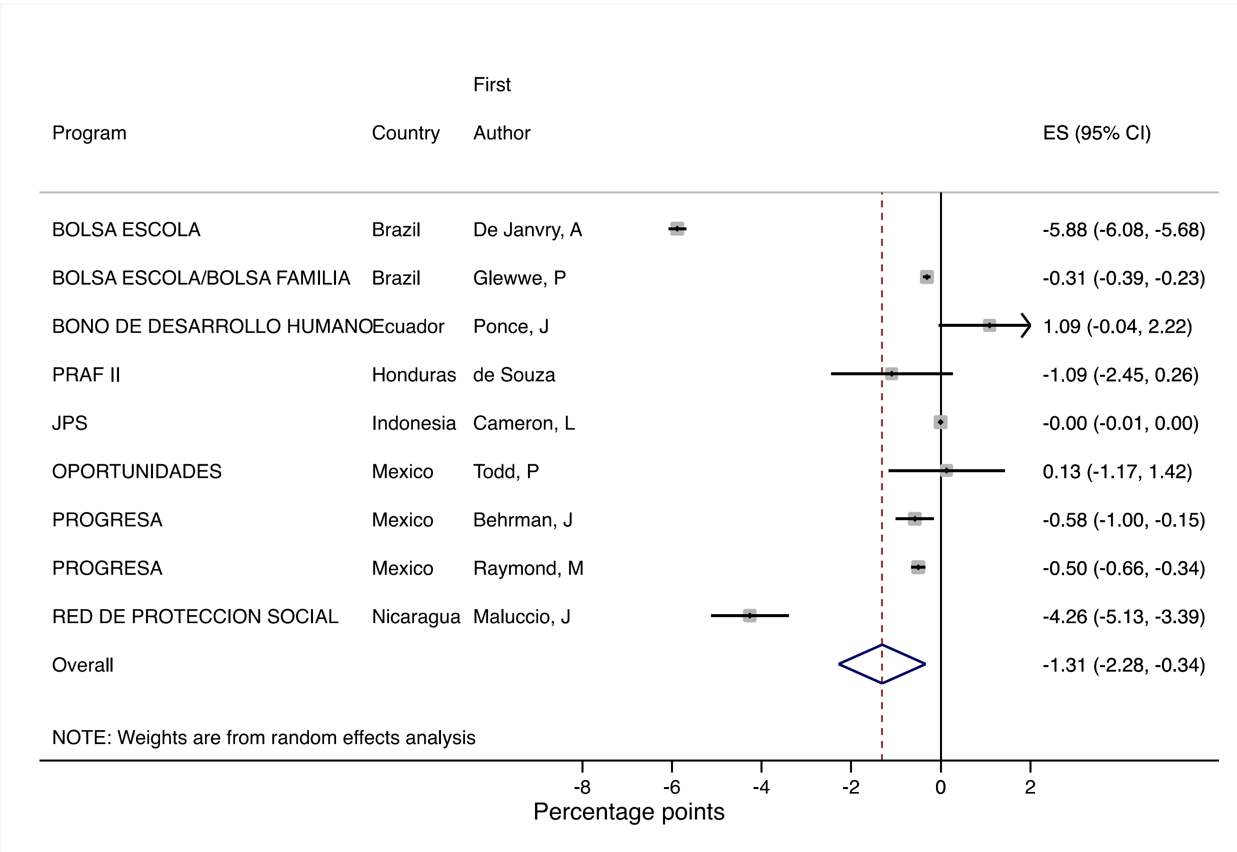
Notes: For each study we compute one average effect size using a fixed effects model to combine all estimates in the study. The overall average effect estimate is from a random effects methods of moments model. Skoufias and Parker (2001) (First Author Skoufias for Progresa) reports effects for three 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 113.4 (p-value 0.000). Mean baseline primary attendance computed from studies in the sample reporting it is 80%.

Figure 5. Forest plot of effect sizes on secondary attendance (with one fixed effects average estimate per paper)



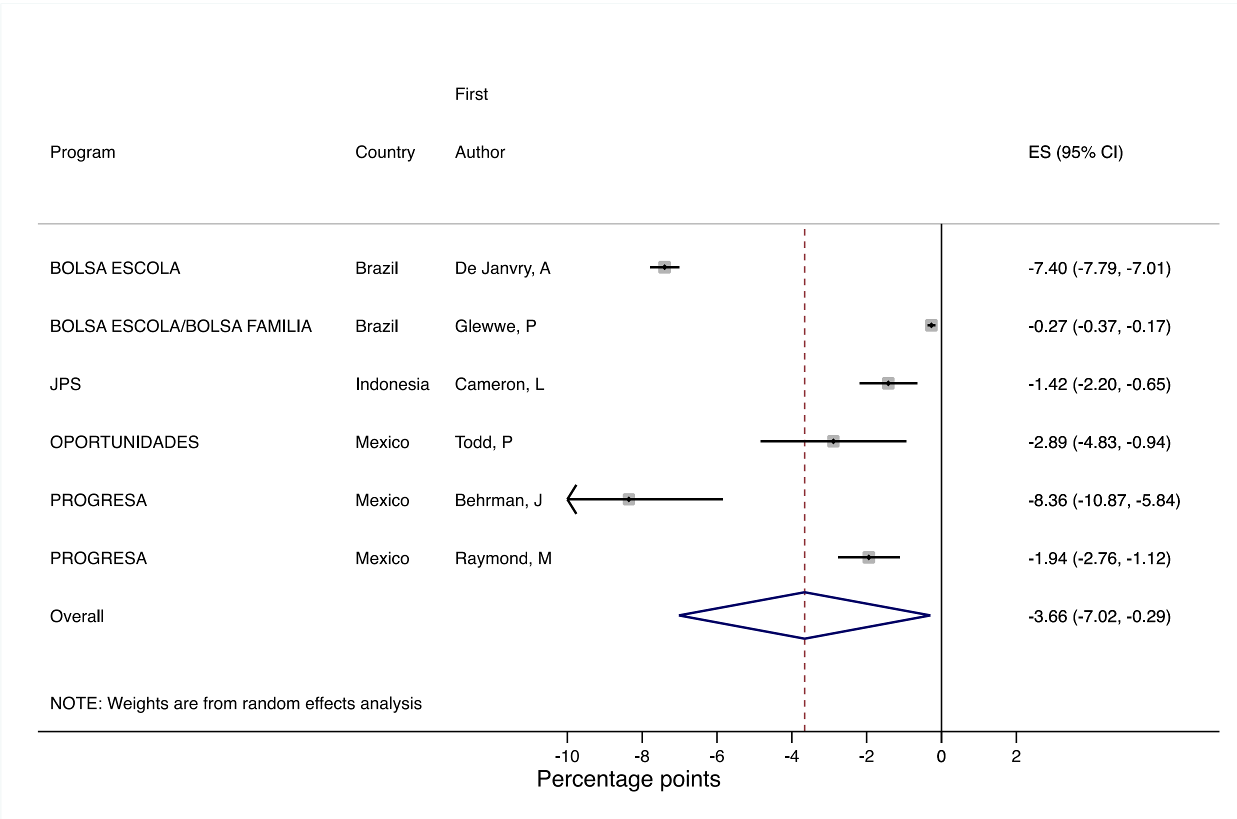
Notes: For each study we compute one average effect size using a fixed effects model to combine all estimates in the study. The overall average effect estimate is from a random effects methods of moments model. Duryea and Morrison (2004) reports effects for two post-treatment periods and we compute one 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 4050 (p-value 0.000).Mean baseline secondary attendance computed from studies in the sample reporting it is 68%.

Figure 6. Forest plot of effect sizes on primary dropout (with one fixed effects average estimate per paper)



Notes: For each study we compute one average effect size using a fixed effects model to combine all estimates in the study. The overall average effect estimate is from a random effects methods 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).

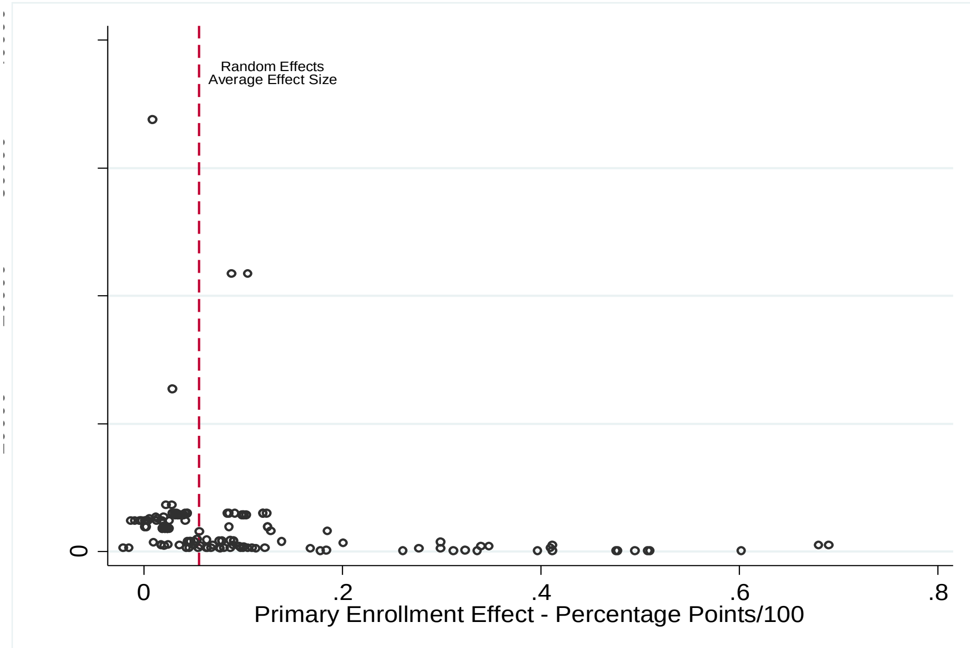
Figure 7. Forest plot of effect sizes on secondary dropout (with one fixed effects average estimate per paper)



Notes: For each study we compute one average effect size using a fixed effects model to combine all estimates in the study. The overall average effect estimate is from a random effects methods 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 8. Funnel plot of sample size on reported enrollment effect size estimate (all estimates)

a. Primary enrollment



b. Secondary enrollment

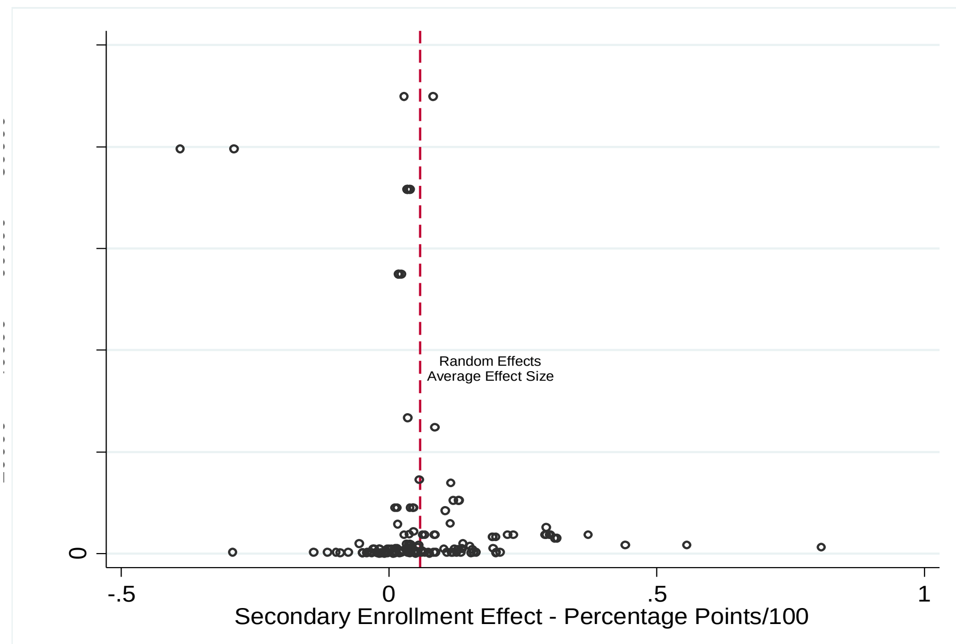
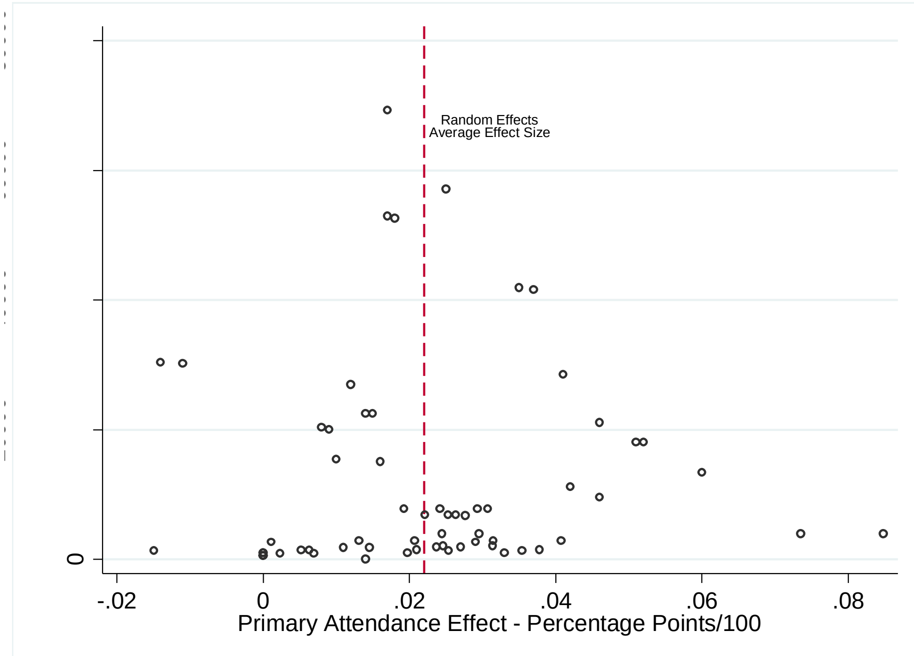


Figure 9. Funnel plot of sample size on reported attendance effect size estimate (all estimates)

a. Primary attendance



b. Secondary attendance

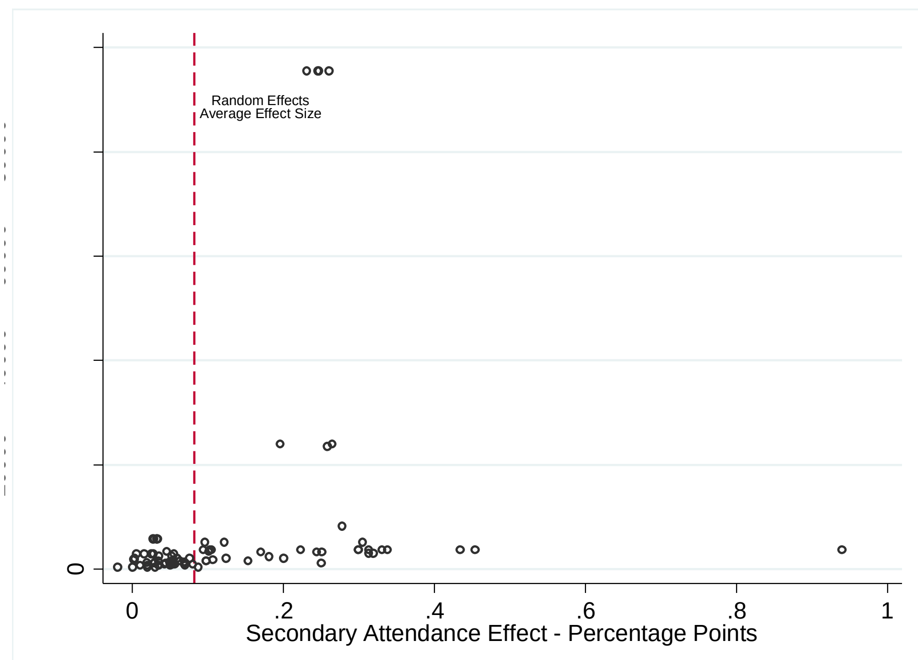
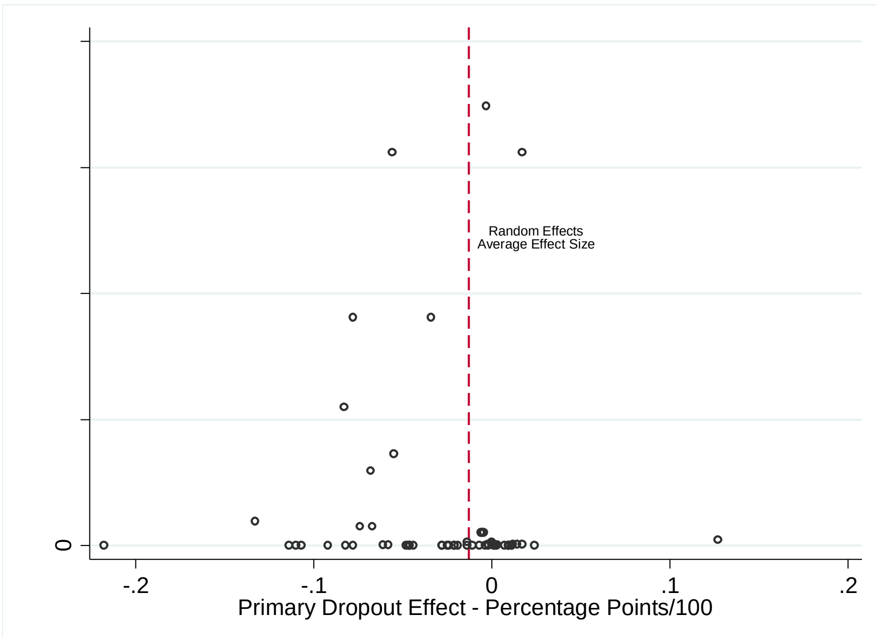
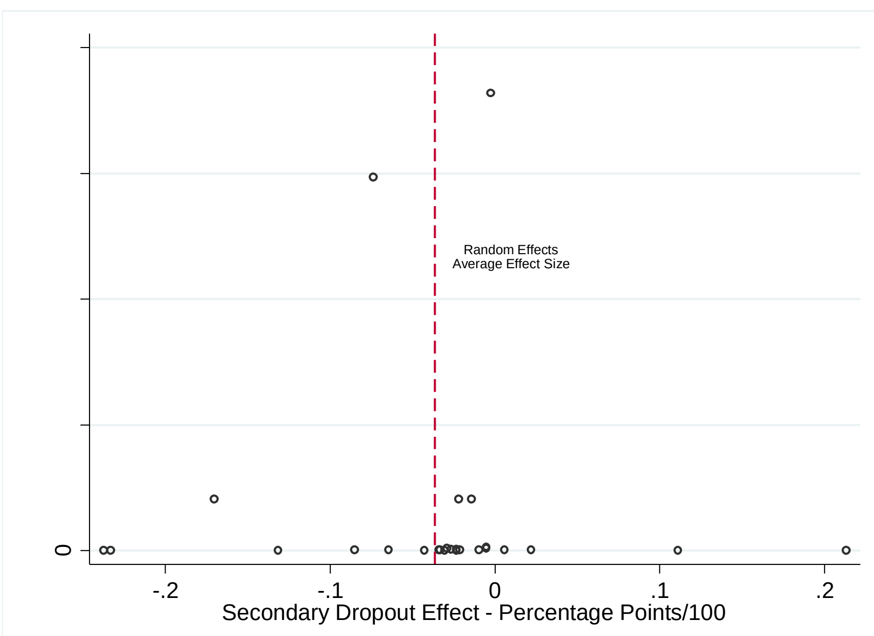


Figure 9. Funnel plot of sample size on reported dropout effect size estimate (all estimates)

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.	2008	Conference paper	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.	2009	Working paper	Program survey	3,225	Yes	Yes	No
Cambodia	CESSP	Filmer, D.	2009	Working paper	Program survey	95,493	No	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	2,691	No	Yes	No
Colombia	Familias en Acción	Attanasio, O.	2004	Government report	Program survey	3,935	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.	2009	Working paper	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.	2010	Working paper	Program survey	1,832	Yes	Yes	No
Mexico	Oportunidades	Behrman, J.	2004	Technical report	Program survey	1,796	Yes	No	No
Mexico	Oportunidades	Behrman, J.	2005	Working paper	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.	2005	Working paper	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.	2002	Working paper	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.	2001	Working paper	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.	2009	Working paper	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.	2006	Working paper	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	Subsidy received by	Subsidy varies by	Supply compo-nent	Random Assign-ment
						Primary	Secondary					
Bangladesh	Female Stipend Program	1994	Attendance, academic proficiency and remain unmarried	75	Yes	Not applicable	1.42	Monthly	Student	Grade	Yes	No
Brazil	Bolsa Escola	2001	Attendance	85	N/A	0.77	0.77	Monthly		None	No	No
Brazil	Bolsa Escola/Bolsa Familia	1995	Enrollment and attendance	85	N/A	1.05	1.05	Monthly	N/A	None	N/A	No
Cambodia	CESSP	2004	Enrollment, attendance and grade promotion	N/A	Yes	Not applicable	10.01	3 times per year	Parents	Dropout risk	No	No
Cambodia	JFPR Scholarship Program	2005	Enrollment, attendance and grade promotion	N/A	N/A	Not applicable	8.95	3 times per year	Parents	None	No	No
Colombia	Familias en Acción	2001	Enrollment and attendance	80	N/A	1.10	2.21	Bimonthly	Mother	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]	Parents	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		None	No	Yes
Honduras	PRAF II	1998	Enrollment and attendance	85	No	2.06	Not applicable	Monthly	Parents	None	Yes	Yes
Indonesia	JPS	1998	Enrollment and passing grades	N/A	N/A	0.39	0.98	3 times per year	Student	Grade	No	No

Country	**Program name**	**Year program started**	**Conditionality**	**Minimum attendance rate (%)**	**Conditions verification**	**Transfer amount[a]**		**Payment frequency**	**Subsidy received by**	**Subsidy varies by**	**Supply compo-nent**	**Random Assign-ment**
						Primary	**Secondary**					
Jamaica	PATH	2001	Attendance	85	Yes	1.11	1.11	N/A	Parents	None	No	No
Malawi	CCT for schooling	2007	Enrollment and attendance	75	Yes	Not applicable	17.3	Monthly	Parent or guardian and student[c]	Randoml y	No	Yes
Mexico	Oportunidades	2002	Attendance	85	N/A	1.21	3.92	Bimonthly	Mother	Gender and grade	No	No
Mexico	Progresa	1997	Attendance	85	N/A	1.05	2.49	Monthly	Mother	Gender and grade	Yes	Yes
Nicaragua	Red de Protección Social	2000	Enrollment and attendance	85	Yes	5.23	Not applicable	Bimonthly	N/A	None	Yes	Yes
Punjab	Pakistan	2004	Attendance	80	N/A	Not applicable	2.28	Monthly	Student	None	No	No
Turkey	SRMP	2004	Attendance and not repeating a grade more than once	80	Yes	1.56	2.62	Bimonthly	N/A	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).