NBER WORKING PAPER SERIES

EDUCATIONAL IMPACTS AND COST-EFFECTIVENESS OF
CONDITIONAL CASH TRANSFER PROGRAMS IN DEVELOPING COUNTRIES:
A META-ANALYSIS

Sandra García
Juan Saavedra

ABSTRACT

We meta-analyze for impact and cost-effectiveness 94 studies from 47 conditional cash transfer
(CCT) programs in low- and middle-income countries worldwide, focusing on educational
outcomes that include enrollment, attendance, dropout and school completion. To conceptually
guide and interpret the empirical findings of our meta-analysis, we present a simple economic
framework of household decision-making that generates predictions, all else constant, for the
association between certain program context and design characteristics, and impact estimates. We
also present a simple model for the analysis of program costs, using it to compute cost-
effectiveness estimates for a subsample of CCT programs. For all schooling outcomes, we find
strong support for heterogeneity in impact, transfer-effectiveness, and cost-effectiveness estimates.
Our meta-analytic results of impact and transfer-effectiveness estimates provide support to some –
but not all – of the predictions from the household decision-making model.

Sandra García
Escuela de Gobierno
Universidad de los Andes
Cra 1 # 18A -12
Bogota, Colombia
sagarcia@uniandes.edu.co

Juan Saavedra
Dornsife Center for Economic and Social Research
University of Southern California
635 Downey Way
Los Angeles, CA 90089
and NBER
juansaav@usc.edu

Educational Impacts and Cost-Effectiveness of Conditional Cash Transfer Programs in

Developing Countries: A Meta-analysis

Conditional cash transfers (CCTs) are one of the most prevalent social assistance

programs in low- and middle-income countries today. In Latin America and the Caribbean, for

example, there are currently 26 CCT programs in operation, benefiting more than 135 million

people (Stampini & Tornarolli, 2012). Also, the use of CCTs is expanding; today, more than 50

countries worldwide operate CCTs, greater than twice as many in 2008 (World Bank, 2014).

In CCT programs, target households receive cash payments on the condition that they

invest in the human capital of children. Health and nutrition conditions typically require medical

visits and vaccinations for children less than 5 years of age, and attendance by mothers at

periodic health information talks. Schooling conditions usually require enrollment, continued

attendance, and, occasionally, some measure of performance (Fiszbein et al., 2009)[1].

CCTs are complex programs and many decisions – typically made by in-country officials

– determine their final shape and size. Some of these decisions include how much to pay

families, how and when to pay them, how stringent the conditions for compliance should be, and

who is eligible to participate. But to what extent does the variation in program characteristics

map onto heterogeneity in impact and cost-effectiveness estimates? And to what degree, if any,

can we empirically validate theoretically-derived predictions for the relationship between

---

[1] Recent review studies of CCT evaluations (Baird, Ferreria, Özler, & Woolcock, 2014; Fiszbein et al., 2009; Independent Evaluation Group, 2011) conclude that, on the whole, these programs have positive effects on schooling (enrollment, attendance, dropout). There is limited evidence, however, on the impact of CCTs on educational achievement outcomes such as school completion, school promotion, or learning. Snilstveit et al. (2015) find the average effect of CCTs on math and language test scores is close to zero. Molina-Millan et al. (2016) review evidence of CCT impacts on educational outcomes in Latin America and report effect sizes ranging from 0.5 to 1 additional years of schooling. However, some of these estimates have a high risk of internal validity due to attrition bias (Molina-Millan et al., 2016). Also, evidence on other long-term outcomes such as access to tertiary education, participation in the labor market, and earnings is still very limited. Molina-Millan et al. (2016) suggest that evidence on the long-term effects of CCTs remains inconclusive.

heterogeneity in program characteristics and heterogeneity in impact and cost-effectiveness

estimates?

To address these research questions, we meta-analyze for impact and cost-effectiveness

94 studies from 47 CCT programs in low- and middle-income countries worldwide, focusing on

educational outcomes that include enrollment, attendance, dropout, and school completion. To

conceptually guide and interpret the empirical findings of our meta-analysis, we present a simple

economic framework on household decision-making that generates predictions, all else constant,

for the association between program context and design characteristics, and impact estimates.

We also present a simple model for the analysis of program costs, using it to compute cost-

effectiveness estimates for nine CCT programs for which we have data on both transfer-

effectiveness (impact units per dollar of transfer) and cost-transfer ratios (administrative costs

per dollar of transfers).

Our framework for analysis and data allow us to significantly advance – in numerous

ways – our understanding of the interplay between design characteristics, educational impacts,

and cost-effectiveness of CCT programs in developing countries. First, relative to recent

systematic reviews, we analyze a much larger array of CCT programs.[2] Second, our study is the

first to quantitatively meta-analyze educational impact and cost-effectiveness estimates for

multiple educational outcomes separately for primary and secondary schooling. Third, our

conceptual framework also predicts empirical relationships between educational impact

estimates for a broader set of program characteristics relative to those analyzed in recent prior

---

[2] We include 47 unique CCT programs in our meta-analysis. Snilstveit et al. (2015) includes 38 unique cash transfer programs – 35 of which are CCTs. Baird et al. (2014) includes 30 cash transfer programs – 26 of which are CCTs. Petrosino et al (2013) includes 13 CCT programs. IEG (2011) reports schooling effects for 30 CCT programs. Filmer & Schady (2009) reports schooling effects for 13 programs. Rawlings & Rubio reports schooling effects for six programs.

reviews (e.g., Snilstveit et al., 2015; Baird et al., 2014).[3] Fourth, our review is alone in computing educational cost-effectiveness ratios for multiple CCT programs. Fifth, while some of our findings confirm results from earlier reviews and experimental studies, other discoveries are new – particularly those stemming from the analysis of cost-effectiveness data and meta-regressions – and highlight potential important avenues for future research.

With regards to the first research question, we find unambiguous empirical support for the prediction that variation in program characteristics maps onto heterogeneity in impact and cost-effectiveness estimates across programs in our sample. For all educational outcomes (enrollment, attendance, dropout, completion) and schooling levels (primary, secondary) we strongly reject homogeneity in impact, transfer-effectiveness, and cost-effectiveness estimates. This conclusion is robust to including in the sample only studies meeting standards of evidence.[4]

For educational cost-effectiveness estimates, in particular, heterogeneity arises both from variation in transfer-effectiveness as well as variation in cost-transfer ratios across programs. With the caveat of a limited sample size, however, it is unclear whether programs that are most cost-efficient (i.e., programs with low cost-transfer ratios) are also the most transfer-effective (i.e., programs producing the greatest educational impact per dollar of transfer). We find, for instance, a negative correlation, albeit weak, between a program's cost-transfer ratio and its primary enrollment transfer-effectiveness estimate, suggesting that more cost-efficient programs (weakly) produce greater primary enrollment impacts per dollar of transfer. In contrast, we find

---

[3] We add two new predictions for the relationship between educational impact estimates and program characteristics. These characteristics are whether a CCT program has a complementary supply-side intervention, and if the program imposes on students school attainment requirements such as test scores or grade completion in addition to the standard conditions for enrollment and minimum attendance.
[4] Snilstveit et al. (2015); Baird et al. (2014); Petrosino, et al. (2013); IEG (2011); Fiszbein & Schady (2009); and Rawlings & Rubio (2005) also suggest that 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). All of these studies, however, have a more limited number of programs in their sample.

a strong positive correlation between a program's cost-transfer ratio and its primary attendance transfer-effectiveness estimate, suggesting that more cost-efficient programs (strongly) produce smaller primary attendance impacts per dollar of transfer.

With regards to our second question, the empirical evidence only partially validates theoretically-derived predictions for the relationship between heterogeneity in program characteristics and heterogeneity in impact and cost-effectiveness estimates. Inconsistent with the theoretical prediction, we find that, all else constant, greater transfer amounts are not statistically correlated with greater educational effect sizes for any outcome or schooling level.[5] We find no empirical support for the prediction that expected behavioral effects stemming from a rise in permanent income differ from those emerging from a rise in current income. Specifically, all else constant, educational effect size estimates from long-standing national CCT programs are statistically indistinguishable from those of pilot one- or two-year CCT programs. We also do not find evidence consistent with stronger effects for transfers that target mothers. This result leaves open the possibility that unitary household decision-making is perhaps an adequate benchmark for decision-making in target CCT households for programs in our sample.[6]

For some educational outcomes, we find that effect sizes are greater when other schooling conditions, such as grade promotion or test scores, are imposed on beneficiaries, beyond the typical requirements of enrollment and minimum attendance. This result is consistent with the notion of additional conditions possibly causing steeper reductions in the marginal cost of

---

[5] This finding, however, is consistent with other systematic review evidence on CCTs from Baird at al. (2014) and Snilstveit et al. (2015) and with single CCT program evidence from Cambodia and Malawi. In Cambodia's CESSP program, increases in school enrollment are not related to transfer size (Filmer & Schady, 2011). In Malawi's CCT program, impacts on educational outcomes such as dropout do not vary with transfer amounts (Baird et al., 2009).

[6] This result, however, is empirically consistent with prior review evidence from Baird et al. (2014) and with experimental evidence from Morocco's Tayssir program (Benhassine et al., 2015).

children's and parent's time, leading to extra time devoted to schooling activities. We also find

support for the prediction that educational program impacts – particularly on school enrollment

and attendance outcomes – should be stronger in settings with low baseline enrollment levels and

in settings with excess school capacity in which supply constraints are not binding.[7] Consistent

with this, for primary schooling – for which access to is close to universal in most low- and

middle-income countries – we find that, all else constant, primary enrollment impact estimates

are greater in CCT programs that complement cash transfers to families with supply-side

interventions such as school grants.

The remainder of the paper proceeds as follows: Section 2 outlines a simple theoretical

framework to organize our empirical predictions regarding program impacts and to conceptualize

our cost-effectiveness analyses. Section 3 describes the literature search and sample. Section 4

describes our approach to code study quality, effect size estimates, and program characteristics.

Section 5 describes the characteristics of CCT programs in our analytic sample. Section 6

describes the methods to combine and statistically meta-analyze impact and transfer-

effectiveness estimates. Section 7 presents the distributions of impact, transfer-effectiveness,

and cost-effectiveness estimates for the educational outcomes of analysis as well as meta-

regression results of impact and transfer-effectiveness estimates. Section 8 concludes.

Conceptual Framework

This section contains two subsections. The first subsection describes a simple economic

model of household decision-making (Skoufias, 2006; Filmer & Schady, 2009). We use this

---

[7] For more educational outcomes, with a larger number of studies and controlling for other factors in a meta-regression framework, this finding confirms earlier descriptive evidence by Filmer & Schady (2009) for school attendance. For school enrollment, we find that everything else equal, baseline enrollment levels are not associated with effect sizes, which is consistent with Baird et al (2014).

model to motivate our meta-analytic empirical predictions on how effect sizes may correlate with certain program and contextual characteristics. The second subsection lays out a simple model of CCT program costs to guide our cost-effectiveness analyses, drawing on earlier work by Caldés and Maluccio (2005) and Caldés, Coady, and Maluccio (2006).

An economic model of household decision-making and predictions for distribution of educational impact estimates the relationship between mediators and CCT program effects

In this model, household welfare is assumed to depend on consumption and the number of children as well as their adult earnings. Households face three constraints: a budget constraint, a human capital production function constraint and an adult earnings production function. In budget constraint, households are assumed to receive income from four sources: non-employment income, which includes labor income of adult men in the household; child labor; mother labor; and a fraction of future earnings of children as adults. Households equate income to expenditure on goods and services that contribute to human capital production and consumption.

The human capital production function constraint has three key household inputs: the time of the child (e.g., in school or studying at home); the time of the mother; and goods and services that contribute to human capital production. Human capital production also depends on exogenous factors including child gender and birth order, biological endowments, and contextual factors (e.g, parental education, knowledge about the human capital production function, and community characteristics). Earnings of children as adults depend on biological endowments and accumulated human capital, both of which have exogenous market rates of return.

Households seek to maximize household welfare subject to these three constraints. In doing so, households equate the marginal rate of substitution between children's earnings as

adults and household consumption with the marginal cost of investing in the human capital of a child. Households also allocate child time, parental time, and market resources to equalize the marginal costs associated with each. This conceptual framework yields clear predictions on the association between program and contextual characteristics, and expected educational impacts. In particular, to the extent there is variation in characteristics and contexts, this conceptual framework predicts strong heterogeneity in educational impact estimates across the CCT programs in our sample.

First, the cash transfers by themselves, ignoring the conditionality, increase non-employment household income. They do not affect the determinants of the marginal costs of children's time, parental time, or goods and services that contribute to human capital production. As such, cash transfers alone represent a pure income effect. To the extent that adult earnings of children and household consumption are normal goods, cash transfers increase human capital production and consumption. Therefore, the first prediction is that a greater transfer, all else constant, should lead to greater increases in human capital production through increased school enrollment and attendance.

In a single period model, there is no distinction between increases in permanent income or transitory income. However, in a model with multiple time periods, standard consumption theory distinguishes between the expected behavioral effects from a rise in permanent income versus that from a rise in current income[8]. In our sample, some national, well-established programs raise permanent income, while other in a pilot stage do not. With that in mind, the second prediction is that programs raising permanent income should have larger educational impacts than those raising only transitory income.

---

[8] For example, a worker receiving a pay increase behaves differently from a lottery winner.

Simple economic reasoning based on revealed preference establishes that household choice sets are larger for an unconditional transfer than for a conditional one, as the conditionality may induce households to make choices they otherwise would not make with an unconditional transfer. Compliance with CCT program conditions, therefore, likely changes the marginal cost of household investment in human capital. The condition of minimum school attendance, for example, affects the marginal cost of children and parents' time in terms of lost wages. This substitution effect reinforces the income effect of the transfer in incentivizing greater human capital production. Our third empirical prediction is, therefore, that more stringent attendance and achievement conditions lead to, all else constant, steeper reductions in the marginal cost of children and parent's time, leading to extra time devoted to schooling activities (e.g., greater attendance or effort).

Differences in non-earned income and market opportunities across households could explain why, in the absence of a CCT program, some children are enrolled in school while others are not. For a household with children with very high school attendance and very little work time, the conditionalities will not be binding. For this household, the program does not produce a substitution effect, only an income effect, so while the impacts may increase the time devoted to schooling, they will likely not affect enrollment or attendance. Such a household may be representative of those in settings with very high enrollment rates in the absence of the program, as is the case, for example, with primary schooling in many countries. This income-schooling gradient prediction may also stem from regional differences in participant households' income. Our fourth empirical prediction is, therefore, that educational program impacts, particularly on school enrollment and attendance, should be stronger in settings with low baseline enrollment

levels, low levels of participant's household income, and in settings with excess school capacity

in which supply constraints are not binding.

The basic framework thus far outlined assumes a household with unitary preferences.

This implies the household is assumed to maximize a single welfare function, without accounting

for potentially divergent preferences between the adult male and mother of the household or

between parents and children. Most evidence rejects the unitary household model in favor or an

alternative collective bargaining household model (e.g., Doss, 2011; Haddad, Hoddinott, &

Alderman 1997). When the preferences of individual household members are allowed to differ,

conflicts of interest may arise because, for instance, parents may value consumption over human

capital investment more than children (e.g., through higher discount rates). Alternatively, the

mother's objectives may be more closely aligned with those of her children. This potential

conflict of interest within household members serves as the justification for delivering transfers

to mothers, as is common in many CCT programs. Therefore, our fifth prediction is that, to the

extent that targeted households behave according to a bargaining model and the mother's

objectives are more closely aligned with those of children in the household, distributing transfers

to mothers should lead, all else equal, to greater human capital investments in children.

The basic framework collapses household decision-making into a single period. Thus, it

cannot shed light on intertemporal choices stemming from liquidity, credit or savings constraints.

In the single-period model, for example, whether households receive a dollar-equivalent payment

in monthly or bimonthly installments makes no difference. In practice, though, households may

face constraints of liquidity or savings that limit the optimal allocation of resources. For

example, if households face savings constraints while expenditures on goods and services that

contribute to human capital production are "lumpy," then the timing and frequency of payments

may make a difference. Therefore, our sixth and final prediction is that, all else constant, payment frequency will be related to CCT program impacts if households face credit or savings constraints for investments in human capital production.

CCT program costs and cost-effectiveness comparisons

One primary contribution of our meta-analysis is the comparative analysis of cost-effectiveness estimates for various CCT programs. This section lays out a simple model of costs to frame our empirical analyses (e.g., Caldés & Maluccio, 2005; Caldés et al., 2006).

There are three sources of costs associated with CCT program implementation: administrative costs, transfer costs, and private costs. Administrative costs arise from aspects related to program design and implementation, and include operations, personnel, equipment, targeting of beneficiaries, and program auditing. Transfer costs are the sum of demand-side transfers to families (i.e., the subsidies themselves) and supply-side transfers (e.g., infrastructure improvements and materials). Private costs are those incurred by users as a result of program conditions. For example, recipients in many contexts incur costs related to transportation and waiting time to cash payments made through the banking system.

To date, cost analyses of CCTs have focused chiefly on comparative cost-efficiency by estimating program cost-transfer ratios. A program's cost-transfer ratio (CTR) is the administrative and private dollar cost associated with a one-dollar transfer to a beneficiary in a given time period (Caldés et al., 2006; Lindert, Skoufias, and Shapiro, 2006). In other words, CTR is the ratio of non-transfer program costs to total program transfers. For the discussion that follows, it is useful to express CTR in per-beneficiary terms; namely, the ratio of non-transfer program costs per beneficiary, C, to program transfers per beneficiary, T:

$$(1) \quad CTR = \frac{C}{T}$$

We define educational cost effectiveness, *CE*, as the ratio of average educational effect-size per

beneficiary, *ES,* to non-transfer program costs per beneficiary, *C* (Dhaliwal, Duflo, &

Glennester, 2009):

$$(2) \quad CE = \frac{ES}{C}$$

Simple algebraic manipulations yield the following decomposition of *CE:*

$$(3) \quad CE \equiv \frac{ES}{T} * \frac{T}{C} = TE * \frac{1}{CTR}$$

where *TE ≡ ES/T* is transfer effectiveness; in other words, the ratio of educational effect size per

beneficiary to dollar of transfer costs.   As (3) shows, educational cost-effectiveness can be

expressed as transfer effectiveness, *TE,* multiplied by the reciprocal of the cost-transfer ratio.

Since *ES* is estimated at the margin, ideally so should *T*. This implies that for an Intent-

to-Treat estimate of *ES*, *T* should be the difference in the mean transfer payment given to the

treatment group and the mean transfer in the control group. This includes actual transfer

amounts to individuals in the treatment group already attending school, as well as actual transfer

amounts in the control group, which may be non-trivial in programs with substantial benefit

crossover. Due to data limitations, our actual measure of T does not exactly match the

conceptual ideal. Specifically, we know only from each CCT program's description, the nominal

transfer amount per child. While this nominal figure reflects amounts given to individuals

already attending school, we lack program-level information about the difference, at the margin,

between actual transfer amounts received by a participating household and actual amounts

received by non-participants in programs in which there is benefit crossover. Therefore, given

these data limitations, our empirical measure of transfer amounts will be a closer approximation

to the conceptual ideal for CCT programs with close to perfect compliance with the

implementation protocol, in the sense of having almost perfect program take-up among eligible households and no crossovers to non-eligible households.

Empirically, we focus most of our attention on estimating *TE* for programs in our sample. This focus is entirely data driven due to lack of data on administrative costs for most programs. While we can estimate *TE* for all programs in our sample under the outlined program implementation assumptions, we can only estimate *CE* for a subsample of nine programs. Specifically, for three programs in our sample we have CTR estimates from Caldés et al. (2006).[9] For six additional programs, we have administrative costs from Grosh et al. (2008) in the form of administrative costs as a percent of total (administrative + transfers). With data on administrative costs as a percent of total costs, and assuming no private costs,[10] for these six programs, we follow Lindert et al., (2006) and estimate *CTR* as:

$$(4) \qquad CTR = \frac{\% \, Administrative \, Costs}{100 - \% \, Administrative \, Costs}$$

With estimates of *CTR* from Caldés et al. (2006) for three programs and from equation (4) for six others, we can estimate *CE* for a total of nine programs in our sample using equation (3).[11]

Note that the economic model of household decisionmaking also yields predictions about heterogeneity in estimates of transfer-effectiveness and cost-effectiveness. These predictions are particularly sharp for program characteristics that are revenue-neutral and, therefore, do not affect the program administrator's cost ledger. For example, to the extent that more stringent school attendance conditions do not represent additional administrative program costs, more

---

[9] These programs are Mexico's PROGRESA, Honduras' PRAF II and Nicaragua's Red de Protección Social.

[10] The assumption of no private costs is a requirement due to data limitations. However, as Caldés and Maluccio (2005) show for Nicaragua's Red de Proteccion Social (RPS) CCT program, accounting for private costs can increase cost-transfer ratio estimates substantially – by about 30 percent in a given year.

[11] Ultimately, it would be far more desirable to conduct a full comparative cost-benefit analysis, especially since CCTs have multiple objectives related education, health, nutrition and redistribution. This is not feasible within the scope of the present analysis and given the current available data.

stringent attendance conditions lead to, all else constant, steeper reductions in the marginal cost

of children and parent's time, leading to greater educational and cost-effectiveness estimates.

Similarly, to the extent that targeting mothers leads to greater educational impacts at no

additional cost, variation in this characteristic would predict, all else constant, heterogeneity in

transfer-effectiveness and cost-effectiveness estimates in our sample. We also test these

predictions in the results section.

Method

Literature Search and Sample

We limit our review to CCTs in developing countries, which we determine to be

countries in the "low-income," "lower-middle income," or "upper-middle income" categories for

2014, according to World Bank classification.[12] We then search for whether a CCT program

currently exists or previously existed in these countries as of 2015.[13] To do so, we employ three

information sources.

First, we used the World Bank's Conditional Cash Transfers Country Overview and

Project Information database.[14] Next is the World Bank's 2014 social safety net inventory, which

contains information on all social assistance programs for 155 countries worldwide, with detailed

descriptions of the type of program and program name. We retrieve all programs under the

"Conditional Cash Transfer" label for the countries in our sampling frame. This World Bank

---

[12] These income classifications are based on Gross National Income calculated using the World Bank Atlas Method. Of the countries that met the income threshold to be classified as low, lower-middle or upper-middle income in 2014, Argentina and Hungary were re-classified in 2015 to be "high-income." We include these two countries in the analysis.

[13] 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).

[14] Available from:
http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTSOCIALPROTECTION/EXTSAFETYNETSANDTRANSFERS/0,,contentMDK:22055402~pagePK:148956~piPK:216618~theSitePK:282761,00.html, retrieved June 21, 2016.

classification in based on the bank's operational needs. From a theoretical standpoint, however, CCT programs provide income subsidies to households in poverty with conditions attached to schooling and health care. The income subsidy raises permanent income, while the conditions ensure human capital investment. As noted in the conceptual framework section, standard consumption theory distinguishes between expected behavioral effects from a rise in permanent income versus a rise in current income. Therefore, the World Bank classification will mostly identify established programs, namely those that raise permanent income. To identify CCT programs raising current income, we searched the Web using Google to identify CCT programs in these countries that we may have missed under the first two approaches. This omission typically occurred when a country had a pilot CCT project in place, which we include in the analysis (e.g., China's Pilot CCT program or Morocco's Tayssir program).

Under this sampling procedure, we identify 75 CCT programs in developing-country settings. Once we identify our universe of CCT programs, we proceed to determine whether an eligible study was available for each program. [15] We define an eligible study as an impact evaluation reporting effect estimates for a given CCT program and outcome of interest *at a given point in time*. This definition of study implies that if a reference reports one- and two-year impacts for primary school enrollment, we code it as two separate studies. Defining a study using a temporal dimension implies potential clustering of effect estimates within program/outcomes over time. We explain in the methods section below how we handle such clustering for analysis. Study eligibility consists of two criteria:

---

[15] Columns C and D of Suppplemental Table S1 (online only) describe whether we retrieve an impact study meeting the outcome restriction and research design criteria.

1.   *Outcome restriction*: Must report CCT program effect estimates for at least one outcome of interest (school enrollment, attendance, dropout or school completion) and an associated standard error or t-statistic.

2.   *Research design*: Must use a treatment-comparison research design. The comparison group can be experimental (i.e., random assignment) or observational (e.g., matched), and can be wait-list or no treatment at all. One group pre-post designs are not eligible.

To retrieve all relevant studies, in 2015 we searched Google and Google Scholar for references using the following search term variations "COUNTRY CCT impact evaluation" "PROGRAM NAME impact evaluation" or "COUNTRY CCT PROGRAM NAME impact evaluation." 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 CCT review (2009), Milazzo's annotated bibliography on CCT programs (2009), Independent Evaluation Group (2011) and Baird et al.'s review (2014). We found a total of 94 eligible studies (from 59 references) for 47 CCT programs in our universe of developing countries (Table 1).[16]

Study Quality, Effect Estimates and Program Characteristics

Defining and coding study quality. To define and code study quality, we follow the Institute for Educational Sciences' What Works Clearinghouse (WWC) version 2.1 Evidence Standards Protocol (Institute of Educational Sciences, 2011). According to these guidelines, a study's quality rating can be affected by its design, sample attrition and covariate equivalence. Depending on these factors, studies either meet evidence standards, meet standards with reservation, or do not meet evidence standards.

---

[16] As a benchmark, Baird et al. (2014) include 61 reports of 30 CCT programs in developing countries (four of which also include unconditional transfers). This benchmark suggests that our literature search was exhaustive. Supplemental Table S2 (online only) provides details of the programs, references, and studies in the final analysis sample.

Study design can be of two types: randomized or observational (quasi-experimental). Correctly implemented, a randomized study design receives the highest quality rating of meeting evidence standards (which we code as one). Correct implementation of a randomized study meeting evidence standards requires actual randomization of participants and low attrition. Low attrition is defined by a combination of overall attrition and differential attrition between randomized participants and non-participants, and is based on a model developed by WWC statisticians (Institute of Educational Sciences, 2011). For example, a study with an overall attrition rate at follow-up of 20 percent, to meet evidence standards, can have no more than a 4 percentage point differential attrition between randomized participants and non-participants.

Randomized study designs in which the combination of overall and differential attrition exceeds WWC parameters can, at most, meet standards with reservations (which we code as two). A randomized study design meets standards with reservations if – despite having high attrition – it demonstrates baseline equivalence across treatment and control groups. Baseline equivalence requires reported mean differences in characteristics – chiefly mean outcome variables at baseline – to be less than 0.25 standard deviations based on the variation of that characteristic in either the pooled sample or the control group, as reported. Without baseline equivalence in all reported characteristics, a randomized study design does not meet evidence standards (which we code as three). Note, however, that the equivalence condition only applies to randomized study designs with high attrition, not those with low attrition.

Observational (quasi-experimental) study designs can, at best, meet standards with reservations if they show baseline equivalence between treatment and comparison groups, as defined above. If an observational study does not demonstrate equivalence, it does not meet evidence standards. Among observational studies, however, WWC uses a separate protocol for

regression discontinuity designs (RDDs). To be reviewed as an RDD, a study must meet three criteria: (a) treatment assignment uses a cutoff rule based on a pre-determined forcing variable; (b) the forcing variable is cardinal with enough values; and (c) there is no factor confounded with the cutoff rule (i.e., another program using the same cutoff).[17]

For an eligible RDD study to meet evidence standards, it must demonstrate four conditions: (a) integrity of the forcing variable, demonstrated through a detailed institutional description of the scoring and treatment assignment process *and* a formal statistical test to establish smoothness in the density of the forcing variable; (b) low attrition, as defined by the WWC attrition standards on overall and differential attrition; (c) continuity of baseline covariates at the assignment cutoff point; and (d) appropriate functional form and bandwidth. These conditions are demonstrated by a regression framework controlling for the forcing variable, graphical analyses depicting the non-parametric relationship between the forcing variable and outcomes, an optimal bandwidth choice or evidence of sensitivity to alternative bandwidths, and, if there are multiple sites, separate reporting of site impacts.

An RDD study meets standards with reservations if it meets conditions a, b and d above. An RDD study does not meet evidence standards if it fails to meet conditions a, b and d above. An RDD that does not report information on overall and differential attrition (condition b) is reviewed as an observational study if equivalence is demonstrated on key baseline covariates, including continuity of the forcing variable at the cutoff. Three additional methodological aspects are worth mention in the context of RDD studies. First, while estimators based on experimental samples typically identify an average treatment effect, RDD-based estimators identify a local average treatment effect for a subgroup of households near the eligibility cutoff.

---

[17] One study (Cameron, 2009) uses an instrumental variable research design. Since it provides evidence on the validity of the exclusion restriction, we coded the study as meeting standards with reservations.

Second, evidence by Buddelmeyer & Skoufias (2004) suggests that in the context of Mexico's

PROGRESA program, RDD methods may approximate well benchmark estimates using

experimental data.[18] Third, according to the WWC review protocol, RDD studies can, at best,

meet standards with reservations, whereas experimental studies with low attrition meet evidence

standards. Controlling for study quality, therefore, controls for differences in underlying

statistical parameters.

Research design quality is, overall, high among studies included in the sample. Two-

thirds of the studies meet evidence standards or meet standards with reservations; 43 percent

received the highest rating of meeting evidence standards and 24 percent received the rating of

meeting standards with reservations. Only 33 percent of the studies do not meet evidence

standards (Supplemental Table S2, online only).

Coding program effect estimates. We retrieve the "best" effect estimate for each

program for six separate outcome domains, when available. The outcome domains are: primary

school enrollment, primary school attendance, primary school dropout, secondary school

enrollment, secondary school attendance, secondary school dropout and school completion. For

some programs, no impact estimate exists for a given outcome domain, and for those we simply

have no estimate at all. Four cases arise when determining the best impact estimate for a

program/outcome domain combination.

Case 1: This is the case when one study reports an overall program effect for a given

outcome domain and time period. For example, for Cambodia's Education Sector Scholarship

Program, one study was available (Filmer & Schady 2009), focusing on secondary school

---

[18] Cook and Wong (in press) highlight that the Buddelmeyer & Skoufias (2004) replication of PROGRESA experimental estimates in the RDD framework performs better in the second round of follow-up data collection due to the presence of spillover effects among comparison households in the first follow-up round.

attendance impacts after one year. In this case, the best estimate is the estimate from models with the most comprehensive set of control variables in the sample with the best baseline equivalence.

Another instance of Case 1 is programs for which multiple studies report an overall program effect for a given outcome domain and period. In this case, the best program effect estimate on a given outcome domain is the one from the study with the highest quality rating. If the studies have the same quality rating, the best estimate is the most recent study's. As with programs with only one study per outcome domain, the best estimate comes from models with the most comprehensive set of control variables and best baseline equivalence.

Case 2: This case arises when a study reports program effects on multiple non-overlapping subgroups within a given outcome domain and period. For example, the Attanasio et al. (2010) report for Colombia's Familias en Accion impact estimates for primary enrollment separately rural girls, rural boys, urban girls, and urban boys – but not an aggregate primary enrollment effect. In this case, the best estimate is obtained by computing a fixed-effect, precision-weighted, average-effect size for the corresponding domain. To do this, we only use one estimate per subgroup; the subgroup estimates included in the fixed-effect estimator are those with the most comprehensive set of controls and sample with best baseline equivalence.

Another instance of Case 2 is programs for which multiple studies report subgroup impact estimates for a given outcome domain and period. As with Case 1, the best program estimate is drawn from the study with the highest quality rating or, if the studies are rated the same, the most recent. Once the study is chosen, the best estimate is obtained by computing a fixed-effects estimator using subgroup estimates from models with the most comprehensive set of controls and sample with best equivalence.

Case 3: This occurs when there is more than one study per reference because the reference reports impacts for a program and outcome domain for multiple time periods. For example, for Nicaragua's Red de Proteccion Social, Maluccio and Flores (2005) reported impacts on primary enrollment one year and two years after randomization. Since we define a study as an impact evaluation reporting effect estimates for a given outcome of interest at a given point in time, this is a case in which there is more than one study in a reference for a given program. Case 3 studies raise issues related to clustering of effects within programs; in the methods section below, we explain how we deal with them.

Case 4: This case arises when there are multiple independent programs in a study. For example, for Colombia's Subsidios Condicionados a la Asistencia Escolar, Barrera-Osorio, Bertrand, Linden and Perez-Calle (2011) reported impact estimates for a basic CCT treatment, a savings treatment, and a tertiary treatment from separate samples in the same time frame. In this case, we code the best estimate separate for each program. The best estimate, as before, comes from models with the most comprehensive set of controls in the sample with best equivalence.

PIs independently determined the corresponding case for each of the 59 references in our sample (94 studies due to references under Case 3 or Case 4). PIs also separately coded the quality rating, the outcome domain, the domain's best estimate choice (or subgroup estimates), associated standard errors (or t-statistics, as reported), and type of parameter reported (e.g., intent-to-treat). PIs then crosschecked each decision; in all but two instances, the decision was unanimous. In these two instances, PIs revisited together once more each corresponding reference and easily reached a unanimous decision.

In the full sample, 66 percent of parameter estimates are intent-to-treat, typically from randomized study designs or RDDs. Twenty-four percent are average treatment effects, typically

from matched treatment comparison studies. The best estimate in the few remaining studies is either an average treatment effect (6 percent) a local average treatment effect (LATE, 3 percent), or the impact of an extra year of treatment (1 percent). Among studies with a quality rating of 1 or 2, 83 percent of parameter estimates are intent-to-treat, 14 percent are average treatment on the treated parameters and 3 percent are local average treatment effects (LATE).

Coding program characteristics. We capture and code contextual and program characteristics that include: baseline primary and secondary enrollment; nature of 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); member of the household receiving the subsidy (child, mother, father, or both parents); schooling subsidy amount; payment frequency of both schooling and health subsidies; whether the subsidy amounts vary by gender, grade, age, or other characteristics; and whether the program provides supply incentives for education.

We collected data on program characteristics from the World Bank's Conditional Cash Transfers Country Overview and Project Information database[19], ECLAC's Conditional Cash Transfer Database[20] and, for programs not appearing in these two databases, from the relevant reference in the sample. We obtained information on baseline school enrollment from the World Bank Development Indicators database[21], UNECO's data center[22], and, when appropriate, the sample's relevant study.

---

[19] Available from:
http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTSOCIALPROTECTION/EXTSAFETYNETSANDTR
ANSFERS/0,,contentMDK:22055402~pagePK:148956~piPK:216618~theSitePK:282761,00.html, retrieved
June 21, 2016.
[20] Available from: http://dds.cepal.org/bdptc/, retrieved June 21, 2016
[21] Available from: http://data.worldbank.org/data-catalog/world-development-indicators, retrieved June 21,
2016.
[22] Available from: http://data.uis.unesco.org/Index.aspx?DataSetCode=EDULIT_DS, retrieved June 21, 2016.

Characteristics of CCT Programs in Analysis Sample

Table 2 summarizes the CCT programs in our analysis sample. Our sample contains 47 programs in 31 countries. Fifty-three percent are from Latin America, 32 percent are from Asia and 15 percent are African. Sixty-four percent of programs are full-scale and implemented at the national level, 32 percent are pilot programs, while 4 percent of programs are implemented at the state or local level.

Table 2 illustrates the degree of variation in program characteristics. For example, 68 percent of programs condition transfer-receipt on school attendance and 32 percent impose additional conditions such as satisfactory grade promotion or grades. Payment frequency varies; 19 percent pay transfers monthly, 45 percent pay bimonthly, and 36 percent pay transfers less frequently (quarterly, bi-annually, or annually). In 68 percent of programs, all children regardless of age, grade, or gender are entitled to the same transfer amount. In 23 percent of programs, the transfer amount varies by age, grade or gender. In 37 percent of programs, only the mother can receive the payment.

For comparability across countries, we constructed measures of monthly-equivalent average transfer amounts in 2015 U.S. dollars.[23] Payments vary considerably across programs and schooling levels. The average per-pupil monthly transfer for a primary school program is $18.90 (s.d. = $29.60); in secondary school, the average is $23.20 (s.d. = $29). In 91.5 percent of CCT programs in the sample, the transfer is unaccompanied by supply-side interventions. In

---

[23] First converted to nominal dollars in the base year (the year the program begins) using base-year exchange rates and then inflated to dollars in the year of analysis (2015) using the U.S. consumer price index (CPI).

8.5 percent of programs, CCTs to families are complemented with a supply incentive to schools or teachers[24] (Table 2).

Methods to Combine and Analyze Heterogeneity in Impact and Cost-effectiveness Estimates

Combining effect size estimates. Our definition of a study – an impact evaluation reporting effect estimates for a given program and outcome of interest at a given point in time (one-year post-treatment, two-years post-treatment, etc.) – implies potential statistical dependence (i.e., clustering) of effect estimates within a given program over time. Specifically, this occurs for studies in the Case 3 category described in the previous section. We otherwise assume independence of effect size estimates across programs.

We take two approaches to address effect-size dependence within programs and outcome domains over time. One approach is to include only one effect size estimate per program and outcome domain, achieved by including only the effect size estimate for the lengthiest time of program exposure reported in the overall mean calculation. This is the approach we favor in the meta-analytic forest plots presented in the results section. With only one effect estimate per program and outcome domain, we calculate the overall mean effect size for a given outcome domain (e.g., primary enrollment) using a standard intercept-only random effects model.

Our second approach to deal with effect-size dependence explicitly accounts for the block-diagonal structure of the meta-analytic variance matrix using the methods of Hedges,

---

[24] Most of these incentives are in the form of cash transfers to schools: Indonesia's Jarig Pengamanan Social program offered block grants to schools (Fiszbein et al. (2009), Honduras' PRAF II program offered payments to parent-teacher associations in primary schools where a quality-improvement plan was required (Galiani & McEwan, 2013), and Nicaragua's Red de Protección Social program offered a supply cash transfer ("*bono a la oferta*") distributed between teachers and schools (Maluccio & Flores, 2005). In the case of Bangladesh's Female Secondary School Stipend Program, the supply component consisted of complementary investments on infrastructure, instructional materials, and teacher training (Khandker, 2003).

Tipton, and Johnson (2010). In this approach, the mean effect size for a given outcome domain is:

$$(5) \quad ES = \frac{\sum_{j=1}^{m} \sum_{j=1}^{k_j} (w_{ij} ES_{ij})}{\sum_{j=1}^{m} \sum_{j=1}^{k_j} w_{ij}}$$

where m is the number of independent clusters (in our case, programs), k_j is the number of effect size estimates within a cluster, ES_{ij} are the effect size estimates within cluster m, and w_{ij} is the associated weight equal to the reciprocal of the variance.

Assuming all estimates in a program and outcome domain take equal weights, as recommended by Hedges et al. (2010), the robust variance estimate is:

$$(6) \quad v^R = \frac{\sum_{j=1}^{m} w_j^2 (\overline{ES}_j - ES)^2}{(\sum_{j=1}^{m} w_j)^2}$$

Throughout the analysis, we report effect estimates in probability units (i.e., percentage points). This is the most natural unit given that all outcome domains are dichotomous (enrollment, attendance, dropout, and completion) and that the overwhelming majority of studies in sample report effect estimates in probability units.

Computing program transfer-effectiveness estimates. In addition to overall mean effect estimates, we also construct transfer-effectiveness estimates for all programs following the conceptual approach described above. We compute program-level transfer-effectiveness ratios for a given outcome domain as the effect size estimate for that program $\overline{ES}_j$ per dollar of annual transfer amount T_j, or $\overline{ES}_j / T_j$.[25] We compute the variance of $\overline{ES}_j / T_j$ using the delta method,

---

[25] Note that this calculation does not control for how long people are exposed to the program. For example, holding impact constant, a program could have a favorable cost ratio if it pays little annually but exposes participants for a long duration.

assuming that the only source of uncertainty arises from estimation of program impacts, not from transfer amounts.[26]

Note that for many programs, we have effect estimates for various outcome domains. In the program's transfer-effectiveness ratios computed above, we divide effect sizes for all outcome domains available by the transfer amount. Therefore, we 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 takes the perspective of an implementing government interested in judging the cost-effectiveness of an education CCT program against all schooling outcomes and not just, for instance, enrollment.

Analyzing heterogeneity in effect size and transfer-effectiveness estimates. To test the predictions borne out by the economic model of household decision-making outlined above, we explore how specific contextual and program characteristics explain heterogeneity in effect size and cost-effectiveness estimates, running separate meta-regression models for each outcome domain whenever feasible. Specifically, given the number of program-level effect size estimates per domain, this is only feasible for primary enrollment, primary attendance, secondary enrollment, and secondary attendance. For these four domains, we report meta-regression results on both effect sizes and cost-effectiveness.

We pursue two meta-regression approaches. First, we estimate random effects meta-regression models with only one effect estimate per outcome domain as a function of program characteristics. As before, for programs and outcome domains in which multiple effects for

---

[26] This assumption is simply due to data limitations on actual transfer amounts households receive at the margin. Conceptually, there also is uncertainty around mean transfer amounts, so with detailed household level data one could compute uncertainty in transfer amount estimates. We defer this extension for future work.

different lengths of program exposure are reported, we choose the estimate for the longest time of exposure.

In the second approach, we include all time-period estimates for a given program and outcome domain when available, and use the robust random effects model proposed by Hedges et al. (2010) that accounts for clustering of estimates within program over time.

As moderating variables in the meta-regression models, we focus primarily on those for which the theoretical model yields a clear prediction. These moderators are: transfer amounts; whether the program is a national program vs. a pilot (i.e., an increase in income that is either permanent or transitory); minimum schooling attendance conditions and whether the program imposes conditions on achievement beyond the standard attendance conditions; baseline enrollment and the provision of educational supply-side supplements; whether the program targets mothers; and payment frequency. In addition, we control for region, study quality, and publication characteristics. Note that this exercise assumes independence in impact from the different characteristics in the design of the program. To the extent that program characteristics are a bundle designed by policy makers in order to maximize some welfare function, it is possible that effects may depend on the interactions of these design issues. In our analysis we do not consider these possible interaction effects.

Results

We present results in three subsections. First, we discuss heterogeneity in the distribution of educational impacts and transfer effectiveness estimates. Second, we present meta-regression results that empirically test the predictions for the association between impact and transfer-effectiveness estimates and CCT program characteristics borne out by the conceptual framework described above. Third, we present cost-effectiveness estimates.

Heterogeneity in the Distribution of Educational Impact and Transfer-effectiveness Estimates

As documented earlier (Table 2), among programs in our sample there is considerable variation in CCT program characteristics. Our model predicts that heterogeneity in program characteristics that, theoretically, mediate educational impacts should be reflected in heterogeneity in educational impact estimates. We find evidence strongly supporting this prediction. As demonstrated below, for all educational outcomes and schooling levels, we strongly reject the null hypothesis of homogeneity in impact estimates across programs.

School enrollment (Figure 1). The overall random-effects average primary enrollment impact is 3.0 percentage points, statistically different from zero.[27] Relative to the mean baseline primary enrollment of 87 percent, the average impact represents an enrollment increase of 3.4 percent. We strongly reject homogeneity in primary enrollment impact estimates.[28] Average impact estimates or conclusions about heterogeneity in primary school enrollment estimates are robust to excluding studies with the lowest quality rating (Figure S1, online only).

The average secondary enrollment impact is more than twice as large to that of primary enrollment – 7.1 statistically-significant percentage points. Baseline secondary enrollment is, on average, 51 percent, implying that the average secondary enrollment impact estimate represents a 14 percent secondary enrollment increase. The secondary enrollment plot exhibits considerable effect-size variation. For instance, evaluations of Cambodia's CESSP and Mexico's

---

[27] In these forest plot figures, we show only one report impact estimate per program. For programs with impact estimates for multiple time-periods of exposure to treatment, we plot the effect for the lengthiest time of program exposure.

[28] The *chi-square* test-statistic for the null hypothesis of homogeneity in primary enrollment impact estimates is 229 (*p-value* 0.000). Bangladesh's Primary Education Stipend Program program estimate is the largest (17 percentage points, although not statistically significant), followed by Nicaragua's Red de Protección Social (13 percentage points). In contrast, estimates for Macedonia's CCT Project, Ghana's Livelihood Empowerment Against Poverty and Indonesia's Keluarga Harapan are negative and not statistically significant.

Oportunidades report secondary enrollment estimates close to 20 percentage points, and Cambodia's JFPR Scholarship reports an estimate of 31 percentage points.[29] Only one program (Indonesia's Keluarga Harapan) reports a negative impact estimate, which is not statistically significant. Results about mean impact and heterogeneity are robust to excluding the lowest quality studies (Figure S1, online only).

The overall transfer-effectiveness mean for primary enrollment is 0.016 statistically-significant percentage points per dollar/year/child (Figure S2, online only). There is considerable heterogeneity in transfer-effectiveness estimates across programs.[30] With respect to primary enrollment, Indonesia's JPS Scholarship and Grant Program has the highest transfer-effectiveness ratio. Also, Nicaragua's RPS, Morocco's Tayssir, Ecuador's Bono de Desarrollo Humano, the Philippines' Pantawid Pamilyang and Mexico's Oportunidades appear to be more cost-effective than the average program.

The overall transfer-effectiveness mean for secondary enrollment is 0.035 statistically-significant percentage points per dollar/year/child, about twice as large as the average transfer-effectiveness estimate for primary enrollment (Figure S3, online only). In other words, a dollar of transfer allocated to subsidizing secondary schooling appears to be more effective than a dollar allocated to subsidizing primary schooling – even though secondary transfers in the typical program are larger than for primary. We strongly reject homogeneity in secondary enrollment transfer-effectiveness estimates.[31] Cambodia's CESSP Scholarship and JFPR Scholarship

---

[29] The *chi-square* test-statistic for the null hypothesis of homogeneity in secondary enrollment impact estimates is 538 (*p-value* 0.000).

[30] The *chi-square* test-statistic for the null hypothesis of homogeneity in primary enrollment cost-effectiveness estimates is 242 (*p-value* 0.000). We only show transfer-effectiveness estimates' forest plots for enrollment outcomes. For other outcomes, these plots are available upon request.

[31] The *chi-square* test-statistic for the null hypothesis of homogeneity in secondary enrollment cost-effectiveness estimates is 730 (*p-value* 0.000).

programs, and Pakistan's Punjab Female School Stipend have the greatest secondary enrollment transfer-effectiveness ratios; Turkey's SRMP and Honduras' Bono 10,000 have the smallest.

School attendance (Figure 2). The average random-effects primary attendance effect is 2.7 statistically-significant percentage points. Relative to average attendance at baseline of about 80 percent, this represents an increase in attendance of about 3 percent for the average CCT program. We strongly reject the null hypothesis of primary attendance estimate homogeneity (*chi-square*=127, *p-value* 0.000). Two programs have substantially larger impact estimates than average: Burkina Faso's Orphans and Vulnerable Children and Nicaragua's Red de Protección Social, with reported average attendance effect of close to thirteen percentage points.

The CCT average secondary attendance impact estimate is a statistically-significant 5.75 percentage points. There is considerable heterogeneity in secondary attendance impact estimates across programs.[32] At the most impactful extreme stands the JFPR Scholarship Program and CESSP of Cambodia, with average secondary attendance impact estimates of close to 30 percentage points. At the other extreme, we find Tanzania's Community Based CCT with a negative secondary attendance impact (although not statistically significant) and Macedonia's CCT Project with a very small secondary attendance impact estimate that is not statistically different from zero.

The overall transfer-effectiveness mean for secondary attendance is 0.026 statistically-significant percentage points per dollar/year/child. This mean estimate is over three times larger than the transfer-effectiveness mean for primary attendance, which is 0.008 attendance percentage points per transfer dollar/year/child. For both primary and secondary attendance

---

[32] We reject the null hypothesis of homogeneity in impact estimates, with a homogeneity test *chi-square* statistic of 620 and associated *p-value* of 0.000.

transfer-effectiveness, we reject the null hypothesis of homogeneity in transfer-effectiveness estimates. For primary attendance, the programs with the highest primary attendance transfer-effectiveness estimates are Burkina Faso's Orphan and Vulnerable Children and Ghana's Livelihood Empowerment Against Poverty, while the highest transfer-effectiveness ratio for secondary attendance are Cambodia's CESSP and JFPR Scholarship programs.

Dropout (Figure 3). Few CCT programs have evaluations reporting school dropout outcomes. The relatively low number of studies reflects on the uncertainty with which we calculate average impacts. The overall primary dropout impact estimate is -1.2 percentage points. We reject the null hypothesis of homogeneity in primary dropout impact estimates (*chi-square*=156, *p-value* 0.000). Evaluations for Brazil's Bolsa Escola and Nicaragua's Red de Protección Social report the largest dropout reduction impact estimates, while the evaluation of Honduras' PRAF II suggests increases in dropout because of program participation.

The average secondary dropout effect of -2.9 percentage points is more than twice as large than that for primary dropout. There is considerable heterogeneity, yet all reported secondary dropout impact estimates are negative and statistically different from zero. The evaluation from China's Pilot CCT Program stands out as the largest effect in secondary dropout reduction.[33]

The overall transfer-effectiveness mean for primary dropout is a -0.01 statistically-significant percentage point reduction per transfer dollar/year/child. For secondary dropout, it is a -0.02 statistically-significant percentage point reduction per dollar/year/child. The program with the highest transfer-effectiveness ratio for primary school dropout is Brazil's Bolsa Escola. In the case of secondary dropout, Indonesia's Jaring Pengamanan Social program and China's

---

[33] *Chi-square* statistic for homogeneity test=47,000, (*p-value* 0.000).

Pilot CCT Program have transfer-effectiveness ratios that are substantially larger than the average (respectively, -0.06 and -0.04).

School completion (Figure 3). Only 10 programs from six countries have evaluations on school completion. Due to the limited sample size, we pool completion outcomes for various schooling levels. The average school completion impact estimate is 3.28 percentage points and is statistically significant. We reject the null hypothesis of homogeneity in school completion estimates across programs (*chi-square*=30.8, *p-value* 0.001). Cambodia's Scholarship Pilot Poverty Targeting and Merit Targeting have the largest impact estimates (respectively, 18 and 12 percentage points). Tanzania's Community Based CCT program also has a large impact estimate for secondary school completion, but is not statistically significant. Most impact estimates (6 of 10) are not statistically different from zero.

The overall transfer-effectiveness mean for school completion is 0.012 percentage points per transfer dollar/year/child and is statistically significant. Cambodia's Pilot Scholarship Poverty Targeting and Merit Targeting programs have the largest transfer-effectiveness ratios (respectively, 0.82 and 0.55).

Meta-regressions of Impact and Transfer-effectiveness Estimates

Program impacts' meta-regression results. Table 3 reports meta-regression estimation results. Columns (1) to (4) present, for each outcome domain, robust random effects estimates accounting for clustering of estimates within program over time, using all time-period estimates. Columns (5) to (8) present random effects meta-regression estimates with only one effect estimate per outcome domain. Due to the limited number of estimates available for dropout and completion, analyses only include enrollment and attendance outcomes (separately for primary and secondary schooling).

We present six meta-regression findings stemming from predictions borne out by the economic model of household decision-making outlined above. The first prediction from the model is that, all else constant, a greater transfer should lead to greater increases in human capital production through increased school enrollment and attendance. We do not find evidence in support for this prediction. All else constant, transfer amounts are not statistically correlated to effect sizes for any outcome or schooling level. In some specifications, the meta-regression point estimate is even negatively-signed. While inconsistent with the theoretical prediction, this finding is consistent other systematic review evidence on CCTs from Baird et al. (2014) and Snilstveit et al. (2015), and with single CCT program evidence from Cambodia and Malawi. In Cambodia's CESSP program, increases in school enrollment are not related to transfer size (Filmer & Schady, 2011). In Malawi's CCT program, impacts on educational outcomes such as dropout do not vary with transfer amounts (Baird et al., 2009).

The second prediction of the model distinguishes between the expected behavioral effects from a rise in permanent income versus a rise in current income. National, well-established programs in our sample raise permanent income, while those in a pilot stage do not. Following with this logic, in our empirical analyses we distinguish between programs that (arguably) raise permanent income from those only raising transitory income. We find no evidence consistent with this prediction. All else constant, educational effect size estimates from national CCT programs are statistically indistinguishable from those of pilot CCT programs.

The third prediction of the model is that, all else constant, more stringent attendance and achievement conditions lead to, all else constant, steeper reductions in the marginal cost of children and parent's time, leading to extra time devoted to schooling activities (e.g., greater attendance or effort). We find some evidence consistent with this prediction. Specifically, all

else constant, effect sizes for primary attendance are statistically significantly larger when, in

addition to school enrollment, other conditions are required such as grade promotion or test

scores. However, for primary and secondary enrollment, and for secondary attendance, we do

not find a consistent association between estimates and programs imposing additional schooling

conditions.[34]

The fourth model prediction is that educational program impacts, particularly on school

enrollment and attendance, should be stronger in settings with low baseline enrollment levels and

in settings with excess school capacity in which supply constraints are not binding. We find

strong evidence in support of this prediction, particularly for attendance. Enrollment and

attendance impact estimates are smaller in settings with high baseline enrollment. For both

primary and secondary attendance impact estimates, the (negative) association with baseline

enrollment is statistically significant. This result, in particular, is also consistent 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, & Todd,

2005). Once we control for baseline enrollment and other program characteristics, however, the

geographic location of programs as captured by the Latin America vs other regions indicator is,

in general, not predictive of stronger program impacts.[35]

Moreover, impact estimates for primary enrollment and attendance are larger in programs

that, all else constant, complement cash transfers to families with supply-side interventions such

---

[34] It is important to underscore that we are not comparing UCTs vs CCTs (as all programs in our sample are CCTs) or whether conditions are monitored or enforced, but rather whether in addition to school enrollment or attendance, other conditionalities are imposed to students (such as passing a grade or having a minimum test score). Baird et al. (2014) found that transfer programs in which conditions are more rigorously monitored and enforced have larger enrollment impacts.

[35] To the extent that average household income is greater in Latin America than in other regions, the theoretical model would have predicted a negative and statistically significant sign for the Latin America indicator variable. In contrast with this prediction, coefficient estimates from all specifications are positive and with the exception of one specification (Column 8, Table 3) not statistically significant.

as school grants or direct payments to parent-teacher associations or to teachers (see Table 2).

This result is robust to alternative specifications (Row 5, Columns 1, 2, 5, 6, Table 3). The

association between the magnitude of primary enrollment and attendance impact estimates, and

the existence of a supply-side intervention, however, needs to be cautiously interpreted, as only

three CCT programs for primary schooling have a supply component (Honduras' PRAF II,

Nicaragua's Red de Protección en Crisis [RPS], and Indonesia's Jaring Pengamanan Social

[JPS]). This caveat notwithstanding, impact estimates for two out of these three programs with a

supply component are substantially larger than the overall effect size (RPS and JPS for primary

enrollment, and PRAF II and RPS for primary attendance).[36]

The fifth prediction from the economic model is that, to the extent that target households

behave according to a bargaining model and the mother's objectives are more closely aligned

with those of children in the household, targeting transfers to mothers (when present) should

lead, all else equal, to greater human capital investments in children. We do not find evidence

consistent with stronger effects for transfers that target mothers. This pattern of results is

empirically consistent with prior review evidence from Baird et al. (2014) and with experimental

evidence from Morocco's Tayssir program (Benhassine et al., 2015). This result leaves open the

possibility that unitary household decision-making is perhaps an adequate benchmark for

decision-making in target CCT households for programs in our sample.

The sixth prediction from our theoretical model is that, all else constant, payment

frequency will be related to CCT program impacts if households face credit or savings

constraints for investments in human capital production. While the sign of the coefficients is

---

[36] We do not take a strong stance on the coefficient for having a supply-side component in models for secondary enrollment and attendance because, for the former, only two programs have a supply component (Indonesia's JPS and Bangladesh's Female Secondary School Stipend Program), whereas for the latter only one program does so (Indonesia's JPS program).

consistent with greater enrollment and attendance impact estimates in programs with less frequent payments, all else constant, these coefficients are largely statistically insignificant, preventing us from making strong inferences.

There are two additional results not directly related to predictions from our conceptual framework. The first is that, based on the sign of the coefficients, lengthier program exposure is associated with greater impact estimates, all else constant. The association, however, is only statistically significant in one out of eight models, casting doubt on the robustness of this result.[37]

Our second result unrelated to predictions from the model is methodological. Study quality is, all else constant, unrelated to impact estimates for any outcome or schooling level. This finding highlights the fact that, existing evaluations of CCT programs using treatment-control contrasts are, in general, rigorous.

Transfer-effectiveness meta-regression results. Table 4 presents moderators of transfer-effectiveness estimates, defined as impact per transfer dollar/year/child.[38] Columns (1) to (4) present, for each outcome domain, robust random effects estimates accounting for clustering of estimates within programs over time, using all time-period estimates. Columns (5) to (8) present random effects meta-regression estimates with only one effect estimate per outcome domain.

The economic model of household decisionmaking produced sharp predictions about heterogeneity in transfer-effectiveness estimates for program characteristics that, in principle, are

---

[37] We also estimated whether effect sizes varied for one-year of exposure versus two years or longer (including a dummy for one year against more than one year instead of a continuous variable) and find similar results. (Results available upon request.)

[38] These results should be interpreted under the caveats outlined in the Conceptual Framework section, in the sense that, due to data limitations, we are unable to estimate actual transfer amounts at the margin, which is the conceptually ideal transfer metric.

revenue-neutral and, therefore, do not affect the program administrator's cost ledger. We argued in the conceptual framework section this was particularly the case for payment frequency, the stringency of school attendance conditions, and the targeting of mothers. But we find limited empirical support for these predictions.

Specifically, we find some evidence in support for the prediction that less frequent payments are associated with greater transfer-effectiveness estimates. In all specifications, the coefficient estimate is negatively-signed and is statistically significant for one secondary enrollment specification and for one secondary attendance specification..

We find no evidence that more stringent school attendance conditions are associated with greater transfer-effectiveness estimates. This is not even the case for primary attendance, for which we find that more stringent school attendance conditions are associated with greater impact estimates (Table 3). We also find no evidence suggesting that targeting mothers is associated, all else constant, with greater transfer effectiveness estimates (Row 3, Table 4), which is consistent with our finding that targeting mothers is also not associated with greater educational impact estimates.

Consistent with the economic model and with our earlier findings on educational program impacts, transfer-effectiveness estimates are, all else constant, stronger in settings with low baseline enrollment levels and in settings with excess school capacity in which supply constraints are not binding. In all specifications, the coefficient estimate is negative. The coefficient on baseline enrollment is statistically significant for all primary attendance specifications and for one secondary attendance specification.

Moreover, transfer-effectiveness estimates are, in some primary enrollment and attendance specifications, statistically greater when programs provide a supply-side

complementary intervention. Note, however, this result is not robust to alternative

specifications. In addition, this result does not imply that increased impacts because of

complementary supply side interventions outweigh additional supply-related costs, as we only

account for direct transfer amounts in the cost ledger. But it does suggest that a dollar of transfer

is potentially more cost-effective in programs targeting primary schooling and that include a

supply-side intervention as well.

Cost-effectiveness estimates

For nine CCT programs, we can compute cost-effectiveness estimates. Following our

conceptual framework, we obtain these by combining our own data on transfer-effectiveness

with data from ancillary sources on program cost-transfer ratios (Caldés et al., 2006; Grosh et al.,

2008). We present cost-effectiveness findings in Table 5.

Among the nine programs, there is considerable cost-efficiency heterogeneity, as

measured by CTR. In Ecuador's Bono de Desarrollo Humano and Bangladesh's Primary

Education Stipend Program, for instance, it takes about four cents of administrative costs to

deliver one dollar of transfers to beneficiaries – well below the (arithmetic) average of 21 cents

of administrative costs/year for each dollar of transfer. At the other extreme, in Honduras'

PRAF II and in Nicaragua's Red de Protección Social it takes, respectively, 50 cents and 63

cents to deliver one dollar of transfers to beneficiaries.

As documented earlier, there also is considerable heterogeneity in program impact and

transfer-effectiveness estimates. It is, therefore, not surprising to find considerable heterogeneity

in cost-effectiveness estimates. If one were to focus solely on primary enrollment – the outcome

for which we have the most transfer-effectiveness estimates – one would conclude that

Ecuador's Bono de Desarrollo Humano and Bangladesh's Primary Education Stipend Program

are the most cost-effective CCT programs, as they deliver, respectively, 1.3 and 13 percentage

points of enrollment impact per dollar of administrative cost/year. Programs like Mexico's

PROGRESA and Brazil's Bolsa Familia, on the other hand, deliver between 0.4 and 0.5

percentage points of enrollment impact per dollar of administrative cost/year.

  With the caveat of only having data on transfer-effectiveness for secondary enrollment in

three out of these nine CCT programs for which we can make cost estimates, it appears that

Mexico's PROGRESA and Colombia's Familias en Acción are more cost-effective than Brazil's

Bolsa Familia. Mexico's PROGRESA delivers a 0.21 percentage point increase in secondary

enrollment per dollar of administrative cost/year. Colombia's Familias en Acción delivers a 0.12

percentage point increase in secondary enrollment per dollar of administrative cost/year. For our

most distal outcome of secondary school graduation, we only have data on CTR and transfer-

effectiveness for Colombia's Familias en Acción. This program delivers a 0.10 percentage point

increase in secondary graduation per dollar of administrative cost/year.

Again with the caveat of a limited sample size, it is unclear whether the most cost-efficient

programs (i.e., those with low cost-transfer ratios) are also the most transfer-effective (i.e.,

producing the greatest educational impact per dollar of transfer). We find, for instance, a

correlation of -0.13 between a program's cost-transfer ratio and its primary enrollment transfer-

effectiveness estimate, suggesting that more cost-efficient programs (weakly) produce greater

primary enrollment impacts per dollar of transfer (N=8). In contrast, we find a correlation of

0.87 between a program's cost-transfer ratio and its primary attendance transfer-effectiveness

estimate, suggesting that more cost-efficient programs (strongly) produce smaller primary

attendance impacts per dollar of transfer (N=6).

Conclusion

Conditional cash transfer (CCT) programs are one of the most prevalent type of social-assistance programs in the developing world. This study integrates, quantitatively and in a systematic manner, the most recent and comprehensive available evidence on the educational impacts and cost-effectiveness of CCT programs.

In this study, we meta-analyze for impact and cost-effectiveness up to 94 studies from 47 CCT programs in low- and middle-income countries worldwide, focusing on educational outcomes that include enrollment, attendance, dropout, and school completion. To conceptually guide and interpret the empirical findings of our meta-analysis, we present a simple economic framework on household decision-making that generates predictions, all else constant, for the association between program context and design characteristics, and impact estimates. We also present a simple model for the analysis of program costs, using it to compute cost-effectiveness estimates for nine CCT programs for which we have data on transfer-effectiveness (impact units per dollar of transfer) and cost-transfer ratios (administrative costs per dollar of transfers).

Our framework for analysis and data allow us to significantly advance – in numerous ways – our understanding of the interplay between design characteristics, educational impacts, and cost-effectiveness of CCT programs in developing countries. Relative to recent systematic reviews, for example, we analyze a much larger array of CCT programs. Our study also is the first to quantitatively meta-analyze educational impact and cost-effectiveness estimates for multiple educational outcomes separately for primary and secondary schooling. In addition, our conceptual framework predicts empirical relationships between educational impact estimates for a broader set of program characteristics relative to those analyzed in recent prior reviews (e.g.,

Snilstveit et al., 2015; Baird et al., 2014). Finally, our review is the first to compute educational cost-effectiveness ratios for multiple CCT programs.

We find unambiguous empirical support for the prediction that variation in program characteristics maps onto heterogeneity in impact and cost-effectiveness estimates across our sample. For educational cost-effectiveness estimates, in particular, heterogeneity arises both from variation in transfer-effectiveness as well as in cost-transfer ratios across programs.

Our empirical evidence validates some – but not all – theoretically-derived predictions for the relationship between heterogeneity in program characteristics and heterogeneity in impact and cost-effectiveness estimates. We find that, all else constant, having a mother or female head of household as required payee is not correlated with effect sizes in enrollment or attendance for any school level. This finding is consistent with the meta-analysis of Baird et al. (2014), but is not consistent with the meta-analysis of Snilstveit et al. (2015). One possible explanation for this inconsistency is that Snilstveit et al. (2015) had a smaller number of CCT programs than included in our study and their moderators analysis included a smaller number of moderator variables.

Consistent with Baird et al. (2014), we find that neither payment frequency (that is, whether cash transfers are payed on a monthly, bimonthly, quarterly, or annual basis) nor whether a program is full-scale or pilot are associated with enrollment effect sizes. In addition, we show this also holds for attendance impact estimates.

Previous narrative reviews (Fiszbein et al., 2009) reported larger impacts on schooling outcomes when baseline enrollment rates are lower. Our study – which includes more than twice as many programs as Fiszbein & Schady (2009) and measures quantitatively this association – finds that this only holds for attendance. Everything else equal, baseline enrollment levels are

not associated with enrollment effect sizes. This is consistent with Baird et. al (2014), which showed follow-up enrollment rate in the control group is not associated with effect sizes on enrollment.

A novel contribution of this paper is the inclusion of two moderator variables: if the program has a supply component, and whether other conditionalities are required in addition to school enrollment and attendance. We find that primary enrollment and attendance impact estimates are larger – on an absolute and per-dollar of transfer basis – in CCT programs that complement cash transfers to families with supply-side interventions such as school grants or cash transfers to teachers or parent-teacher associations. Consistent with the idea of binding educational capacity constraints, we also find some evidence suggesting how impacts on a per-dollar of transfer basis are smaller in contexts with high baseline enrollment.

We also do not find a robust association between effect sizes in primary enrollment, and secondary enrollment and attendance, and whether programs impose conditions on school achievement beyond the standard attendance requirements. This result is consistent with meta-analytic evidence indicating that CCTs and unconditional transfer programs have, on average, comparable impacts on school enrollment (Baird et al., 2014). We do find, however, that imposing achievement conditions, in addition to school enrollment or attendance, is associated with larger primary attendance effect sizes. This is a novel contribution from previous studies that examined the strictness of enforcement and monitoring – but not the type of condition. Baird et al. (2014) found no differences in impact estimates between UCTs and CCTs, but did find larger effects sizes on enrollment and attendance when conditions are monitored and enforced. Here, we show that for attendance, it matters if additional conditions are required, on top of school enrollment or attendance requirements.

It is important to underscore that meta-regression results are associations and do not imply causation. Although we include a broad set of program and contextual characteristics in the multivariate analysis, there may be other characteristics of a program or country associated to effect sizes not included in the analysis. Therefore, results should be interpreted with caution. Also, effect sizes may vary according to interactions between design features such as transfer size and conditionality type, or transfer recipient and payment frequency. Further research on this direction would entail a larger sample of studies to estimate interactions between design characteristics.

Another avenue of future research could refine and extend our analysis of educational cost-effectiveness. For example, we could only retrieve data on cost-transfer ratios from nine CCT programs and so only estimated educational cost-effectiveness ratios for a subsample of nine programs for selected outcomes. Ultimately, it would be far more desirable to conduct a full comparative cost-benefit analysis, especially because CCTs have multiple objectives related education, health, nutrition, and redistribution. This is not feasible with the current available data. Improving on this domain is also an area of potentially promising future research.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Ahmed, A. U., Gilligan, D., Kudat, A., Colasan, R., Tatlidil, H., & Ozbilgin, B. (2006). *Interim Impact Evaluation of the Conditional Cash Transfers Program in Turkey: A Quantitative Assessment.* Washington, D.C.: International Food Policy Research Institute.

*Akresh, R., de Walque, D., & Kazianga, H. (2013). *Cash Transfers and Child Schooling: Evidence from a Randomized Evaluation of the Role of Conditionality.* (Policy Research Working Paper No. 6340) Washington, D.C.: World Bank. doi: 10.1596/1813-9450-6340.

*Alam, A., Báez, J., & Del Carpio, X. V. (2011). *Does Cash for School Infl uence Young Women's Behavior in the Longer Term? Evidence from Pakistan*. (Policy Research Working Paper No. 5669). Washington D.C.: World Bank.

*Alatas, V., Cahyadi, N., Ekasari, E., Harmoun, S., Hidayat, B., Janz, E., … Wai-Poi, M. (2011). *Program Keluarga Harapan: Main Findings from the Impact Evaluation of Indonesia's Pilot Household Conditional Cash Transfer Program*. Jakarta: World Bank.

*Armand, A., & Carneiro, P. (2013). *Impact Evaluation of the Conditional Cash Transfer Program for Secondary School Attendance in Macedonia: Results from the First Follow-Up*. Republic of Macedonia: Ministry of Labour and Social Policy.

*Arraiz, I., & Rozo, S. (2011). *Same Bureaucracy, Different Outcomes in Human Capital? How Indigenous and Rural Non-Indigenous Areas in Panama Responded to the CCT*. (Working Paper OVE/WP-03/11). Washington, D.C.: Inter-American Development Bank.

*Attanasio, O., Fitzsimons, E., Gomez, A., Gutiérrez, M. I., Meghir, C., & Mesnard, A. (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. doi: 10.1086/648188

* Barrera-Osorio, F., Linden, L., & Saavedra, J. E. (2016). *Medium Term Educational Consequences of Alternative Conditional Cash Transfer Designs: Experimental Evidence from Colombia.*   Unpublished manuscript.

*Báez, J., & Camacho, A. (2011). *Assessing the Long-term Effects of Conditional Cash Transfers on Human Capital: Evidence from Colombia.* (Policy Research Working Paper 5681). Washington D.C.: World Bank.

Baird, S., McIntosh, C., & Özler, B. (2009). *Designing Cost-Effective Cash Transfer Programs to Boost Schooling Among Young Women in Sub-Saharan Africa* (Policy Research Working Paper No. 5090). Washington, D.C.: World Bank. doi: 10.1596/1813-9450-5090

*Baird, S., McIntosh, C., & Özler, B. (2011). Cash or Condition? Evidence from a Cash Transfer Experiment. *Quarterly Journal of Economics, 126*(4), 1709-1753. doi:10.1093/qje/qjr032

Baird, S., Ferreira, F., Özler, B, and Woolcock, M (2014). Conditional, Unconditional and Everything in Between: A Systematic Review of the Effects of Cash Transfer Programs on Schooling Outcomes. *Journal of Development Effectiveness*, *6*(1), 1-43. doi: 10.1080/19439342.2014.890362

*Barrera-Osorio, F., Bertrand, M., Linden, L. & Perez-Calle, F. (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. doi:10.1257/app.3.2.167

*Barrera-Osorio, F., & Filmer, D. (2016). Incentivizing Schooling for Learning. *Journal of Human Resources*, *51*(2), 461-499. doi: 10.3368/jhr.51.2.0114-6118R1

*Baulch, B. (2011). The medium-term impact of the primary education stipend in rural

   Bangladesh. *Journal of Development Effectiveness*, *3*(2), 243-262. doi:

   10.1080/19439342.2011.570449

*Behrman, J. R., Gallardo-García, J., Parker, S. W., Todd P.E., & Vélez-Grajales, V. (2012). Are

   conditional cash transfers effective in urban areas? Evidence from Mexico. *Education

   Economics, 20*(3), 233-259. doi:10.1080/09645292.2012.672792

Behrman, J. R., Parker, S. W., & Todd, P. (2005). Long-term impacts of the Oportunidades

   conditional cash transfer program on rural youth in Mexico. *Discussion Paper 122*.

   Goettingen, Germany: Ibero-American Institute for Economic Research.

*Behrman, J. R., Parker, S. W., & Todd, P. E. (2004). *Medium-Term Effects of the

   Oportunidades Program Package, Including Nutrition, on Education of rural children Age

   0-8 in 1997* (Technical Document No. 9). Mexico DF: Instituto Nacional de Salud Pública.

*Benedetti, F., Ibarrarán, P., & McEwan, P. (2015). *Do Education and Health Conditions Matter

   in a Large Cash Transfer? Evidence from a Honduran Experiment* (IDB Working Paper

   Series No. IDB-WP-577). Washington, D.C.: Inter-American Development Bank.

*Benhassine, N., Devoto, F., Duflo, E., Dupas, P. & Pouliquen, V. (2015). Turning a shove into

   a nudge: A "labeled cash transfer" for education, *American Economic Journal: Economic

   Policy, 7*(3), 86-125. doi: 10.1257/pol.20130225

Buddelmeyer, H. & Skoufias, E. (2004). *An Evaluation of the Performance of Regression

   Discontinuity Design on PROGRESA*. (Policy Research Working Paper No. 3386)

   Washington, D.C.: World Bank. doi: 10.1596/1813-9450-3386

Caldés, N. & Maluccio, J.A. (2005). The Cost of Conditional Cash Transfers. *Journal of

   International Development, 17:* 151-168.

Caldés, N., Coady, D. & Maluccio, J.A. (2006). The Cost of Poverty Alleviation Transfer

Programs: A Comparative Analysis of Three Programs in Latin America. *World*

Development, 34(5): 818-837.

*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. doi:10.1016/j.econedurev.2007.09.013

*Cardoso, E., & Souza, A. P. (2004). *The Impact of Cash Transfers on Child Labor and School*

Attendance in Brazil (Working Paper 04-W07). Department of Economics, Vanderbilt

University.

*Chaudhury, N., Friedman, J., & Onishi, J. (2013). *Philippines Conditional Cash Transfer*

Program Impact Evaluation 2012 (Report Number 75533-PH). Washington, D.C.: World

Bank.

Cook, T. & Wong, V. (in press). Empirical Tests of the Validity of the Regression Discontinuity

Design. *Annales d'Economie et de Statistique.*

*Dammert, A. C. (2009). Heterogeneous Impacts of Conditional Cash Transfers: Evidence from

Nicaragua. *Economic Development and Cultural Change, 58*(1), 53-83. doi: 10.1086/605205

*De Brauw, A., & Gilligan, D. (2011). *Using the Regression Discontinuity Design with Implicit*

Partitions: The Impacts of Comunidades Solidarias Rurales on Schooling in El Salvador

(IFPRI Discussion Paper 01116). Washington D.C.: International Food Policy Research

Institue.

*De Brauw, A., Gilligan, D. O., Hoddinott, J., & Roy, S. (2015). The Impact of Bolsa Familia on

Schooling *World Development, 70*, 303–316. doi: 10.1016/j.worlddev.2015.02.001

*De Janvry, A., Finan, F., & Sadoulet, E. (2006). *Evaluating Brazil's Bolsa Escola Program: Impact on Schooling and Municipal Roles*. Berkeley: University of California at Berkeley.

*Departamento Nacional de Planeación. (2006). *Programa Familias en Acción. Impacto del Programa a un Año y Medio de su Ejecución* (Vol. 4). Bogotá D.C.: Departamento Nacional de Planeación - DNP.

*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.

Doss, C. (2013). Intrahousehold Bargaining and Resource Allocation in Developing Countries. *World Bank Research Observer, 28*, 52–78.

*Evans, D., Hausladen, S., Kosec, K., & Reese, N. (2014). *Community-Based Conditional Cash Transfers in Tanzania: Results from a Randomized Trial.* Washington, D.C.: World Bank.

*Ferré, C., & Sharif, I. (2014). *Can Conditional Cash Transfers Improve Education and Nutrition Outcomes for Poor Children in Bangladesh? Evidence from a Pilot Project*. (Policy Research Working Paper No. 7077). Washington, D.C.: World Bank.

*Filmer, D. & Schady, N. (2008). Getting Girls into School: Evidence from a Scholarship Program in Cambodia. *Economic Development and Cultural Change, 56*(3), 581-617. doi: 10.1086/533548

*Filmer, D. & Schady, N. (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. doi:10.1016/j.jdeveco.2010.05.006

*Filmer, D. & Schady, N. (2009). *School Enrollment, Selection and Test Scores* (Policy Research Working Paper 4998). Washington, D.C.: World Bank.

Fiszbein, A., Schady, N., Ferreira, F., Grosh, M., Kelleher, N., Olinto, P., & Skoufias, E. (2009). *Conditional Cash Transfers: Reducing Present and Future Poverty*. Washington D.C.: World Bank.

*Gajate-Garrido, G. (2013). *Assessing the Differential Impact of "Juntos" Conditional Cash Transfer on Indigenous Peoples.* Unpublished manuscript.

*Galiani, S., & McEwan, P. (2013). The heterogeneous impact of conditional cash transfers. *Journal of Public Economics*, *103*, 85-96. doi: 10.1016/j.jpubeco.2013.04.004

*Gutiérrez, J. P., Abreu, R., & Charles, A. (2011). *Evaluación Externa de Impacto del Programa de Transferencias Monetarias Condicionadas: Efectos en Indicadores de Educación.* México: Institución Nacional de Salud Pública - INSP.

Haddad, L. J., Hoddinott, J. F., & Alderman, H. (1997). *Intrahousehold resource allocation in developing countries* Baltimore, MD Johns Hopkins University Press.

*Handa, S., Park, M., Osei, R., Osei-Akoto, I., Davis, B., & Diadone, S. (2013). *Livelihood Empowerment Against Poverty Program Impact Evaluation.* Chapel Hill, NC: Carolina Population Center University of North Carolina at Chapel Hill.

*Hasan, A. 2010. *Gender-targeted Conditional Cash Transfers: Enrollment, Spillover Effects and Instructional Quality.* (Policy Research Working Paper No. 5257). Washington, D.C.: World Bank.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust Variance Estimation in Meta-regression with Dependent Effect Size Estimates. *Research Synthesis Methods, 1*(1), 39-65. doi: 10.1002/jrsm.5

*Heinrich, C. (2005). *Demand and Supply-Side Determinants of Conditional Cash Transfer Program Effectiveness: Improving the First-Generation Programs*. (Working Paper OVE/WP-05/05). Washington, D.C.: Inter-American Development Bank.

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

*Khandker, S. R., Pitt, M. M., & Fuwa, N. (2003). *Subsidy to Promote Girls' Secondary Education: The Female Stipend Program in Bangladesh.* Unpublished manuscript.

*Levy, D., & Ohls, J. (2010). Evaluation of Jamaica's PATH conditional cash transfer programme. *Journal of Development Effectiveness*, *2*(4), 421-441. doi: 10.1080/19439342.2010.519783

Lindert, K., Skoufias, E. & Shapiro, J. (2006). Redistributing Income to the Poor and the Rich: Public Transfers in Latin America and the Caribbean. Social Protection Discussion Paper 605, Washington, D.C.: World Bank.

*Macours, K., & Vakis, R. (2014). Changing Household's Investment Behavior Through Social Interactions with Local Leaders: Evidence from a Randomised Transfer Programme. *The Economic Journal*, *124*(May), 607-633. doi: 10.1111/ecoj.12145

*Maluccio, J., Murphy, A., & Regalia, F. (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. doi: 10.1080/19439340903584085

*Maluccio, J. A., & Flores, R. (2005). *Impact Evaluation of a Conditional Cash Transfer Program.* (Research Report No. 141). Washington, D.C.: International Food Policy Research Institute.

Milazzo, A. (2009). *Conditional Cash Transfers: An Annotated Bibliography*. Retrieved from

http://siteresources.worldbank.org/SAFETYNETSANDTRANSFERS/Resources/281945-

1131738167860/CCT_Biblio_6Feb2009.pdf

*Mo, D., Zhang, L., Yi, H., Luo, R., & Brinton, S. R. C. (2013). School Dropouts and

Conditional Cash Transfers: Evidence from a Randomised Controlled Trial in Rural China's

Junior High Schools. *Journal of Development Studies*, *49*(2), 190-207. doi:

10.1080/00220388.2012.724166

Molina-Millan, T., Barham, T., Macours, K., Maluccio, J. A., & Stampini, M. (2016). *Long-

Term Impacts of Conditional Cash Transfers in Latin America: Review of the Evidence* (IDB

Working Paper Serial No. IDB-WP-732). Washington, D.C.: Inter-American Development

Bank.

* No author(2008). *Impacto del Programa Solidaridad en Asistencia Escolar y Niveles

Nutricionales: Un Enfoque Empírico.* Unpublished manuscript.

*Núñez, F., Llanos, O., Rivas, G., Pedraza, S., Villegas, M., Bateman, A., . . . Castañeda, C.

(2011). *Evaluación del programa Familias en Acción en grandes centros urbanos: Informe

Final*. Bogotá: Centro Nacional de Consultoría.

*Parker, S., Todd, P. E., & Wolpin, K. I. (2006). *Within-family Treatment Effect Estimators: The

Impact of Oportunidades on Schooling in Mexico*. Unpublished manuscript.

*Perez-Ribas, R. P., Veras-Soares, F., Teixeira, C., Silva, E., & Hirata, G. (2011). *Externality

and Behavioural Change Effects of a Non-Randomised CCT Programme: Heterogeneous

Impact on the Demand for Health and Education* (PEP-PIERI Working Paper No. 2011-19).

Poverty and Economic Policy Research Network.

*Perova, E., & Vakis, R. (2012). 5 Years in Juntos: New Evidence on the Program's Short and Long-Term Impacts. *Economía*, 35(69), 53-82.

Petrosino, A., Morgan, C., Fronius, T. A., Tanner-Smith, E. E., & Boruch, R. F. (2013). *Interventions in Developing Nations for Improving Primary and Secondary School Enrolment of Children: a systematic review, 3ie Systematic Review*. London: International Initiative for Impact Evaluation (3ie).

*Ponce, J. (2006). *The impact of a conditional cash transfer program on school enrolment: the 'Bono de Desarrollo Humano' of Ecuador.* (Working Paper 06/302). Quito, Eduador: Facultad Latinoamerica de Ciencias Sociales - Sede Ecuador.

Rawlings, L. B., & Rubio, G. M. (2005). Evaluating the Impact of Conditional Cash Transfer Programs. *The World Bank Research Observer, 20*(1), 29-55. doi: 10.1093/wbro/lki001

*Robertson, L., Mushati, P., Eaton, J. W., Dumba, L., Mavise, G., Makoni, J., . . . Gregson, S. (2013). Effects of unconditional and conditional cash transfers on child health and development in Zimbabwe: a cluster-randomised trial. *Lancet*, *381*(9874), 1283–1292. doi: 10.1016/S0140-6736(12)62168-0.

*Salvia, A., Tuñón, I., & Poy, S. (2015). Asignación Universal por Hijo para Protección Social: impacto sobre el bienestar económico y el desarrollo humano de la infancia. *Población & Sociedad*, *22*(2), 101-134.

*Schady, N., & Araujo, M. C. (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*(1), 199–250. doi:10.1016/j.jdeveco.2003.12.009

*Sinha, N., & Yoong, J. (2009). *Long-Term Financial Incentives and Investment in Daughters: Evidence From Conditional Cash Transfers in North India* (Working Paper WR-667). Santa Monica, CA: RAND.

*Skoufias, E., & Parker, S. (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.

Snilstveit, B., Stevenson, J., Phillips, D., Vojtkova, M., Gallagher, E., Schmidt, T., . . . Eyers, J. (2015). *Interventions for improving learning outcomes and access to education in low- and middle-income countries: a systematic review: International Initiative for Impact Evaluation (3ie)*. . London: International Initiative for Impact Evaluation (3ie).

*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. doi: 10.1111/j.1468-0084.2006.00438.x

Stampini, M. and Tornarolli, L. 2012. The Growth of Conditional Cash Transfers in Latin America and the Caribbean: Did They Go Too Far? IDB Policy Brief n. 185. Washington D.C., USA: Inter-American Development Bank.

*Todd, P. E., Gallardo-Garcia, J., Behram, J. R., & Parker, S. W. (2005). Impacto de Oportunidades sobre la Educación de Niños y Jóvenes de Áreas Urbanas Después de un Año de Participación en el Programa. In B. Hernández-Prado and M. Hernández-Ávlia (Eds.). *Evaluación Externa de Impacto del Programa Oportunidades* (pp. 165-227). Mexico D.F.: Instituto Nacional de Salud Pública.

*Vera, D. (2011). Matriculación y Trabajo Infantil: Un Análisis Quasi Experimental. *Serie de Documentos de Trabajo sobre Desarrollo* (# 11/2011). La Paz, Bolivia: Instituto de Estudios Avanzados en Desarrollo - INESAD

*Ward, P., Hurrell, A., Visram, A., & Riemenschneide, N. (2010). *Cash Transfer Programme for Orphans and Vulnerable Children (OT_OVC), Kenya: Operational and Impact Evaluation, 2007-2009. Final Report*. Oxford: Oxford Policy Management.

Institute of Educational Sciences (2011). *What Works Clearinghouse: Procedures and standards handbook version 2.1.*. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v2_1_standards_handbook.pdf.

World Bank. (2012). *World Development Indicators*. Retrieved from http://data.worldbank.org/data-catalog/world-development-indicators.

World Bank. (2014). *The state of social safety nets*. Retrieved from https://openknowledge.worldbank.org/bitstream/handle/10986/18376/879840WP0FINAL00Box385208B00PUBLIC0.pdf

Table 1

Countries, CCT Programs and Studies in Analysis Sample

(A) Region/ Country	(B) CCT Program Name	(C) Number of Eligible References Retrieved
Africa		
Burkina Faso	Orphans and Vulnerable Children [a]	1
Ghana	Livelihood Empowerment Against Poverty	1
Kenya	Cash Transfer for Orphans and Vulnerable Children	1
Malawi	CCT for Schooling in Malawi [b]	1
Tanzania	Community-Based Conditional Cash Transfer Pilot [e]	1
Zimbabwe	Manicaland HIV/STD Prevention Project	1
East Asia and Pacific		
Cambodia	Cambodia Education Sector Support Project Scholarship Program	2
Cambodia	Japan Fund for Poverty Reduction Scholarship Program	1
Cambodia	Pilot Primary School Scholarship Program - Poverty Targeting	1
Cambodia	Pilot Primary School Scholarship Program - Merit Targeting	1
China	China Pilot CCT Program	1
Indonesia	Jaring Pengamanan Social	2
Indonesia	Program Keluarga Harapan (Hopeful Family Programme)	1
Phillipines	Pantawid Pamilyang Pilipino Program	1
Eastern Europe and Central Asia		
Macedonia, FYR	Conditional Cash Transfer Project	1
Turkey	Social Risk Mitigation Project	1
Latin America and the Caribbean		
Argentina	Asignación Universal por Hijo para Protección Social	1
Argentina	Programa Nacional de Becas Estudiantiles (PNBE)	1
Bolivia	Bono Juancito Pinto	1
Brazil	Bolsa Escola	1
Brazil	Bolsa Familia/Bolsa Escola	1
Brazil	Programa de Erradicacão do Trabalho Infantil	1
Colombia	Familias en Acción	1
Colombia	Familias en Acción (Urban Expansion)	1
Colombia	Subsidios Condicionados a la Asistencia Escolar – Basic	2
Colombia	Subsidios Condicionados a la Asistencia Escolar – Savings	2
Colombia	Subsidios Condicionados a la Asistencia Escolar – Tertiary	2

(A) Region/ Country	(B) CCT Program Name	(C) Number of Eligible References Retrieved
Dominican Republic	Programa Solidaridad	1
Ecuador	Bono de Desarrollo Humano	3
El Salvador	Comunidades Solidarias Rurales [f]	1
Guatemala	Mi Familia Progresa	1
Honduras	Bono 10000: Educación, Salud y Nutrición	1
Honduras	Programa de Asignación Familiar (PRAFII)	2
Jamaica	Programme of Advancement through Health and Education (PATH)	1
Mexico	Oportunidades	4
Mexico	Progresa (Educación, Salud y Nutrición)	3
Nicaragua	Red de Protección Social	3
Nicaragua	Sistema de Atención a Crisis	1
Panama	Red de Oportunidades	1
Paraguay	Tekoporâ	1
Peru	Juntos	2
Middle East and North Africa		
Morocco	Tayssir	1
South Asia		
Bangladesh	Female Secondary School Stipend Programme [h]	1
Bangladesh	Primary Education Stipend Project	1
Bangladesh	Shombhob Project	1
India	Apni Beti Apna Dhan (Our Daughter, Our Wealth)	1
Pakistan	Punjab Female School Stipend Program	2
Total		59

Note. Total references reported in Column (E) does not correspond to the sum of all eligible references retrieved because references that report more than one study (for example, Barrera & Filmer, 2016 for Cambodia's Pilot Primary School Scholarship Program - Poverty Targeting and Pilot Primary School Scholarship Program - Merit Targeting) are only counted once.
[a] Also known as Nahouri Cash Transfer Program. [b] Also known as Zomba or SIHR. [c] Also known as Bourses Maman. [d] Also known as Conditional Cash Transfer to Support Vulnerable Children and Households in the Context of HIV/AIDS and Poverty. [e] Also known as Tanzania Social Action Fund (TASAF). [f] Previously Red Solidaria. [g] Previously Tekoporâ. [h] Also known as Female Secondary School Assistance Project (FSSAP II).

Table 2

Characteristics of CCT Programs in Analysis Sample

	Max *n*	%	mean	*SD*
Region	47			
Sub-Saharan Africa	6	12.8		
East Asia and Pacific	8	17.0		
Eastern Europe and Central Asia	2	4.3		
Latin America and the Caribbean	25	53.2		
Middle East and North Africa	1	2.1		
South Asia	5	10.6		
Scale	47			
National	30	63.8		
Pilot	15	31.9		
State	2	4.3		
CCT type	47			
Primary only CCT	9	19.2		
Secondary only CCT	11	23.4		
Primary and secondary CCT	27	57.5		
Education conditionality requirements	47			
School attendance	13	27.7		
School enrollment and attendance	19	40.4		
Grade promotion or achievement	15	31.9		
Minimum school attendance requirement	47			
No	5	10.6		
Yes	42	89.4		
% school attendance required (conditional on requirement)	41			
75%	2	4.9		
80%	20	48.8		
85%	15	36.6		
90%	2	4.9		
95%	2	4.9		
Payment frequency	47			
Monthly	9	19.2		
Bimonthly	21	44.7		
Quarterly/trimesterly	10	21.3		
Other	7	14.9		
Limits on total amount of family transfer	47			
No	22	46.8		
Yes, maximum amount per family	12	25.5		
Yes, maximum number of beneficiaries	13	27.7		
Monthly average subsidy amount per child – 2015 U.S. dollar [a]				
Primary	36		$18.90	$29.60

	Max n	%	mean	*SD*
Secondary	38		$23.20	$29
School subsidy amount varies by	47			
Grade or age	8	17.0		
Family composition/family size	4	8.5		
Gender and grade	3	6.4		
None (flat transfer)	32	68.1		
Who receives the transfer	46			
Guardian, parent or head of household	21	45.7		
Mother or female head of household	17	37.0		
Student (or student and guardian)	8	17.4		
Savings component [b]	46			
Yes	5	10.9		
No	41	89.1		
Supply incentives for education [c]	47			
Yes	4	8.5		
No	43	91.5		
School enrollment at baseline (mean, SD)				
Primary	46		87.1	14.8
Secondary	44		50.9	19.6
Total number of programs	46			

[a] Average subsidy amount is the program's average monthly subsidy amount per child, averaged over the different subsidy amounts, when appropriate (e.g., girls and boys, or grades), separately by primary and secondary schooling. Subsidy amounts in U.S. dollars of 2015. For comparability, we converted subsidy amounts from original currency to nominal dollars in the base year (the year the program begins) using base-year exchange rates and then inflated to dollars in the year of analysis (2015) using the U.S. consumer price index. [b] Savings component refers to programs that in addition to a regular subsidy during the year, provide to beneficiaries a lump-sum at the end of the school year or at the end of secondary schooling. [c] Supply incentives refers to CCT programs that are accompanied by a supply-side intervention such as school grands for infrastructure textbooks of other school inputs. See text for additional details on data sources.

Table 3

Multivariate Meta-regressions of CCT Effect Size Correlates

	All time-period estimates				One time period estimate (lengthiest exposure)			
	PE	PA	SE	SA	PE	PA	SE	SA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Program characteristics								
1. Average annual subsidy amount/child in 2015 dollars according to grade level	0.06	-0.02	0.06	0.08	0.09	-0.02	-0.07	0.08
	(0.05)	(0.03)	(0.12)	(0.08)	(0.08)	(0.03)	(0.21)	(0.06)
2. National scale (1=yes, 0=pilot or state)	-2.12	0.01	-0.96	2.61	-3.09	-0.52	-1.97	3.69
	(2.75)	(2.15)	(6.08)	(5.58)	(2.84)	(2.13)	(8.81)	(5.66)
3. Additional conditions on grade promotion or test scores (1=yes)	-1.46	8.57*	1.87	0.27	-3.15	8.04*	-1.65	1.34
	(1.33)	(4.73)	(4.06)	(5.79)	(5.18)	(4.09)	(4.80)	(4.93)
4. Baseline enrollment (primary or secondary accordingly)	-3.90	-27.92***	-4.38	-40.94*	-6.36	-24.30**	2.76	-47.21**
	(8.10)	(9.55)	(9.82)	(22.05)	(12.08)	(10.10)	(15.54)	(19.17)
5. Supply-side supplement (1=yes)	7.08***	6.62**	-18.68*	-15.61*	3.11	5.53*	-20.60*	-14.77
	(2.68)	(2.90)	(10.41)	(9.07)	(5.02)	(2.70)	(10.76)	(9.62)
6. Payment to mother (1=yes, 0=any parent/guardian, or student)	-1.00	-1.85	-6.22	-7.45	-0.29	-2.13	-9.29	-7.26
	(2.11)	(2.33)	(6.16)	(6.16)	(1.98)	(2.62)	(8.96)	(5.03)
7. Payment frequency (1= monthly or bimonthly, 0=less frequent)	2.00	-1.61	-4.03	-14.07	0.51	-2.45	-4.44	-14.03*
	(1.91)	(3.07)	(4.39)	(9.21)	(2.00)	(2.94)	(6.09)	(6.42)
8. Latin America (1=yes)	1.13	4.87	0.49	10.11	0.50	5.12*	3.96	11.39
	(1.69)	(3.10)	(3.66)	(11.03)	(2.24)	(2.38)	(6.51)	(8.28)
9. Years of program exposure	0.14	0.36	2.01*	N/A	0.13	0.49	1.32	0.36
	(0.60)	(0.46)	(1.10)	N/A	(0.63)	(0.44)	(2.06)	(0.98)
Study characteristics								
10. Meets evidence standards in full or with reservations (1=yes)	-1.17	-4.40	4.59	-2.69	-1.32	-4.23	4.72	-3.28
	(1.30)	(2.87)	(3.43)	(7.57)	(1.99)	(3.33)	(6.13)	(6.25)
11. Publication year	-0.19	0.91*	-1.41***	-0.48	-0.43	0.93*	-2.06*	-0.26
	(0.26)	(0.51)	(0.48)	(1.10)	(0.73)	(0.43)	(0.94)	(1.09)
Observations	31	27	34	23	25	24	21	21

Note. PE = Primary Enrollment; PA = Primary Attendance; SE = Secondary Enrollment; SA = Secondary Atendance. Standard errors in parentheses. Estimates in columns (1)-(4) are from robust random effects model proposed by Hedges, Tipton, and Johnson (2010) that accounts for clustering of estimates within program over time and use all time-period effect sizes. Estimates in columns (5)-(8) are from random effects meta-regression models with only one time period effect (the lengthiest effect size). Estimate for years of program exposure in column (4) is not available because matrix has missing values.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4

Multivariate Meta-regressions of CCT Transfer-effectiveness Correlates

	All time-period estimates				One time period estimate (lengthiest exposure)			
	PE	PA	SE	SA	PE	PA	SE	SA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Program characteristics								
1. Payment frequency (1= monthly or bimonthly, 0=less frequent)	0.004	-0.047	-0.100**	-0.132	-0.007	-0.053	-0.174	-0.261***
	(0.019)	(0.036)	(0.045)	(0.117)	(0.019)	(0.049)	(0.099)	(0.079)
2. Additional conditions on grade promotion or test scores (1=yes)	-0.007	0.066	0.019	-0.002	-0.010	0.073	0.028	0.076
	(0.011)	(0.045)	(0.052)	(0.044)	(0.031)	(0.061)	(0.083)	(0.065)
3. Payment to mother (1=yes, 0=any parent/guardian, or student)	0.002	-0.026	0.021	-0.058	0.005	-0.030	0.004	-0.137*
	(0.018)	(0.025)	(0.073)	(0.070)	(0.019)	(0.036)	(0.148)	(0.063)
4. National scale (1=yes, 0=pilot or state)	-0.031	-0.007	-0.095	N/A	-0.042	-0.006	-0.157	0.097
	(0.026)	(0.021)	(0.061)		(0.029)	(0.025)	(0.144)	(0.071)
5. Baseline enrollment (primary or secondary accordingly)	-0.042	-0.295**	-0.118	-0.301	-0.092	-0.318*	-0.083	-0.740**
	(0.161)	(0.142)	(0.115)	(0.244)	(0.184)	(0.179)	(0.271)	(0.253)
6. Supply-side supplement (1=yes)	0.056**	0.034*	-0.176	-0.096	0.016	0.032	-0.176	-0.196
	(0.027)	(0.020)	(0.147)	(0.109)	(0.040)	(0.036)	(0.184)	(0.125)
7. Latin America (1=yes)	0.007	0.026	0.004	0.084	0.012	0.025	0.009	0.195*
	(0.018)	(0.026)	(0.041)	(0.111)	(0.020)	(0.032)	(0.093)	(0.101)
8. Years of program exposure	0.000	-0.002	0.023	0.007	0.002	-0.002	-0.001	0.015
	(0.004)	(0.004)	(0.022)	(0.010)	(0.005)	(0.005)	(0.034)	(0.012)
Study characteristics								
9. Meets evidence standards in full or with reservations (1=yes)	-0.007	-0.050	-0.009	0.004	-0.010	-0.057	0.035	-0.039
	(0.009)	(0.033)	(0.037)	(0.073)	(0.013)	(0.046)	(0.108)	(0.072)
10. Publication year	-0.001	0.005	-0.007	-0.003	-0.002	0.005	-0.017	0.002
	(0.002)	(0.003)	(0.008)	(0.007)	(0.004)	(0.005)	(0.016)	(0.013)
Observations	31	27	34	23	25	24	21	21

Note. PE = Primary Enrollment; PA = Primary Attendance; SE = Secondary Enrollment; SA = Secondary Atendance. Standard errors in parentheses. Estimates in columns (1), (3,) (5), and (7) are from robust random effects model proposed by Hedges, Tipton, and Johnson (2010) that accounts for clustering of estimates within program over time and use all time-period effect sizes. Estimates in columns (2), (4), (6), and (8) are from random effects meta-regression models with only one time period effect (the lengthiest effect size).

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5

Cost-effectiveness estimates of selected CCT programs

Country	Program	Cost Data Year	Annual Cost-Transfer Ratio (CTR)	Cost Effectiveness Estimates (CE)						
				Primary			Secondary			
				Enr.	Att.	Do.	Enr.	Att.	Do.	Grad.
Mexico	PROGRESA	1997/00	0.106	0.044	0.063		0.211	0.143		
Honduras	PRAF II	1999/02	0.499	0.153	0.088	0.243				
Nicaragua	Red de Protección Social	2000/02	0.629	0.175	0.139	-0.055				
Brazil	Bolsa Familia	2003	0.14	0.049		-0.005	0.042		-0.004	
Colombia	Familias en Accion	2000/04	0.117	0.137	0.293		0.115	0.387		0.104
Ecuador	Bono de Desarrollo Humano	2005	0.043	1.326		0.184				
Jamaica	PATH	2004/05	0.149		0.177			0.177		
Peru	Juntos	2006	0.131	0.063	0.000	0.000				
Bangladesh	Primary Education Stipend Program	2002	0.042	13.002						

Note. Enr. = Enrollment; Att. = Attendance; Do. = Dropout; Grad.=Graduation. For PROGRESA, PRAF II, and Red de Protección Social programs, cost-transfer ratio data is from Caldés, Coady, and Maluccio (2006: 823, Table 2 column total). For all other programs, administrative cost data is from Grosh et al. (2008: 411-12, compiled from various sources). Cost-transfer ratios for these programs are obtained from administrative cost data using equation (4) in text. Cost-effectiveness estimates (CE) are annual transfer-effectiveness (TE) divided by the annual cost-transfer ratio (CTR), see equation (3) in main text. Annual transfer-effectiveness is educational impact in percentage points per annual transfer amount. For PROGRESA, educational impacts are measured after one year. For PRAF II, primary enrollment impact is measured after one year and primary attendance impact is measured after two years. For Red de Protección Social, impacts are measured after one year. For Bolsa Familia, impacts are measured after four years. For Familias en Acción, enrollment impacts are measured after two years; attendance impacts after one year. For Bono de Desarrollo Humano, impacts are measured after two years. For PATH, impacts are measured after one year. For Juntos, enrollment impacts are measured after three years; attendance and dropout impacts are measured after one year. For Primary Education Stipend Program, impacts are measures after six years.

Figure 1. Forest plot of impact estimates for primary and secondary enrollment.

Note. For each program, we plot the effect size estimate for the lengthiest time of program exposure reported. The overall mean effect size is from a standard intercept-only random effects model. The *chi-square* test statistic for the null hypothesis of homogeneity in impact estimates in the random effects model is 229 (*p-value* 0.000) for primary enrollment, and 538 (*p-value* 0.000) for secondary enrollment. Mean baseline primary enrollment is 87.1 percent and mean baseline secondary enrollment is 50.9 percent (from the World Development Indicators data source for the year the program began or closest available).

School Attendance-All Studies

WWC Quality	Country	Years	Program	ES (95% CI)
Primary school				
1	Burkina Faso	2	Orphans and Vulnerable Children	13.40 (3.80, 23.00)
1	Cambodia	2	CSP (Merit Targeting)	6.85 (-25.88, 39.58)
1	Cambodia	2	CSP (Poverty Targeting)	32.45 (-8.55, 73.44)
1	Honduras	2	Programa de Asignación Familiar II	4.60 (2.25, 6.95)
1	Indonesia	2	Keluarga Harapan	-0.50 (-2.46, 1.46)
1	Morocco	2	Tayssir	1.40 (-1.48, 4.28)
1	Nicaragua	2	Red de Protección Social	12.94 (9.68, 16.20)
1	Nicaragua	1	Sistema de Atención en Crisis	5.00 (1.28, 8.72)
1	Tanzania	3	Tanzania Community Based CCT	-1.00 (-8.84, 6.84)
1	Zimbabwe	2	Manicaland HIV/STD Prevention Project	7.60 (1.13, 14.07)
2	Argentina	2	Asignación Universal por Hijo para Protección Social	2.20 (0.83, 3.57)
2	Dominican Rep	5	Programa Solidaridad	1.10 (-4.98, 7.18)
2	Ghana	2	Livelihood Empowerment Against Poverty	10.00 (4.90, 15.10)
2	Indonesia	1	Jaring Pengamanan Social	1.20 (0.02, 2.38)
2	Jamaica	1	PATH	2.50 (0.34, 4.66)
2	Kenya	2	Cash Transfer for OVC	-1.82 (-4.05, 0.41)
2	Mexico	1	Oportunidades	1.70 (0.72, 2.68)
2	Mexico	2	Progresa	1.11 (-0.01, 2.23)
2	Peru	0	Juntos	0.00 (-1.39, 1.39)
2	Phillipines	3	Pantawid Pamilyang Pilipino Program	3.80 (0.47, 7.13)
3	Bangladesh	1	Shombhob Project	3.65 (-1.81, 9.11)
3	Brazil	10	Programa Erradicacao do Trabalho Infantil	3.05 (2.77, 3.33)
3	Colombia	1	Familias en Acción	3.87 (-1.74, 9.48)
3	Paraguay	1	Tekopora	3.30 (-2.58, 9.18)
Subtotal				2.73 (1.74, 3.72)
Secondary school				
1	Colombia	1	Subsidios Condicionados Bogota (Basic)	3.20 (1.83, 4.57)
1	Colombia	1	Subsidios Condicionados Bogota (Savings)	2.70 (1.33, 4.07)
1	Colombia	1	Subsidios Condicionados Bogota (Tertiary)	5.60 (1.68, 9.52)
1	Indonesia	2	Keluarga Harapan	0.60 (-1.36, 2.56)
1	Malawi	2	CCT for Schooling in Malawi	8.00 (1.14, 14.86)
1	Nicaragua	1	Sistema de Atención en Crisis	5.00 (1.28, 8.72)
1	Tanzania	3	Tanzania Community Based CCT	-1.00 (-8.84, 6.84)
1	Zimbabwe	2	Manicaland HIV/STD Prevention Project	10.40 (9.75, 11.05)
2	Argentina	2	Asignación Universal por Hijo para Protección Social	11.70 (3.86, 19.54)
2	Cambodia	1	CESSP Scholarship Program	27.80 (16.24, 39.36)
2	Cambodia	1	JFPR Scholarship Program	31.30 (23.46, 39.14)
2	Dominican Rep	5	Programa Solidaridad	14.50 (2.94, 26.06)
2	Ghana	2	Livelihood Empowerment Against Poverty	5.00 (-1.27, 11.27)
2	Indonesia	1	Jaring Pengamanan Social	1.80 (0.82, 2.78)
2	Jamaica	1	PATH	2.50 (0.34, 4.66)
2	Kenya	2	Cash Transfer for OVC	5.07 (-4.93, 15.07)
2	Mexico	1	Oportunidades	1.70 (0.72, 2.68)
2	Mexico	2	Progresa	7.65 (4.74, 10.56)
2	Phillipines	3	Pantawid Pamilyang Pilipino Program	6.18 (3.35, 9.02)
3	Brazil	10	Programa Erradicacao do Trabalho Infantil	3.05 (2.77, 3.33)
3	Colombia	1	Familias en Acción	17.15 (5.67, 28.64)
3	Macedonia	2	Macedonia CCT Project	0.10 (-0.88, 1.08)
Subtotal				5.75 (3.94, 7.55)
Overall				4.01 (3.07, 4.95)

NOTE: Weights are from random effects analysis
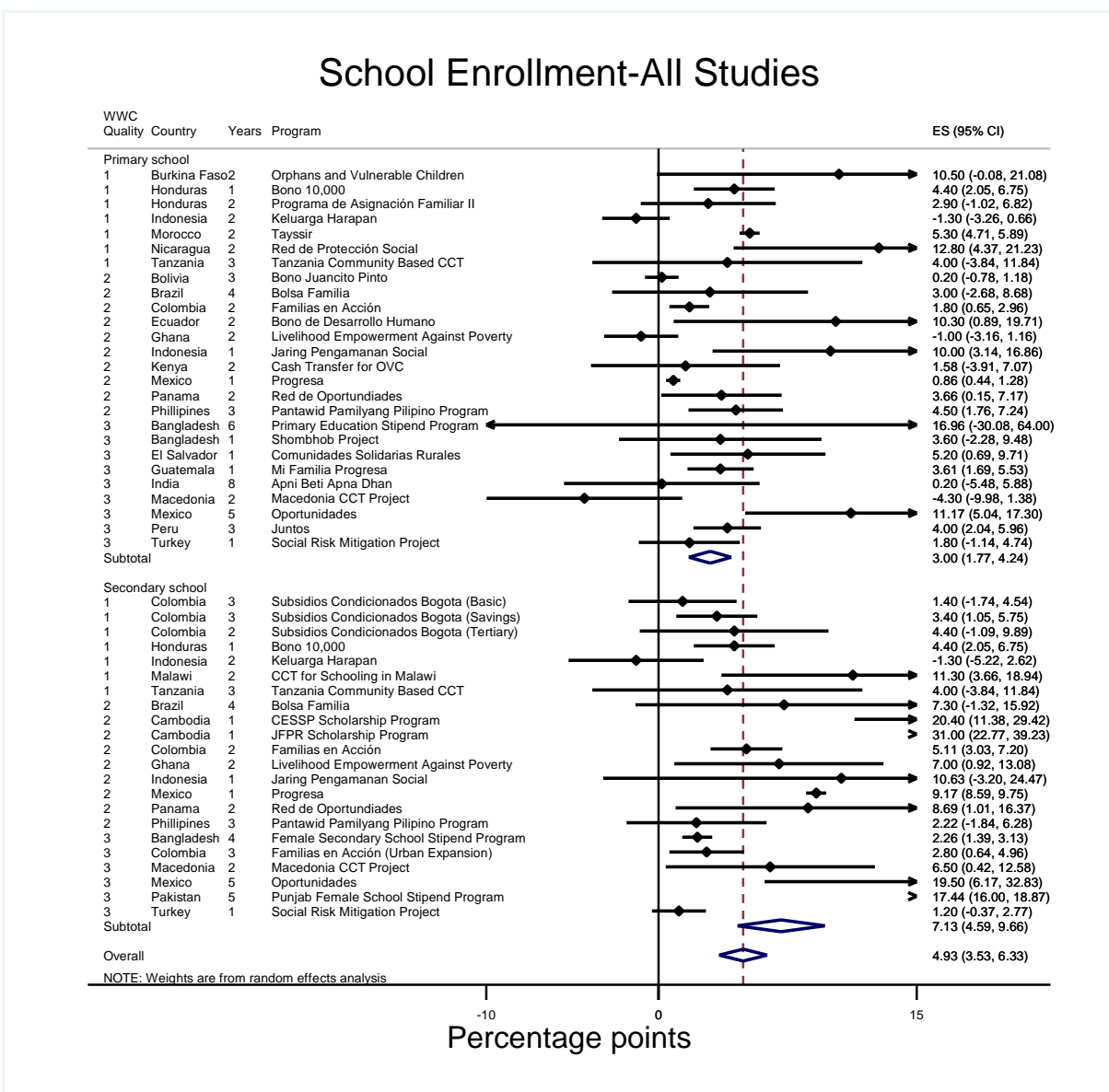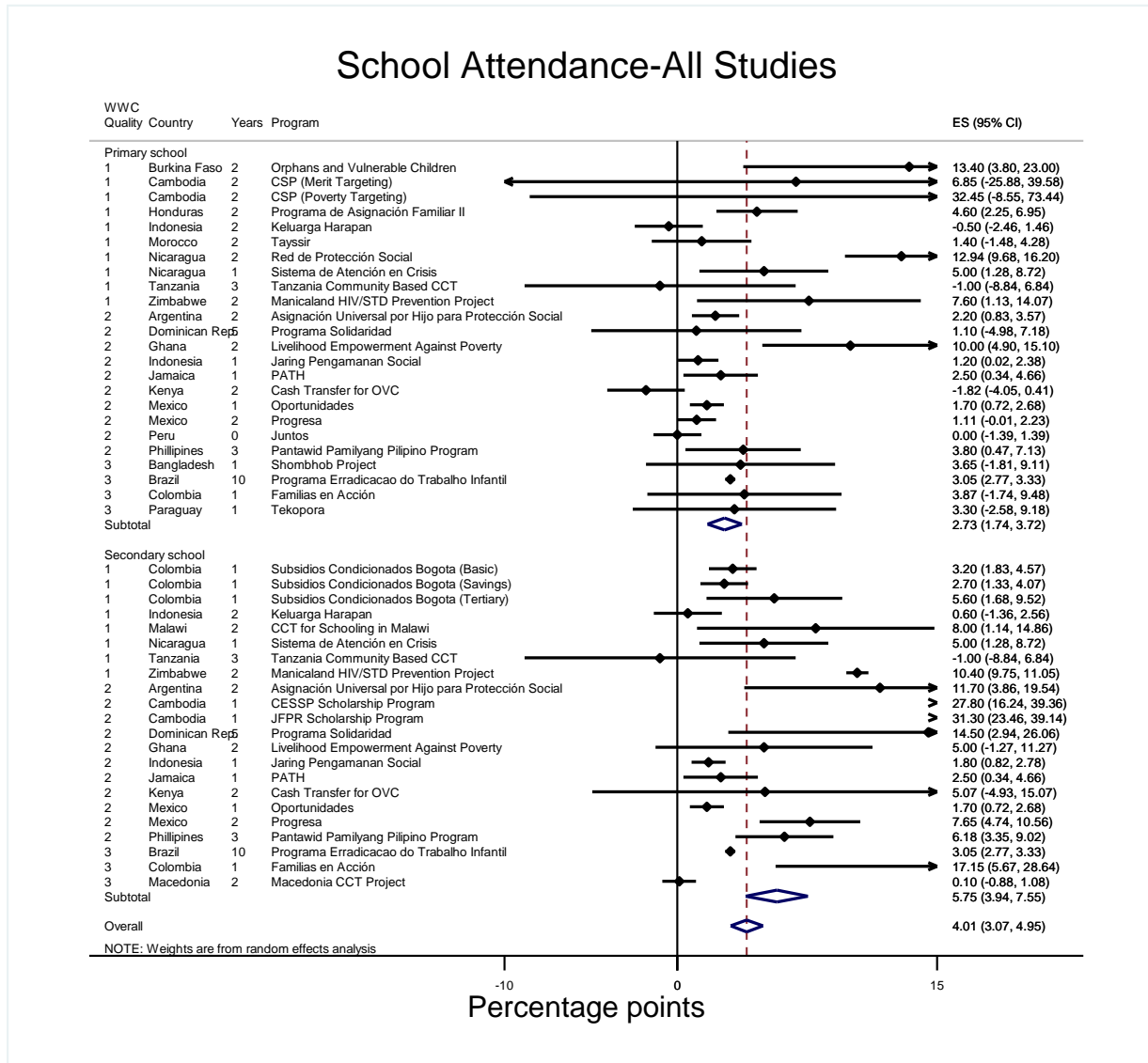
Percentage points (axis: -10, 0, 15)

Figure 2. Forest plot of impact estimates for primary and secondary attendance.

Note. For each program, we plot the effect size estimate for the lengthiest time of program exposure reported. The overall mean effect size is from a standard intercept-only random effects model. The *chi-square* test statistic for the null hypothesis of homogeneity in impact estimates in the random effects model is 127 (*p-value* 0.000) for primary attendance and is 620 (*p-value* 0.000) for secondary attendance.
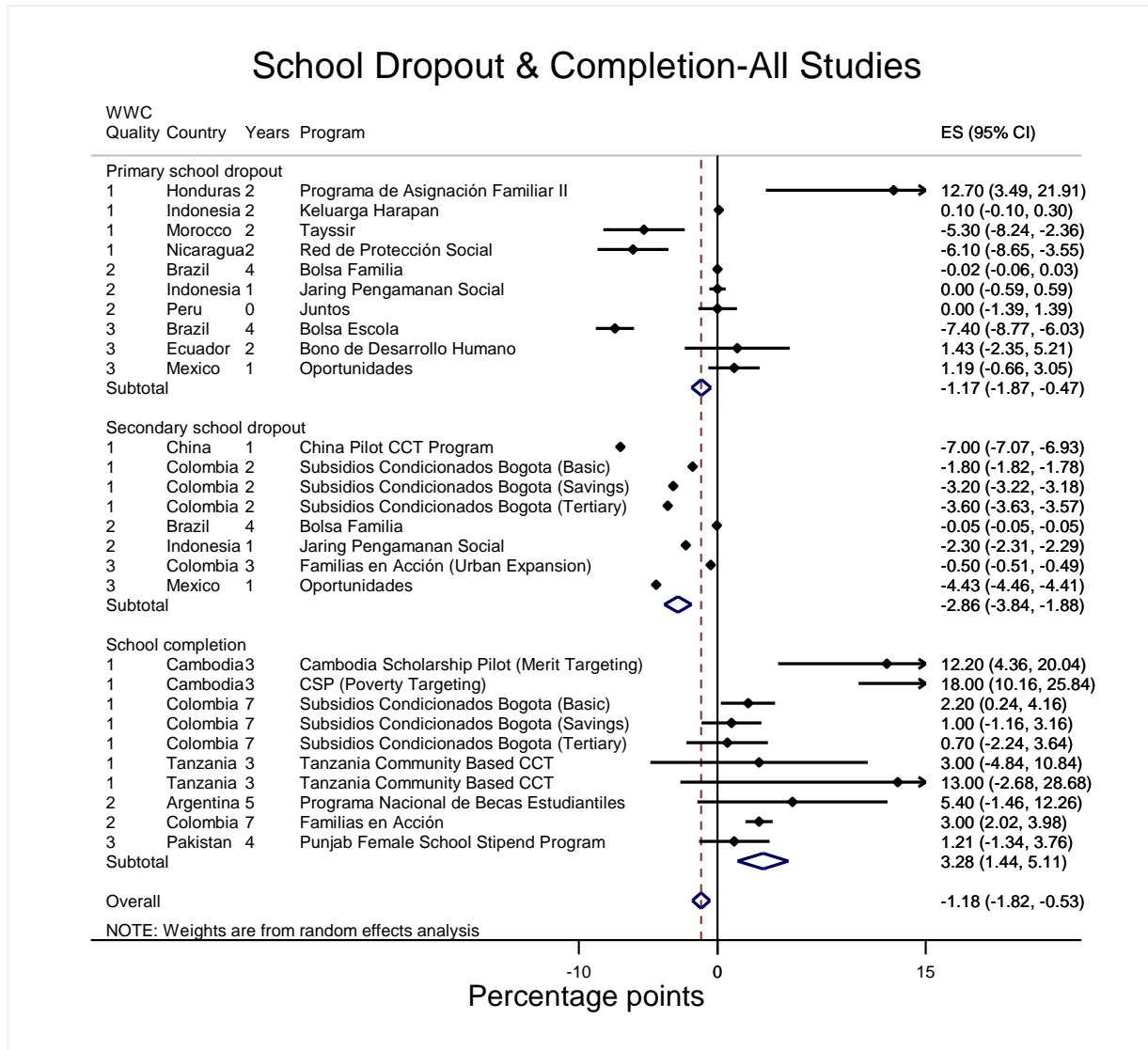
Figure 3. Forest plot of impact estimates school dropout and completion.

Note. For each program, we plot the effect size estimate for the lengthiest time of program exposure reported. The overall mean effect size is from a standard intercept-only random effects model. The *chi-square* test statistic for the null hypothesis of homogeneity in impact estimates in the random effects model is 156 (*p-value* 0.000) for primary dropout and 47,000 (*p-value* 0.000) for secondary dropout, and 30.8 (*p-value* 0.001) for school completion.