Analytical framework for evaluating the productive impact of cash transfer programmes on household behaviour

Methodological guidelines for the From Protection to Production Project

Solomon Asfaw, Katia Covarrubias, Benjamin Davis[1], Josh Dewbre, Habbiba Djebbari, Alessandro Romeo and Paul Winters

Version: 10 August 2012

1. Introduction

Cash transfer programmes have become an important tool of social protection and poverty reduction strategies in low- and middle-income countries. In the past decade, a growing number of African governments have launched cash transfer programmes as part of their strategies of social protection. Most of these programmes have been accompanied by rigorous impact evaluations (Table 1). Concern about vulnerable populations in the context of HIV/AIDS has driven the objectives and targeting of many of these programmes, leading to an emphasis on the ultra-poor, labour-constrained, and/or caring for orphan and vulnerable children (OVC) (Davis, et al, 2012). As a result, the objectives of most of these programmes focus on food security, health, nutritional and educational status, particularly of children, and so as would be expected, the accompanying impact evaluations concentrate on measuring these dimensions of programme impact.[2]

[1] Corresponding author: benjamin.davis@fao.org

[2] One important exception would be the PSNP in Ethiopia, and the accompanying impact evaluation (see Behane, et al, 2011).

Table 1. Recent and ongoing cash transfer impact evaluations in Sub Saharan Africa

- Malawi SCT
  - Mchinji pilot, 2008-2009
  - Expansion, 2012-2013
- Kenya
  - CT OVC, Pilot 2007-2011
  - CT OVC, Expansion, 2012-2014
  - HSNP, Pilot 2010-2012
- Mozambique PSA
  - Expansion, 2008-2009
- Zambia
  - Monze pilot, 2007-2010
  - Child Grant, 2010-2013
- South Africa CSG
  - Retrospective, 2010
- Burkina Faso
  - Experiment, 2008-2010
- Sierra Leone
  - Pilot, 2011-2012
- Ethiopia
  - PNSP, 2006-2010
  - Tigray SPP, 2012-2014
- Ghana LEAP
  - Pilot, 2010-2012
- Lesotho, CSP
  - Pilot, 2011-2013
- Uganda, SAGE
  - Pilot, 2012-2014
- Zimbabwe, SCT
  - Pilot, 2013-2015
- Tanzania, TASAF
  - Pilot, 2009-2012
  - Expansion, 2012-2014
- Niger
  - Begins in 2012

Most of these accompanying impact evaluations in Sub Saharan Africa pay little attention to livelihoods per se, or to the current economic and productive activities of beneficiary households, in terms of either data collection or analysis. Cash transfer programme induced investments in health and education have both short and long term economic benefits through improvements in human capital, which lead to an increase in labour productivity and employability. Indeed, such effects constitute the underlying rationale for many of the pioneer CCT programmes in Latin America and the Caribbean, including the PROGRESA programme, which had a long term vision of exiting poverty through labor markets. Consequently, most impact evaluations of CCTs in Latin America paid attention to wage labour participation, but relatively little to self employment activities.

And indeed, there is good reason to believe that cash transfer programs will influence the productive dimension of beneficiary households. In the case of many beneficiaries in Sub Saharan Africa, livelihoods are still based in agriculture, and particularly in subsistence agriculture, and will continue to be for the foreseeable future. The exit path from poverty is not necessarily the formal (or informal) labor market, but self employment generated by beneficiary households themselves, whether in agriculture or outside agriculture. Most beneficiaries live in places where markets for financial services (such as credit and insurance), labour, goods and inputs are lacking or do not function well. Cash transfers typically represent about 20 percent of per capita expenditure, and when provided in a regular and predicable fashion, may help households in overcoming the obstacles that limit their access to credit or cash. This, in turn, can increase productive and other income-generating investments, influence beneficiaries' role in social networks, increase access to markets and inject resources into local economies. These impacts come through changes in household behaviour (labour supply, investments, risk management) and through impacts on the local economy of the communities (social networks, labour and good markets, multiplier effects) where the transfers operate.

The study of the economic and productive impacts is also important for policy. The perception exists among many officials in ministries of finance and the economy that cash transfer programmes do not have economic impacts. These programmes are often seen as welfare, charity and/or handouts. In a number of countries, such as Rwanda and Ethiopia, cash transfer households are specifically separated from potentially productive households, who receive cash for work, bundled with complementary productive oriented components of the programme. Such perceptions are not surprising since the transfers are targeted towards the ultra poor, bottom 10%, labour constrained, elderly, infirm and households headed by children. Such mistaken perceptions may be buttressed by the fact that beneficiaries are primarily women—despite the fact that women are as economically active as men.

Moreover, cash transfers can be an important complement to a broader rural development agenda. The importance of a pro-poor growth strategy focusing on agriculture, and particularly the need for a new Green Revolution in Sub-Saharan Africa, has been widely discussed (World Bank, 2008; Binswanger-Mkhize, McCalla and Patell, 2010; Diao, Heady and Johnson, 2008; Toenniessen, Adesina and DeVrie, 2008). Such a strategy would imply a combination of increased access to a diverse package of modern agricultural technologies, including an initial fertilizer subsidy, and investment in rural infrastructure and agricultural research and extension (World Bank, 2008). Yet, a lack of access to agricultural assets, markets and institutions, and in particular credit, is constraining potential engagement in agriculture (Zezza et al., 2011). One mechanism to overcome such constraints, especially among poor farmers who are most likely to be credit constrained, is through the provision of cash transfers. Thus, cash transfers can serve not just as a means of social protection but as a means of promoting farm-level production gains.

The Food and Agriculture Organization of the United Nations (FAO) has signed a 3-year agreement with the research programme at DFID—the From Protection to Production Project (PtoP)—to study the impact of cash transfer programmes on household economic decision making and the local economy[3]. This research project seeks to understand the potential economic development impacts of cash transfers on the rural poor in Sub Saharan Africa. It aims at contributing to the understanding of how social protection interventions can contribute to sustainable poverty reduction and economic growth at household and community-levels. This will be documented by the production of country case studies of economic impacts for each country and by comparison papers across different types of countries. The project is using a mixed method approach, combining econometric analysis of impact evaluation data, local economy SAM/CGE models, and qualitative methods.

The project is implemented jointly by FAO and UNICEF, and the research will build on ongoing or planned impact evaluations in seven countries (Ethiopia, Ghana, Kenya, Lesotho, Malawi,

[3] http://www.fao.org/economic/ptop/en/. Additional funding has been provided by the European Commission and the World Bank. The PtoP project forms part of the larger Transfer Project (http://www.cpc.unc.edu/projects/transfer).

Zambia and Zimbabwe) which are being implemented and managed by the respective governments, UNICEF country offices and other development partners (Table 2).

Table 2. Country programmes participating in the PtoP project.

Country	CT Programme	Baseline survey	Follow-up(s) Survey	Status
Lesotho	Child Grant Program (CGP)	2011	2013	Pilot
Kenya	Cash Transfers for Orphan and Vulnerable Children (CT-OVC)	2007	2009, 2011	Scale-up
Ethiopia	Tigray Minimum Social Protection Package	2012	2013, 2014	Pilot
Malawi	Social Cash Transfer (SCT)	2012	2013	Scale-up
Ghana	Livelihood Empowerment Against Poverty (LEAP)	2010	2012	Scale-up
Zambia	Child Grant programme	2010	2012, 2013	Pilot
	Monze cash transfer	2007	2010	Pilot
Zimbabwe	Social Cash Transfer (SCT)	2013	2014, 2015	Pilot

The PtoP project has five main areas of work, all geared towards strengthening data collection and analysis in ongoing impact evaluations.

- The first area is to finance, design, pilot and supervise implementation of additional modules in household surveys, including information on asset accumulation, productive activities and labor allocation; risk coping strategies and time use; social networks and if possible, climate change adaptation.
- Taking advantage of experimental and non experimental design and panel data across countries, the second area of work, led by a team at FAO, is to promote and carry out analysis of the impact of cash transfer programmes on household and individual decision making regarding productive activities, including adult and child labour.
- The third area is to model the impacts of cash transfer programmes on the local economy. This involves constructing village SAM/CGE models for cash transfer programme areas in each country, by a team led by Prof Ed Taylor at UC Davis. Such modelling requires collection of a business enterprise survey in programme communities as well as minor modification of impact evaluation household questionnaires. The methodological guide for this area of work can be found in Taylor (2012).
- The fourth area involves the integration of qualitative/quantitative design and methods in each country, which will be lead by Oxford Policy Management (OPM) consultants. The concept note describing this area of work can be found in OPM (2012).
- The fifth area focuses on feeding back the analytical results into the policy process, and increasing the capacity of programme managers and policy-makers in terms of impact evaluation, design and implementation of cash transfer programmes.

The purpose of this paper is to describe the methodology that will be utilized for the household level analysis of economic and productive impacts under the PtoP project. We will first review the conceptual framework underlying our analysis, then delve into the analytical framework, with detailed sections on the methods we may use in the different contexts of each impact evaluation: difference in difference estimators, propensity score matching and regression discontinuity design. This is followed by a discussion of the specific evaluation design of each of the seven countries participating in the project.

2. Conceptual framework

The concept of cash transfer programs leading to economic and productive impacts is built around the hypothesis that the provision of regular and predictable cash transfers to very poor households in the context of missing or malfunctioning markets has the potential to generate economic and productive impacts at the household level and to stimulate the local economy through the networks that link individuals, households, businesses and institutions. In rural areas most beneficiaries depend on subsistence agriculture and/or rural labour markets and live in places where markets for financial services (such as credit and insurance), labor, goods and inputs are lacking or do not function well. Cash transfers often represent a significant share of household income, and can be expected to help households overcome the obstacles that block their access to credit or cash. This, in turn, can increase productive and other income-generating investments, influence beneficiaries' role in social networks, increase access to markets and inject resources into local economies.

The impact of cash transfer programmes on the economic decision process is thus potentially manifested through changes in household behaviour and on the communities and local economies where the transfers operate.[4] This may occur through the following five channels:

i. *Human capital*. By facilitating the accumulation and improvement of human capital, cash transfer programmes may enhance productivity and increases employability in the long term.
ii. *Income generating strategies*. By relaxing credit, savings and/or liquidity constraints, cash transfer programmes can facilitate changes in income generating strategies. This may include changes in labour allocation (to and/or from labour off farm and on farm); changes in productive activities (use of inputs); accumulation of productive assets (such as farm implements, land or livestock); changes in productive strategies (such as new crops, techniques or natural resource conservation); and introduction of new lines of products or services or new activities.
iii. *Risk Management*. Through the regular and predictable provision of financial resources, cash transfer programmes may improve the ability to manage risk and shocks. This includes the avoidance of detrimental risk coping strategies (distress sales of productive assets, children school drop-out, risky income-generation activities); the avoidance of risk averse production strategies (safety or eat first); increased risk taking into more profitable crops and/or activities.

[4] See both Hoddinott (2008) and Barrientos (2012) for alternative and more detailed, though broadly similar, frameworks.

iv. *Social Network*. Through the regular and predictable provision of financial resources to the poorest and most vulnerable, cash transfer programmes may reduce pressure on informal insurance mechanisms such as social networks of reciprocity, which have been particularly stretched in the context of HIV/AIDS and economic crisis, and allow beneficiaries to actively participate in these networks.

v. *Local Economy*. The injection of a significant amount of cash into the local economy can stimulate local product and labour markets and create multiplier effects.

This conceptual framework needs to fit within a behavioural model of the household and a given socio-economic context, to understand how a given cash transfer programme might impact beneficiary households in the short, medium and long term.[5] For very poor households of the type targeted by these programmes, who typically spend 60 to 70 percent of their household budget on food, the first, immediate impact of a cash transfer programme is almost always on food expenditure and composition—which given that most beneficiary households are subsistence agricultural producers, has implications for on farm productive activities. A second level of impact is less direct, but perhaps strongly associated with the program during the implementation, either as a message or a formal conditionality, such as spending on school uniforms. A third level of impacts is again less direct, with more mediation by other external factors, and it may be less surprising if we see little or no impact. These may include school attendance (mediated by supply, etc), nutrition (mediated by sanitation, information, etc), investment or changes in certain productive activities (mediated by access to relevant goods, services and markets), and so on.

A number of potential outcome variables emerge from this conceptual framework. While ultimately we are interested in seeing whether cash transfer programmes lead to increased returns from household income generation strategies, we need to focus on more direct and intermediate impacts for two reasons. First, we are interested in understanding the mechanism of impact—we do not just want to know if cash transfer increases production, but how they increase production—investment, more and/or different labour allocation, different use of inputs, shift in activities, etc. Second, given that income and agricultural production are mediated by factors outside the control of the programme and the producer—such as prices, weather and access to input and output markets—we may not see impact on the final outcome, but we may see impact among the intermediate outcomes. Impacts may vary by subgroups of the population, such as by gender, household size, and previous access to productive assets. We discuss heterogeneity of impact in more detail below. Finally, a finding of no impact does not necessarily mean that a given programme is ineffective. Few of the cash transfer programmes under study have as an objective the strengthening of beneficiary livelihoods, and thus we do not necessarily expect to find impacts. By comparing results across the seven countries we hope to highlight where and under what conditions social cash transfer programmes have productive and economic impacts.

[5] Here we are assessing transfers which are delivered in a timely and consistent fashion. Transfers which by design or by logistical complication (such as the LEAP programme in Ghana) arrive in a lumpy fashion may induce a different response from households.

Note that this framework is geared towards cash transfer programs, conditional or unconditional, not linked to any kind of labour requirements. Public works or cash for work programs, which are also increasingly popular in Sub Saharan Africa, in both emergency and development contexts, would require a modified framework with two additional dimensions. First, the labour requirement would alter the labor allocation decision at the household level, and possibly effect the functioning of local labor markets, thus altering the generation of multipliers. Second, community assets produced by public works (such as irrigation) may alter household-level returns to own production as well as again altering economic linkages within a given community, thus affecting the income multiplier.

While ample evidence exists from the conditional cash transfer (CCT) impact evaluation literature (and increasingly from the cash transfer literature in Sub Saharan Africa) in terms of the first channel of human capital improvement,[6] relatively few studies have looked at the labour supply or productive impacts of cash transfer programs.[7] Labour supply has been somewhat more studied; from impact evaluations employing a valid counterfactual, most find little evidence that cash transfer programmes reduce adult work.[8] Covarrubias, Davis and Winters (2012) find that the Malawi SCT programme led to decreased agricultural wage labour, as adults switched from *ganyu* labour of last resort, and children from all types of wage labour, to on farm agricultural production. Gilligan, Hoddinott and Taffesse (2009) find that households with access to both the Productive Safety Net Programme (PSNP) as well as complementary packages of agricultural support showed no indication of disincentive effects on labour supply.

In terms of production, despite the lack of available information, most of those studies that do exist point to potential productive impacts, as well as potential conflicts between social objectives and livelihood activities. Todd, Winters and Hertz (2010) and Gertler, Martinez and Rubio-Codina (2012), for example, find that the Mexican PROGRESA program led to increased land use, livestock ownership, crop production and agricultural expenditures and a greater likelihood of operating a microenterprise. Yet, Handa et al (2010) find that agricultural households benefiting from PROGRESA were less likely to comply with conditionality due to time conflicts with their livelihood activities.

From Sub Saharan Africa, Covarrubias, Davis and Winters (2012) and Boone, et al (2012) found that the Malawi SCT programme led to increased investment in agricultural assets, including crop implements and livestock and increased satisfaction of household consumption by own production. For Ethiopia, Gilligan, Hoddinott and Taffesse (2009) find that households with access to both the Productive Safety Net Programme (PSNP) as well as complementary packages of agricultural support were more likely to be food secure, to borrow for productive purposes, use improved agricultural technologies, and operate their own nonfarm business activities. In a later study, Berhane et al. (2011) found that the PSNP has led to a significant improvement in food security status for those that had participated in the programme for 5 years versus those who

[6] See reviews in Djebbari and Hassine (2011), Fisbein and Shady (2009) and Handa and Davis (2006).

[7] See reviews in Barrientos (2012) and Davis et al (2010).

[8] See, for example, Parker and Skoufias (2000), Galasso (2006), Skoufias and di Maro (2006), Edmonds and Schady (2008), Amarante and Vigorito (2010) and Alzua, Cruces and Ripani (2010).

only received one year of benefits. Moreover, those households that participated in PNSP as well as the complementary programmes had significantly higher grain production and fertilizer use.

Analytical framework

The objective of an impact evaluation is to attribute an observed impact to the programme intervention. The identification of the counterfactual is the organizing principle of an impact evaluation; that is, it tells us what would have happened to the beneficiaries if they had not received the intervention. As a household cannot be both a participant and non participant of the same programme, the counterfactual is identified by selecting a control group. A group of control households should be chosen from non-beneficiaries to be representative of the group of participants with one key difference: the control households did not receive the intervention. If the two groups are dissimilar in other dimensions, the outcomes of non-beneficiaries may differ systematically from what the outcomes of participants would have been without the programme, producing selection bias in the estimated impacts. This bias may derive from differences in observable characteristics between beneficiaries and non-beneficiaries (e.g., location, demographic composition, access to infrastructure, wealth, etc.) or unobservable characteristics (e.g., natural ability, willingness to work, etc.). Some observable and unobservable characteristics do not vary with time (such as natural ability) while others may vary (such as skills). Furthermore the existence of unobservables correlated with both the outcome of interest and the program intervention can result in additional bias (i.e., omitted variables).

Let D_i denote a dummy variable equal to one if a household receives a cash transfer and equal to zero if a household does not receive a cash transfer. Similarly, let Y_i denote an outcome of interest such that potential outcomes are defined as $Y_i(D_i)$ for every household. The treatment effect of the programme for household i, τ_i, is then the change in the outcome measure caused by the transfer:

$$\tau_i = Y_i(1) - Y_i(0) \tag{1}$$

Equation (1) formalizes the question posed above; i.e., what would have happened to treated households in absence of the programme? As mentioned, only one outcome is observable—either the household receives the transfer or it does not, leaving the counterfactual component in equation 1 unknown. The implications are twofold. First, the success of any impact evaluation relies on identifying a suitable counterfactual sample. And second, it is not possible to measure unit specific treatment effects, but rather average treatment effects incorporating information from the counterfactual.

The most direct way of ensuring a comparable control group is via an experimental design (randomized control trial), in which households are randomly allocated between control and treatment groups. This guarantees that the treatment status is uncorrelated with other (observable and unobservable) variables, and as a result the potential outcomes will be statistically independent of the treatment status. On average the groups will be identical, except for the fact that one of them received the treatment. Under these conditions, the average treatment effect (ATE) of the cash transfer can be identified simply as the mean difference in outcomes between the two groups:

$$E(\tau) = ATE = E[Y(1)] - E[(Y(0)] \tag{2}$$

A large number of average treatment effects can be estimated[9]. In addition to the ATE, perhaps the most commonly reported is the average treatment effect on the treated (ATT), which measures the average impact of the cash transfer programme on those that received treatment. This is defined as:

$$ATT = E[\tau \mid D=1] = E[(Y(1) \mid D=1) - E[Y(0) \mid D=1] \tag{3}$$

Again, the counterfactual mean for those being treated is not observed, rendering crucial the choice of a proper substitute in order to estimate the ATT. In an experimental setting the ATE equals the ATT. However, in a non-experimental setting they usually differ and in addition, using the mean outcome of untreated individuals, $E[Y(0) \mid D=0]$, runs the risk of comparing apples and oranges if factors that determine the participation decision also influence the outcome variable of interest (i.e., if there is selection bias).

The validity of experimental estimators relies on the assumption that the control group units are not affected by the programme; this is also referred to as the Stable Unit Treatment Value Assumption (SUTVA) (Rubin, 1980; Djebbari and Hassine, 2011). However there are two possibilities where control households can be affected: market interactions and informal transaction and risk sharing (also known as non-market interaction).

Experimental designs are often difficult to implement in practice, however, for political, ethical, institutional and/or logistical reasons, particularly when programmes are owned by national governments (as opposed to researchers). Non experimental design methods are often used when a randomized experiment is not possible, or when the experimental design fails to achieve observable balance among groups, due to chance or when for example the number of units of randomization is relatively small.

In non-experimental studies one has to invoke some identifying assumptions to solve the selection problem. The same is true also when differences between treatment and control groups at baseline emerge despite randomisation. More systematic differences at baseline between treatment and control groups require econometric techniques to create a better counterfactual by removing pre-existing significant differences in key variables.

In the rest of this section we present the methodologies that the PtoP project is using, or is planning to employ, in the household level analysis of the productive impacts of cash transfer programmes. We begin with difference in difference (DD) estimators, which can be employed

[9] See Ravallion (2008) for a concise overview of the different types of average treatment effects.

using data from an experimental design, and then we move on to techniques that help us deal with weakened experimental designs or non experimental settings: propensity score matching methods (PSM) and regression discontinuity design (RDD).

Difference-in-Difference Estimators (DD)

As discussed above simple mean comparisons (equation 2) identify treatment impacts in successful experimental designs. Nevertheless, impact estimates can be verified and in some cases improved upon by applying a DD methodology. The latter might occur if, for example, randomization (by chance) produces baseline differences between treatment and control groups. Similarly, when the data do not come from a randomized design the DD estimator may be used; often in conjunction with other approaches.

When panel data are available with pre and post intervention information, which will be the case in most of our impact evaluation studies, the estimator in equation 3 can be improved by subtracting off the difference in pre-programme outcomes between participants and non-participants. This can be seen in equation 4:

$$\begin{aligned} ATT = E[(\tau_t - \tau_{t-1}) \mid D=1] &= E[(Y(1)_t - Y(0)_t) - (Y(1)_{t-1} - Y(0)_{t-1}) \mid D=1], \\ &= E(Y(1)_t - Y(1)_{t-1}) \mid D=1) - E(Y(0)_t - Y(0)_{t-1} \mid D=1) \end{aligned} \tag{4}$$

where $t-1$ and t represent time periods before and after the introduction of the cash transfer programme.

By taking the difference in outcomes for the treatment group before and after receiving the cash transfer, and subtracting the difference in outcomes for the control group before and after the cash transfer is disbursed, the DD estimator controls for unobserved heterogeneity that may lead to selection bias (Woodridge, 2002). DD is able to control for pre-treatment differences between the two groups, and in particular the time invariant unobservable factors that cannot be accounted for otherwise. The key assumption is that differences between treated and control households remain constant through the duration of the project. If prior outcomes incorporate transitory shocks that differ for treatment and comparison households, the DD estimate will interpret these shocks as representing a stable difference, and thus contain a transitory component that does not represent the true programme impact.

When differences between treatment and control groups at the baseline exist, the DD estimator with conditioning variables has the advantage of minimizing the standard errors as long as the effects are unrelated to the treatment and are constant over time (Wooldridge, 2002). Control variables are most easily introduced by turning to a regression framework which is convenient for the DD and is our preferred approach. Equation 5 presents the regression equivalent of DD with covariates;

$$Y_{it} = \beta_0 + \beta_1 D_{it} + \beta_2 R_t + \beta_3 (R_t * D_{it}) + \sum \beta_i Z_i + \mu_{it} \tag{5}$$

where Y_{it} is the outcome indicator of interest; D_{it} is a dummy equal to 1 if household i received the treatment; R_t is a time dummy equal to 0 for the baseline and to 1 for the follow up round; $R_t * D_{it}$ is the interaction between the intervention and time dummies and, μ_{it} is an error term. To control for household and community characteristics that may influence the outcome of interest beyond the treatment effect alone, we add in Z_i, a vector of household and community characteristics to control for observable differences across households at the baseline which could have an effect on Y_{it}. These factors are not only those for which some differences may be observed across treatment and control at the baseline, but also ones which could have some explanatory role in the estimation of Y_{it}. As for coefficients, β_0 is a constant term; β_1 controls for the time invariant differences between the treatment and control; β_2 represents the effect of going from the baseline to the follow-up period, and β_3 is the double difference estimator, which captures the treatment effect.

When panel data are not available or when there is additional need to account for baseline differences between treatment and control groups, propensity score matching or propensity score weighting can be applied. The details for applying these techniques are developed in the next section.

Propensity Score Methods : propensity score matching (PSM) and inverse probability weighting (IPW)

Propensity score methods attempt to simulate the conditions of an experiment in which recipients and non-recipients are randomly assigned, allowing for the identification of a causal link between treatment and outcome variables.

Let $P(Z) = \Pr(D = 1 | Z)$ be the probability of participating in a cash transfer programme where Z is a vector of observed control variables measured (ideally) before programme implementation. Propensity score matching constructs a statistical comparison group by matching individual treatment households with control households based on similarities in $P(Z)$. A closely related alternative involves weighting control households using this score, such that the mean of each Z variable is approximately equal across participants and non participants (Khandker, Koolwal and Samad, 2010).

There are two fundamental assumptions of these models which pertain to the estimation of the propensity model, $\hat{P}(Z)$. The first is the conditional independence assumption (CIA), which implies that potential outcomes are independent of treatment conditional on Z,

$$E[(Y(1)_{t=0} \mid Z, D = 1)] = E[(Y(0)_{t=0} \mid Z, D = 0)] \tag{6}$$

Equation 6 indicates that conditional on observable characteristics, non-participants of the cash transfer programme have the same mean outcomes as participants, had they not received treatment. The second main assumption of propensity models is the common support condition, which requires that the propensity score lie strictly between zero and one.

$$0 < \hat{P}(Z) < 1 \tag{7}$$

Equation 7 requires that the proportion of treated and untreated households must be greater than zero for every possible value of Z. The overlap condition ensures that treatment observations have comparison observations 'nearby' in the propensity score distribution (Heckman et al., 1998; Rosenbaum and Rubin, 1983). This implies that the effectiveness of propensity methods also depends on having a large number of non beneficiaries so that a substantial region of common support can be found.

In addition to these two basic assumptions, analysis by Heckman et al. (1998) suggests that it is equally important that (i) the same data source is used for participants and non-participants and; (ii) participants and non-participants have access to the same markets. The seminal explanation of the PSM method is provided by Rosenbaum and Rubin (1983), and its strengths and weaknesses are elaborated, for example, by Dehejia and Wahba (2002), Heckman et al. (1998), Caliendo and Kopeinig (2008), and Smith and Todd (2005).

For propensity score matching, participants are matched to non-participants using $\dot{P}(Z)$. Several matching methods have been developed to match participants with non-participants of similar propensity scores. These include Nearest Neighbor Matching, Stratification and Interval Matching, Caliper and Radius Matching and Kernel Matching, among others. Asymptotically, all matching methods should yield the same results. However, in practice, there are trade-offs in terms of bias and efficiency with each method (Caliendo and Kopeinig, 2008). The basic approach is to numerically search for "neighbours" of non participants that have a propensity score that is very close to the propensity score of the participants.

Sometimes matches are conducted on a 'one to one' basis, in the sense that one treatment case is matched to one and only one control case. In practice however it is more common to match one treatment household with several control households. A closely related distinction is whether to perform matching 'with' or 'without' replacement. With replacement, a control household can be matched with several treatment households. Without replacement, a control observation is taken out of the sample once it is matched, and cannot be used for other comparisons.

The most straightforward matching method is Nearest Neighbor Matching, which pairs control households to treatment households that have the closest propensity score. Nearest Neighbor Matching runs the risk of poor matches if the closest neighbours are far away. In such cases a tolerance level on the maximum propensity score distance (caliper) can be imposed. Stratification and Interval Matching approaches involve partitioning the common support of the propensity score into strata and calculating the impact within each strata by taking the mean difference in outcomes between treated and control observations. This method is also known as interval matching, blocking and sub classification (Rosenbaum and Rubin, 1983).

These approaches have in common that only a few control observations are used to construct the counterfactual outcomes. Kernel Matching on the other hand is a non-parametric matching estimator that uses weighted averages of all control group households to construct the counterfactual outcome. Thus, one advantage of this approach is the lower variance that is attainable because more information is used (Caliendo and Kopeinig, 2008).[10]

For propensity score weighting, control observations are assigned weights equal to the inverse of their propensity score; i.e., $w = \frac{\dot{P}(Z)}{(1-\dot{P}(Z))}$, and treatment observations receive a weight equal to one. Applying these weights[11] to control households effectively reweights the distribution of observable characteristics included in $P(Z)$ to be like that of the treatment group. Said differently, control observations that are dissimilar to the treatment group will have a $\dot{P}(Z)$ near zero and a weight near zero; and conversely, control observations who are similar to the treatment group will receive a higher weight.

A regression of an outcome on treatment and Z variables thus amounts to a comparison of means and produces an estimate of the ATT.[12] One advantage of the weighting approach is that it is considered to be "doubly robust": if *either* the propensity model or the outcome equation is correctly specified the estimator will be consistent. As with matching, ensuring that a region of common support exists is necessary to avoid observations with extremely large weights, which can yield estimates with high variance and undue influence on results (Imbens and Wooldridge, 2009).

There are two common methods to determine the region of common support more precisely. The first one is based on comparing the minima and maxima of the propensity score in both groups. This approach entails deleting all observations whose propensity score is smaller than the minimum and larger than the maximum of the treatment group. The second method to overcome this problem is based on estimating the density distribution in both groups and using a trimming procedure to determine the common support region (Smith and Todd, 2005). The downside of this trimming method is that it can exclude treated and control observations in that propensity score range, and thus result in a different sample than the original one. If equations 6 and 7 are valid, then PSM provides a plausible method for estimating $E[(Y(0)\,|\,D=1)]$ and obtaining unbiased estimates of the ATT.

When restricting analysis on some range of propensity scores, it is important to assess how the treatment and control groups differ, and which sub-population is being studied. Trimming risks changing the sample in a way that makes estimated treatment impacts difficult to generalize. Moreover, there is likely to be heterogeneity in the treatment impact across groups. When this is

[10] Note however that lower variance is likely to be associated with a larger bias in the estimates – especially if there are bad matches or the choice of the bandwidth parameter overly smoothes the estimates.

[11] Propensity weights may also be multiplied by survey sampling weights, if appropriate.

[12] An estimate of the ATE can be achieved by replacing the weight of one for treatment observations with $w = \frac{1}{\dot{P}(Z)}$.

a concern, two approaches may be taken: either matching on the overall sample and then disaggregating into subgroups and comparing outcomes, or disaggregating into subgroups, and then performing the matching. Matching on the overall sample may be preferable; matching on successive disaggregated groups leads us farther away from the original sample. However, if the balancing properties are not very good on the aggregate-matched, disaggregated groups, when there is some fundamental difference in the two groups making it conceptually desirable to have a new matching model, then disaggregated matching may be appropriate. One example of the latter could be the decision by men and women to participate in labour markets. Note that disaggregated matching runs the risk of having incompatible results between a general model and the disaggregate models. We return to the issue of heterogeneity in more detail in a later section.

Given that the analysis does not condition on all covariates, but on the propensity score, there is a need to check if the matching or weighting procedure is able to balance the distribution of the variables used in the construction of the propensity score. Some type of balancing test is required after matching or weighting to ascertain whether the distribution of the covariates in the two groups has been balanced, in which case, the constructed comparison group can be considered a plausible counterfactual. Alternatively, when dealing with only ex-post data, to ascertain whether among the untreated sample a set of observations with characteristics statistically similar to the treatment group was created. Although several versions of balancing tests exist in the literature, the most widely used is the mean absolute standardized bias (MASB) between participants and non-participants suggested by Rosenbaum and Rubin (1983). Additionally, Sianesi (2004) proposed a comparison of the pseudo R^2 and p-values of the likelihood ratio test of the joint insignificance of all the regressors obtained from the logit analysis before and after matching the samples.

For propensity score methods to be valid, there should be no systematic differences in the distribution of covariates between the two groups. As a result, the pseudo- R^2 should be lower and the joint significance of covariates should be rejected (or the p-values of the likelihood ratio should be insignificant). Of course it is not possible to interpret the results of the impact estimation without estimating the standard errors, which provide an indicator of the importance of sampling error in the generated estimates.

Testing the statistical significance of treatment effects and computing their standard errors is not straightforward with PSM. Imbens and Wooldridge, 2009 provide a detailed overview of the specific approaches that can be taken. In short, the problem is that the estimated variance of the treatment effect should also include the variance due to the estimation of the propensity score, the imputation of the common support, and if possible the order in which treated individuals are matched. Conventionally, one way to deal with this problem is by computing standard errors of PSM using bootstrap methods (Lechner, 2001; Heinrich, Maffioli and Vazquez, 2010). In general, the bootstrap relies on sampling from the analysis sample with replacement, replicating the analysis multiple times. The estimated standard error is the standard deviation of the estimated impact estimate across replications.

The advantage of propensity models is that they do not necessarily require a baseline or panel survey (especially for the outcome variables) although the observed covariates entering the logit model for the propensity score would have to satisfy the conditional mean independence assumption by reflecting observed characteristics that are not affected by participation. The drawback of propensity methods is that they rely on the degree to which observed characteristics drive programme participation, and do not control for unobservable variables that could lead to bias in the estimates. Thus it is important to test the sensitivity of results with respect to such bias and small changes to the matched or weighted samples. With PSM, it is recommended to adopt the Rosenbaum (2002) bounds test, which suggests how strongly an unmeasured variable must influence the selection process in order to undermine implications of matching analysis. In all cases, if the results are sensitive and there are doubts about the CIA, alternative identifying assumptions should be considered.

Regression discontinuity design (RDD)

The final method that may be used by the PtoP team is RDD, which uses programme eligibility rules to exogenously identify the jump—or discontinuity—in the enrolment of households into the cash transfer programme. If a discontinuity is observable based on programme eligibility criteria, it is possible to identify impacts in the neighbourhood of the cut-off point. Most of the programmes we consider have been targeted using some form of "proxy means test"—a generated index conveying the relative well-being of sampled households—to choose who is eligible for the program. Intuitively, households that just barely missed the cut-off can be comparable to households that just barely qualified. That is, the sample of households falling within this range can be treated *as if* treatment was randomly assigned.

In order for RDD to be valid there must be a sufficient number of households around the cut-off range, and there can be no other discontinuity in outcomes or level of treatment outside of the cut-off point. RDD also requires perfect compliance, since unobserved heterogeneity in programme take-up will introduce selection bias (Khandker, Koolwal and Samad, 2010).

Consider the simple impact model presented in equation 8:

$$Y_{it} = \beta_0 + \beta_1 D_{it} + \sum \beta_i Z_i + \mu_{it} \tag{8}$$

where D_i is a dummy equal to 1 if household i received the treatment; Z_i is a vector of conditioning variables, and μ_i is an error term. The parameter β_1 measures the impact of the programme and $\sum \beta_i Z_i$ measures the impact of conditioning variables. Applying OLS to equation 8 generates biased parameter estimates because the error term and the treatment variable might be correlated as a result of the targeting procedure. RDD removes this selection bias by making use of the discontinuity in the eligibility criterion around the programme eligibility threshold.

Following Buddelmeyer and Skoufias (2003) and Khandker, Koolwal and Samad (2010), the approach of estimating the effect of cash transfers on individual outcomes using RDD can be modelled by substituting D_i with an indicator of programme eligibility, M_i, that is for example a function of a proxy means test, S_i, for the i[th] household:

$$Y_i = \beta_0 + \beta_1 M_i(S_i) + \mu_i \tag{9}$$

where the influence of control variables M_i on Y_i is controlled for by including M_i in the construction of proxy mean score. So to estimate effectively the effect of a cash transfer programme using an RDD approach, one needs a variable S_i that determines programme eligibility with an eligibility cut-off of S_i^*.

When the criterion that S_i must lie equal to or above the programme eligibility threshold (S_i^*) is strictly applied to determine eligibility, then M_i is a deterministic function of S_i that is discontinuous at S_i^*,

$$M_i = \begin{cases} 1 \; if \;\; S_i < S_i^*, \\ 0 \quad otherwise \end{cases} \tag{10}$$

This is knows a "sharp" RD design. Because in practice the determination or enforcement of the eligibility may not be "sharp" (as in a randomized experiment), M_i can be replaced with a probability of participating $P(M_i) = E(M_i \mid D_i)$, where $D_i = 1$ if treatment is received and $D_i = 0$ otherwise (Ravallion, 2008; Khandker, Koolwal and Samad, 2010). In this case, the discontinuity is stochastic or "fuzzy" and instead of measuring the difference in outcomes above and below the cut-off point, S_i^* , the impact estimator would measure the difference around a neighbourhood of the threshold. Often this occurs when the eligibility criterions are not strictly adhered to or when certain geographic localities are targeted but with no well defined boundaries. If the eligibility threshold is exogenously determined by the programme and highly correlated with treatment, one might also use the cut-off (S_i^*) as an instrumental variable for participation.

Assuming a constant treatment effect, the average treatment effect can be estimated by comparing the unconditional mean values of Y_i for households with S_i just below and just above the S_i^*. Denoting these means by Y^- and Y^+, respectively, the RDD estimator is given by

$$\beta_1 = Y^- - Y^+ = \lim_{S_i \uparrow S_i^*} E(Y_i \mid S_i = S) - \lim_{S_i \downarrow S_i^*} E(Y_i \mid S_i = S) \tag{11}$$

In order to obtain estimates of the unconditional means of the outcome measures of interest, denoted by Y^{-} and Y^{+}, one can use one-sided kernel regressions[13]. In cases where the treatment effect is allowed to be heterogeneous across households, this approach would identify the local average treatment effect (LATE).

The advantages of RDD are that it yields an unbiased estimate of the treatment effect at the discontinuity and that a group of eligible households need not be excluded from the treatment as controls. Nevertheless the challenges with RDD are that it produces LATE that are not always generalizable, and that the specification can be sensitive to functional form. Besides, the effect is also estimated at the discontinuity and as a result fewer observations are available when compared with a randomized experiment of the same sample size. Finally, RDD is undermined in the presence of spillover effects, as the relevant controls are usually located in the same village or community.

Spillover effects

Through market and non-market interactions, benefits from cash transfer programmes may pass to the rest of the population. If this is the case, impacts found for beneficiaries represent only a part of the overall effect of cash transfer programmes. Depending on the context, documenting these spillover effects may be crucial in understanding the overall contribution of cash transfer programmes to poverty reduction and their cost-effectiveness. Moreover, spillover effects from beneficiaries to non-beneficiaries may complicate the evaluation design by violating the internal validity of estimates of programme impacts. As we mentioned above, the validity of experimental and non-experimental estimators relies on the assumption that the comparison or control group units are not affected by the programme (SUTVA) (Djebbari and Hassine, 2011).

For this reason, as well as for ethical and political considerations, one common solution is to employ village-level randomization, as opposed to household-level randomization, which puts some distance between treatment and control households. This comes at a cost: with clustering the sample size is likely to increase and analysis must account for the intra-cluster correlation to provide consistent estimates of standard errors (Bertrand, Duflo and Mullainathan, 2004).[14] As we will see below, all of the experimental design impact evaluations of our cash transfer programmes in Sub Saharan Africa are randomized at the village level[15].

However, even with village-level randomization, an exclusive focus on programme-eligible treatment and control households will not capture potential spillover effects. However, a sampling strategy that samples both eligible and non eligible households in both treatment and control communities can help recover these spillover effects. The direct impact on beneficiaries can be estimated as the difference in average outcomes between the eligible in treated villages and the eligible in control villages, while the indirect impact on the non beneficiaries can be

[13] Detailed expositions of this can be found in Buddelmeyer and Skoufias (2003); Khandker, Koolwal and Samad (2010)

[14] Sample size calculations are discussed in the country and programme specific sections below.

[15] See Pellerano (2011) for a good example and explanation.

estimated as the difference in average outcomes between the non eligible in treated villages and the non eligible in control villages. The PROGRESA experiment in Mexico employed this strategy, which is exploited in a number of studies to capture spillover effects.[16] Among the cash transfer impact evaluations covered under the project, funding permitting, sampling of both eligible and non eligible households in treatment and control communities may be possible in Malawi, Lesotho, Zimbabwe and Ethiopia.

Moving from spillover effects to ascertain the impact from a fully scaled up version of a given programme is yet another issue. Market linkages transmit the benefits of a transfer from those directly affected to others in the rural economy. When scaled-up, the effects of a cash transfer programme on beneficiaries and non-beneficiaries may differ from the effects obtained in the experimental setting, as general equilibrium or macro feedback effects kick in and change the environment in which the programme is operating (Manski and Garfinkel, 1992). The PtoP project will use the data from the impact evaluation to build village economy models (SAM/CGE) to explore impacts in a general equilibrium setting (Taylor, 2012).

Heterogeneity of Programme Impact Effects

All of the presented estimators assume that the impact of a cash transfer is constant, irrespective of who receives it. The mean impact of a programme or policy based on this assumption is a concise and convenient way of evaluating impacts. Heckman, Ichimura and Todd (1997) justify this approach if researchers and policy makers believe that (a) total output increases total welfare and (b) detrimental effects of the programme or policy on certain parts of the population are not important or are offset by transfers—either through an overarching social welfare function or from family members or social networks.

Within the context of cash transfer programmes, a number of dimensions of heterogeneity may be relevant. Even if the mean programme effect were significant, whether the programme had a significant beneficial or detrimental effect might vary across the subgroups of beneficiaries (Khandker, Koolwal and Samad, 2010). For those programmes with fixed transfers, impact is likely to vary by household size, as the value per capita of the transfer is greater for smaller households. Labour allocation decisions at the individual level are likely to vary between males and females, and between adults and children. Production decisions may vary by the availability of household labour, geographic location, and/or by access to key assets, such as land. But in addition to examining the overall mean impacts it is important to understand how cash transfers affect different types of individuals and households.

There are a number of ways to present the heterogeneous impacts of a cash transfer programme. For example, one could divide the sample of households and individuals into different demographic groups (e.g., by gender or age cohort) and perform separate analysis on each group, and test to see if estimated impacts are different. Interacting the treatment with different household socioeconomic characteristics is another way to capture differences in programme

[16] See, for example, Handa et al. (2001); Angelucci and di Giorgi (2009); Lalive and Cattaneo (2009); Barrientos and Sabates-Wheeler (2011).

effects. Another way to present distributional impacts of a cash transfer programme is by using a quintile regression approach. One could assess, for example, whether poorer or better-off households experienced larger gains from cash transfer programs.

3. Data requirements and outcome indicators

In order to identify the economic impacts of cash transfer interventions at the household level, data collection must extend beyond the topics normally collected by impact evaluations whose objectives align with poverty, food security and human capital improvements. This is one of the main challenges of the project—convincing managers of the impact evaluation process to collect more detailed information in the context of an already overcrowded survey instrument. The difficulties are particularly acute in that, as mentioned above, most of the programmes are located in ministries of social welfare or the like and often the firms contracted to carry out the analysis have scant experience in the livelihood dimension of rural households.

A second key challenge to resolve in the design of the household surveys is the tradeoff between in country and cross country comparability—that is, how can we approximate as much as possible the modules used in existing, government sponsored national household surveys, while also ensuring cross country comparability. Clearly the first priority should be standardization with existing national instruments—this is in the best interests of a given country. However, when appropriate one should push towards greater details of relevant modules along the lines of emerging international good practices, such as the Living Standards Measurement Study-Integrated Surveys on Agriculture (LSMS-ISA) project in Sub Saharan Africa. Of additional importance to our strategy is that the survey instruments used to elicit information are consistent and similar; making a core set of variables comparable across countries and studies.

Using as a framework the channels described in Section 2 through which cash transfer impacts may operate in the near and medium term, data must be collected on the following subjects:

- Household roster and labour market participation at the level of a typical LSMS survey
- Adult and child time use in terms of household chores and own farm/business activities
- Access to land, land tenure, land use and land quality
- Individual decision making on use of household productive assets
- Crop level information on planting, harvest, sales/barter and other uses of production (own consumption, storage, gifts, etc)
- Flow of livestock stocks (including births, deaths, consumption, theft, sales/barter, etc), by animal
- Livestock by-product production and sales/barter
- Crop and animal input use, intensity of use, and cost (seeds, fertilizer, chemicals; veterinary services, feed, etc), including hired and family labour
- Use and ownership of agriculture implements
- Non agricultural business, including monthly costs and income, and use of hired and family labour
- Access and use of credit, insurance and savings
- The whom and where of all cash transactions

- Standard LSMS household modules on consumption, home characteristics, household durables, shocks
- Reciprocal exchange
- Climate change adaptation (specific agricultural related shocks; changes in productive activities, and why)

Impacts on income generation strategies are captured in part through the investment and agricultural production impacts but must be triangulated with changes in the labour allocation of household members and net income from household non-agricultural enterprises. The experience of the Rural Income Generating Activities (RIGA) project[17] is the basis for identifying the key variables for estimating total household income levels and household participation in income generating activities. Data on the revenues earned from household farm and non-farm enterprises and all input expenses related to those enterprises is required for obtaining net income from each activity. Net income from dependent (wage) activities must also be collected for each household member, while household level earnings and outputs due to public and private transfers provide information on non-labour income sources.

Finally, risk management and coping strategies can be ascertained in part through different modules of the survey, including the diversification of income generating activities, child schooling, and sales of assets. Most surveys also collect basic information on incoming and outgoing transfers, which provide an indication of social networks and informal safety nets. However, the importance of these networks of reciprocity in the African context, and the important role that the community plays in targeting and monitoring cash transfer programmes, calls for more innovative methods in detailing these social networks. A partial example of this type of network module (from Lesotho) appears in Figure 1—not only are incoming and outgoing cash transfers asked, but also in-kind, labour and equipment exchanges with friends, family, and community members.

Figure 1. Example of social networks module (Lesotho CGP)

HH17 Q1	A) During the **last 12 months**, has any other family member, friend or neighbour lent or helped with any **money** to support your household?		B) **Ask the names, in order of importance**: What is [CONTRIBUTOR]'s relationship to the household?	C) What is [CONTRIBUTOR]'s gender?	D) Where does [CONTRIBUTOR] live?	E) ♦ **Interviewer**: Look for [CONTRIBUTOR] NAME in the **village list** and record the MIS household ID code.	F) In the last 12 months how much has [CONTRIBUTOR] given to your household?	G) Did you or will you have to give something back in return?
	01 = Yes 02 = No ► **Next question**	No.	***Interviewer:*** *Use network codes* ***(d)*** **If group assistance ► COLUMN F**	01 = Male 02 = Female	***Interviewer:*** *Use Code Sheet* ***(h)*** Not in the village ► **COLUMN F**	MIS household ID Code ***Use 999999 if not in the list***	Amount in MALOTI	01 = Yes 02 = No
		1					M	
		2					M	
		3					M	
		4					M	

An additional two key parameters affecting individual and household decision making—attitudes towards risk and time preferences—derive from economic theories of choice under uncertainty. For example, household investment in productive activities depends on the household's willingness to postpone current consumption for future consumption. A cash transfer program, if it succeeds in being regular and predictable, can alter a household's willingness to delay present consumption for future consumption. A household's predisposition

[17] www.fao.org/economic/riga/en/.

towards risk will also influence investment behaviour; risk averse households may avoid investments which are perceived to be more risky even though the average return may be higher. Further, subjective risk assessments of life expectancy and future quality of life may also influence the planning horizon and thus affect inter-temporal decisions. Attitudes towards risk and time preferences can be captured through hypothetical questions,[18] which have been included in the surveys in Kenya, Zambia, Lesotho and Ghana described below. Figure 2 is an example from Lesotho.

Figure 2. Example of risk aversion and time preference module (Lesotho CGP)

I am now going to read some hypothetical situations where you need to imagine that someone offers you some money without this implying any commitment, debt or obligation for you.			
HH14 Q1	Imagine that you need to choose between the following two alternatives: **01**: Receive **M 500** **02**: Toss a coin and if it is tails you receive **M 1,000** but if it is heads you don't receive anything Which alternative would you chose?	01 02 ► HH14 Q4	
HH14 Q2	Now imagine that you need to choose between the following two alternatives: **01**: Receive **M 300** **02**: Toss a coin and if it is tails you receive **M 1,000** but if it is heads you don't receive anything Which alternative would you chose?	01 02 ► HH14 Q4	
HH14 Q3	Now imagine that you need to choose between the following two alternatives: **01**: Receive **M 100** **02**: Toss a coin and if it is tails you receive **M 1,000** but if it is heads you don't receive anything Which alternative would you chose?	01 02	
HH14 Q4	Again imagine that someone wants to give you some money, without implying any commitment, debt or obligation for you, but imagine that you have two choose between the following two alternatives: **01**: Receive **M 1,000** now **02**: Receive **M 1,050** in a month Which alternative would you chose?	01 02 ► **Next Section**	
HH14 Q5	Now imagine that you need to choose between the following two alternatives: **01**: Receive **M 1,000** now **02**: Receive **M 1,200** in a month Which alternative would you chose?	01 02 ► **Next Section**	
HH14 Q6	Now imagine that you need to choose between the following two alternatives: **01**: Receive **M 1,000** now **02**: Receive **M 1,500** in a month Which alternative would you chose?	01 02 ► **Next Section**	

4. Programme design, sampling framework and estimation strategy in the countries included in the PtoP project

Lesotho Child Grant Programme (CGP)

The Lesotho Child Grants Programme provides an unconditional cash transfer to poor and vulnerable households. The primary objective of the CGP is to improve the living standards of orphan and vulnerable children (OVC) including nutrition and health status and increased school enrolment (Pellerano et al, 2012). The CGP is targeted at poor households with children (<18), including child-headed households—there is no other special definition of "vulnerable" children. As of February, 2012 the program reached 9915 households (covering 28,000 children), with a planned scale up of 5000 additional families annually, reaching 25,000 in 2014. The monthly value of the transfer is 120 Maloti (approximately 14 USD) and is disbursed on a quarterly basis.

[18] See Delavande, Gine and McKenzie (2011) for an overview on the measurement of subjective expectations in household surveys; the particular formulation of the questionnaire shown in Figure 2 is still somewhat experimental.

The quantitative analysis for the Lesotho study is an experimental design impact evaluation. Participation in the program was randomized at the level of the electoral district (ED). First, all 96 EDs in four community councils were paired based on a range of characteristics. Once these 48 pairs were constructed, 40 pairs were randomly selected to be included in the evaluation survey. Within each selected ED, 2 villages (or clusters of villages) were randomly selected, and in every cluster a random sample of 20 households (10 potentially called to enrolment and 10 potentially non-called to enrolment) were randomly selected from the lists prepared during the targeting exercise. After the baseline survey data were collected in all evaluation EDs, public meetings were organized where a lottery was held to assign each ED in each of the pairs (both sampled and non-sampled) to either treatment or control groups. Selecting the treatment electoral districts after carrying out the baseline survey helped to avoid anticipation effects (Pellerano, 2011). Sample and power calculations can be found in the impact evaluation inception report (Hurrel, et al, 2011).

The baseline household survey was carried out (fieldwork finished as of August 30, 2011) prior to the distribution of the first transfers to treatment households; a follow up panel survey will take place two years later in 2013. A total of 3102 households were surveyed; 1531 programme eligible households (766 treatment and 765 control) to be used for the impact evaluation analysis, with the remaining 1571 programme non eligible households to be used for targeting analysis and spillover effects. Besides the household survey, community and business enterprise questionnaires were implemented.[19] The impact evaluation is being implemented by Oxford Policy Management (OPM) and Sechaba Consultants.

The method of randomization described above, including the relatively large number of units of randomization, reduced the likelihood of systematic differences between treatment and control households. The baseline analysis report (Pellerano et al., 2012) shows that the randomization was quite successful, with few significant differences between the households in treatment and control groups on key selection indicators. Dewbre (2012) showed the same for the PtoP related variables. As to be expected in any randomization process, a small number of significant differences did emerge between treatment and control households, including some dimensions of food security, social networks, land cultivation, and livestock and land ownership. The household and individual level impact analysis will thus use DD estimators and, when necessary, propensity score methods.

Kenya Cash Transfers for Orphan and Vulnerable Children (CT-OVC)

The Kenya Cash Transfer Programme for Orphans and Vulnerable Children (CT-OVC) is the Government's flagship social protection programme, reaching over 125,000 households as of the end of 2011, with the ultimate goal of providing coverage to 300,000 households or 900,000 OVC (ages <18) (Kenya CT-OVC Evaluation Team, 2012a). The monthly value of the transfer over the course of the evaluation was Ksh 1500 (approximately 20 USD) and is disbursed on a bi-monthly basis.

[19] The questionnaires can be found on the Transfer Project website (http://www.cpc.unc.edu/projects/transfer/countries/lesotho).

The impact evaluation used a randomized cluster longitudinal design, with the baseline quantitative survey fieldwork carried out in 2007. The underlying justification for the design was that the programme could not be spread out to all eligible locations at the same time, and as a result sites whose entry was expected to happen later were considered as control sites. Approximately 2750 households were surveyed in seven districts across Kenya (Nairobi, Kwale, Garissa, Homa Baye, Migori, Kisumu and Suba). Two-thirds of these households later began receiving cash transfers while the remaining households served as the control group. Within each district, two locations were chosen randomly to benefit from the programme intervention and two were selected as controls (Ward et al., 2010; Kenya CT-OVC Evaluation Team, 2012a). These households were re-interviewed (first round) two years later in order to assess the impact of the programme on key welfare indicators (Ward et al., 2010). The second round follow up study was conducted in 2011. The initial stage of the evaluation was implemented by OPM, while the four year follow-up survey was conducted by the University of North Carolina (UNC), the Government of Kenya, Research Solutions Africa, and the FAO. Attrition is significant, particularly between baseline and first follow up (18 percent, and five percent between 2009 and 2011), though analysis has shown that this attrition was random (Handa, 2012).

Unfortunately the 2007 and 2009 rounds of data collection were very weak on economic activities, including labour market participation. The 2011 second round follow up added a more detailed economic activity module (including wage labour, self-employment, crop and livestock activities, risk preferences, etc) to capture the potential investment and productive activity benefits of the programme.[20] For some of the household level analysis, such as ownership of livestock and agricultural implements, baseline data were collected. Thus for many of the outcome variables of interest to the PtoP project, we have only one data point (no baseline).

The randomization in Kenya was not as robust as in the case of Lesotho due to the fewer units over which the randomization took place; DD alone (Kenya CT-OVC Evaluation Team, 2012a and 2012b) and combined with propensity score techniques when necessary (Ward, et al, 2010; Asfaw, et al, 2012) have been used to analyze these data. In cases where the outcome variables are not available at the baseline, we will be forced to rely exclusively on the PSM technique using cross sectional data (with some baseline predictors) to identify the effect of treatment.

Ethiopia Tigray Social Cash Transfer Pilot Programme

This cash transfer programme initiated by Tigray regional state and UNICEF aims to reduce poverty, hunger and starvation in extremely poor, labour constrained households, with particular attention on elderly and child headed households, female headed households, and households with disabled. The programme also seeks to enhance household access to essential social welfare services such as health care and education via access to schools. A total of 169,540 (4% of the regional population) are expected benefit directly or indirectly during the pilot phase (BOLSA-Tigray, 2011). The initial disbursement took place in September, 2011, and monthly transfers

[20] The questionnaires can be found on the Transfer Project website (http://www.cpc.unc.edu/projects/transfer/countries/kenya).

will continue for two years as a pilot. The basic household grant is 155 ETB (approximately 9 USD),[21] per beneficiary household, plus 35 ETB per child in school, 50 ETB per disabled child in school, 25 ETB per child out of school, 40 ETB per disabled child out of school and 60 ETB per elderly dependent. Entitlements are provided for a maximum of 4 dependents (children and elderly persons) per household, and the average household grant is 300 ETB ($17) per month.

One urban and one rural woreda were purposefully selected for the pilot: Abi Adi town and Hintalo-Wajirat woreda, respectively. All residents of Abi Adi that met selection criteria were eligible for the programme, while in Hintalo-Wajirat only 7 of 22 purposefully selected tabias were included. Subsequently an 8th tabia was included, where payments had not yet begun as of June, 2012. The programme will be targeted via local community care coalitions (CCCs) at the Tabia level, which act as an entry point and support mechanism for especially vulnerable households in the community. The CCCs first identify households that meet the extreme poverty ("poorest of the poor" and not able to meet essential needs) and labor constrained (no able-bodied labour aged 19-64, or dependency ratio of greater than 3) criteria, and then rank these households by neediness. Beneficiaries are then selected from this ranked list.

The objective of the impact evaluation, which will be implemented by a consortium including the International Food Policy Research Institute (IFPRI), the Institute of Development Studies (IDS), and Mekelle University, is to provide evidence as to the viability of the expansion of the programme to other regions of the country. The evaluation strategy aims to evaluate the impacts of the cash transfers on the socioeconomic conditions of the beneficiary families, such as food and non-food consumption, as well as on a number of human development indicators such as school enrolment and attendance, health and nutrition outcomes, and child labour, and employment effects and its contribution to economic growth. The evaluation will also focus on targeting and operational effectiveness.

As is evident from the operational implementation of the pilot, the impact evaluation design is non-experimental. A detailed description of the impact evaluation design can be found in Berhane, et al (2012). Treatment households will be randomly selected from beneficiary lists. Comparison households will be taken from treatment communities, due to IFPRI-led consortium's argument that sufficiently comparable tabias are not to be found in the study area. These will be randomly drawn from those eligible households not selected into the program. For both these treatment and comparison groups, elderly-, child- and female-headed households, as well as households with a disabled member, will be oversampled. A final group of non-eligible household will be randomly drawn from tabia lists.

Sample size calculations were based on data from the PSNP impact evaluation from Tigray, carried out by IFPRI, using the following criteria: power of 80 percent; 5 percent significance level; and minimum detectable effect sizes ranging from 10-15 percent for five outcomes (months of food security, livestock holdings, net transfers received, fertilizer use, and access to credit). Calculations yielded 1500 treatment, 1500 comparison, and 470 non eligible households for a total of 3470 households. Separate calculations were made for the 8th tabia (Bahr Tseba) totalling 405 households (Behane, et al, 2012).

[21] 10ETB is equivalent to .56 USD as of July 24, 2012.

Besides the non random assignment of the programme, a second challenge of this impact evaluation is that that transfers began in September, 2011, nine months prior to baseline. Where possible, this will be addressed by the inclusion of retrospective questions in the baseline survey. Overall, given these challenges, the consortium has proposed using DD combined with PSM as well as possibly RDD to identify impact. After the baseline in June, 2012, shorter monitoring surveys will be fielded later in 2012, 2013 (three times) and 2014, with a final household survey in June, 2014.

Malawi Social Cash Transfer (SCT) Programme

The Malawi Social Cash Transfer Programme (SCT) was initiated in 2006 in the pilot district of Mchinji, providing small cash grants to ultra-poor, labour-constrained households. The SCT programme objectives include reducing poverty and hunger in vulnerable households and increasing child school enrolment. The SCT is currently operational in 7 districts and reaches over 26,000 ultra-poor and labor-constrained households and is expected to serve 300,000 households by 2015. At present, approximately 103,000 individuals benefit from the programme of which nearly two-thirds are children and nearly half are orphaned children. The programme is executed by the Government of Malawi through the District Councils by Social Welfare Officers (UNICEF-Malawi, 2010). On average 12 USD per month are paid quarterly. The transfers vary according to household size: 4, 7, 10 and 13 USD will be given to households with 1, 2, 3 and 4 members, respectively. In addition, 1.5 USD will be given on a monthly basis to households for each child enrolled in primary school and 3 USD for each child enrolled in secondary school.

An impact evaluation of the programme pilot in Mchinji in 2007-08 was implemented by Boston University and the Center for Social Research at the University of Malawi (Miller, Tsoka and Reichert, 2008). A new impact evaluation is planned to accompany the next round of programme expansion. The final impact evaluation design is awaiting the tendering of the contract for the evaluation, which will be finalized by August, 2012. As such, it is unclear what level of randomization will be possible. The baseline will take place in the fourth quarter of 2012, with a follow up survey one year later.

Zimbabwe Social Cash Transfer (SCT) Programme

To address household poverty as a key driver of child vulnerability in Zimbabwe, the revised National Action Plan for Orphans and Vulnerable Children (NAP II) 2011-2015 and its accompanying pooled funding mechanism (the Child Protection Fund) have included social cash transfers as a major programme component, accompanying other interventions in child protection and access to social services. The Fund is a multi-donor pooled funding mechanism managed by UNICEF in partnership with other partners (UNICEF-Zimbabwe, 2012).

District selection for the prioritization of the phased cash transfer programme has been based on a cross-analysis of the Poverty Assessment Survey (2003), the Nutrition Survey (2010) and the ZIMVAC (2010) to determine a proxy for prioritizing specific districts. The overall objective of the new impact evaluation is to generate policy-relevant evidence on the impact of the cash transfer scheme of the NAP II on key child health, education, HIV, equity, nutrition, protection and livelihood (or household economy) outcomes (UNICEF-Zimbabwe, 2012). The first transfers were made in December; the impact evaluation will be applied on successive rounds of implementation of the pilot phase of the programme. The average volume of transfers is set as USD 20 per household per month. For individual households the transfers vary in accordance with the size of the household. Households with one member get 10 USD whereas households with two, three and four (or more) members get 15, 20 and 25 USD respectively.

The tender for the impact evaluation was awarded to AIR in collaboration with the University of North Carolina, Ruzivo Trust and the University of Zimbabwe, and the inception workshop was held in late June, 2012. The impact evaluation is a two year, 3000 household longitudinal design, comparing cash transfer recipient households from 45 wards in three Phase 2 districts (Binga, Mwenzi, and Mudzi) to eligible households in 45 wards in three Phase 4 districts (UMP, Chiredzi, and Hwange) that will not receive the transfers during the period of the study. The design is non experimental since Phase 2 districts had already been told that all eligible district residents would receive the transfers as the district enters the program.[22] The steps for selecting the sample are the following:

1. The three treatment districts from Phase 2 and three matching comparison districts from Phase 4 have been selected by the Ministry of Labor and Social Services (MoLSS).
2. The MoLSS, with oversight from UNICEF and the evaluation team, will randomly select 45 wards from the three treatment districts.
3. The evaluation team will work with the MoLSS to select 45 wards from the comparison districts that match the selected wards from the treatment districts. Wards will be matched by geography, climate, over-all development level, availability of services, access to other development programs, and culture, with an emphasis on making sure that the agro-ecological environment of the treatment wards is similar to that of the comparison wards.
4. After selecting the 90 study wards, the MoLSS will conduct targeting in these 90 wards to identify eligible households. The targeting process will be conducted in exactly the same way in both the treatment and comparison wards.
5. The evaluation team will randomly select 34 households that have been identified through the targeting process as eligible for the program from each of the 90 wards. These randomly selected households will comprise the sample for the impact evaluation.

The baseline data collection will take place from February to March 2013, allowing MoLSS time to complete targeting of households in the study wards. The two rounds of follow up data collection will occur the same time of year, in one year intervals, in order to control for seasonal effects. The study includes five quantitative instruments: the household survey, an adolescent survey for those aged 14-23; anthropometric measurement for children under 5; a community survey; and a business enterprise survey. The study also includes four qualitative components:

[22] For more details, see Matondi, et al, 2012.

1. Focus group and key informant interviews during the last quarter of 2012, focusing on the impacts of the HSCT on economic activities and social networks, as well as general impressions of the implementation of the first phase of the transfer. The two main objectives of the field work would be to first provide some early insight into the impacts of the program, and second help guide the development of the household questionnaire that will be used in the baseline survey in early 2013.
2. Focus groups at baseline to assess experience, beliefs, and perceptions of targeting in Phase 2 districts.
3. Embedded longitudinal interviews that track Phase 2 recipients over time, including adolescents, women, and men.
4. Focus groups at 12 month follow up to learn about experiences, perceptions, social networks, and economic activities of Phase 2 beneficiaries.

Ghana Livelihood Empowerment Against Poverty (LEAP) programme

The Ghanaian Livelihood Empowerment Against Poverty (LEAP) programme is a social cash transfer program which provides cash and health insurance to extremely poor households to improve short-term poverty and encourage long-term human capital development. LEAP started a trial phase in March 2008 and then began expanding gradually in 2009 and 2010, and currently reaches over 35,000 households across Ghana with an annual expenditure of approximately USD11 million (Handa and Park, 2012). The programme is fully funded from general revenues of the Government of Ghana, and is the flagship programme of its National Social Protection Strategy. On average 22.5 USD per month is paid quarterly. The transfers vary according to household size: 15, 18, 21 and 27 USD will be given to households with 1, 2, 3 and 4 members, respectively. In addition, beneficiary households are linked free of charge to Ghana's national health insurance programme.

The LEAP programme operates in all 10 regions of rural Ghana. Within regions, districts are selected for inclusion based on the national poverty map; within districts, local DSW offices choose communities based on their knowledge of relative rates of deprivation.

The impact evaluation, implemented by the Carolina Population Center at UNC and the Institute for Statistical, Social and Economic Research (ISSER) at the University of Ghana, takes advantage of a nationally representative household survey implemented during the first quarter of 2010. The initial treatment sample of 700 households were randomly drawn from the group of 13,500 households that were selected into the programme in the second half of 2009, and are located in 7 districts across 3 regions (Brong Ahafo, Central, Volta). These households were interviewed prior to receiving any indication that they had been selected for LEAP (Handa and Park, 2012). The baseline survey instrument was a redued version of the national household survey instrument, and the national survey sample and the treatment household sample were surveyed at the same time by ISSER. The strategy is to draw the control households from the national survey using PSM techniques. A comparison group of 'matched' households was selected from the ISSER sample and re-interviewed after 2 years, in April-May, 2012, along with LEAP beneficiaries to measure changes in outcomes across treatment and comparison groups.[23]

Handa and Park (2012) carried out PSM analysis to identify the comparison group from the Yale/ISSER sample and assessed the appropriateness of this comparison group for the evaluation strategy. Their results show that for the outcome variables of the original evaluation design (demographic and child welfare), the PSM strategy works well. For the purposes of the PtoP project, however, this sample of control households is not optimal, as can be seen in Table 3. There is little that can be done to change this sample, since the PSM corresponds to the original objectives of the evaluation. One practical solution would be to rerun the PSM expost once we have obtained the collected second round data, hoping that there is an improvement in the matching. A second strategy will be to re-estimate the matching using alternative methods, such as propensity score weights.

Table 3: Comparison of outcomes by matched samples

			Matched Samples		
Outcome Variables	**Leap**	**Yale Rural**	**All Rural Regions**	**Six Regions**	**Three Regions**
Use fertilizer/pesticides (Binary)	0.15	0.36	**0.23**	**0.24**	**0.24**
Sell crops (Binary)	0.31	0.51	**0.41**	**0.46**	**0.44**
Have non-farm enterprise (Binary)	0.29	0.29	0.28	0.3	**0.34**
Total number of hoes	1.72	2.17	1.9	1.6	1.86
Total number of axes	0.37	0.34	**0.29**	**0.28**	**0.3**
Total number of rakes	0.06	0.05	0.04	**0.02**	**0.03**
Total number of shovels	0.09	0.24	**0.19**	0.11	**0.17**
Total number of pick	0.1	0.18	0.13	0.1	0.09
Total number of sickle	0.07	0.33	**0.24**	0.09	**0.02**
Total number of cutlass	1.42	2.06	**1.7**	**1.87**	**1.94**
Total days of casual labour	4.94	5.45	3.95	4.75	4.06
Total hours of casual labour	4.84	5.62	4.22	4.83	4.97
Total number of casual labourers	2.41	4.19	2.86	2.89	**3.05**

Note: 2012 Handa and Park (2012) analysis of the LEAP baseline data. Column 1 and 2 report mean outcomes of interest for treatment (LEAP) sample and control (YALE) sample before matching. Columns 3, 4 and 5 report mean outcomes for control households when the matching exercise is run over different subsamples of the YALE sample. Means from the matched samples that are statistically different from LEAP are in bold.

[23] The baseline questionnaire can be found on the Transfer Project website (http://www.cpc.unc.edu/projects/transfer/countries/ghana).

Zambia Social Cash Transfer Programmes

In 2010, Zambia's Ministry of Community Development and Social Services (MCDSS) began implementing the Child Grant cash transfer programme (CGP) in three districts—Kalabo, Kaputa, and Shongombo—the three districts with the highest rates of mortality, morbidity, stunting, and wasting among children under five years of age. All households in these three districts with a child under five years of age are eligible to receive benefits. Eligible households receive 55,000 kwacha a month (equivalent to 11 USD) irrespective of household size, an amount considered sufficient to purchase one meal a day for everyone in the household for one month. The goal of the programme is to reduce extreme poverty and the intergenerational transfer of poverty. The objectives of the programme are to (1) supplement and not replace household income; (2) increase the number of children enrolled in and attending primary school; (3) reduce the rate of mortality and morbidity among children under 5 years old; (4) reduce stunting and wasting among children under 5 years old; (5) increase the number of households owning assets such as livestock; and (6) increase the number of households that have a second meal a day (Seidenfeld and Handa, 2011a).

UNICEF-Zambia contracted AIR to design and implement a randomized controlled trial for a three-year impact evaluation of the programme. The baseline was carried out in September-October 2010, with follow ups planned for 2012 and 2013. Baseline data collection occurred in Zambia's lean season (September through February), when people have the least amount of food left from the previous harvest and hunger is at its greatest. The study includes 2,515 households, 1228 treatment and 1287 control. Communities were randomly assigned to treatment (incorporated into the programme in December 2010) or control (to be brought into the programme at the end of 2013). The baseline data collection began before communities were randomly assigned to treatment or control groups. Analysis of the baseline data shows that randomization appears to have worked well; greater detail on the randomization process, sample design and power calculations can be found in Seidenfeld and Handa (2011a).

The AIR-led impact evaluation also followed up on an earlier cash transfer program in Monze District. Also implemented by the MCDSS, the Monze cash transfer programme delivers bi-monthly cash transfers to destitute and labor constrained households. Under the Monze scheme beneficiary households receive ZMK 40,000 (if they have no children) and ZMK 50,000 (if they have children)—roughly 8 USD and 10 USD, respectively (Seidenfeld and Handa, 2011b).

Targeting for the programme was based on the identification and ranking of poor households in each of the 105 communities comprising Monze district. The strategy consisted first of identifying needy and destitute households by community leaders. Then, basic information was collected on these households by community council teams (CWACs), and rankings were assigned to every household. The bottom 10% of these households were selected for potential enrolment into the program, and final decision making was undertaken by the District Social Welfare Office (DSWO).

The baseline survey took place in August, 2007 with follow-up data collection implemented by AIR in August, 2010. Treatment and control status was initially randomised at community level, with eligible households in control communities slated to enter the program after 3 years.

Unfortunately, the targeting procedure was not perfectly mimicked in control communities, leading to significant differences between treatment and control households at baseline (Seidenfeld and Handa, 2011b). The quantitative analysis will have to rely on quasi-experimental approaches to construct suitable comparison groups.

Final words

Ascertaining the economic impacts of cash transfer programs is a new area of research among impact evaluations of cash transfer programmes. It has important implications for policy—given the increasing popularity of cash transfer programmes in Sub Saharan Africa, the time is ripe for understanding the full impacts – intended and unintended – of these programmes. This can help inform the ongoing policy debate by documenting the full contribution of cash transfer programmes to hunger reduction, poverty reduction and inclusive growth in addressing concerns about sustainability and analyzing the productive and economic contribution of social assistance. Understanding household behavior and local dynamics can help sharpen programme design and implementation by highlighting potential synergies and constraints and strengthening programmes' graduation strategies, which in rural Africa will come primarily through agricultural and rural non-farm activities.

References

Alzúa, M., Cruces, G. and Ripani, L. 2010. Welfare Programs and Labor Supply in Developing Countries. Experimental Evidence from Latin America. CEDLAS Working Papers 0095, Universidad Nacional de La Plata.

Amarante, V. and A. Vigorito (2010) "Conditional cash transfers, labor supply and informality: the case of Uruguay", Unpublished manuscript, Universidad de la Republica.

Angelucci, M. and Di Giorgi G. (2009) "Indirect effects of an aid program: how do cash injections affect ineligibles' consumption," *American Economic Review*, 99(1), 486-508.

Barrientos, A. 2012. Social Transfers and Growth: What do we know? What do we need to find out? *World Development* 40(1): 11-20.

Barrientos, A., and R. Sabates-Wheeler, R. 2011. Strategic complementarities and social transfers: How do PROGRESA payments impact nonbeneficiaries." *Applied Economics* 43.

Berhane, G., Hoddinott, J., Kumar, N. and Taffesse, A.S., (2011). The impact of Ethiopia's productive safety nets and household asset building programme: 2006-2010. IFPRI, Washington DC. USA.

Berhane, G., Devereux, S., Hoddinott, J., Tegebu, F., Roelen, K., Schwab, B. and Taffesse, A.S., (2012). Evaluation of the Social Cash Transfers Pilot Programme, Tigray Region, Ethiopia. Inception Report. IFPRI, Washington, DC. USA.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). “How Much Should We Trust Differences in differences Estimates?”, *The Quarterly Journal of Economics* 119(1):249-275.

Binswanger-Mkhize, H., McCalla, A. and Patel, P. (2010). “Structural Transformation and African Agriculture.” *Global Journal Of Emerging Market Economics*, Vol.2, Issue 2, Chapter 1, pp.113-152.

Boone, R, Covarrubias. K., Davis. B. and Winters, P., (2012). “Cash Transfer Programs and Agricultural Production: The Case of Malawi”, Mimeo, FAO, Rome.

Buddelmeyer, H. and Skoufias, E., (2003). An evaluation of the performance of regression discontinuity design on *PROGRESA.* Bonn, Germany, IZA.

BOLSA-Tigray, 2011. Manual of Operations for the Tigray Social Cash Transfer Pilot Program, April 2011 – March 2014. Tigray Bureau of Labour and Social Affairs and UNICEF, Mekelle, August.

Caliendo, M. and Kopeinig, S., (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys,* 22 (1), pp. 31–72.

Covarrubias. K., Davis. B. and Winters, P., (2012). “From Protection to Production: Productive Impacts of the Malawi Social Cash Transfer Scheme.” *Journal of Development Effectiveness*, 4(1): 50-77, March.

Davis, B., Carletto G. and Winters, P. 2010. *Migration, Transfers and Economic Decision Making among Agricultural Households*, introduction to Special Issue, *Journal of Development Studies*, 46(1), January.

Davis, B., Gaarder, M., Handa S. and Yablonski, J., (2012). “Evaluating the impact of cash transfer programs in Sub Saharan Africa: an introduction to the special issue.” *Journal of Development Effectiveness,* 4(1): 1-8, March.

Dehejia, H.R. and Wahba, S., (2002). “Propensity score matching methods for non-experimental causal studies.” *Review of Economics and Statistics,* 84 (1), pp. 151–61.

Delavande, A., Gine, X. and McKenzie, D. 2011. Measuring Subjective Expectations in Developing Countries: A Critical Review and New Evidence, *Journal of Development Economics*, 94(2):151-163.

Dewbre, J. 2012. “A Note on the 2011 Lesotho Child Grants Program (CGP) Baseline Data”, FAO, Rome.

Djebbari, H. and Hassine, N.B., (2011). Methodologies to analyze the local economy impact of SCTs. Report prepared for UNICEF-ESARO.

Diao, X., Heady, D. and Johnson, M., 2008. Toward a green revolution in Africa: what would it achieve, and what would it require? *Agricultural Economics* 39(S1), 539-550.

Edmonds, E. V., and N. Schady (2008). "Poverty Alleviation and Child Labor." Policy Research Working Paper 4702, World Bank, Washington, DC.

Fisbein, A. and Schady, N., (2009). Conditional Cash Transfers for Attacking Present and Future Poverty. Policy Research Report, World Bank, Washington DC

Galasso, E. (2006). ""With their effort and one opportunity": Alleviating extreme poverty in Chile" Unpublished manuscript, World Bank.

Gertler, P., Martinez, S., and Rubio-Codina, M., (2012). "Investing cash transfers to raise long-term living standards." *AEJ: applied economics*, 4 (1), 164–192.

Gilligan, D., Hoddinott J. and Taffesse, A. 2009. "The Impact of Ethiopia's Productive Safety Net Program and Its Linkages." *Journal of Development Studies* 45(10): 1684- 1706.

Handa, S. 2012. Kenya Cash Transfer for Orphans and Vulnerable Children Evaluation: Sample Design and Description, mimeo, University of North Carolina.

Handa, S., Huerta, MC., Perez, R. and Straffon, B. 2001. Poverty, inequality, and spillover in Mexico's education, health, and nutrition program, FCND Discussion Paper #101, IFPRI, Washington, DC.

Handa, S., Davis, B., Stampini, M. And P. Winters., (2010). "Heterogeneous Treatment Effects in Conditional Cash Transfer Programs: Assessing the impact of Progresa on Agricultural Households." *Journal of Development Effectiveness*, 2(3):320-335, September.

Handa, S. and Davis, B., (2006). The experience of conditional cash transfers in Latin America and the Caribbean. *Development policy review*, 24 (5), 513–536.

Handa, S. and Park, M. (2012). Livelihood Empowerment Against Poverty Program. Report, University of North Carolina at Chapel Hill.

Heckman, J., Ichimura, H., Smith, J. and Todd, P., (1998). "Characterizing selection bias using experimental data." *Econometrica*, 66 (5), pp. 1017–1098.

Heckman, J.J., Ichimura, H. and Todd, P.E., (1997). "Matching as an econometric evaluation estimator: evidence from evaluating a job training program." Review of Economic Studies, 64, pp. 605–654.

Heinrich, C., Maffioli, A. and Vazquez, G., (2010). A primer for applying propensity score matching. Impact evaluation guidelines, Technical notes IDB-TN-161. Inter-American Development Bank.

Hirano, K., Imbens, J. and Ridder, G., (2003) "Efficient Estimation of Average Treatment Effects using the Estimated Propensity Score", *Econometrica*, 71(4); 1161-1189

Hoddinott, J. 2008. Social safety nets and productivity enhancing investments in agriculture, IFPRI, Washington, DC.

Hurrell, A., Pellerano, P., MacAuslan, I., Merttens, F and Kardan, A. 2011. CGP Impact Evaluation—Inception Report. OPM.

Imbens, G. and Wooldridge, J. (2008) "Recent Developments in the Econometrics of Program Evaluation", *Journal of Economic Literature*, 47(1): 5-86.

Kenya CT-OVC Evaluation Team (Palermo, T., Alviar, C., Davis, B., Handa, S., Hurrell, A., Hussein, A., Musembi, D., Ochieng, S., Pearson, R., Pellerano, L., Visram, A. and Ward, P.). (2012a). "The impact of the Kenya Cash Transfer Program for Orphans and Vulnerable Children on household spending." *Journal of Development Effectiveness*, 4(1): 9-37, March.

Kenya CT-OVC Evaluation Team (S Handa, C Alviar, B Davis, A Hurrell, A Hussein, D Musembi, S Ochieng, R Pearson, L Pellerano, A Visram, P Ward). (2012b). "The impact of Kenya's Cash Transfer for Orphans and Vulnerable Children on human capital," *Journal of Development Effectiveness*, , 4(1): 38-49, March.

Khandker, R.K., Koolwal, G.B., and Samad, H.A., (2010) *Handbook on impact evaluation: quantitative methods and practices*. The World Bank, Washington DC., USA.

Lalive R. and Cattaneo, A. (2009) "Social Interactions in Schooling Decisions", *Review of Economics and Statistics*, 91(3): 457–477.

Lechner, M., (2001). "Program heterogeneity and propensity score matching: an application to the evaluation of active labour market policies." *The Review of Economics and Statistics* 84(2): 205-220.

Manski, C.F. and Garfinkel, I., eds, 1992. *Evaluating welfare and training programs*. Cambridge: Harvard University Press

Miller, C., Tsoka, M. and Reichert, K., (2008). Impact Evaluation Report: External Evaluation of the Mchinji Social Cash Transfer Pilot. Boston: Center for International Health and Development, Boston University and Zomba: Centre for Social Research, University of Malawi.

Matondi, P., Zanamwe, L., Seidenfeld, D. and Handa, S. 2012. Inception Report for the Impact Evaluation of Zimbabwe's Harmonized Social Cash Transfer Program, AIR, July.

OPM, (2012). Qualitative research and analyses of the economic impacts of cash transfer programmes in sub Saharan Africa: Research Guide. Prepared for the Food and Agriculture Organization of the United Nations, March.

Parker, S. W., and E. Skoufias (2000). "The Impact of PROGRESA on Work, Leisure, and Allocation." Unpublished manuscript, International Food Policy Research Institute, Washington, DC.

Pellerano, L., (2011). CGP impact evaluation: sampling design and targeting evaluation research. OPM, June 30.

Pellerano, L., Hurrell, A., Kardan, A., Barca, V., Hove, F., Beazley, R., Modise, B., MacAuslan, I., Dodd, S. and Crawfurd, L., (2012). CGP impact evaluation: targeting and baseline evaluation report. OPM, January 10.

Ravallion, M., (2008). Evaluating anti-poverty programs. Working paper 3625, Development Research Group, World Bank, Washington DC, USA.

Rosenbaum, P.R. and Rubin, D.B., (1983). "The central role of the propensity score in observational studies for causal effects." *Biometrika*, 70(1), pp. 41-55.

Rubin, D. B., (1980). Discussion of "Randomization Analysis of Experimental Data in the Fisher Randomization Test," *Journal of the American Statistical Association*, 75, 591–593.

Schady, N., and Araujo M. C., (2008). "Cash Transfers, Conditions, and School Enrolment in Ecuador." *Economia* 8 (2): 43–70.

Seidenfeld, D. and Handa, S. (2011a). Zambia's Child Grant Program: Baseline Report, American Institutes for Research, Washington, DC., November.

Seidenfeld, D. and Handa, S. (2011b). Results of the Three Year Impact Evaluation of Zambia's Cash Transfer Program in Monze District: Final Report, American Institutes for Research, Washington, DC., June.

Sianesi, B. (2004). "An evaluation of the Swedish system of active labour market programmes in the 1990s." *Review of Economics and Statistics*, 86 (1), pp. 133–155.

Skoufias, E. and V. Di Maro (2006). "Conditional Cash Transfers, Adult Work Incentives, and Poverty". World Bank Policy Research Working Paper No. 3973.

Smith, J.A and Todd, P. E., (2005). "Does matching overcome LaLonde's critique of on experimental Estimators?" *Journal of Econometrics*, 125, pp. 305–353.

Todd, J, P. Winters and T. Hertz. (2010). “Conditional Cash Transfers and Agricultural Production: Lessons from the Oportunidades Experience in Mexico”, *Journal of Development Studies*, 46(1), January.

Taylor, E., (2012). A Methodology for Constructing Local Economy-wide Models for Project Impact Evaluation, Prepared for the Food and Agriculture Organization and The World Bank. March.

Toenniessen, G., Adesina, A. and DeVrie, J., 2008. Building an Alliance for a Green Revolution in Africa. Annals of the New York Academy of Science 1136, 233–242.

UNICEF-Malawi, (2010). Malawi Protection for the most vulnerable households. Available from: http://www.unicef.org/infobycountry/malawi_56675.html. [Accessed 27 October 2011]

UNICEF-Zimbabwe, (2012). Request for Proposal: RPF /ZIMA/2011/013, Harare, January.

Ward, P., Hurrell, A., Visram, A., Riemenschneider, N., Pellerano, L. and MacAuslan, I. (2010). Kenya CT-OVC programme operational and impact evaluation 2007-2009. Oxford Policy Management, Oxford, UK.

Wooldridge, J.M., (2002). *Econometric analysis of cross-section and panel data*. Cambridge, MA: The MIT Press.

World Bank, 2008. *World Development Report 2008: Agriculture for Development*. World Bank, Washington, DC.

Zezza, A., Winters, P., Carletto, C., Covarrubias, K., Davis, B., Tasciotti, L. and Quinones, E., 2011. “Rural Household Access to Assets and Agrarian Institutions: A Cross Country Comparison,” *European Journal of Development Research* 23, 569–597.