Methodologies to analyze the local economy impact of SCTs

June 27th, 2011

Final report prepared for UNICEF-ESARO

Habiba Djebbari
IZA, CIRPÉE and Laval University

Nadia Belhaj Hassine
Economic Research Forum

Executive summary

The objective of this review is to provide a framework to understand and document the full range of impacts from Social Cash Transfer (SCT) programs. Local economy impacts of SCT programs may be generated through market and/or non-market interactions between the claimants of the transfers and other people in the community. Our review of the micro approach to local economy effects of SCT programs provides a few lessons for data collection and analysis. The modern econometric framework for impact evaluation based on the definition of a counterfactual is a useful tool to start comprehending the full impacts of a SCT intervention. An impact evaluation study can be designed to capture spillover effects at the local economy level. Two-stage experimental design is appropriate to account for spillovers and measure impacts at the village level. But, evidence from the few empirical studies available so far is incomplete. We need good data on agricultural activities and outcomes to assess the impacts of SCTs on these dimensions. We need a theory to explain spillover effects. Insights from theory can help guide data collection to test competing hypothesis of the sources of spillover effects. Investigating the heterogeneity in impacts is a promising approach for uncovering the mechanisms at play within this reduced-form framework.

There are, however, a number of limitations intrinsic to the micro approach. This approach can only account for impacts on *equilibrium* outcomes. The evaluation of the effects of SCT programs on local or national economies can be approached by drawing on analytical frameworks that capture how policy changes and economic shocks affect key macroeconomic balances, and how the repercussions are transmitted to households, particularly via factor employment and incomes. But studies using meso/macro methods to assess economy-wide effects of SCT programs are scarce. This may be because CGE models typically explore effects of policy reforms that take place at a macro level in a top-down approach. In contrast, cash transfers enter the economy at the household level. Village SAM/CGE offer a bottom-up approach that fits better with the level at which cash transfers enter the economy. Village SAM/CGE are well-rooted in household agricultural models. This modeling tool seems appropriate to investigate village-level effects of SCTs. But village SAM/CGE models consider villages as closed economies within the national economy. National-level CGE modeling may complement village-level analysis in a number of ways. They may better allow to model flows between villages (e.g., migration). They may also complement the analysis by accounting for indirect effects arising from the need to finance the SCT program domestically. A national-level CGE model may also be more appropriate for the analysis of poverty and distributional issues, especially when combined with disaggregated survey data.

Introduction

The objective of this review is to provide a framework to understand and document the full range of impacts from Social Cash Transfer (SCT) programs. These programs are part of a new policy agenda for the poor promoting a move from universal to targeted programs and from publicly funded service delivery to direct income support to poor households. These programs are currently operating in a large number of countries, many of them in Latin America. This new policy agenda is also remarkable in terms of the importance it places on conducting credible evaluations of the programs to support informed decision-making with regards to scale-up or continuation. Empirical evidence on the impacts on current poverty and long-term welfare are widely discussed. The emerging consensus is that these programs have proved effective (see for instance a review by Fiszbein and Schady 2009). A number of countries in Sub-Saharan Africa are now launching similar interventions. The framework is meant to help provide a basis for analyzing a number of these programs. Specifically, this report is intended to inform the design and data collection for evaluations conducted under the larger research programme "From protection to production: the role of social cash transfers in fostering broad-based economic development", and which include, among others: the Orphan and Vulnerable Children Social Transfer in Kenya, the Social Cash Transfer in Malawi, and the Food Subsidy Program (Programa de Subsídios de Alimentos) in Mozambique.

Welfare programs in Sub-Saharan countries are currently targeted at different population groups (the elderly, households caring for orphaned and vulnerable children and labor-constrained households) but most of these populations share a commonality: they are extremely poor. Yet, these groups are not completely segregated from the rest of the population and often form part of the local economy.

Through market and non-market interactions, benefits from SCT programs may locally trickle down to the rest of the population. If this is the case, impacts recorded for beneficiaries are only a part of the overall effect of SCTs. Why should this matter? Documenting spillover effects is crucial in understanding the contribution of SCTs to poverty reduction and conducting a thorough cost-effectiveness analysis.

Our objective with this review is three-fold: (1) discussing channels through which local economy effects are generated; (2) reviewing findings on local economy effects from the existing literature; (3) discussing data requirements and methods to further strengthen analysis of local economy impacts.

We present theoretical arguments and find some empirical evidence in support of indirect and equilibrium effects. We argue for the need for better data to allow accounting for these spillover effects using both micro and meso/macro approaches. On the micro side, randomization at the aggregate level of the local economy has proven helpful (Angelucci and de Giorgi 2009). Collection of social network data can also help to identify structural social features that foster/inhibit the propagation of the benefits of SCT programs. In addition to identifying these enabling factors, social networks can be considered as one of the outcomes of interest, likely to be affected by SCT programs. Collection of detailed data on agriculture activities is also required to analyze effects on productive activities. On the meso/micro side, village Social Accounting Matrix (SAM)-based analysis seems to be a useful tool to start capturing local economy impacts. Village SAM also requires the design of new data instruments. Several extensions for improving modeling are proposed.

The rest of the paper is organized in two parts. Part I presents the micro approach for evaluating the effects of SCTs and is divided in three sections. We first discuss the general identification framework for impact assessment of SCT programs. In the second section, we discuss the (expected and estimated) effects on SCTs on the local economy by investigating two important market-based mechanisms, *i.e.*, those generated by changes in labor supply of beneficiaries and those associated with investments in productive activities. In the third section, we discuss community-wide effects generated through informal exchanges. Part II presents the modelling tools for analyzing the impacts of SCTs at the meso- and macro-economic levels, and is also divided in three sections. We first describe the role of Social Accounting Matrix (SAM) in economy-wide and village modeling, and examines the relevance of SAM and village SAM multiplier models. We then discuss CGE, village general-equilibrium modeling techniques and their main extensions. Finally, we review recent work that attempts to bring together micro-simulation and CGE to perform poverty and distributional analyses, and discusses the suitability of these approaches for evaluating SCT programs. In the last section, we conclude and discuss the possibility of reconciling the micro empirical approach that yields reduced-form effects to the macro structural approach which is mostly simulation based.

Part I. Micro approach for evaluating of the effects of SCTs

1. General framework

Experimental and non-experimental techniques based on the construction of a counterfactual are commonly employed to identify effects of SCTs. A valid counterfactual represents the situation that would have experienced program beneficiaries had they not participated in the program. These methods help produce reliable estimates of program effects, free of confounding effects.[1]

Spillovers from beneficiaries to non-beneficiaries may complicate matters and lead to a violation of the internal validity of estimates of program impacts. The validity of experimental and non-experimental estimators relies on the assumption that the comparison or control group units are not affected by the program (an assumption referred to as the Stable Unit Treatment Value Assumption or SUTVA in the statistical literature, see e.g., Rubin 1980), among other hypotheses.

A simple solution is to redefine the unit of randomization in such a way that we avoid contamination. Village-level randomization, as opposed to individual-level randomization, is typically used to produce reliable estimates of the effects of SCT

[1]See Apppendix 1 for a review of the various approaches to impact evaluation.

programs.[2] The absence of interaction between treated and non-treated units is simply more plausible when the unit is redefined to be a larger entity (village, classroom, labor market). A village-level randomization design helps when villages are sufficiently distant from each other to avoid contamination between experimental units. This comes at a cost: power analysis to determine the minimum sample size to detect an effect must take into account this new level of aggregation. The sample size (and cost) is likely to be higher for this type of evaluation study than in the setting where randomization is at the individual level. Analysis must account for the intra-cluster correlation to provide consistent estimates of standard-errors (Bertrand, Duflo and Mullainathan 2004).

When randomization takes place at the group level, estimation can be carried out by Generalized Least Square, allowing for a group random effect. This model assumes no heteroscedasticity and common covariance structure. Alternatively, one can relax the assumption of common covariance structure and use a cluster-correlated Huber-White covariance matrix estimator. However, this approach is found to work poorly when the number of clusters is small. This will be often the case when the study is not well powered. In this case, a Fisher randomization test can be performed. This approach does not allow estimating the magnitude of the effects, but can at least allow us to test the null hypothesis of no effect. In practice, it works by simulating a large number of possible random assignments and estimate for each of them a "placebo" impact using the data from the experiment. Then, the test is based on comparing locating where the impact estimate based on the actual assignment vector is placed in the distribution under the null of no effect obtained from the "placebo" assignments.

Well-designed randomized experiments have proved useful as a way to identify program effects *on beneficiaries* (see e.g., Duflo, Glennerster and Kremer 2008). They allow us to simply attribute the difference in average outcomes between program and control units to the program (see Appendix 2 for technical details). The evaluation of the Mexican social cash transfer program (PROGRESA) which is based on a random assignment of villages in two experimental groups continues to stimulate much literature in the development field (see e.g., Parker, Rubalcava and Teruel 2008).

Local economy effects may also be investigated using randomized experiments. (Philipson 2000; Hahn and Hirano 2010). A two-stage randomized experiment that delivers the intervention to a subset of the population in treated villages allows us not only to identify the average effect on the beneficiary population, but also the effect on the non-beneficiaries (see Appendix 2 for technical details). 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. The indirect impact on the non-beneficiaries can be estimated as the difference in average outcomes between the ineligible in treated villages and the ineligible in control villages. The PROGRESA experiment features a two-stage randomization design, which is exploited in a number of

[2] Randomization at the village-level is also operationally simpler to conduct. It may avoid conflicts within communities. When administrative data is not available at the household-level, it may be less onerous in information. In addition, even if eligible households are more numerous than program possible beneficiaries, programs generally rank eligible households by priority rather than randomly allocating benefits to them.

studies to learn about the impact of this program on non-beneficiaries (see, e.g., Angelucci and de Giorgi 2009; Lalive and Cattaneo 2009). In non-experimental settings, local economy level effects can be investigated using partial population intervention designs (Moffitt 2001, Manski 2010).

Questioning the relevance of experimental estimates of program effects to learn about the impacts from the fully scaled-up version of the program is legitimate. Estimates from an evaluation study may not provide much information on the expected impact for the scaled-up version of the intervention. This is very much related to the question of local economy impacts. When scaled-up, the effects of SCTs program on beneficiaries and non-beneficiaries may differ a lot from the effects obtained in the experimental setting because of the mere effect of size. At an experimental or pilot stage, a program may not generate much local economy effects and effects may not spill over to non-beneficiaries. But when the program gets scaled-up to cover the entirety of the population of interest (e.g., all the poor, which is many of Sub-Saharan countries, represent more than half the total population), then market equilibria and non-market arrangements can be expected to change: violation of SUTVA are more likely and general equilibrium effects are more likely. This is because the impact of an intervention does vary with the size of the population enrolled in the program.

Philipson (2000) and Manski (2010) provide a simple framework for discussing how expected impacts from scaling-up the intervention may differ from the experimental effects. Let $\mu(s) = E(Y \mid s)$ be the expected outcome (Y) when the program is scaled-up to a share s of the population. In the experimental setting, where a share s of the population is assigned to the treatment group ($d = 1$) and a share ($1 - s$) to the control group ($d = 0$), we can define $\mu_d(s) = E(Y \mid d, s)$. When the program is scaled-up, the mean outcome is thus a weighted average of the two outcomes provided in the experimental setting, where weights are simply the share of the population benefitting from the fully scaled-up version of the program. We thus have:

$$\mu(s) = s\ \mu_1(s)\ +\ (1 - s)\ \mu_0(s).$$

Under SUTVA we would expect $d\mu_d(s)/ds=0$, and the marginal effect of raising the share of the treated by 1 percent to be equal to the treatment effect:

$$\frac{d\mu(s)}{ds} = \mu_1 - \mu_0.$$

But the potential outcome of an individual in a given treatment may depend on the treatment status of other individuals, thus violating SUTVA. Then,

$$\frac{d\mu(s)}{ds} = [\mu_1 - \mu_0]\ +\ [s\ \frac{d\mu_1(s)}{ds}\ + (1 - s)\ \frac{d\mu_0(s)}{ds}\ ].$$

The experimental effects could thus understate or overstate the economy wide effects when implementing the program on a large scale.

To sum up, the modern econometric framework for impact evaluation based on the definition of a counterfactual is a useful tool to start comprehending the full impacts of a SCT intervention. If spillover effects are expected, the impact evaluation study can be designed to consider the unit of analysis at the level of aggregation for which SUTVA is satisfied. A two-stage design where only a subset of the population is treated may offer the opportunity to measure spillover effects at the local economy level. Adapting the standard framework for impact evaluation can then help to start accounting for the full range of impacts of these programs. However, this micro approach may still overstate or understate the economy wide effect of a fully scaled-up program.

2. Local economy impacts through market exchanges

In this section, we investigate the possible channels through which local economy impacts may be generated and review the evidence from microeconometric studies to assess these predictions. We focus on two possible mechanisms: (1) changes in local labor markets, (2) changes in investment in agricultural activities. Both channels are likely to be crucial for the income-generating process of the poor.

a. Labor supply channel

We can expect social cash transfers (SCTs) to directly impact beneficiaries' labor supply. In addition, second-round effects on beneficiaries and non-beneficiaries may arise from changes in wage rates at the local economy level.

Just like the demand for goods may be expected to rise as a result of the cash influx, labor supply can be expected to decrease as a result of the pure income effect on the demand for leisure of beneficiary household members. When households are required to maintain school-age children in school, conditionalities could reinforce the pure income effect and lead to a stronger decline in child labor. This is because school attendance requirements embedded in some cash transfer programs may constrain children's time allocation and lower their opportunity cost of time. A lower opportunity cost of time, in turn, may generate its own income and substitution effects, which would reinforce the negative effect of SCTs on child labor (Skoufias and Parker 2000). Conditionalities may further affect labor supply of adults in two opposite ways: if it frees time previously allocated to childcare, reducing their cost of time (Alzua, Cruces and Ripani 2010); it reduces children's contribution to household income, and may offset the disincentive effects on labor supply of adults.

If labor demand is fixed at the local level, then a change in wage rates may be expected. Through the change in wage rates, non-beneficiaries may, in turn, be affected by the program. Whether the new equilibrium on the local labor market is characterized with a higher or lower quantity of labor supplied and wage depends on labor supply elasticities of beneficiaries and non-beneficiaries, and is ultimately an empirical question.

Evaluation techniques based on the construction of a counterfactual can help identify the effect of SCTs on labor supply. Comparing mean participation and hours worked of SCTs beneficiaries and would-be beneficiaries can provide an unbiased estimate of the effect of SCTs on labor supply for SCTs beneficiaries. Similarly, one can compare labor supply of people in the same labor markets as beneficiaries to those of similar individuals who belong to labor markets where no one benefits from SCTs. This comparison of means would identify spillover effects to the ineligible individuals through the workings of the local labor market.

However, it is important to note that the micro approach can only inform us on the effects of the program on *equilibrium* prices and quantities on the child and adult labor markets and not on their effect on the whole schedule of labor demand and supply. Browning (1971) provides an early critique, while Deaton (2009) and Imbens (2010) are recent contributions on this question. Experimental and non-experimental approaches allow us to identify *reduced-form* effects but do not allow us to distinguish first-round direct effects from second-round indirect effects through markets. Moreover, the estimated effects can be consistent with the theoretical discussion on the workings of the labor market. But they can also be consistent with effects through non-market interactions, e.g. neighborhood, peer group or herd effects. If spillover effects are found, then SCT programs generate local economy effects which are consistent with market and/*or* non-market interactions. Finally, and as discussed in the previous section, effects measured in pilot settings may understate or overstate effects from a *fully scaled-up* program.

From studies employing a valid counterfactual, we find little evidence that SCTs reduce adult work (Parker and Skoufias, 2000; Galasso 2006; Skoufias and di Maro, 2006; Edmonds and Schady, 2008; Amarante and Vigorito 2010; Alzua, Cruces and Ripani 2010), except in Nicaragua (Maluccio and Flores, 2005) where Red de Proteccion Social results in a reduction in labor supply (coming mainly from a reduction in the total number of hours worked in agricultural activities). Bolsa Familia is found to have had no effect at the extensive margin and a small positive effect at the intensive margin on women workers, those unpaid and in the informal sector (Teixeira, 2010).

Doran (2011) actually finds that adult labor participation and adult labor wages on the daily labor market both increase as a result of PROGRESA. He also finds that this program leads to a new equilibrium characterized by a lower quantity of child labor supplied and a higher child wage rate, as well as an increase in both quantity and wages of adult daily work. This reduction in child labor is also found in Skoufias and Parker (2000).

Empirical findings on the impacts of SCT programs on labor supply and wages of adult and children are thus mixed at best. Mixed results suggest that local initial conditions in which these programs operate, as well as their precise schedule of benefits, matter. This, in turn, indicates that exploring the heterogeneity in program impacts may reveal why SCTs sometimes result in a change in equilibrium prices and quantities on the child and adult labor markets, and other times do not (Djebbari and Smith 2008).

b. Productive activities

Cash transfers may relax liquidity constraints enabling poor rural households to invest in productive activities. SCTs may also increase production as households become more able to avoid detrimental risk-coping strategies.

When beneficiaries invest in productive activities, we may expect a *multiplier effect* from "putting the cash transfer money to work". There is a multiplier effect when the marginal propensity to consume out of transfers is greater than 1. Sadoulet, de Janvry and Davis (2001) are the first to empirically investigate the multiplier effect of a cash transfer program, the PROCAMPO transfer to farmers in Mexico that was introduced in the mid-1990s. This study is also original in that it considers a fully-scaled intervention rather than a pilot program. Pre-intervention data allows correcting for unobserved effects that may have explained cropping patterns and the receipt of the benefits. Variability in transfers allows estimating the marginal effect of transfers. The authors find that a one peso increase in transfer results in 1.5-2.6 increase in income.

What is the evidence on the effects of SCTs on productive activities? Using data from PROGRESA, Gertler, Martinez and Rubio-Codina (2006) find an increase in land use and livestock ownership. They find that for each peso transferred, beneficiary households consume 88 cents directly, and invest the rest. This means a 1.8 cent increase in consumption for each peso of transfers received. Using the same data, Todd, Winters and Hertz (2010) also find an increase in variable inputs for crop production per hectare. Interestingly, they find that changes in agricultural activities are larger for small landholders and the landless.

Bianchi and Bobba (2010) find that, besides relaxing current liquidity constraints, the PROGRESA program promotes entrepreneurship by helping beneficiaries to self-insure against risk. They thus focus on the impact of PROGRESA at the extensive margin and cross-household variation in the timing of these transfers. Their outcome of interest is occupational choice and they contrast the effect of current transfers and the effect of transfers expected for the future. They find that the choice of becoming an entrepreneur (mainly becoming self-employed) is more responsive to future transfers than to current transfers.

Based on data for the Nicaraguan RPS social cash transfer program, Maluccio (2010) finds little effect on investment and no multiplier effect. Given the difference in contexts and program implementation details between the Mexican and Nicaraguan cases,[3] these findings point to the importance of investigating the variation in impacts. Another potential explanation for the difference in findings resides more simply in the difference in terms of the data from the agriculture modules that are collected for these studies. This points to the need to harmonize the quality of data collected to describe agricultural activities and outcomes.

Looking at evidence on the effects of other sources of income on productive activities may also be interesting. Davis, Carletto and Winters (2010) compare the effects of

[3] For instance, RPS was only meant to last for 3 years, whereas Progresa's benefits may have lasted longer.

private transfers (remittances from migrants) and those of public transfers. De Mel, McKenzie and Woodruff (2008) design a field experiment that randomly assigns cash grants to micro-entrepreneurs in Sri Lanka in order to estimate returns to capital. They find evidence that returns are higher than market interest rates. Interestingly, they find that returns are higher the more liquidity-constrained entrepreneurs are.

Pfeiffer, López-Feldman, and Taylor (2009) present a theoretical discussion of the expected effects of non-farm income. If markets are perfect and production and consumption decisions are separable, increase in off-farm income should lead to an increase in leisure and in consumption of normal goods. But when credit markets are imperfect, off-farm income may be expected to increase input use and production. In addition, if labor markets are imperfect, loss of labor to off-farm may reduce production when there are no perfect substitutes to the off-farm laborer in the family. Using data from Mexico, the authors find that production and input use differ between households with and without access to off-farm income. More specifically, they find a negative effect of off-farm income on production and a positive effect on input use.

But, it is important to note that off-farm income is different from SCT benefits in many aspects. First, for the household to get off-farm income, it is likely that at least one household member works less on the farm. So it is less family labor for the household. There is no such direct mechanical effect of SCTs, though the presence of conditionalities in SCT programs may result in a reduction in child labor on the family farm. Second, off-farm income is not received with the same regularity as SCT income.

If beneficiary households are better able to avoid detrimental risk-coping strategies, they may also be better able to invest in productive activities. De Janvry, Finan, Sadoulet and Vakis (2006) investigate a related question using PROGRESA data but on children's enrollment to school. A similar approach could be applied to examine whether the program may help farmers to buffer the effects of shocks on agricultural investments. This could simply be done by adding an interaction term between shock variables and the treatment variable in an equation describing the effect of the program on livestock/agricultural equipment ownership.[4]

In summary, there are a number of potential channels through which SCT programs may affect labor market outcomes and productive activities at the local economy level. The evidence on impacts on labor market outcomes is mixed and more work is needed to understand the heterogeneity in response. Assessment of impacts of SCT programs on productive activities is scarce. This is partly because promoting agriculture has not been the primary objective of most SCT to date, and thus limited effort was made to collect precise information on productive activities (input, prices, assets, production). Collecting better data should help fill this gap. In terms of methodology, even the best-designed evaluation studies that collect information on both beneficiaries and non-beneficiaries can only produce reduced-form estimates of the impacts on market equilibrium outcomes. Furthermore, the external validity of these estimates for understanding impacts of a fully scaled-up program may be limited.

[4] This is another instance where looking at the heterogeneity of impacts may provide clues on how the program works in generating its effects.

3. *Local economy effects through informal exchanges*

Beneficiaries and non-beneficiaries may not only interact through local market interactions, but also via informal arrangements. As a result of SCT programs, households may be more likely to enter/exit informal exchanges, and inter-household transfers may also be affected. It is thus interesting to investigate the effects on informal exchanges at the extensive and intensive margin.

a. Do beneficiaries disengage from risk-sharing arrangements?

If SCTs programs act as a safety net (de Janvry, Finan, Sadoulet and Vakis, 2006), they may undermine existing informal insurance schemes. The presence of a SCT program may increase the value of autarky relative to being in the informal insurance scheme and thus affect the degree of risk-sharing (Dercon and Krishnan 2003).

Albarran and Attanasio (2003) propose a model of risk sharing with imperfect enforceability. In this model, SCTs reduce the amount of risk-sharing in equilibrium. The model yields two testable implications when a SCT program is introduced: (1) a reduction in private transfers, (2) a greater reduction in villages with lower income variability. The authors find evidence that the PROGRESA cash transfer does crowd out private transfers and that this effect is stronger in villages with lower variance in income. These results are consistent with the expected effects from the model of risk-sharing with imperfect enforceability.

This hypothesis – disengagement from existing networks of informal exchanges – can also be tested directly. This would require collecting detailed data on the existing links. A dyadic model would allow modeling formation of links (see e.g., Fafchamps and Gubert 2007 and technical details in Appendix 3). This framework can help to test if treatment status influences links formation and to evaluate this effect.

b. Do beneficiaries share transfers with other households? with those who are closer socially?

Several papers find that SCTs benefit the local economy at large, not only beneficiaries. This can be readily done using experimental data where the unit of randomization is the village and villages include both the eligible and the non-eligible. Such a design allows us to attribute the difference in average outcome between the non-eligible in treated communities and the non-eligible in control communities to the program.

Spillover effects to non-beneficiaries are investigated in a paper by Angelucci and de Giorgi (2010) using the PROGRESA experimental design. This study is particularly interesting in that it also aims at understanding why ineligible households in treated villages have higher consumption than ineligible households in control villages.

In addition to this indirect effect on consumption, the authors find that, on average, ineligible households in treated villages receive more transfers, borrow more and reduce precautionary savings compared to ineligible households in control villages.

They propose a model of perfect risk-sharing in which an increase in aggregate income due to cash transfer is expected to generate an increase in income for all risk-sharing partners, including non-poor households. They also rule out that higher consumption is due to higher labor earnings, higher income from sales from higher prices or higher demand in treated communities compared to control communities.

In terms of methodology, there are two potentially interesting development within this framework. One would be to draw up a 2-stage randomization design where communities would first randomly be assigned to treated and control communities and then beneficiary households would be randomly selected within treated communities (Hahn and Hirano, 2010).

Such a design would allow us to identify the overall effect on beneficiaries from comparing beneficiary households to their counterparts in control communities. It would also allow us to identify the indirect (second-round) effect on beneficiaries from comparing non-beneficiary households in treated communities to their counterparts in control communities. This is because beneficiaries in treated communities are similar in all respects to non-beneficiaries in treated communities apart from the fact that the former benefitted from the program while the latter did not. Therefore, the indirect effect on non-beneficiaries is equal to the indirect effect on beneficiaries. In this setting, the direct effect on beneficiaries may be estimated as the difference between the overall effect and the indirect effect.

In addition, this double randomization setting allows estimating the error associated with projections that assume externalities away and extrapolating the effect of a program from an experiment. The design also has a number of drawbacks: as noted in footnote 2, the design may be difficult to implement in close-knit communities and program managers tend to prefer to use other rules for assigning benefits at the local level.

In the context of an intervention in the U.S. that provides cash transfers conditional on exercising, Babcock and Hartman (2011) find evidence that beneficiaries are affected by other beneficiaries among their peers. They do not find evidence of an indirect effect on non-beneficiaries. They also find that if all the target population had been treated, the estimated effect would have been 64% larger than the one obtained from comparing mean outcomes in the two experimental groups.

The other methodological development could come from collecting better data as in Bandiera, Burgess, Gulesci, Rasul (2009). Finer data on social interaction networks may help to test whether peers help to insure against risk (Helmers and Patnam 2010) and identify peer group effects (Bramoullé, Djebbari, Fortin, 2009). They may also be useful for addressing two broad sets of questions: (1) Can SCTs help reduce social exclusion of the poor? More generally, how are local market and non-market institutions affected by SCTs? (2) Do better connected households benefit more from SCTs than the less-connected ones? And, how do impacts depend on local institutions?

Collecting social network data may be challenging. Both sampling design and questionnaire design issues must be addressed. Ideally, sampling design would allow us to collect data on social networks by surveying every household in treated and control communities (if, for the behavior that is examined, important links are those outside the communities, then even this approach would only provide a partial understanding of the impacts of SCTs).

Alternatively, targeting surveys may provide a timely opportunity to collect information on all households in the study villages, while social network modules may only be applied to a sample of households.[5] Using modelling, one may then predict unobserved links between households using the sample of links and the information collected on all village households.

Survey design must clearly specify the type of relationships that matter for the behavior under consideration (e.g., informal insurance network; family network). Bandiera, Burgess, Gulesci, Rasul (2009) collect data on these types of networks. They ask each respondent to name insurance network members: two households are sharing risk if they exchange (borrow/lend) food items, transfer cash/in-kind transfers or provide assistance to each other in times of crisis.

Networks whose members are linked through family relations identify family networks. Family networks are likely to be fixed in the short-term and spillover effects of SCTs on non-beneficiaries are likely to be greater when non-beneficiaries are related to beneficiaries through family links, as found by Angelucci, de Giorgi, Rangel, Rasul (2010) in the case of PROGRESA.

c. Social network effects

Social interaction effects may reflect a broad array of behavior, from conformism and social pressure, to information-sharing and social learning. Some studies examine the extent to which SCTs may generate peer effects on enrollment of children to school (e.g., Bobba and Gignoux 2010, Lalive and Cattaneo 2009).

Because they may generate a social multiplier, social interactions are typically found to amplify the direct effect of SCTs at the local level. Given that SCTs may have an impact on agricultural investments of beneficiaries, it may be interesting to examine whether this direct effect can be amplified through a social multiplier at the village level, or within the networks of interactions between farmers (e.g., neighbors, those farming the same crops).

Likely social processes include information sharing and social learning. Conformism may also matter. In order to be able to tell which on these mechanisms is at play, we would ideally need information on whether people exchange information on the topics we want to study, whether they learn from each other (e.g. on what to do to start a business), on whether they compare their experiences in ways to keep up with the Joneses.

Social network modules that focus on a very specific type of interaction (risk-sharing in the case so far for Lesotho) may be too restrictive. One (costly) approach consists of collecting data on as many networks of interaction as there are types of interactions (networks of information, networks of learning, networks of conformism). The alternative (less costly) is to add a social network section gathering data on your

[5] Personal communication with Arun Chandrasekhar.

friends/those with whom you interact on a regular basis. Analyzing the structure of these networks should give us some clues on the type of social processes that are more likely to be present (more on that by the end of the section).

Assuming that we can identify the “right” network of interactions, identifying the social multiplier raises a number of challenges (Manski 1993). Agricultural investments are often highly correlated within groups of farmers. Yet, this correlation may be due to various causes besides peer group effects. We expect, for instance, that farmers with better unobserved managerial abilities may interact more with each other. This selection effect is clearly not a consequence of the interactions between farmers. More able farmers may also have better land, and this would explain why these farmers make higher investments. When the quality of land is not observed, we may incorrectly credit this effect to peer interactions.

Various strategies have been developed to address these problems of identification. A structural model of peer effects in networks, which allows us to exploit the heterogeneity in the structure of interaction, can help identify peer effects if the unobserved correlated effects are fixed at the level of the network (Bramoullé, Djebbari and Fortin 2009). When the last assumption is not plausible, random assignment of a treatment that directly affects investment decisions helps identify the social multiplier associated with the decisions to invest in the group of peers. This would also allow us to distinguish between the direct effect of the cash transfer on investment decisions and the indirect effect from peer interactions.

A more reduced-form approach may also provide useful information regarding the underlying mechanisms through which the social multiplier is generated. A network is a NxN dimensional object describing all the links in a population of size N. But some network features are more relevant to describe certain mechanisms than others (e.g., clustering for risk sharing; eigenvalue distribution for diffusion of information and social learning). In a simple regression framework, one would include an interaction term between the treatment indicator and a village or individual-specific network parameter. This analysis would inform us on the heterogeneity of impacts according to the underlying nature of the social structure, and may also inform us on the mechanism at play.

The previous sections offer a review of micro approaches for studying local economy effects of SCT programs. These approaches prove useful for framing the micro-level behavioral responses, at the individual level and local market level. A few empirical studies provide evidence of impacts beyond those experienced by the residual claimant for the program. Within this framework of analysis, more can be done by *(1)* collecting better data (on agricultural activities, on informal arrangements), *(2)* improving the design of the studies to capture spillover effects on the ineligible (2-stage randomization design), and *(3)* investigating the heterogeneity in impacts to uncover the mechanisms at play. There are, however, a number of limitations related to the approach. Most of the econometric work is reduced-form and only accounts for impacts on equilibrium outcomes. Although of high value in terms of internal validity, experimental estimates obtained at the pilot stage of a SCT program may have limited external validity for understanding impacts of a fully scaled-up program.

Part II. Modeling tools for analyzing the impacts of SCT programs at the meso- and macro-economic levels

Cash transfers affect consumption, income and employment, and are likely to induce behavioral changes, which in turn can generate meso level and economy-wide changes through general equilibrium effects. In assessing the potential of SCT programs for poverty reduction, it is useful to examine the changes brought about by the impact of cash transfers on the level and composition of demand and supply, and to investigate the economic linkages which transmit the impacts of cash transfers from beneficiary households to others in the local or national economy.

The evaluation of the effects of SCT programs on local or national economies can be approached by drawing on analytical frameworks that capture how policy changes and economic shocks affect key macroeconomic balances, and how the repercussions are transmitted to households, particularly via factor employment and incomes.

Social Accounting Matrix (SAM) multiplier and Computable General Equilibrium (CGE) models provide a rigorous and yet practical framework for local and economy-wide analyses, given their ability to link the macro and micro levels and to account for the effects of policy changes on incomes and consumption of different household groups (Robinson and Lofgren, 2005*).*Although these models can provide a comprehensive framework for evaluating economic policy choices, they do not allow us to consider some complex aspects of SCTs or capture the full impact of such transfers on economic activities and agents. Social cash transfers are quite complex programs, given that they aim to reduce current levels of poverty, and to stimulate investment in human capital for sustained poverty decline. Designing a framework that accommodates all the underlying interactions and pathways through which cash transfers affect current and future poverty and inequality levels is a particularly challenging task.

Part II below reviews the various tools presently available and the data required to evaluate the poverty and distributional effects of economic policies at the local and national level, and from an *ex ante* point of view.

This review focuses on village and economy-wide SAM multiplier and CGE models, and explores how these models can be extended to incorporate the various mechanisms and linkages through which cash transfers contribute to poverty and inequality alleviation, particularly in the short term. In particular, this part reviews methods and data required to evaluate the impacts of SCTs on goods and labor markets at the local and national levels; and discusses how the economic linkages among households that transmit the impacts of cash transfers from eligible households to ineligible households can be captured in a general equilibrium setting.

We first describe the role of Social Accounting Matrix (SAM) in economy-wide and village modeling, and examines the relevance of SAM and village SAM multiplier models for the analysis of the contribution of SCT programs to the local and national economy. We then discuss CGE and village general-equilibrium modeling techniques and their relevance as a tool for the evaluation of SCT programs, focusing on the theoretical underpinnings of these models and their treatment of production, demand, and labor markets. This section also considers extensions of these models to capture inter-household linkages and general-equilibrium feedbacks of policy changes. Finally, we review recent work that has attempted to bring together micro-simulation and CGE to perform poverty and distributional analyses, and discusses the suitability of these approaches for evaluating SCT programs.

1. *SAM and SAM-based multiplier models*

SAM multiplier and CGE models are appealing frameworks for evaluating local and economy-wide impacts of policy reforms. For these types of modeling approaches, the SAM provides the statistical underpinnings and the logical framework (Taylor and Adelman, 1996; Robinson and Lofgren, 2005).

The SAM offers a comprehensive representation of the macro and meso economic accounts of a region or nation, which traces out the circular income flow including production activities, commodities, factors, domestic institutions –households, enterprises and government – and the rest of the world. The notable strengths of the SAM lie in its comprehensiveness and flexibility in adapting to different economic structures and institutional settings. SAMs provide a useful starting point for local and economy-wide economic analyses and has proved very helpful for addressing income distribution issues (Round, 2003).

SAMs are generally built to represent the national economy, but they can be constructed at the regional or village levels as well (Taylor and Adelman, 1996). Village SAMs and national SAMs have the same overall conceptual structure, but village SAMs differ in that they are designed to reflect the specific characteristics of village institutions and production activities. The village SAM captures the complex inter-linkages between village production activities, village institutions and the rest of the world which includes the regional and national economies outside the village and the world economy. While national-level SAMs build, in general, on ready-made data sources coming from input-output tables, national accounts, government budgets, and commodity trade data as well as on information on total factor payments, total household income (by income category) etc.; village-level SAMs require more detailed data, which need to be collected through village household surveys and participatory rural appraisal. The challenge for building these SAMs is that a comprehensive knowledge of the economic structure of the village and of the functioning of the markets and institutions is required before setting up and carrying out the survey (Taylor and Adelman, 1996; Davis et al., 2002). To ensure consistency of the village-SAM, it is essential to collect detailed information about the origin and destination of most economic transactions within the village. If the SAM is built on sample surveys which capture only part of the transactions (to the extent that not all households or businesses are included), a potential sample bias could occur, causing unbalanced markets. Tracking transactions outside the village remains another difficulty.

The village SAM typically includes village production activities, village commodities, village factors of production, village institutions, village capital accounts and the rest of the world. The level of disaggregation in a village SAM depends on the objectives of the research and the availability of data. The production activity accounts typically include village farming activities (various crops and livestock for own-consumption and sale), and off-farm activities (services, retail activities etc.). The commodities account captures household consumption structure and product markets. The factor accounts typically include family labor, hired labor, land and physical capital (animal traction, machinery etc.). Village institutions consist of households (disaggregated by income groups, landownership, occupation etc.), government and local administrative structures. It is important to specify as many representative household groups as there are households in the village population to capture the complex linkages among diverse agents that characterize small communities. This involves identifying the heterogeneity of factor endowments or preferences at the level of each representative household (Bourguignon *et al.*, 2003). The village SAM accounts should therefore include disaggregated household and production accounts, in order to be able to trace the process of income generation, its distribution and redistribution across the various household groups, and the structure of production. The capital accounts contain investments and savings. Depending on the level of market development in the village, the rest of the world account includes exchange with the rest of the zone of influence, other regional, national and international markets. The rest of the world account provides transfers to households, buys the village's exports and sells its imports. The rest of the world account can be disaggregated according to the specificity and intensity of the exchanges with the outside markets. A high level of disaggregation rises the issue of handling the imbalances between inflows and outflows of taxes, transfers, factor payments etc.

Estimation of the village SAM thus entails designing the SAM framework, gathering data and expanding the sample results to the universe under study. The design of the SAM should reflect the realities of the local economy as well as the objectives of the research. For example if the focus is SCTs impact it is important to distinguish, in the SAM, beneficiary and non beneficiary households and to capture details on income sources and expenditure destinations in order to be able to establish the spending patterns of all economic actors in the village and to identify the secondary and higher beneficiaries of the program. Village household surveys serve as the basis for constructing the village SAM, but these surveys may be insufficient and need to be complemented by enterprise surveys. Household surveys gather data on farm budgets and consumption mainly and often miss information on the origin and destination of transactions, revenue and expenditures. Also, many businesses within the village or local economy are often under-sampled or missing. It is sometimes necessary to conduct a separate enterprise survey to capture all the transactions within (and outside) the local economy. Depending on the level of diversity of the local economy and the degree of heterogeneity of the of business activities across clusters, different approaches can be adopted to carry out the survey. One way would be to draw up a sample based on the different kinds of activities existing in each cluster, skipping those already captured by the household survey. Stratification ensures the inclusion of all types of activities which may employ individuals within the village. When possible, the questionnaire needs to be designed in a way that suits the village SAM and needs to capture detailed information on the revenue and expenditures of the units included in the sample. The origin and destination of household and business transactions market- by- market (nearby local markets and distant markets) together with the transaction costs need to be gathered. The analysis of the impact of SCTs on the local economy requires accurate information on the spending of the transfers. It is important to collect for each recipient where and from whom each purchase is made, in order to determine which economic actors, apart from the beneficiaries, profit from the program. In all cases, the design of the SAM and the surveys requires a thorough knowledge of the functioning of the local economy.

The SAM framework can serve as a base for multiplier analysis. SAM-based multiplier models, pioneered by Richard Stone, have been an important advance in village, regional and national modeling because they highlight the economic linkages among households that transmit exogenous changes and injections through the local or national economy. The purpose of these models is to assess the multiplier effects of exogenous policy actions on the whole system, in general and on the incomes of households (or socio-economic groups of households), in particular (Round, 2003). The models assume excess capacity, fixed prices and that the responses of economic actors to income changes are strictly proportional to the total level of activity in each account. Therefore, on the production side there is a Leontief-type production technology, and on the expenditure side, marginal expenditure propensities equal average expenditure propensities. Developing a multiplier model requires partitioning the village SAM into endogenous and exogenous accounts. The endogenous accounts are frequently limited to production activities, commodities, factors, and village institutions. They capture the response of the village economic agents to exogenous changes. The government accounts and the accounts of rest of the world are generally considered exogenous. The village capital market can also be treated as exogenous if it is fully integrated with outside capital markets (Taylor and Adelman, 1996). The exogenous accounts may be aggregated into a single account, which records an aggregate set of injections into the system and the leakages from it.

SAM multiplier analysis can be a useful tool in evaluating the effects of SCT programs. It can be used to simulate the impact of cash injections on the level of output and input use of different production activities, and on the incomes of various factors and household groups within a village. The impact of the exogenous injection on the village economy is shaped by the economic linkages between production factors and households which are, in turn, determined by the structural characteristics of the local economy. The size of the multiplier depends upon these characteristics and the ability of the village to retain revenues generated locally within the region. For example, if locally produced goods and services in the village are an important share of households' consumption demand, then increasing household incomes will benefit local producers whose incomes would rise accordingly. The indirect linkage effects will be larger and the leakages out of the village will be smaller (Taylor and Adelman, 1996; Round, 2003; Breisinger et al., 2009).

Social cash transfers are frequently targeted geographically and cover a large proportion of the population in poor communities and remote areas. Cash injections into the local economy from the introduction of SCT programs could stimulate local consumption and production and could also have effects upon employment. Cash transfers are likely to have impacts beyond the direct beneficiaries, and can positively affect non-eligible groups and the local economy through multiplier effects**.** The potential positive external effects of cash transfers on non-beneficiaries might be more important in smaller local economies where household linkages are more pronounced (Barrientos and Scott, 2008). Since villages are partially integrated into regional and national markets, an ever-widening circle of economic actors, inside and outside the village or local community then becomes influenced by the exogenous income increase, even if they do not directly benefit from the transfer (Taylor et al., 2005). Village and regional SAM-based multiplier models can be used to investigate these multiplier and spillover effects of SCTs.

Davis and Davey (2008) use a regional multiplier approach to estimate the total contribution to the regional economy of the Dowa Emergency Cash Transfer (DECT) program, carried out in Dowa, Malawi during the 2006/7 agricultural season. They construct a *Reduced Social Accounting Matrix* which permits to identify which economic actors, apart from the recipients, benefit from the program. The construction of the SAM is challenging as it requires accurate data on the consumers' reported spending of their transfer. Based on reported consumer expenditure patterns, interviews are also conducted with the secondary beneficiaries to evaluate the spending patterns of all relevant actors in the local economy. The analysis enables to follow the exogenous cash injection, resulting from the DECT program, around the local economy and to identify secondary and other higher order beneficiaries.

The simplest version of the SAM multiplier models assumes that the economy's factor resources are unconstrained. So that any increase in demand induced by additional income can be matched by a corresponding increase in output without having any effect on prices. The direct and indirect effects of the exogenous cash injections on total outputs of various economic sectors and the incomes of different factors and socioeconomic household groups can be estimated through the multiplier process and transmitted through the interdependent SAM system. No behavioral change is assumed, meaning that the structural relationships between households and sectors in the economy are not affected by the exogenous changes (Breisinger et al., 2009).

While the hypothesis of unlimited supply may be realistic for some production activities, it is hard to think of all the producing sectors, particularly the agricultural ones, to be without supply constraints. Ignoring these constraints would overstate the impacts of linkage effects. The incorporation of the supply constraints would permit a more realistic evaluation of the multiplier effects. This can be done by partitioning the SAM accounts into supply constrained and unconstrained. The structural constraints in some sectors means that output responses are permitted only in the supply unconstrained sectors. In sectors facing perfectly inelastic supply, an increase in demand by households or firms is generally assumed to be satisfied through a decrease in village exports, or marketed surplus (Yunez-Naude et al. 2006; Ferede, 2009).

SAM multiplier models can also be extended to consider the impacts of cash transfers on human capital. In this case, the human capital accounts, comprising health and education inputs and outputs, need to be considered as endogenous and the other capital accounts as exogenous. Inclusion of human capital helps to evaluate the sustainability of SCT programs in terms of long term poverty reduction as well as to identify the activities that best promote household income, human capital and growth (Ferede, 2009).

The strength of the SAM-based modeling approach comes from tracing out chains of linkages from changes in demand to changes in production, factor incomes, household incomes, and final demands (Robinson and Lofgren, 2005). However, it has a number of shortcomings and is likely to give a distorted picture of the possible impacts of SCTs and policy changes on the village economy. First, it assumes fixed prices. While this assumption is less limiting in a village economy where the transaction prices are generally determined by markets outside the village, the presence of market imperfections may cause village prices to diverge from market prices. In these models, expenditure patterns in beneficiary households convert the higher income, induced by the cash injection, into increased demand for goods produced in the village and for imports. Higher demand for village products stimulates village production creating a new round of income increase for beneficiaries and non beneficiaries with part of this income leaking out of the village in form of import demand (from regional, national or international markets). Resource constraints in household farms result in a less than perfectly elastic supply response in village production. Unless, there is a surplus of village land and labor resources, part of the impact of higher demand translates into higher prices for goods and factors within the village and introduces a series of complex effects into the SCT multiplier. The increase in wages and opportunity costs of other resources dampens the positive impact of SCTs on village production. The increased infusion of income into the village could benefit the village production of non-tradables, while the production of tradables may be dampened. Taking into account the price effects in modeling the impact of SCTs on village economy would therefore greatly change the impacts predicted the SAM multiplier model. The assumption of linear, fixed-proportion technologies is also another limitation (Taylor and Adelman, 1996). A third major limitation is related to the use of SAM-based multipliers in the context of poverty analysis. No matter the level of disaggregation of the SAM accounts, multiplier effects are confined to assessing the income effects of socio-economic household groups. The intra-group income distributions are not generated directly. This shortcoming can be overcome by using the multiplier decomposition method proposed by Thorbecke and Jung (1996). This approach can be used to capture the various economic mechanisms and linkages through which the exogenous change contributes to poverty alleviation within household groups (Round, 2003).

2. *CGE and Village general equilibrium models*

CGE models can help address some of the previous modeling caveats. These models portray the operation of an economy through mathematical representation of the behavior of agents in different commodity and factor markets, and allow specification of the economy with the desired level of detail with regards to sectors and factor markets linkages and in relation to the behavior of economic agents. They preserve the advantages of SAM-based models and overcome their principal limitations by adding the supply side, endogenizing commodity and factor prices and incorporating nonlinearities and resource constraints (Taylor and Adelman, 1996).

Economy-wide CGE models are commonly employed to capture the aggregate impacts of economic shocks or policy changes on all economic activities, prices, factor markets, households and other institutional groups. National CGE models are constructed from the top down, using ready-made aggregate data. The parameter and elasticity values that feed the equations of the model cannot be estimated econometrically and are generally calibrated from the SAM or borrowed from other studies for countries with a similar structure to that of the economy being modeled. The results of these models depend critically on these parameters and assumed functions which can barely be tested individually, let alone in combination. Another drawback is that economy-wide CGE models generally neglect the complex interactions among heterogeneous households. Households in small communities and villages are heterogeneous in terms of their income sources, expenditure patterns and factor ownership. Members of these households often have multiple occupations. Also, different households are frequently engaged in the same activity but may use different technologies. The linkages among households in these communities are complex and may be important in transmitting the impacts of policy reforms to other households and firms in the local economy (Taylor et al., 1999). Using aggregate CGE models tends to blur these microeconomic interactions that are critical for the analysis of the local economy effects of policy changes.

Village-wide modeling techniques presented in Taylor and Adelman (1996) and Taylor *et al.* (1999) allow for heterogeneous interactions among economic actors in a general-equilibrium context and offer a robust framework for the analysis of the direct and indirect impacts of policy changes on the local economy. Village-wide CGE models integrate micro-models of household behavior into a village general equilibrium framework, exploiting the advantages of each method. This approach occupy a middle ground between household-farm models and aggregate national CGE models and has much appeal because it enables to capture the complex production and expenditure linkages among heterogeneous households as well as the general equilibrium feedbacks when exploring the impacts of policy reforms. Village-wide models are built from the bottom up using micro data from surveys on farm household and non-farm businesses, living standards and living conditions, village censuses etc. The model parameters and elasticities are generally calibrated from the village SAM or estimated econometrically on the basis of a minimal number of assumptions.

Village-wide models can be particularly suitable for assessing local economy effects of SCT programs. Cash injections into the local economy have complex direct and indirect effects on income and production. The direct effects of cash payments on the expenditure and behavior of the recipients are transmitted inside and outside of the local economy through interactions across households in commodity and factor markets. These effects may be even more important as the household linkages are strong and the village integrated into regional and national markets. While the direct effects of SCT programs can be assessed using partial equilibrium or micro-modeling approaches, evaluating the second and higher-round feedbacks require an economy-wide modeling approach which nests micro-models of household behavior within CGEs for larger economic spaces in order to capture both the heterogeneity of households and the diversity of activities in which these households are involved (Dyer et al., 2006).

A prototype village-wide economic model incorporates all the flows in the village SAM and consists of several blocks of equations for production, income, expenditure, prices and general equilibrium closures.[6] Its key actors are households and producers. Producers maximize net income from the production activities subject to available technology, market or shadow prices for output, intermediate inputs and factors of production, whereas households maximize utility defined on own produced and purchased consumption commodities and leisure. Prices and wages can be assumed to be exogenous when the village economy is closely integrated with factor and goods markets outside the village. Local demand variations do not influence local production and only affect marketed surplus. Local production and consumption decisions can however also be guided by endogenous prices in the presence of high transaction costs or missing markets. Local wage is endogenous in the presence of a local labor market. Endogenous local prices are incorporated into the model through general-equilibrium constraints for village non-tradables. The production technology in the model is homogeneous of degree zero in prices. Demand for intermediate inputs is determined through the use of fixed input–output coefficients. Physical capital and land inputs are generally considered fixed across activities in the short run, while labor is assumed to be variable. The total capital and land demands, across village activities, equal total village endowments of these factors. The key parameters needed to estimate the model, are the production function coefficients, the distribution of village value added across household groups, and the household expenditure shares. All the variables and parameters of the model are calibrated to reproduce the village SAM as the base solution of the model. The solution to village CGE gives quantities and prices for all commodities and factors, incomes for all household groups, savings, and net exports for tradable commodities.

The ability of the village CGE models to simulate the village wide effects depend on their general equilibrium closures. These include local market clearing conditions for factors and goods, a village savings–investment balance, and a village trade balance equation. Various closure equations can be specified depending on the level of integration with outside markets and on the nature of factor markets. If the village is perfectly integrated with regional/national factor and goods markets, wages and commodity prices are exogenous to the village and the supply of factor and goods are perfectly elastic. Under this assumption of perfect markets, where the village economy is price taker and faces no transaction costs, the general equilibrium closure determines the village (or household) net marketed surplus as the difference between output supply and consumption demand for produced goods, and labor supply minus labor demand or net wage-labor supply for labor. Household farms and enterprises within the local economy face resource constraints (fixed land and capital). Family labor, considered a key input in rural village economies, cannot be assumed tradable, and family wage is valued by shadow price. Under these assumptions, an exogenous increase in household incomes, affects consumption but does not influence production in the village, the first-order conditions for profit maximization does not change. Income changes affect trade with outside markets and net village marketed surplus and do not generate any income multipliers within the local economy as the potential multiplier leaks out through trade with outside markets.

6 More details on the modeling approach can be found in Adelman and Taylor (1996) and Taylor et al. (1999).

If the access to labor markets is geographically concentrated and segmented by family networking and if transacting in extra-village labor markets is costly, these market imperfections will induce wage rigidities and even local wages. Local wage for non family labor can no longer be considered exogenous. Also, some producers (or farm households) may be isolated from the regional (or national side) for their products by high transaction costs for their products related to poor transportation systems, or insufficient marketing infrastructures and communication (Taylor et al. 1999, 2005). Given the large heterogeneity of rural communities it is unlikely that all producers would face the same constraints in accessing outside markets. The result is a heterogeneous local rural economy in which commercial producers facing low transaction costs interact, in factor and commodity markets, with subsistence producers who are isolated from outside selected markets by high transaction costs. Prices for labor and non tradable goods are endogenous and determined by interaction of supply and demand in village markets. For tradables, prices are exogenously determined by markets outside the village and marketed surplus is endogenous, while for non tradables prices are determined endogenously through local market clearing conditions. An increase in household incomes from cash payment leads to an increase in demands for local non-tradables as well as tradable goods. Excess demand for tradable goods is satisfied through imports from nearby or national markets outside the local economy. On the other hand, higher village demand for non-tradable commodities drives-up local output and factor prices and thus provokes an expansion in activities that supply local demand, generating an increase in their demand for intermediate and factor inputs and stimulating new rounds of production and income increases. Changes in demand induced by cash transfers are transmitted to other production activities in the village economy through local prices. The determination of the potential impacts of SCTs on the village economy is fairly complex because different households, in diverse regions and market settings, will be affected in different ways by the policy change.

The economic interactions among heterogeneous households are captured in village-wide models by incorporating differences in the source and size of households' income and in the marginal propensity of their consumption. The most common approach relies on specifying a number of representative household groups for the village population. All groups participate in local land and labor markets, buying and selling factors from one another in order to carry out their various production activities. Households are generally assumed to share the same form of preferences and technologies, but the parameters in the utility and production functions may vary across groups. A policy change affecting one group is transmitted to the other groups via factor and goods markets. The use of these models for the evaluation of the impacts of SCT programs requires putting the targeted households into a separate group. An important shortcoming of this approach is that the analysis of the distributional consequences of these programs focuses on changes in inequality between representative household groups. The within-group inequality can be captured by assuming that income distribution within each household group follows an exogenous law, which may lead to an underestimate of the true distributional impacts. This is quite limiting given that the main objectives of SCT programs are to improve income distribution.

The village-wide models can be extended to accommodate the full heterogeneity of households. These disaggregated models integrate individual household models into a village general equilibrium framework by incorporating all households from microsurveys (see Taylor et al., 2005 and Dyer et al., 2006). The disaggregated micro economy-wide models take into account the village-specific diversity of production activities, technologies and demand patterns, as well as the market structures that transmit policy changes among individual households. They consider the different activities in which individual households can be engaged and allow for the specification of different technologies for dissimilar households engaged in the same activity. This approach can be used to simulate impacts of SCT programs on each household as well as the aggregate impacts on the whole local economy. It helps us to understand micro responses and interactions among economic agents within the local economy in a general-equilibrium context, and to explore how household heterogeneity combines with market equilibrium mechanisms to produce more or less poverty and inequality as a consequence of policy changes. These models would be appropriate for exploring the complex poverty and income distribution effects of SCTs in heterogeneous rural communities.

The procedure for estimating such models consists in constructing separate SAMs for each household or small homogenous household groups. These models are relatively intensive in data and require detailed information, generally obtained from micro surveys, on socio-demographic characteristics, production, income sources, and expenses destination for each household in the model. Missing data such as information on family inputs and value-added can be estimated econometrically. The construction of these models entails estimating models of household farm activity for each household (or household group) integrated in the model. This includes estimating production functions for each production activity in the village, and expenditure functions for each household group and expenditure category (Taylor, 1995).

Dyer and Taylor (2004) use a village general-equilibrium modeling approach to investigate the responses of Mexican agricultural households to two alternative cash-transfer programs, namely PROCAMPO and PROGRESA. They simulate a cash payment of $161 to all household heads in the village of Zoatecpan situated in East Central Mexico. This simulation resembles the existing PROGRESA welfare program. The model allows us to explore the impact of the policy change on every household in the sample. The findings reveal significant additional benefits of cash transfers besides achieving income increase of the targeted population. The program seems to play an important role in sustaining household expenditure and ameliorating local economic contraction after the decline of commercial maize agriculture following NAFTA.

Taylor *et al.* (2005) use a disaggregated rural economy-wide model to explore the impacts on West-Central Mexico of removing PROGRESA subsidies to the rural poor. Direct income payments under the PROGRESA program compensated rural households for negative income effects of lower maize prices. The findings indicate that PROGRESA has a progressive effect on rural incomes, and eliminating the payments would induce a decrease of incomes of landless and small landholding households. This would also negatively affect demand in subsistence households. Wages and land rents appear to slightly decrease without PROGRESA, while migration remains unchanged.

Despite the usefulness of these models in assessing the consumption, employment and local general equilibrium effects of policy changes, in general and cash transfers in particular; they are limited in their ability to explore the SCT impacts on investment in productive activities, health, education outcomes and long-term poverty. Another shortcoming comes from the inability of these models to capture efficiency and distributional impacts arising from the financing strategy of SCT programs. Also, the basic tenet of these models is that local outcomes of changes in national policies depend upon heterogeneous local conditions. Therefore, findings in the local economy can hardly be generalized to the national economy. A more appropriate framework for investigating impacts of SCT programs at the country level may be that of national CGE models.

3. Combining CGE and micro-simulation modeling

Social cash transfers vary a great deal in scope and coverage. Some programs are targeted to small communities or villages, and others work at regional and national levels. Large SCT programs are likely to generate nationwide economic changes through general equilibrium effects. Evaluating the effects of these programs on national employment, consumption, poverty and inequality entails the use of countrywide CGE models. The macroeconomic and distributional outcomes of CGE models depend, among other things, on the level of disaggregation used for household types, factors, and production activities. The treatment of households in these models is crucial for the analysis of poverty and distributional issues.

Conventional CGE models distinguish a limited number of representative household groups. While simple, this approach fails to capture heterogeneity within household types and may result in misleading poverty and inequality outcomes. The distribution of income within household groups is assumed as exogenous, and distributional changes arise only as a result of redistribution between groups. On the other hand, poverty variation can be due only to within-group changes (Bourguignon et al., 2003; Davies, 2009).

A rather simple way to address this shortcoming is to combine the results from the general equilibrium model with the information in household surveys to provide a more accurate evaluation of the inequality and poverty consequences of cash transfers. The method consists in estimating the proportionate changes in representative household incomes and commodity prices resulting from simulation change (i.e., income for beneficiary households) using the CGE models. These changes are then mapped onto the individual households to compute poverty and inequality measures.

Coady and Harris (2004) have used a similar approach to evaluate transfer programs. The authors combined results from a CGE model with disaggregated household data for evaluating the welfare impacts of domestically financed transfer programs. They illustrate their approach using data from Mexico to evaluate the recent shift of Mexico's poverty alleviation strategy from universal food subsidies to targeted cash transfers. Their CGE model is calibrated to a SAM for Mexico compiled using the 1996 household income and expenditure survey. The model includes five regions (four rural regions and one urban) and three income groups. Households are grouped into poor, medium or rich according to the income tercile which they belong to. This disaggregation leads to fifteen household groups and helps to capture important regional diversity in terms of income shares and expenditure patterns. The authors employ a two step procedure to evaluate the direct and indirect distributional effects of cash transfer programs to capture the redistribution, reallocative, and distortionary effects of the programs. The first step consists of simulating an increase of the incomes of poor rural households, resulting from the cash transfer, and estimating direct and indirect income changes, arising from factor price changes, using the CGE model and considering various budget closures. Then, the direct transfers and the indirect effects simulated in the CGE (the proportional income and commodity price changes) are applied to the household survey data, where households are mapped to one of the representative households in the CGE.

Few studies have investigated the general equilibrium effects of SCT program. The study conducted by Coady and Harris helps fill some of the gap and provides insightful results on the general equilibrium welfare impacts associated with domestically financed programs and shows the importance of well-targeted direct transfer schemes for inequality alleviation. This stresses the importance of general equilibrium frameworks in the analysis of the welfare consequences of redesigning social safety nets. The main weakness of this analysis is related to the difficulty of capturing the heterogeneity of households and behavioral responses of beneficiaries to the policy change.

One convenient way to deal with the distributional problems in CGE modeling is to include as many representative households as there are households in the representative national survey. This entails identifying the heterogeneity of factor endowments or preferences at the single household level, which may complicate the writing and estimation of the model (Bourguignon *et al.*, 2003).

The fully integrated model is based on structural econometric modeling, where the underlying household heterogeneity is treated by incorporating fixed effects. This is done by estimating behavioral equations on the cross section of households using microeconomic data. The econometric estimation is complicated and frequently involves controversial identification assumptions. This generally leads to the integration of structural models focusing on small dimensions of household behavior, which may limit the model's capacity to capture the full complexity of household living standard inequality and the way it may be influenced by policy changes (Bourguignon, 2003; Cogneau and Robillard, 2006). The fully integrated models have the advantage of considering all the observed heterogeneity of the population of households and providing an accurate picture of the magnitude and direction of the poverty and distributional effects of policy shocks. However, the implementation of this approach raises many difficulties and remains quite challenging.

A more simple option would be to work with a layered approach which links CGE and micro-simulation models in a top-down fashion. At the top, a conventional CGE model is used to simulate policy changes and estimate prices, factor rewards, and macro changes. These changes are then associated with changes in the set of coefficients of the micro simulation model, which incorporates reduced-form econometric modeling of occupational choice and income determinants, to yield estimates of the change in the individual households income. The full distribution of real household income corresponding to CGE model counterfactual simulations can then be estimated. The main difficulty of this approach is to ensure consistency between the micro-simulation and the CGE results. Consistency is generally achieved by judiciously adjusting parameters in the micro-simulation model, which remains less satisfactory than modeling behavior identically in both models (Davies, 2009). Also, this approach does not allow feedback from the micro side to the macro side.

Conclusion

Local economy impacts of SCT programs may be generated through market and/or non-market interactions between the claimants of the transfers and other people in the community. We review potential channels through which SCT may affect labor market outcomes and productive activities and generate local economy impacts.

Our review of the micro approach to local economy effects of SCT programs provides a few lessons for data collection and analysis. The modern econometric framework for impact evaluation based on the definition of a counterfactual is a useful tool to start comprehending the full impacts of a SCT intervention. An impact evaluation study can be designed to capture spillover effects at the local economy level. Two-stage experimental design is appropriate to account for spillovers and measure impacts at the village level. But, evidence from the few empirical studies available so far is incomplete. We need good data on agricultural activities and outcomes to assess the impacts of SCTs on these dimensions. We need a theory to explain spillover effects. Insights from theory can help guide data collection to test competing hypothesis of the sources of spillover effects. Investigating the heterogeneity in impacts is a promising approach for uncovering the mechanisms at play within this reduced-form framework.

There are indeed a number of limitations intrinsic to the micro approach. This approach can only account for impacts on *equilibrium* outcomes. These impacts are most often obtained at the pilot stage of the SCT program and may differ from those of a fully scaled-up program which is more likely to produce general equilibrium effects between villages.

Studies using meso/macro methods to assess economy-wide effects of SCT programs are scarce. This may be because CGE models typically explore effects of policy reforms that take place at a macro level in a top-down approach. In contrast, cash transfers enter the economy at the household level. Village SAM/CGE offer a bottom-up approach that fits better with the level at which cash transfers enter the economy. Village SAM/CGE are well-rooted in household agricultural models. This modeling tool seems appropriate to investigate village-level effects of SCTs. But village SAM/CGE models consider villages as closed economies within the national economy. National-level CGE modeling may complement village-level analysis in a number of ways. They may better allow to model flows between villages (e.g., migration). They may also complement the analysis by accounting for indirect effects arising from the need to finance the SCT program domestically. A national-level CGE model may also be more appropriate for the analysis of poverty and distributional issues, especially when combined with disaggregated survey data.

Finally, there is some scope for integrating the micro and meso/macro approaches to overcome their respective shortcomings. The meso/macro based approaches are based on the structural modeling of individual behavior aggregated at the market level. Models may be augmented to account for informal markets. An important weakness of the approach, when compared to the micro approach, is that impact estimates obtained from these models are based on simulations. Results are thus as credible as the assumptions underlying the modeling. In contrast, the micro approach takes great care at ensuring the internal validity of impact estimates but it only produce reduced-form estimates of village-level impacts. Validation of village-level models may be obtained by comparing simulation-based results to the impacts obtained from the micro evaluation study. A valid village-level model can then be a useful tool to understand how to increase SCT impacts on the local economy.

Appendix 1: Impact evaluation research: the quest for a valid identification strategy.

Impact evaluation research is a complement to other types of program evaluation efforts (e.g., studies based on beneficiary satisfaction and participants' self-evaluation). Its one distinctive feature is the weight researchers in this field put on the search for a valid identification strategy for isolating the effects of the program from those of other contemporaneous factors.

A simple comparison of the welfare of program beneficiaries to that of non-participants would often yield an erroneous measure of the impact of the program. Participants and non-participants usually differ in important ways over a range of characteristics besides their participation in the program. In order to determine the true impact of a program, one would ideally want to compare what happens when the person is exposed to the program with what would have happened to her/him in the absence of the program. Clearly, one cannot observe the same person in the two states (exposed and unexposed). So one instead compares program beneficiaries to non-beneficiaries who are as similar as possible except for the fact that they are not enrolled in the social program. This can be done in a variety of ways.

Various strategies exist to address this "missing data" problem. They can be classified into two broad categories: experimental and non-experimental. The experimental method forcibly constructs the comparison (or control) group by randomly postponing the incorporation to the program of a selected set of people who will consist of the control group (Skoufias, 2001). As a consequence, individuals in the treatment group (those incorporated earlier on) and individuals in the control group have similar observable and unobservable characteristics.

The timing of the evaluation *vis-à-vis* the start of the program is an important factor driving the choice of approach for the impact evaluation study. Setting-up an experimental design is more natural before the start of the program. For on-going programs, experimental impact evaluation studies are usually not feasible, simply because it becomes impossible to define a control group. At one extreme, each and every eligible person is already a program recipient. But, even when some of the eligible have not yet joined the program, it is likely that these later comers would differ in a systematic way from those who are already receiving benefits.

There are three situations where setting-up an experimental design can still be feasible (desirable) to assess the impacts of an on-going program. First, program managers and policy-makers may be considering extending the program to a new population (e.g., through an increase in the threshold income defining the poverty line when the program is targeted to those below a poverty line; through an increase in the age threshold when the program is targeted to children below a certain age). In this case, one may design an experiment to learn about the impact for this new population of beneficiaries.

Secondly, program managers and policy-makers may have identified issues in the program design and bottlenecks in its implementation that would explain the lack of results. Then, an experiment can be designed to assess the effects of various innovations on the initial design.

Thirdly, it may be that the program is experiencing some issues with enrolling the target population. This could happen in the beginning of the operations of a program, before every eligible person learns that he/she is entitled to receiving benefits. It could also happen if some stigma is attached to receiving benefits, so that a large fraction of the target population remains outside the program. Then, a “promotion” experiment can be designed to offer more information on the program or nudge the reluctant towards enrolling. This experiment should exogenously increase take-up in the sub-sample of the eligible population targeted by the promotion effort. Instrumental variable methods (see below for a discussion) should then prove useful for analyzing data from the promotion experiment. To sum up, the experimental approach may still prove useful, even if the program has already started, especially when combined with non-experimental analysis.

Conducting a social experiment needs planning. It also requires the collection of data for the two experimental groups. A thorough power analysis is required to determine the size of each of the groups. In the absence of a well-powered design, the impact study will not find any statistically significant effects, but we will not be able to tell if this is a problem with the size of the experimental sample or if the program indeed has no effect. There are several points worth discussing when thinking about the appropriate sample size for the evaluation study.

First, in two-stage randomization designs, the level at which the randomization occurs (usually the village) is key. In contrast, adding hundreds or even thousands of households in each village will have very little influence on the power of the design. The units of randomization are the villages, and power is almost completely determined by the number of villages in the experimental sample. A similar point can be made for interventions for which the level of intervention differs from the level at which outcomes are observed (e.g., the program is delivered at the school level, and outcomes are measured for students). A correction for the fact that outcomes tend to be correlated at the cluster level (in our example, schools) will result in a large sample size for the experimental study than in the absence of a cluster effect.

Second, researchers are often interested in a multitude of outcomes. The power analysis should be based on the main outcome that the program aims to influence. When more than one outcome is identified by policy-makers, a conservative power analysis should focus on the one outcome which is highly variable or a relatively rare event. For instance, if one of the objectives of the program is to reduce child mortality, then one must focus on having a sufficiently large sample of children below age 5.

Third, when non-compliance may be an issue, sample size should be increased accordingly to allow detecting an effect. Fourth, the minimum sample size required to look at the heterogeneity in impacts by subgroup should also expected to be larger since impacts must be detected for a smaller population. In all cases, a number of few parameters must be estimated to conduct a power analysis: unexposed mean outcome and variance, intra-cluster correlation (when applicable), minimum detectable effect, power level. Estimates for the first two are typically obtained from existing data source (a correction must be made to account for the change through time). Minimum detectable effect should be set at a value that would justify the intervention. Information on intervention cost, the opportunity cost of the intervention or simply the level of impact required for the intervention to be judged successful is critical in setting the minimum detectable effect. Finally, the choice of power level depends on the risk at which we are willing to fail to identify an effect when there is one.

With enough power and a larger sample size, the researcher can look at the heterogeneity in impacts by subgroups (e.g., male-headed households *vs.* female-headed households, landowners *vs.* the landless). This usually helps to go beyond determining if the program works and directly investigating the causes for success or failure.

Although costly and difficult to implement, experimental evaluations rely on weaker assumptions than non-experimental evaluations. Thus, they provide the most credible estimates of the true impact of the program, when properly conducted. A related advantage is that they are readily understandable by policy-makers (Heckman and Smith 1995): a simple difference in the average outcomes between the treatment group and the control group yields a consistent estimate of the average impact of the program on the beneficiaries.

Experimental impact evaluations require a steady support from many stakeholders, including program managers, from the very start of the process. Although this could be said of all impact evaluations, the question is even more salient for experimental studies. For instance, people in the control group should be sheltered from any intervention in the sector to which the one that is assessed belongs; otherwise, we end up comparing the situation experienced by participants to the situation that control group people are experiencing: the benchmark is flawed. Local officials and sector specialists in the areas where the experimental sample is drawn must provide support and monitoring. In the absence of support and monitoring, even the best-designed experimental study does not yield valid estimates. Besides being difficult to get, this close collaboration raises problems of its own. The "right distance" has to be found between the persons in charge of the evaluation and those conducting the intervention to ensure the independence of the former and the credibility of the results. Yet, a close collaboration also has its own virtue, if well led: program managers and policy makers are well aware of the existence of the study, they are engaged with the researchers, expecting the results. In this sense, experimental studies have the potential to influence policy.

It is also important to note that the success of an experimental impact evaluation is not measured in terms of the size of the effect it finds: at the extreme, "no-effect" results, when based on a strong design, provide a clear message to program managers and policy-makers. Through trial and errors, one can find out about what works and what does not work. This is strength of the experimental approach: it helps settle a debate and can be used as a tool for policy design.

In other cases, the treatment cannot be randomized. The internal validity of the impact estimates ultimately depends on the assumptions we make on the factors driving the selection into the program. Here, a note of caution may be necessary. Suppose we are interested in evaluating a program that provides unconditional cash transfers to the poor.

Some may find it questionable that we argue that people chose to participate or not to the program: most of the poor will be glad to receive the extra cash! By selection into the program, we mean that those who we observe getting the training and those who are not differ in many aspects. So the question is whether we observe all of these various factors or not.

Assumptions are required about the processes underlying the selection into the program and the data available. We distinguish between 2 broad types of methods: (*i*) those based on selection on observed characteristics affecting program participation and unexposed outcomes, and (*ii*) those based on selection on unobserved characteristics. The first assumes that selection into the program depends on observable characteristics and, conditional on those characteristics, participation does not depend on outcomes in the unexposed state. Regression (e.g. ordinary least squares, probit and logit) and matching belong to this class of methods.

Matching methods rely on the construction of a comparison group such that, conditional on a set of covariates, participation does not depend on the outcome when not exposed to the program. Intuitively, this means selecting non-participants who are as similar as participants in terms of a set of covariates. The selection bias gets differenced out by comparing the outcomes of participants and "matched" non-participants. Matching is similar to regression but does not impose a functional form on the outcome equation. In contrast to regression, matching highlights the support problem, *i.e.* helps to compare comparable individuals (Heckman, Ichimura, Smith and Todd 1998). Obviously, the validity of this first type of methods relies heavily on the assumption that we observe all the factors that are driving the selection into the program. In turn, this suggests that the better the information we can exploit, the more likely this assumption will hold. Having access to a rich set of variables that can be argued to make participation "as-random" increases the credibility of the estimates. Having access to a set of non-participants who already share many of the characteristics of participants (e.g., non-participants who are in the pipeline to become participants) also helps in finding better "matches". In this sense, convincing program managers to collect more data on program applicants and making use of these administrative datasets could have many advantages.

Can matching techniques be useful for analyzing experimental data? If the experiment is well-designed and well-conducted, then participation *is* independent of the unexposed outcome, and there is no need to condition on observed variables. In terms of identification of the impact, the selection bias (driven by observed and unobserved factors) is balanced in the two experimental groups, so that, by taking the difference in average, the selection bias is cancelled out. Yet, in terms of estimation, one may decrease the variance of the impact estimates by conditioning on observed variables that are not influenced by the program (typically, pre-treatment variables). But there should not be any advantage in using matching to doing so compared to running a simple OLS model.

Selection on unobservables allows unobservable characteristics that affect outcomes and participation to be correlated. Here again, various methods can be used. Longitudinal methods require these unobservables to be time-invariant. For example, before-after comparison requires participation to depend only on time-invariant unobservables. But what if changes other than the implementation of a program happen simultaneously? In this case, the difference-in-difference method may be more appropriate. This method compares mean outcomes before and after the treatment for a treatment and a comparison group. It helps differencing-out changes in outcomes over time that did not occur because of the program. It requires eligible individuals not to change their behavior in anticipation of the program, or at least observing them before they do (Heckman and Smith 1999). A careful look at the trends of the treatment and comparison groups prior to the program helps providing some support to the method. Finally, it is interesting to combine matching and difference-in-difference. Repeated cross-sections are sufficient, but data at a pre-program baseline is necessary. Difference-in-difference can be successfully combined to matching. Difference-in-difference matching estimators have been shown to produce estimates that are close enough to the true value of the impact.

The instrumental variables (IV) method and the bivariate selection model are two additional ways of addressing the problem of selection on unobservables. Both methods only require cross-sectional data. Yet, they require finding a variable that affects outcomes only through its effect on participation. In a heterogeneous impacts world, the IV estimator estimates a local average effect, i.e. the impact for those who change their participation in response to changes in the value of the instrument. This may or may not be a parameter of interest to program managers and policy makers, and relies on this strong and untestable assumption of exogeneity of the instrument.

Finally, although less "*en vogue*" among development economists today, selection models that control for the part of the error in the outcome equation that is correlated with the participation may offer an alternative to the reduced-form approaches that we discussed so far. It usually forces the researcher to build an explicit model of participation and outcome choices. As it provides additional structure compared to the IV method, it makes it possible to examine heterogeneity in program impacts. An interesting development in the literature is the validation of structural models using experimental estimates. The idea is quite simple: if the model is correct, then applying it to the experimental control group participants should yield back an estimate that is close to the experimental one. If validated, it becomes a useful tool for simulations and an alternative to costly trial and error experiments.

Regression discontinuity designs approach is an "old" new method that is gaining support interest among researchers. The regression discontinuity method is useful when there is no common support for participants and non-participants (thus, in a situation where matching cannot be implemented) because treatment is allocated to anyone below (or above) a certain cutoff value. The idea is to compare those who are just above the cutoff point to those just below the cutoff point. This requires having enough observations around the cut-off point. Census data would be a good source of data on which to apply the method. In addition, the method relies on the assumption that expected gains from the program should not incite those above the cutoff point to change their decisions in order to comply with the rule. When these conditions are met, the regression discontinuity estimator makes it possible to recover the mean impact of the program for individuals who are located at the cutoff point. This parameter may be of interest to policy-makers who are considering extending the program benefits to those above/below the cutoff value (e.g., change in the poverty line hen cash transfers are targeted to the poor).

Both experimental and non-experimental approaches to impact evaluation can produce reliable estimates of the impact or fail to do so. Experiments face many challenges, at the design stage, at the implementation stage, because of its reliance of the goodwill of donors (experiments are expensive), its reliance of the support of local politicians and the monitoring of activities on the field by program managers and sector specialists. Non-experimental methods require assumptions, and more importantly, support for assumptions for which we usually do not have statistical tests to rely on. Finally, however difficult to obtain, scientific rigor is but a first step towards policy influence.

Appendix 2: Impact evaluation research: the quest for a valid identification strategy.

The fundamental problem in assessing the impact of a program is that we cannot observe the same person in two states of the world: one where this person is experiencing the program (is "treated") and one where she is not ("untreated"). For each treated person ($D_i = 1$), one does not observe the counterfactual outcome Y_0 she would have experienced had she not participated in the program ($D_i = 0$). One can only observe Y_1. A selection bias in estimating program impacts is likely to arise when comparing participants to non-participants. A well-designed randomized experiment can help: treated and untreated participants have similar observed and unobserved characteristics on average. This design allows us to balance the selection bias between the two experimental groups ($T_i = 1$ for the treated participants, 0 for the untreated participants). More formally, let potential outcomes for participants and non-participants be determined by different processes:

$$Y_0 = g_0(X) + U_0,\ E(U_0 \mid X) \neq 0.$$

$$Y_1 = g_1(X) + U_1,\ E(U_1 \mid X) \neq 0.$$

The difference in expected outcomes between participants and non-participants introduces a selection bias:

$$\begin{aligned} &E(Y_1 \mid D=1, X) - E(Y_0 \mid D=0, X) \\ &\quad = E(Y_1 \mid D=1, X) - E(Y_0 \mid D=1, X) + \{E(Y_0 \mid D=1, X) - E(Y_0 \mid D=1, X)\} \\ &\quad = E(Y_1 - Y_0 \mid D=1, X) + \{E(U_1 - U_0 \mid D=1, X)\} \neq E(Y_1 - Y_0 \mid D=1, X). \end{aligned}$$

Expected value of the outcome in each experimental group is:

$$\begin{aligned} E(Y \mid D=1, T=1, X) &= E(Y_1 \mid D=1, X) = g_1(X) + E(U_1 \mid D=1, X). \\ E(Y \mid D=1, T=0, X) &= E(Y_0 \mid D=1, X) = g_0(X) + E(U_0 \mid D=1, X). \end{aligned}$$

The difference is expected outcomes between the treated and untreated in a randomized experiment yields the average effect on the treated:

$$\begin{aligned} &E(Y \mid D=1, T=1, X) - E(Y \mid D=1, T=0, X) \\ &\quad = E(Y_1 \mid D=1, X) - E(Y_0 \mid D=1, X) = E(Y_1 - Y_0 \mid D=1, X). \end{aligned}$$

The average effect on the treated is identified because the experimental design helps balance the bias between the two experimental groups:

$$\begin{aligned} &E(Y \mid D=1, T=1, X) - E(Y \mid D=1, T=0, X) \\ &\quad = g_1(X) - g_0(X) + E(U_1 \mid D=1, X) - E(U_0 \mid D=1, X). \end{aligned}$$

A similar argument can be made for two-stage randomization designs. Now, groups are the unit of randomization. Participants within treated groups benefit from the treatment. Participants in untreated groups are excluded from the treatment. Thus, comparing participants in treated groups to participants in untreated groups provides the impact of the treatment on participants. Spillover effects on non-participants can be obtained from a comparison of the realized average outcome among non-participants in the treated groups and in the untreated groups. Indeed, non-participants in treated and untreated groups are similar in all aspects expect than the former belong to the same local economy as treated participants, whereas the latter interact with untreated participants. Thus, the difference in these two average outcomes for non-participants can be attributed to the program. Formally, the average effect on the participants (resp. average effect on the non-participants) is obtained as follows:

$$\begin{aligned} &E(Y \mid D=1, T_g=1) - E(Y \mid D=1, T_g=0) = E(Y_1 - Y_0 \mid D=1) \\ &E(Y \mid D=0, T_g=1) - E(Y \mid D=0, T_g=0) = E(Y_1 - Y_0 \mid D=0) \end{aligned}$$

The average effect on non-participants can be interpreted as a spillover effect of the program.

Appendix 3: Dyadic model of link formation.

Let L_{ij} be the propensity to form a link between two individuals i and j, X_{ij} (e_{ij}) observed (unobserved) attributes of the link and g_{ij} a variable indicating if the link is observed. We have:

$$\begin{aligned} &L_{ij} = X_{ij}'b + e_{ij}, \\ &g_{ij} = 1 \; if \;\; L_{ij} = 0 \text{ and 0 otherwise.} \end{aligned}$$

The model should typically include X_i, X_j and $|X_i - X_j|$ if links are directed (X_i, X_j and $|X_i - X_j|$ and $(X_i + X_j)$ to allow for undirected links). Treatment status can be included in the set of X. See Bramoullé and Fortin (2010) for a review on these models.

References

Albarran, P. and Attanasio, O. P. (2003), “Limited commitment and crowding out of private transfers: evidence from a randomized experiment.” Economic Journal, 113: C77–C85.

Alzua, M.L., G. Cruces and L. Ripani (2010). “Welfare programs and labor supply in developing countries. Experimental evidence from Latin America”, CEDLAS Discussion paper no.95.

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 De Giorgi G., (2009) “Indirect effects of an aid program: how do cash injections affect ineligibles' consumption,” American Economic Review, 99(1), 486-508.

Angelucci, G. De Giorgi, M. A. Rangel and Imran Rasul. (2010). “Insurance and investment within family network.” Unpublished manuscript, University of Michigan.

Babcock, P. and J. Hartman (2010). Networks and Workouts: treatment size and status specific peer effects in a randomized experiment” NBER working paper 16581.

Bandiera O., Burgess R., Gulesci S., Rasul I. (2009). “Community Networks and Poverty Reduction Programs: Evidence from Bangladesh”, EOPP working paper No 2009/15.

Barrientos, A. and J. Scott (2008). Social Transfers and Growth: A Review. BWPI Working Paper 52, Manchester, UK.

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

Bianchi, M. And M. Bobba (2010). “Liquidity, Risk, and Occupational Choices”, Unpublished manuscript, Université Paris-Dauphine.

Bobba, M. and J. Gignoux. (2010). “Spatial externalities and social multipliers of schooling interventions.” Unpublished manuscript, Paris School of Economics.

Bourguignon, F., A.-S. Robillard and S. Robinson (2003). Representative versus real households in the macroeconomic modelling of inequality. DIAL Working Paper N.2003-10, Paris-France.

Breisinger, C.. M. Thomas and J. Thurlow (2009). Social Accounting Matrices and Multiplier Analysis: An Introduction with Exercises. International Food Policy Research Institute, Washington, D.C., USA.

Bramoullé, Y., H. Djebbari and B. Fortin (2009). “Identification and estimation of social interaction effects in networks” Journal of Econometrics, 150: 41-55.

Bramoullé, Y. and B. Fortin (2010). “The econometrics of Social Networks” New Palgrave dictionary of economics.

Browning, E. (1971). “Incentive and Disincentive Experimentation for Income Maintenance Policy Purposes: Note”, American Economic Review, Vol. 61, No. 4.

Coady, P. D. and R. L. Harris (2004). Evaluating transfer programmes within a general equilibrium framework. Economic Journal, 114 (October): 778–799.

Cogneau, D. and A.-S. Robillard (2006). Growth, distribution and poverty in Madagascar: learning from a microsimulation model in a general equilibrium framework, Chapter 3 in Spadaro A (Ed.), Microsimulation as a tool for the evaluation of public policies: methods and applications, Bilbao: Fundacion BBVA.

Davies, B. J. (2009). Combining Microsimulation with CGE and Macro Modelling for Distributional Analysis in Developing and Transition Countries. International Journal of Microsimulation: 49-65

Davies, S. and J. B. Davey (2008). A Regional Multiplier Approach to Estimating the Impact of Cash Transfers on the Market: The Case of Cash Transfers in Rural Malawi. Development Policy Review, 26 (1): 91-111.

Davis, B., T. Reardon, K.G. Stamoulis, and P. Winters (2002). Promoting Farm/Non-farm Linkages for Rural Development: Case Studies from Africa and Latin America. Food and Agriculture Organization of the United Nations, Rome, Italy.

Davis, B. , G. Carletto, and P. Winters (2010) “Migration, Transfers and Economic Decision Making among Agricultural Households: an Introduction”, Journal of Development Studies, 46: 1, 1-13.

de Janvry, A., E. Sadoulet, and B. Davis. 2001. “Cash Transfer Programs with Income Multipliers: PROCAMPO in Mexico”. World Development 29(6): 1043-56.

de Janvry A., F. Finan , E. Sadoulet and R. Vakis (2006), “Can conditional cash transfer programs serve as safety nets in keeping children at school and from working when exposed to shocks?” Journal of Development Economics 79: 349– 373.

De Mel S., McKenzie D. and Woodruff C., (2008) “Returns to capital in microenterprises: evidence from a field experiment.” Quarterly Journal of Economics 123(4): 1329-1372.

Deaton, A. (2009). “Instruments of development: randomization in the tropics, and the search for the elusive keys to economic development”, NBER Working Paper 14690.

Dercon, S. and Krishnan, P. (2003), "Risk sharing and public transfers.", Economic Journal, 113: C86–C94.

Djebbari, H. and J. Smith (2008). "Heterogeneous program impacts of the PROGRESA program" *Journal of Econometrics*, 145(1-2): 64-80.

Doran, K. (2011). "How Does Child Labor Affect the Demand for Adult Labor? Evidence from Rural Mexico", Unpublished manuscript, University of Notre-Dame.

Duflo, E., R. Glennerster, and M. Kremer (2008). "Using Randomization in Development Economics Research: A Toolkit". In T.P. Schultz and J. Strauss (eds.): Handbook of Development Economics, Vol. 4, 3895-3962.

Dyer, A.G. and J.E. Taylor (2004). Rethinking the Supply Response to Market Reforms in Agriculture: Household Heterogeneity in Village General Equilibrium Analysis from Mexico. 7th Annual Conference on Global Economic Analysis, "Trade, Poverty and the Environment." The World Bank and Center for Global Trade Analysis, Washington, DC.

Dyer, A.G., S. Boucher and J.E. Taylor (2006). Subsistence Response to Market Shocks. American Journal of Agricultural Economics 88(2): 279–291.

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

Fafchamps, M. and Gubert, F. (2007). Risk sharing networks in rural Philippines. Journal of Development Economics 71, 261–287.

Fiszbein, A. and N. Schady (2009). Conditional cash transfers. A World Bank policy research report.

Ferede, T. (2009). In Search for Broad-Based Growth: Linkages and Policy Effects in a Village Economy in Rural Ethiopia. Journal of Sustainable Development in Africa, 11(3): 44:83.

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

Gertler, P., S. Martinez and M. Rubio-Codina. (2006). “Investing Cash Transfers to Raise Long-Term Living Standards.” World Bank Policy Research Working Paper No. 3994.

Helmers, C. and M. Patnam (2011) “Does the Rotten Child Spoil His Companion? Spatial Peer Effects Among Children in Rural India.” Unpublished manuscript, University of Oxford.

Hirano K and J. Hahn (2010). “Design of randomized experiments to measure social interaction effects”, Economics Letters, Volume 106, Issue 1, Pages 51-53.

Imbens, Guido W.. (2010). "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)." Journal of Economic Literature, 48(2): 399–423.

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

Maluccio, J. (2010). “The Impact of Conditional Cash Transfers on Consumption and Investment in Nicaragua.” Journal of Development Studies 46(1): 14-38.

Manski, C. (2010). “Identification of Treatment Response with Social Interactions.” Unpublished manuscript Northwestern University.

Moffitt, R. (2001). “Policy Interventions, Low-Level Equilibria, and Social Interactions”. In: S.N. Durlauf and P. Young, editors. Social Dynamics 45-82. MIT Press.

Parker, S. W., L. Rubalcava. and G. Teruel (2008). "Evaluating Conditional Schooling and Health Programs," In T.P. Schultz and J. Strauss (eds.): Handbook of Development Economics, Vol. 4, 3963-4035.

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.

Pfeiffer L., A. Lopez-Feldman, J. E. Taylor (2009). “Is off-farm income reforming the farm? Evidence from Mexico” Agricultural Economics 40, 125-138.

Philipson, T. (2000). “External Treatment Effects and Program Implementation Bias," NBER Technical Working Paper No. 250.

Robinson, S. and H. Lofgren (2005). Macro Models and Poverty Analysis: Theoretical Tensions and Empirical Practice. Development Policy Review, 23 (3): 267-283.

Round, J. (2003). Social Accounting Matrices and SAM-based Multiplier Analysis. Chapter 14 in Bourguignon F. and L. A. Pereira da Silva (eds): Techniques and Tools for Evaluating the Poverty Impact of Economic Policies. World Bank and Oxford University Press, p. 301-324.

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

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

Skoufias, E., and S. Parker (2001). “Conditional Cash Transfers and their Impact on Child Work and Schooling: Evidence from the PROGRESA Program in Mexico”, Economía, 2(1).

Taylor, J.E., and I. Adelman (1996). Village Economies—The Design, Estimation, and Use of Village wide Economic Models. Cambridge: Cambridge University Press.

Taylor, J.E., A. Yunez-Naude and S. Hampton (1999). Agricultural Policy Reforms and Village Economies: A Computable General-Equilibrium Analysis from Mexico. Journal of Policy Modeling 21(4):453–480.

Taylor, J.E., A.G. Dyer, and A. Yunez-Naude (2005). Disaggregated Rural Economy wide Models for Policy Analysis. World Development, 33 (10): 1671–1688.

Taylor, J. Edward (1995). Micro economy-wide models for migration and policy analysis: an application to rural Mexico. OECD center, Paris.

Teixeira, C. G. (2010). “A Heterogeneity Analysis of the Bolsa Família Programme Effect on Men and Women’s Work Supply.” Working Paper 61. UNDP International Policy Center for Inclusive Growth, Brasília, Brazil.

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): 39-67.

Yunez-Naude, A., A.G. Dyer and J.E. Taylor (2006). Regional Growth Linkages between Villages and Towns in Mexico: A Village-wide Modeling Perspective. International Association of Agricultural Economists Conference, Gold Coast, Australia.