Combining Survey and Ethnographic Methods to Evaluate Conditional Cash Transfer Programs

Michelle Adato
International Food Policy Research Institute, Washington, D.C.
Email: M.Adato@cgiar.org

Q-Squared Working Paper No. 40
November 2007

Q-squared • Centre For International Studies • University Of Toronto
1 Devonshire Place, Toronto ON M5S 3K7 Canada
t: 416-885-7721 • f: 416-652-1678 • e: info@q-squared.ca

1. Introduction[1]

Increasingly recognized as a critical part of poverty reduction strategies, social protection systems have been used to enable individuals, families, and communities to reduce risk and mitigate the impacts of stresses and shocks to their livelihoods. They can also be used to support people who suffer from chronic incapacities to secure basic subsistence. Furthermore, such interventions can contribute to broader development processes through investments in health, nutrition, and education for children and adults, development of productive infrastructure, and support for livelihoods activities. In the last ten years, one type of intervention has swept across Latin America and is now making its way across the globe: conditional cash transfer programs (CCTs). These programs provide a cash transfer to poor households, conditioned on their participation in health and education services. At least twenty countries currently have a CCT or are in the planning stages, with another twenty exploring the idea.[2] With the rise of CCTs has also come a new practice of systematically building rigorous program evaluation into social policy. In many cases, donors and governments have required that these evaluations include both quantitative and qualitative research methods. Drawing on recent experience of the International Food Policy Research Institute's evaluations of conditional cash transfer programs for the governments of Nicaragua and Turkey, this paper explores how ethnographic and survey methods have been combined provide representative measures of impacts on poverty, health, nutrition, education, and other variables, with in-depth, subtle, explanations for those changes (or lack of change), and exploration of social processes and impacts such as effects on gender and other social relations.

Section 2 of this paper provides a background to conditional cash transfer programs, situating them within the wider context of development, social policy, and social protection, and explaining the basic objectives and features of the program. Section 3 then provides detail on two conditional cash transfer programs, in Nicaragua and Turkey, upon which the empirical research in this paper is based. Sections 4 and 5 are the core of the paper. Section 4 begins with a background on mixed method research, and the objectives of CCT program evaluation. It then summarizes the rationale and design of quantitative research in the two studies, followed by a more in-depth discussion of the qualitative and ethnographic methodology, including what they contribute, how they contrast with and complement the survey data, the types of issues they illuminate in the CCT evaluations; and the research design, sampling frameworks, and research methods. Section 5 provides evidence of the value of combining survey and ethnographic methods, providing a few selected findings from the Nicaragua and Turkey CCT evaluations, examples that represent different types of uses of and insights from the

[1] I would like to thank Terence Roopnaraine for his assistance with the design, fieldwork supervision, data analysis, and reports for the qualitative studies in the IFPRI conditional cash transfer program evaluations in Nicaragua and Turkey, and for assistance with this paper.

[2] In June 2006, over 300 representatives attended the "Third Annual International Conference on Conditional Cash Transfer Programs," held in Istanbul, Turkey (World Bank 2006).

different methods. Section 6 concludes with reflections on gaps that remain in the integration of methods for evaluating social policy, other reflections on the process.

2. Social protection and Conditional Cash Transfer Programs

'Social protection' encompasses a broad set of public and private systems for protecting people against risks to their livelihoods, keeping them from falling into poverty. These may be insurance mechanisms that kick-in in the case of a shock such as the illness of a wage-earner, loss of a job, or a natural disaster. It may also take the form of a regular cash or in-kind transfer where people suffer from chronic inabilities to secure livelihoods, due to age or disability, social class, or discrimination due to the economic, social and political systems in which they live. Formal social protection systems can be provided by the state, non-governmental organizations, or private sector employers. Traditionally, they have been and continue to be provided by family or 'community' variously defined, though these informal systems have been overly strained by trends and shocks that affect many family and community members simultaneously. Social protection is often advocated as a right rather than a reactive form of relief. Increasingly, state and donor-designed social protection systems have tried to contribute to long-term, sustainable, development processes, to provide opportunities for people to move out of poverty and achieve a higher standard of living. This can be achieved through interventions that invest in health, nutrition, and education for children and adults, as well as improved social status and rights (Adato et al. 2004; Conway and Norton 2002). "Promotional" and "transformational" approaches to social protection invest in assets, including human capital, and may even build capacities of individuals and organizations to engage in broader social development and political processes (Guhan 1994; Devereux and Sabates-Wheeler 2004; Adato et al. 2005). For decades mainly the domain of richer countries with comprehensive social security systems and benefits provided through formal employment, social protection is increasingly being seen as part of anti-poverty strategies in low-income countries (Norton et al. 2002). It can be advocated on ethical grounds—as a human right and the basic responsibility of the state to protect its citizens from poverty and severe forms of deprivation. However, in order to convince wary economists and finance ministries more swayed by approaches to development that rely on growth and markets, it can be argued that social protection can contribute to growth—through human capital investments, development of infrastructure, strengthening markets, and maintaining political stability.[3]

Using an assets framework that includes financial, human, social, physical, natural, and political capital (see Carney 1998), social protection can be seen to have four types of objectives: 1) protective—e.g. providing relief through cash or food transfers; 2) preventative—e.g. averting deprivation through insurance; 3) promotional—e.g. promoting access to assets including education, health, nutrition, job skills, or credit; and

[3] For example, a conditional cash transfer can increase education levels that increase productivity (Morely and Coady 2003), and public works program can be designed to build roads or structures that promote market activitity, or job training that enhances labor force participation and productivity (Adato et al. 2005).

4) transformative—e.g. using organisational mechanisms that promote women's rights, community organization, and agency. (Guhan 1994; Devereux and Sabates-Wheeler, 2004). Different types of interventions, such as unconditional and conditional food and cash transfers, maternal and child health programs, health or asset insurance, public works, or livelihood support and micro-credit programs are often associated with a particular objective, but they are not confined to that objective. Depending on 1) how the social protection system is designed, and 2) the ability of people to take advantage of the program and the way in which they respond, any of these programs can potentially contribute to any of these objectives. For example, an unconditional cash transfer is most likely to secure basic consumption, but it can also avert asset reduction if it keeps parents from taking their children out of school. A public works program may secure basic consumption and avert asset reduction, but it can also directly build and enhance the use of assets, through construction of physical infrastructure, or transform social relationships through building new institutional structures involving community-based organizations and local government (Adato 2006a)

Conditional cash transfer programs: Program objectives and design

An essential underlying premise of a CCT program is that financial constraints keep parents from sending their children to school, because of the opportunity cost of sending a child to school rather than to work, as well as direct costs for books, supplies, clothing and transportation. Another is that very poor people, for a variety of reasons, do not take advantage of health services that may be available (Davis and Handa 2006). By providing cash assistance conditioned on household participation in education and health services, a conditional cash transfer can play a protective and preventative function—securing basic consumption and averting asset reduction—while simultaneously playing a promotional role—building assets—and even a transformational role, if e.g girls' education alters their relations with male partners in the future. The focus on prenatal, infant and early childhood health and nutrition, is based on the importance of these investments at these early ages: improving nutrition (particularly for children under three years of age) has been found to have positive effects on children's health and physical and cognitive development (Martorell 1995; Walker 2007). More recent evidence indicates a positive longer-term effect of an early childhood nutrition intervention on wages and income earning as future adults (Hoddinott et al. 2007). Behrman et al. (2004) provide a review of numerous studies establishing a relationship between nutrition and education, between education and wages, and between adult cognitive skills and earnings. Such research thus suggests synergies between nutrition, health and education, which CCT programs exploit. The overall objective of the program can be seen as preventing the intergenerational transmission of poverty—though short-term impacts on human capital have thus far been much more firmly established than the achievement of this broader goal.[4]

[4] Samsun (2006) argues that the added value of conditioning the grant, over that of an unconditional transfer, has not yet been demonstrated, as programs have not directly compared the two approaches. Some research from Mexico and Ecuador has more recently used 'natural accidents' to simulate the comparison and found advantages to conditioning, but this has not yet been published. A study underway in Kenya, and one by IFPRI in El Salvador at design stage, are testing and comparing these different approaches directly.

Though CCTs vary in design across countries, there are a number of features that are constant across them, with variations found within broader parameters. First, the programs are targeted to the "poor," using demographic and socioeconomic data and/or data on assets to define a particular poverty threshold. Regions are targeted geographically and households are then targeted within them, though in some programs and places they may decide to include all households within a locality, as was done in some areas of Nicaragua in a later phase of the program.

Second, benefits are conditioned upon children's school enrollment and attendance, usually at around an 80 percent attendance rate, where there are school age children in the household. Where there are pregnant women or children 0-5 years, benefits are conditioned on their participation in preventative health care services such as check-ups, vaccines, and growth monitoring. Some programs also require beneficiary participation in health and nutrition education, a component that can be seen as important to promoting longer-term changes in practices lasting beyond the duration of the cash transfer. Conditions are monitored through a reporting system, with compliance records collected by schools and clinics and then processed at national level. The cash is then delivered to designated pick-up points. Some programs also provide additional in-kind benefits, such as nutritional supplements or school supplies. If conditions are not met over a specified time period, recipients are dropped from the program. Because meeting a program condition requires that services be available, the CCT program is often undertaken in conjunction with an increase in supply of services, for example, extending infrastructure and services into previously under-served areas, or increasing student-teacher or patient-health staff ratios. In some countries there are linkages between CCTs and other development initiatives, related to income generation and infrastructure (Ayala Consulting 2006).

Third, there is a strong gender dimension to CCT programs: the mother of the household designated as the official program 'beneficiary' (with some exceptions), with program implementers telling them that they—rather than their male partners—should keep and control the cash. Women are targeted for health services and health and nutrition education. Reducing discrimination against girls in education is often a major objective, with some programs offering higher transfers for enrolled girls than boys, and higher benefits at the secondary level where girls are more likely to drop out. Some programs also provide opportunities for women to meet collectively for various program-related activities (Adato and Mindek 2000).

3. Conditional Cash Transfers Programs in Nicaragua and Turkey

Nicaragua's conditional cash transfer, *Red de Protección Social* (RPS) was a relatively small program compared to others in the region. Its initial budget in 2000 was about U.S.$11 million for the first phase. The second phase expansion in 2002 was designed for another three years with a budget of $22 million. In 2004, 21,619 families were enrolled in the program. In the first phase, the program was piloted in only two 'departments' (out of 17), Madriz and Matagalpa in the northern part of the Central

Region, on the basis of poverty (80 percent of the population was poor) as well as on their capacity to implement the program. Within these departments, six municipalities were selected based on a governance criteria, but where 78–90 percent of the population was extremely poor or poor. Within these a marginality index further selected the poorest local areas or *comarcas*[5], where all but 6 percent of households were included. The average size of the transfer equaled about 17 percent of annual household expenditures (Maluccio and Flores 2005). In order to receive the cash transfer for food and the nutritional supplement, beneficiaries were required to bring children under five to appointments with health providers for growth monitoring and vaccinations; and to attend a training workshop every two months, covering nutrition, reproductive health, lactation, environmental health and family hygiene.[6] In phase 2, pregnant and lactating women also received check-ups and vitamins, and women in their child-bearing years were given tetanus shots. In addition to growth monitoring and vaccinations, children 0-5 years received vitamin A, iron, anti-parasite treatment, and when necessary oral rehydration. Beneficiaries were given counseling on childraising practices and children 6-9 years received tetanus vaccines. In the second phase the program was redesigned to offer adolescents information, education and communication on topics such as healthy lifestyles, sexual and reproductive health, and prevention of STDs and HIV and AIDS. The health services were provided by NGOs or private health providers, conforming to Ministry of Health rules and standards.

The education benefit included a school attendance transfer, given to households with at least one child in primary school; an in-kind transfer of school supplies and a uniform for each registered child; and a very small cash payment that households turn over to the parent-teacher association, in part to augment the teacher's salary and in part for needed school materials or upgrading (called the *Bono a la Oferta*). These benefits were conditioned on 85% school attendance of children age 7-13 (up to grade 4). The Ministry of Education was responsible for delivery of school services.

The CCT program in Turkey was part of the Social Risk Mitigation Project (SRMP), an initiative of Turkey's Social Solidarity Foundation (SYDTF) that formed part of a broad social safety net reform responding to the earthquake and economic crisis of 2001. The General Directorate of Social Assistance and Solidarity, responsible for the CCT, worked with the Ministry of Health and Ministry of Education for service delivery and monitoring. The program had 1.1 million beneficiaries, and a budget of $360 million (World Bank 2001). Using a proxy means test, the program covered the poorest six percent of the population nationally (it was thus not confined to particular regions, though the highest concentration of beneficiaries was in the poor southeastern region). Beneficiaries receive a cash payment for participation in health services, primary and secondary schooling (there are no in-kind transfers). There is another grant for pregnant

[5] Census *comarcas* are administrative areas within municipalities that typically include between one and five small communities averaging 100 households each. They are determined by the National Institute of Statistics and Censuses and sometimes do not coincide with locally defined areas also referred to as comarcas (Maluccio and Flores 2005).

[6] In phase 1, beneficiaries also had to ensure children did not fall in their percentile ranking in the weight for age distribution during consecutive weighings, a requirement later dropped when it was realized that this may be withholding benefits from children who need it most.

women. Payments are higher for secondary school than primary school, and higher for girls than for boys. Education benefits are conditioned on 80 percent attendance rates, and grades must not be repeated more than once. Health benefits are conditioned on attending check-ups every two months for children 7 to 18 months of age, and every six months for children from 1 ½ to 6 years of age. The pregnancy benefit requires women to attend monthly check-ups while pregnant, give birth in a hospital, and attend post-birth check-ups.

4. Combining Survey and Ethnographic Methods for the Evaluation of CCT programs

As with other uses of mixed method research, combining quantitative and qualitative methods for evaluation of social protection programs enhances the contributions of both methods, providing a richer pool of data and greater analytic power than would have been available with either of these methods alone. The use of quantitative and qualitative methods together and in complementary ways has long been established theoretically and empirically (Brewer and Hunter 1989; Creswell 1995; Tashakkori and Teddlie 1998). Triangulation, where several types of data are used in a single study, and used to cross-check and compare results, enables any weaknesses in one method to be offset by the strengths of another (Denzin 1978; Jick 1979). Discussing the iterative, feedback relationship between ethnographic and survey data in a particular study[7], Bernard (2002:363-354) writes: "The ethnography produced ideas for policy recommendations and for the content of a questionnaire. The questionnaire data illuminated and validated many of the things that the ethnographer learned during participant observation. Those same survey data produced anomalies--things that didn't quite fit with the ethnographer's intuition. Ethnographic and survey data combined produce more insight than either does alone."

A study of 57 mixed method studies from the 1980s identified five purposes for mixing methods (Greene et al. 1989): *triangulation*: seeking convergence of results; *complementarities*: examining overlapping and different facets of a phenomenon; *initiation*: discovering paradoxes, contradictions, fresh perspectives; *development*: using the methods sequentially, such that results from the first method inform the use of the second method; and expansion: adding breadth and scope to a project.

Large-scale evaluations of CCT programs using quantitative and qualitative methods have taken place in Mexico, Nicaragua, Turkey, Colombia and Jamaica (the Brazil and Honduras studies used quantitative methods only). Evaluations have become a standard feature of conditional cash transfer programs, often built into the initial policy and/or loan. Rawlings and Rubio (2005) point out that this systematic, rigorous approach to evaluation of social assistance programs represents a new trend.[8] The purpose of the CCT evaluations are to determine effectiveness of the program design (the cash transfers and conditionalities) and the efficiency of the investment; identify design

[7] This was a mixed-methods study on gender and harassment in the US Army carried out by Laura Miller.
[8] They point to a review of World Bank projects from 1998-2000, where only 10 percent had adequate plans for a rigorous evaluation (World Bank 2001)

and implementation issues requiring change or improvement; understand how people view the program and respond to it, and why they do or do not respond to program incentives; and increase transparency and accountability of government. The primary interest of governments evaluating their CCT programs, and donors and lenders such as the World Bank and Inter-American Development Bank that often fund the evaluations, is the quantitative evaluation—measuring change in the indicators that the program aims to effect: e.g. enrollment and attendance rates, participation rates in health services, and changes in nutritional status. Achieving measurable changes is important to decisions about whether to continue to fund the program. Qualitative methods are used to understand program impacts that are harder to measure through a quantitative survey: for example, changes in social relations, such as intra-household, gender, and community relations, how people interact with institutions; implications of economic, social and cultural attributes on participation and outcomes; how people understand, view, and like the program, and how and why they do or do not respond to the program design, incentives, training, or other aspects.

The Quantitative Methods in the Nicaragua and Turkey CCT Evaluations

In Nicaragua, surveys measured, *inter alia*, participation in growth and development monitoring, vaccinations, composition of expenditures, type of foods consumed (diet quality), nutritional status, school enrollment, attendance, continuation and matriculation, child and adult labor, and targeting accuracy (Maluccio and Flores 2005). In Turkey, surveys collected information on, *inter alia*, school enrollment, continuation, and completion, level of knowledge of the CCT program conditions, sources of program information, targeting accuracy, costs of education, occupation and employment, dwelling characteristics, assets, food and nonfood expenditures, health and immunization, economic shocks, and participation in the CCT program. Both studies collected data on household demographic characteristics and socioeconomic status. They also used records of payments distributed to beneficiary households to establish how much money people actually received. There are also types of quantitative data used in CCT evaluations that were not collected in Nicaragua and Turkey. For example, the IFPRI evaluation of PROGRESA in Mexico[9] used data collected via school, clinic, and nutrition surveys, school and clinic administrative data, and student achievement test scores (Skoufias 2005). The Mexico study also included an operations evaluation to assess the quality of service delivery and administration of the CCT, using surveys and observation checklists (Adato et al. 2000).

The quantitative evaluations undertake to establish the average effect of the program on a number of indicators, at the household level. In order to do this, they must construct a counterfactual that establishes what outcomes would have looked like in the absence of the program. This is best done through an experimental design, randomly assigning otherwise similar households into and out of the program, using statistical

[9] This is mainly due to the fact that the PROGRESA evaluation was much larger in scale and budget than the other two studies.

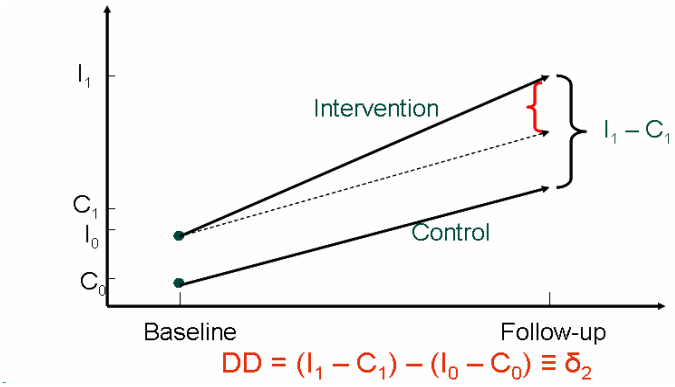
matching of observable characteristics.[10] In the Nicaragua study, a rigorous counterfactual was established through the use of a randomized, experimental design, using a 'double-difference' methodology (Maluccio 2005; see also Ravallion 2001): Out of 42 *comarcas,* 21 were randomly selected into the program, and 21 into the control group. Household and individual level data was collected in 2000, before the intervention began, in both control and treatment localities. Data on the same variables was then collected in the same households in [2002]. Because there are factors that produce changes in the treatment group but have nothing to do with the program, it is important to be able to separate out these effects from program effects. If the treatment and control groups are well-selected, these factors should affect both groups to the same degree. The evaluation measures average program impacts, measuring the changes within the treatment group, and then subtracts the change in the same indicators that occur within the control groups, producing the program effect.

Figure 1 shows the double difference methodology used in the Nicaragua evaluation. I_0 and C_0, denotes the intervention and control groups, respectively, at baseline. $I_1 - C_1$ denotes the intervention and control groups, respectively, at some point after the intervention has been implemented, when it would be expected to have an impact. At baseline, for each indicator measured, on average findings should be the same for the treatment and control groups. Following the program, after some time period, difference should emerge as a result of the program. Figure 1, however, takes into account 1) that there are likely to be some observable or unobservable differences between the two groups at baseline; and 2) that changes are likely to occur in both groups that are not attributable to the program. The double difference program impact is represented by the red bracket, or DD = (I1 – C1) – (I0 – C0) ≡ δ2 (Maluccio and Flores 2005:12-13).[11]

[10] The use of treatment and control groups naturally raises ethical questions, with respect to the possibility that families who might otherwise have had the opportunity to benefit from the program would be purposely denied benefits for the sake of a program evaluation. In practice, this is usually not a realistic problem because these programs do not have the financial or logistical ability to reach the entire target group at once. They are thus rolled out gradually, and those waiting can thus act as a control group.

[11] See Maluccio and Flores (2005) for additional details and caveats on use of this method.

Figure 1: Illustration of double difference estimate of average program effect



Where an experimental design of this type is not possible, as in the case where the program began before the evaluation could be started and there is thus no baseline, statistical methods can be used to establish comparison groups. One technique is propensity score matching, where a comparison group is constructed based on socioeconomic characteristics that would give them the highest probability of participating in the program, were the program available to them (Rawlings and Rubio 2005). Another technique is regression discontinuity design (RDD), used in the quantitative evaluation of the CCT program in Turkey. RDD is a method that compares average outcomes for households that fall just below and just above the program eligibility cut-off line, based on their proxy-means scores applied for the purpose of program targeting. A band is constructed for households that fall just below the line and those just above it. These groups are considered to be very similar, given the lack of precision of the proxy-means test model. Because both groups applied for the program (Turkey uses a system of applications), selectivity bias and unobserved characteristics between the two groups are assumed to be minimized[12] (Ahmed et al. 2006a). The total number of localities and households included in all three rounds of the Nicaragua study was 42 comarcas and 1359 households. In the Turkey study 3 surveys were carried out in 26 provinces (52 districts). First, 2905 households were interviewed for a large cross-sectional household survey (December 2005-April 2006). A further 750 households were interviewed for a two-round panel survey. The first round of panel survey interviews was carried out simultaneously with the cross-sectional household survey, and the second panel survey seven months later..[13]

The Qualitative Methods in the Nicaragua and Turkey CCT evaluations

Qualitative research offers a number of strengths for evaluating conditional cash transfer programs that survey methods do not. While survey methods are essential for

[12] A matching technique was also using some households that met the eligibility criteria but had not been included yet, but the number was small, so this method was considered less reliable.

[13] In the Turkey study, double difference with regression discontinuity design was used for estimating impacts on primary school enrollment, using retrospective questions on enrollment prior to program participation. The data from the second round survey could not be used in this way because there was no control group in the panel.

quantifying impacts on key indicators targeted by the program, they are at a disadvantage for probing issues requiring deeper exploration, due to limitations faced by survey methods in any research context. These include the necessary brevity of questions and the use of proxies that are often blunt measures; respondents' inability to sufficiently express what they mean in selecting among categorical or continuous variables; the limited ability of enumerators to follow up when more information or clarification is needed; and the difficulty of establishing the rapport and trust needed to maximize truthfulness in replies. In contrast, the qualitative research enabled the exploration of social issues and impacts requiring open-ended rather than closed responses; enabled an understanding of people's perceptions, expressed in their own words; raised underlying and less obvious issues, including those that we[14] as researchers had not anticipated; allowed us to probe and challenge responses and internal contradictions, or conflicting responses between respondents, and explore relationships between topics and responses; and finally, to solicit respondents' ideas about solutions to the problems with the program that they raised. Throughout, the qualitative research enabled us to explore the significance of context—social, cultural, political, economic, and historical. The qualitative studies complemented the survey findings by directly providing explanations for them, as well as sometimes confirming, sometimes contradicting, sometimes illuminating those findings. It also suggested new survey questions that should be asked. Conversely, the survey data and analysis suggested questions to be asked or prioritized in the qualitative research.

The qualitative research not only contributed different methods for triangulation of results, but also enabled a social analysis that complemented the economic analysis in the evaluation. As noted above, the economic analysis was better at establishing rates of service participation, changes in education, health and nutrition indicators, levels and types of household consumption, and other quantifiable variables. The social analysis enabled us to understand the reasons why people do or do not participate in the education, health and nutrition services, why we sometimes did not see impacts even where people appear to participate, and the mechanisms through which impacts took place. It also allowed us to analyze impacts of the program on social relationships, and the relevance of social relationships to explaining program processes and outcomes. It allowed us to explore attitudes, culture, politics, and the local meanings that people give to different aspects of the program and the effects on outcomes. By focusing on people's actual lived experience, qualitative methods enable a richer understanding of the meaning that people give to events, processes and structures in their lives.

A basic tension running through social analysis is that of 'relativist' versus 'universalistic' interpretation, the former being an understanding of the world as seen from the respondents' (e.g. the beneficiaries') viewpoint, and the latter being an interpretation of the same world as seen by an external observer. While both the survey and qualitative methods provide the latter, the ethnographic methods are particularly good at taking advantage of the intimacy of the relativist vision. Central to this model is recognition on our part, as evaluators and social scientists, that views, opinions and interpretations of the program held by beneficiaries are important, credible and worth

[14] Throughout this paper, "we" usually used to refer to the team responsible for the qualitative research, but sometimes refers to researchers in general.

listening to. Ultimately, even where we might not believe these local perceptions are 'correct,' they usually have a profound impact on program outcomes. Examples of this are provided later in this paper.

The qualitative research in the Nicaragua and Turkey CCT studies followed similar design principles. The studies aimed to achieve some geographic diversity (regional and/or rural/urban); capture ethnic or religious diversity; include the views of household members of different age, sex, and role in the family (mothers, fathers, young children, adolescents, aunts and uncles, grandparents); gather both individual and group-based responses; obtain the perspectives of a wide range of stakeholders; use mixed qualitative methods, including semi-structured in-depth interviews, participant observation, and focus groups. Both studies drew most heavily on ethnographic methods used in the course of conducting community and households case studies.

Community and household case studies

In both studies, three Nicaraguan or Turkish field researchers, with B.A. or M.A. degrees in sociology or anthropology, conducted research in two communities each (for a total of six communities in each study) over a period of 4-5 months, moving between them at different intervals, and residing with families in the communities.[15] The field researchers were supervised by a senior sociologist and anthropologist[16]. The case studies drew primarily on ethnographic research methods, supplemented by other methods. Ethnography involves the immersion of the researcher in the everyday life of the people or group being studied, providing detailed descriptions and interpretations, with a focus on the *interactions* between different aspects of the social system under study. It employs a number of different research methods in combination, including participant observation, in-depth interviews, and informal conversations. The ethnographic case study approach is particularly suited to gaining a more nuanced understanding of the program's relationship to beneficiaries and nonbeneficiaries from *their* point of view. Sometimes referred to as the "extended case method," it uses intensive interactions and participant observation to understand everyday life, using a reflexive model of science that stresses engagement rather than detachment, establishing "multiple dialogues to reach an understanding of empirical phenomenon" (Burawoy 1998).

A key feature of our research that distinguishes it from more general forms of ethnography is the use of "household-level case studies," and the particular focus on all interactions relating, directly and indirectly, to the conditional cash transfer program. As noted by Mitchell (1987, *italic commentary inserted*): 'What distinguishes case studies from more general ethnographic description is the detail and particularity of the account. Each case study is a description of a specific configuration of events *[in our case, events related to the CCT program]* in which some distinctive set of actors [*mainly household*

[15] In Turkey, one fieldworker whose family was from the region lived with her family. She was also still working on her B.A. at the time of the study.
[16] The author and Terence Roopnaraine, respectively.

members] have been involved in some defined situation *[as beneficiaries or nonbeneficiaries]* at some particular point of time.'

An important element of this case study work is residential fieldwork: researchers live in the study communities for extended periods (in this case, several intervals of several weeks at a time) while they carry out their research.[17] This has many benefits that are unique to this approach: First, it allows the researcher to establish a level of rapport and confidence with households that is simply impossible with other research methods where the researcher is present only for a short time, e.g. a day or a week. For example, in the Turkey CCT study, the researcher located in one culturally conservative and politically volatile region had to spend a month in her study communities, in one helping with work in the fields, before people would begin to talk with her about the CCT program. This level of rapport translates into more reliable, candid and deeper data. Topics that are otherwise difficult to approach become accessible. Initial responses to questions may later be changed as the researcher gains more confidence and respondents become more relaxed. Second, residential fieldwork permits better triangulation and comparison of responses from respondent to respondent. Interviewing multiple family members offers a range of perspectives on the program, along both age and gender axes. Third, multiple visits to study households allow the capture of data at different points in time, rather than the snapshot provided by a single interview.

Case studies were based on a staggered series of household visits, done at different times of day and on different days of the week. During these visits, semi-structured interviews were carried out with different members of the household, capturing variation in age, sex and relationship. Interviews used guides or 'research checklists' reflecting the research questions, and designed to provide ample room for the exploration of emergent topics of interest and for follow-up questioning. Household visits provided an opportunity for direct observation of household and community dynamics and selected program-related topics.

Observation of activities at the household and community levels was also a key method in the research, both participatory—e.g. where researchers helped in the fields, shopped, or prepared meals and ate with household members) and non-participatory, e.g. where researchers observed without engaging in any activity. This allowed the observation of practices, behaviours, and interactions that confirmed or contradicted what people said, or revealed things that people had not mentioned. Participant-observation fieldwork (of which the ethnographic case study method is also a sub-category) has been a cornerstone of anthropological and sociological research since its early stages.[18] Having argued that such fieldwork requires a substantial investment of research time, Bernard identifies five important reasons for 'insisting on participant observation in

[17] For academic research in anthropology or sociology using ethnographic methods, four to six months would be considered a short period of residential fieldwork, not acceptable in some contexts such as thesis work. However, for program evaluation it provides considerable depth of information. While a longer study would provide additional information, e.g. based on seasonality differences or irregular program-related activities, limited timeframes and budgets means that some trade-offs will be necessary.

[18] The professionalization of anthropology and the beginning of the fieldwork era is usually dated to 1922, when the first truly field-based monograph, *Argonauts of the Western Pacific* was published by Bronislaw Malinowski.

conduct of scientific research among cultural groups' (Bernard 2002): First, participant observation opens things up and makes it possible to collect all kinds of data (which would be otherwise inaccessible). Second, it reduces the problem of *reactivity*—of people changing their behaviour when they know they are being studied. As you become less and less of a curiosity, people take less and less interest in your comings and goings (including interviewing, making observations and other research activities). Presence builds trust. Third, participant observation helps the researcher ask sensible questions in the native language. Fourth, it provides an intuitive understanding of what is going on in a culture and allows one to speak with confidence about the meaning of data, allowing one to make strong statements about cultural facts collected. Fifth, it enhances the internal and the external validity of what is learned from interviewing and observing.[19] Bernard concludes that 'many research problems simply cannot be addressed adequately by anything except participant observation' and '...getting a general understanding of how any social institution or organization works...is best achieved through participant observation' (Bernard 2002:335).

With a specific focus on program-related activities and interactions, wherever possible the researchers observed and recorded activities such as interactions between household members; care of children in the household; meal preparation; health and hygiene practices; shopping and other market activities; gatherings and other interactions among community members (including beneficiaries and non-beneficiaries); health service delivery; school activities; interactions between beneficiaries and program officials; interactions at pay points (surrounding delivery of the transfer); health and nutrition workshops.[20] The case studies were then supplemented with other research methods

In-depth Household Semi-structured Interviews: In addition to the more in-depth case studies carried out over time, additional beneficiaries and non-beneficiaries, usually mothers, were interviewed in order to capture the experiences of a larger number and wider range of people that could be covered in the more time-consuming case studies. These interviews took advantage of the trust the researcher had gained through their extended stay in the communities, as well as the efficiencies of conducting shorter once-off interviews (lasting from one to several hours) with a larger number of respondents. Furthermore, both studies had a mid-fieldwork break for data analysis, which revealed priority issues and some new topics of interest identified by the researchers or the respective country program officials. These priority issues could then be explored with a larger number of households in the second phase. For example, in Turkey, following survey results showing that girls' secondary school enrolment rates were still very low despite the CCT (and boys' as well) government policymakers and program implementers were particularly interested in understanding the reasons. Although this was already a focus of the research, the semi-structured interviews enabled us to add the perspectives and experiences of another 46 households to the 41 that had been included in

[19] Because of this strength of understanding and intuitive understanding developed, field researchers continued to play an important role in the analytical stages of the study after the fieldwork period had closed.

[20] Activities observed varied across the two country studies, depending on their relevance within the local context. For example, the health and nutrition workshops only existed in Nicaragua.

the case studies. In Nicaragua, 60 semi-structured interviews were added to the 60 case studies.

Key informant interviews: Many stakeholders have a significant influence on program outcomes, and from their particular vantage point have key insights into processes and impacts with respect to the CCT program. Interviews with these key informants can be particularly revealing, providing new perspectives and revealing information one would not get from beneficiaries. Those interviewed in Nicaragua included promotoras, program management personnel, teachers, health workers, religious officials, and community leaders. Those interviewed in Turkey included Foundation staff,[21] health, education and other service providers, imams (religious leaders), muhtars (local government officials), and other government officials at sub-provincial level. Semi-structured interviewing techniques were used, using the same questions, wherever relevant, that were in the case study 'checklist' to ensure that parallel sets of issues were covered. Most categories of key informants are identified during the study design phase, but individuals were added using a 'chain sampling' method where key informants identified other people of relevance. When needed these informants were interviewed several times, formally or informally, where the interviewer used information from other interviews to clarify or deepen responses, confirm, contradict, or interpret findings from other data sources.

Focus Groups: An advantage to focus group methods is that comments from group participants can trigger recollections and opinions from other participants that might not otherwise emerge. In addition, focus groups enable a larger number of individuals to be interviewed in a shorter period of time than do individual interviews. Focus groups can also be used to confirm or probe, with a larger group, responses received from individual interviews or observations. A possible disadvantage to focus groups relates to the fact that some individuals may be less inclined to speak out due to a variety of social dynamics within the group. Thus, an interviewer must encourage individuals that appear less inclined to speak, hold a minority opinion, or represent a particular social group with different views. For these reasons, and because some of the issues we were exploring were sensitive, focus groups were only carried out in selected circumstances, such as with informal groups gathered in households, or with groups of service providers or government officials.

Site and household selection: Using purposive but systematic criteria for selection sites and households

The qualitative research in these evaluations could not have "representative" samples, because the cost and time involved with qualitative research makes the sample sizes needed impossible. However, it is still important that the sampling procedures for selecting communities and households be done systematically, with careful consideration of criteria and stratification. In both the Nicaragua and Turkey qualitative studies this was

[21] The Foundations refer to the SYDVs (Social Solidarity Foundations) located in each province and sub-province, the local branches of the Social Solidarity Foundation (SYDTF), one of the two main government institutions coping with social risk mitigation, established in 1986.

done, with survey data was used to stratify and select localities and households for the qualitative research.

In Nicaragua, eight communities were selected for the study, across the Matagalpa and Madriz regions. The main study included six of these communities where the program existed. Additional research was then conducted for a short time in two 'comparison communities' where the program did not exist. The six intervention communities were selected according to a set of basic guidelines. These were: 1) participation in the pilot phase of the program; 2) physical safety of fieldworkers; 3) sufficient population; 4) representation of both geographical targeting (four communities) and household targeting (two communities); and 5) reasonable accessibility to Managua.[22]

An average of 20 households were studied in each of the six study communities, for a total of 120 households in the study. Since these communities were small, 20 households represented at least 10% of beneficiary households in all cases. As will be explained below, these households were also stratified to represent a cross-section of the community, using categories of interest to the study. In order to make the selection of households for the case studies more systematic and more closely linked to existing quantitative evaluation data, household selection was stratified according to a set of categories based on their situation as measured at the start of the program. The households selected were stratified across several categories. The first was age of children, so that we included households that had children aged between 0 and 5; households with children aged between 6 and 11; and households with children in both age groups. The second category was health status, so that we had households that entered the program with better health, i.e. all children under 5 years above the 20th percentile in height-for-age z-scores; and households that entered program with worse health than those above. The third category was education status, so that we had households that entered the program with better education, i.e. all children 7-13 years were enrolled in school; and households that entered the program with worse education (children not enrolled). We also applied some secondary stratification criteria in order to understand different types of households and situations, selecting some households with a male beneficiary; some with no children, some households no longer in the program (expelled or withdrawn voluntarily); and unselected households (non-beneficiaries) in household-targeted areas. Close collaboration with the quantitative team was particularly important for the site selection, because existing survey data was used to identify candidate households based on these criteria. However, once in the field, the field researchers had to revise the household selection to some extent based on some differences in actual household conditions.

[22] Accessibility to a major city is not normally a recommended site selection criterion, and will bias the results to some extent. However, since the sample was very small and would always be missing some variation, and based on our knowledge of the different regions, we determined that this would not be that significant in terms of our findings. At the time of our research, before the program expanded, the vast majority of intervention communities which fulfilled the much more important criterion of having participated in the pilot phase of the program were in any case quite accessible to Managua, so the accessibility criterion was not in the end very applicable.

In Turkey, the ethnographic research was carried out in three provinces[23]. Two localities were selected in each of three provinces, for a total of six communities. The following criteria were developed for site selection. Localities selected should be: 1) included in the quantitative survey, to enable the use of quantitative data to select households based on outcome variables derived from survey results, e.g. households performing well or poorly in terms of key impact variables; and to enable comparison of quantitative and qualitative data on these households; 2) from provinces with high levels of poverty and those identified by the government as high priority areas: this resulted in the selection of Eastern Anatolia; Southeastern Anatolia; and Black Sea; 3) capturing geographical and ethnic diversity, including rural and urban areas, and large Kurdish populations (where poverty is concentrated); 4) those with a relatively large number of CCT program beneficiaries included in the quantitative survey in order to enable a large enough sample of qualitative household studies and to ensure the selection of areas with high levels of poverty and thus of greater significance to program operations; 5) within reasonable distance from each other within each province, allowing the field researchers to travel regularly between them.

In each of these communities, the quantitative survey data was used to select beneficiary and non-beneficiary households, stratifying on the basis of high and low performance on selected health and education indicators. The indicators included school enrolment, drop-out rates, and vaccinations, the best survey variables available for making this assessment. In practice this was difficult for two reasons: the first was that households often had some positive and some negative indicators (for reasons that became interesting research findings; see Adato et al. 2007, chapters 4 and 5). We selected as many households as possible with clear positive and negative performance, and used a finer level of purposive selection among the mixed cases to capture a diversity of circumstances. The second problem was that, while we verified the demographic composition and health and education status of each household by reviewing household files at the Foundation offices, the survey data often did not match records in the Foundation offices.[24] Field researchers thus had to do final selections once they had visited homes. This process of household selection was thus very time consuming, but ultimately worthwhile. The purpose of this stratification was to gain an understanding of the conditions, practices, events, and perceptions characterizing households with different outcomes on key variables of interest to the research. Within this sampling design, we also purposively selected on secondary criteria, depending on options available: in all households, we selected those with at least one girl, but wherever possible we selected households with both girls and boys, and as many children as possible, particularly those of secondary school age. We selected households with different ethnicities, though in Diyarbakir and Van a high proportion of households were Kurdish. In Diyarbakir our field researcher was Kurdish and therefore fluent in the language. In Van, in households

[23] In 2005, a 'First Qualitative Assessment" (Kudat 2006) was carried out, with a different set of objectives to the "Second Qualitative and Anthropological Study"(Adato et al. 2007) that is the subject of this paper. The first study used key informant interviews, rapid assessment techniques, and focus groups, covering more regions and localities (15 of the 81 provinces, and 87 localities) but with less depth than the second study. The objective of the furst approach was to provide rapid feedback to policymakers.

[24] This may have been due to inaccurate reporting in the applications or in the survey, as well as different conception of the definition of the household.

where women spoke only Kurdish, daughters mostly translated for mothers. The case studies were begun in the first phase of the fieldwork and continued throughout. Households for the semi-structured interviews were selected in the second phase, using the same criteria.

In total 87 households were included in our sample. Of these, there were 41 full household case studies, and 46 households in which semi-structured interviews were conducted with one or more household members. Within these households 138 adults and 52 children were interviewed. Additionally, 33 key informants were interviewed, individually or in small focus-groups.

5. Exploring the benefits of mixed method research and the contributions of ethnographic approaches in program evaluation: Selected research findings

The use of ethnographic methods to explore issues not explored by the survey, and to find explanations for survey findings, provided insights into social, cultural, and institutional issues, and had significant policy implications (see section 6 below for discussion of policy impact). Since this paper focuses on methods, only a short summary of a few research results are included, to provide examples of the purpose of combining methods. These examples have been chosen to illustrate different types of contributions that can be made by the qualitative research. Full results of the qualitative research in Nicaragua can be found in Adato and Roopnaraine (2004) and in Turkey in Adato et al. (2007). A summary of integrated quantitative and qualitative findings from the Turkey CCT evaluation can be found in Ahmed et al. (2007). There is no integrated report from the Nicaragua evaluation, though some integrated findings are reported in Maluccio et al. (2005).

Selected Findings from the Nicaragua CCT Evaluation

Targeting: Nicaragua's RPS used two targeting approaches for selecting beneficiaries: In most *comarcas* in the first phase, where 80% of households fell below the poverty line, all households were eligible for the program (about 6% were later excluded due to their resources). In a smaller number of *comarcas* where poverty rates were lower, household eligibility was assessed with a proxy means test that identified households above and below the poverty line. In these the average poverty rate was 75%, so 25% of households were excluded, although the children of these households were offered access to the program health services.

Results from the quantitative study found that the RPS program was well-targeted, with 81% of the beneficiaries falling into the poorest 40% of the population. In *comarcas* where geographical targeting was employed, almost all poor and extremely poor households received the benefits, indicating negligible (less than 5%) undercoverage (households that were defined as poor but did not receive benefits). Conversely, only 15% of households in geographically-targeted *comarcas* defined as non-poor received the program benefits, referred to as 'leakage.' (Maluccio et al. 2005). In household-targeted

comarcas, undercoverage was estimated as 3% and 10% among extremely poor and poor households respectively, while leakage was similarly estimated to be 17% and 6% respectively.[25]

Findings from the qualitative study illustrate a fundamental difference between surveys and ethnographic approaches to data collection: while the qualitative findings do not conflict with the numerical results themselves, they do help us to better understand that behind these percentages lie individuals and families who literally *live* the impact of even statistically small targeting problems. The ethnographic research found that targeting was a poorly-understood element of the program: in particular, very few people locally understood the basis for the household targeting, and why they were included or excluded. One of the most difficult concepts for people to understand was the means test: in all study communities, respondents asserted that 'we are all poor here,' and did not perceive the economic differences defined by the targeting system. Even communities where geographical targeting had been employed were not immune to these concerns: because the *comarcas* used by the program to delineate intervention zones were not always coextensive with de facto community boundaries, not all the households in a community were included in the early incorporations. Across all the study communities, the perception of errors of exclusion was widespread, and a source of stress among beneficiaries and non-beneficiaries alike. Thus surveys may establish a finding of successful targeting based on certain 'objective' criteria, but people's reception of those outcomes can differ substantially. Such local opinions matter, because they shape people's attitude toward the program. The ethnographic study identified this issue and some social tensions generated as a result, as indicated by this beneficiary: "*Some of them get very angry when they give us the money because they say that they only give it to us and not to them."* Tensions of this sort arose in most of the study communities, though in only a small number of cases, because most of those who perceived themselves as wrongly excluded did not blame the beneficiaries. It is still important, however, to be aware of how household targeting can create a new type of social differentiation that may have subtle impacts social capital.[26]

A related issue identified in the ethnographic study was an effect on schoolchildren, where some were receiving assistance for uniforms, backpacks and supplies, and others were not. While in theory non-beneficiary households should have enough resources to buy these items for their children, in practice they may not, either because they were non-beneficiaries by error, or they might not have not have the resources or otherwise the inclination to so do. While non-beneficiaries said less about their own responses to exclusion, they were more expressive about the impact on their children: "*One day my son told me that a boy (he didn't say his name) told him 'look, I have a new back pack and you don't', and he started showing him all the new things he*

[25] Note that figures for geographically-targeted *comarcas* were precisely calculated using a baseline survey; because this data does not exist for household-targeted *comarcas*, leakage and undercoverage figures have been estimated using a formula (see IFPRI 2002:28 for further explanation).

[26] See Adato 2000 for research results from the CCT program PROGRESA in Mexico, where targeting-related tensions were more widespread and created more serious divisions in some communities. We did not find such strong findings in Nicaragua, possibly because there were fewer people excluded (Mexico later did a new incorporation and reduced the number of exclusions).

had in his back pack." An interesting finding, indicating how people felt about the beneficiary/non-beneficiary differentiation, is that in two of the six communities, a collection was taken up, in which beneficiary families were all asked to contribute some funds for the purchase of school supplies for non-beneficiary families.

Iron Supplements: The preceding example illustrated the complementarity of applying qualitative and quantitative approaches to the same research issue in order to generate a more holistic and multi-dimensional understanding of the issue. In the following example, qualitative methods were used to explain a survey finding: The survey found that the percentage of children receiving iron in the previous four months increased massively with the program—from under 25 percent to nearly 80 percent. Nevertheless, the very high anemia rates in this population (about 30 percent) were undiminished, i.e., no change was observed (Maluccio and Flores 2005). When asked in the qualitative study interviews whether they gave supplements to their children, a substantial majority of parents said (as in the survey) that they did. However, these assertions contrast with direct observations made by fieldworkers, where across the 60 case study households, only 3 were observed to be doing so[27]). It is highly likely that a higher number than those observed actually do give their children the supplement; however, it was also clear that many do not. Furthermore, in interviews, parents explained reasons why, mainly pertaining to different reasons that their children did not like the supplements. In particular, mothers said that the children do not like the taste of the iron supplement, and that it adversely affect the children's stomachs, and sometimes they throw up, or get diarrhea: "*at the beginning it was bad for him because it gave him diarrhea and made him feel sick, but since they say it is good for them, I kept giving it to him. However, it was also bad for his teeth, now his teeth are damaged.*". In some cases, parents also gave the iron to older children.

Meeting program 'conditions' through overfeeding prior to weighing: Qualitative methods can also be applied to a research question that would simply not be accessible to a survey approach. This might be because an issue is not amenable to quantification, because it is a delicate issue that requires time and rapport to reveal, sometimes because it is about a practice that people know is against the rules or otherwise looked down upon. The question of overfeeding children prior to weighing them is one which falls into both of these categories. On this issue, there are not quantitative findings with which to compare the qualitative: In the first phase of the CCT program, one of the conditions was that children gain weight. If they twice fell below an established rate of weight gain, parents could be sanctioned by suspending benefits. Although the weight gain requirement was dropped in 2003, and was not formally in practice in the communities in the qualitative study, we found that in all these communities many still believed that this requirement was still in effect. This may in part be explained by the fact that the research took place when the program was in transition and there was a lengthy break in the health service delivery, though the change does not appear to have been effectively communicated when the health services resumed.

[27] These statements may reflect the fact that they have given them at some point, whereas they stopped due to problems they encountered. Also, in two of the communities, the health services were not provided during the fieldwork period, so that people did not receive the supplements.

As a result of this belief, beneficiaries were employing last-minute strategies to pass the weight gain test. In five of the six communities, some beneficiaries explained how they gave the child unusually large amounts of food and liquids on the day or days leading up to the weighing. Children and mothers described children being given large quantities of different types of food, and large quantities of liquids, in the days leading up to the weighing, either trying to achieve rapid weight gain or at least water weight. This finding had several implications. One if that there were significant gaps in program communication systems (an important policy change was not communicated). Another is that in addition to penalizing children who most need the benefits, a CCT weight gain requirement causes stress. Finally, the finding provided insight into the strategies that people employ in navigating social programs.

Gender relations and women's 'empowerment': Though these issues can be measured in surveys (Quisumbing and de la Briere 2000; Hallman et al. 2007), there are important aspects that better lend themselves to an ethnographic approach. The latter allows for the expression of people's perceptions and feelings about their own changing place in the world. It gets at the more subtle dynamics of gender relations, and through observation and extended inquiry, may pick up on dimensions of social relations that may be contrary to what they believe or are willing to acknowledge publicly. As discussed above, CCT programs explicitly or implicitly aim to alter gender relationships in several ways: by designating them as the program beneficiary, providing them with an independent source of income, offering them health and nutrition education, increasing girls education, and giving them new opportunities to leave the house and participate in program activities. The qualitative research explored women and men felt about these aspects of the program; whether and how they had changed women's power, status, self-esteem, and intra-household relationships. With respect to relations with male partners, it explored whether this new role led to new tensions or conflict within the household, and/or new decisionmaking roles or other indicators of improvements in women's status.

While only a few of the findings can be mentioned here, the research found women and men supportive of this prominent role for women in the program, because both saw women as more likely to make spending decisions that were better for the household and for children. It was seen as a women's program, which helps to explain why it was not a threat to men's masculinity. New resources in the household appear to have eased social tensions, rather than increase it. Women still adhere to cultural norms associated with securing consent of spouses before making certain purchases, or to general spending patterns that were recommended by the program (e.g. purchase of food). However, they were spending money independently and this was experienced as a source of power: *"At least at home... all of us mothers had a custom that it was men who ran things at home, that if they were the ones who earned the money they had to give us what we were allowed to spend. So we had to be asking for money all the time but not any more…now since they see that we are the ones who get that transfer and we buy what we need for the house they are getting used to that, and now… when they receive the week's transfer they give it to women and now we are the ones who do the shopping.*." Though program effects on gender relations and women's empowerment were more subtle than dramatic, and discourse around women's equality precedes the program, by its gendered design the program appears to have increased it. The time women spent together in

program meetings held by the elected community liaison or in the workshops increased their awareness of women's issues (e.g. women's rights, family planning

Selected Findings from the Turkey CCT Evaluation

Communications: One of the issues studied in the quantitative and qualitative research was program communications—how effectively the SRMP office and the local Foundations had communicated with beneficiaries, and how well aware they were of the program structure and conditionalities. The quantitative evaluation showed that the program had achieved education and health impacts. These are likely to have been greater, however, if there had not been substantial communications gaps. The quantitative and qualitative studies had consistent findings with respect to these gaps, with the quantitative providing the magnitude and the qualitative confirming the strength of the finding and providing explanations. The survey found about 90 percent of the education-beneficiaries and 87 percent of the health-beneficiaries claimed that no one informed them of the program rules. (Ahmed et al. 2006a). Well over half of the households in the qualitative study demonstrated lack of understanding of the conditions. Many of the households were unaware of the difference between education and health benefits, instead referring to the benefit as 'child money.' It also found that people were more aware of the education benefit than the health benefit. The qualitative research identified several explanations: a more detached and negative attitude toward the CCT among health staff than teachers; more contact and association between the program and schools than the program and health facilities (e.g. schools helped with collecting applications); the misconception that the health benefit is an 'immunization aid' only; a better public information campaign for education than for health; and people's generally more attentive attitude to education than to health care. The qualitative research also found sociocultural and class/status-based explanations for why people did not want to participate in formal health services.

Constraints on Schooling: The CCT program's foremost objective was to increase school attendance rates, for the poorest Turkish children in general and for secondary-school girls in particular. Regression analysis based on the survey data found that the CCT program raised secondary school enrollment for girls by 10.7 percent, a significant impact. Despite the program impacts, however, enrollment rates are still very low at the secondary level. The enrollment rate for girls of secondary school age (14-17) is 38.2 percent for beneficiaries, and 46.3 percent for non-beneficiary applicants (the control group). For boys the rates are 57.9 percent and 64.9 percent, respectively. In rural areas the problem is even more pronounced, at 30 percent and 20 percent, respectively. The CCT program had no effect on the progression of girls from primary to secondary school (Ahmed et. al. 2006a). In two of the three provinces where the qualitative research was carried out, Van and Diyarbakir, survey data, while not statistically representative at the provincial level,[28] shows that the rates of girls school enrollment was even lower than the

[28] Given the national sampling frame and smaller numbers at the provincial level, the survey data is not considered statistically significant at the provincial level, but the numbers are considered because they do suggest that enrollment rates are lower than the national average, which is likely.

national average: 64 percent in Diyarbakir and 15 percent in Van. In Van there are even substantial gaps in primary school enrollment for girls, at 81 percent (Adato et al. 2007).

Because of the importance of girls' schooling as a policy issue more broadly and within the CCT program in particular, and these surprising survey findings (even lower enrollment than expected), the ethnographic research focused on explaining the factors constraining girls education, despite the program. It found explanations for why the CCT did have an impact, particularly at the secondary level, e.g. how the money put both financial support and government authority behind the cause: "*Fathers generally do not want to send their daughters to school...Now I can say to my husband that the government is paying me money for my daughters and I am sending them."* However, the research encountered many more constraints that the cash provided by the CCT could not overcome. Most of these were sociocultural, articulated with financial and logistical constraints, particularly in the conservative provinces of Van and Diyarbakir. For boys, these included parents and boys doubt about the value of employment, particularly in the context of high unemployment as well as rural livelihoods, and in rural areas where there is honor associated with working on the land. For girls, work was largely seen as inappropriate, and even counterproductive with respect to their primary role as wives and mothers. The most significant constraints had strong gender dimensions: the primacy of marriage (which also has an economic dimension) and motherhood, and issues of honor, reputation and sexuality—the perceived threats to girls and their families' honor posed by boys at school and men on the street, if girls go to school after they have reached maturity, expressed here by one father in a village in Van: *"the girls have only their honor as a valuable thing in the village and it is my duty to prevent any bad words about that… No one sends their daughters to school anyway. Why should I send mine? They will look at them in a bad way."* A closely related issue identified as highly significant in the qualitative study was transportation constraints—secondary schools are often far from home, and transportation options are not trustworthy with respect to the issues of honor raised above. Physical safety in schools, and children's own preferences and performance in school, were also significant explanations for schooling choices. The cost of school expenses, and the broader state of poverty was also a major factor, and in this case the CCT is responsive. But where the other factors were strong, the cash could not compensate.

Pregnancy incentives: In light of the cash benefit that the program gives to pregnant women conditioned on check-ups, concern developed, particularly among some health providers and Foundation staff, that this component of the program might be creating incentives for families to have additional children, undermining their efforts to promote family planning. The government thus requested that the quantitative and qualitative studies investigate this issue. The survey data and regression analysis found that the CCT program had no statistically significant effect on pregnancy. Rather, it found that receipt of education or health transfers actually reduced the probability of a woman of child-bearing age becoming pregnant by about two to three percent. Besides the RDD estimates, the results of a multivariate regression analysis also suggest that participation in either health or education components of the CCT program has no statistically significant effect on pregnancy (Ahmed et al. 2007). The ethnographic research was well suited to explore this issue because of the sensitive nature of fertility

decisions. The qualitative research findings supported the survey results, and provided three explanations for why the program was unlikely to have an impact on pregnancy decisions. First, there are many sociocultural pressures on women to get pregnant—related to status, social expectations, and economics. If a household has another child it is much more likely to be for reasons other than a cash benefit. As one woman summarized this point: *"I don't think a woman can give birth to get money…If a woman gives birth, it is because first God, second her husband, and third her husband's mother want her to."* Second, many people recognize that it is hard to support many children when poor. Those who do not want more children feel strongly about it, and a small cash grant does not convince them otherwise. However, in some areas rumors had circulated that the size of the grant was much higher (as much as 30 times higher), which might have led to some people considering it worthwhile. Third, very few people understood the differences between education, health and pregnancy benefits and many were not aware that they would get money for being pregnant.

Health care: The national survey found that vaccination coverage was almost universal for health beneficiary and non-beneficiary households, but that the doses were not completed for a considerable share of children from both groups (they were completed for 84 percent and 82 percent of children, respectively) (Ahmed et al. 2006a). The qualitative research contributed explanations for these findings. On the one hand, people tended to see the CCT conditionality as vaccinations only, even though it is not. Second, vaccination was regarded as a potentially harmful practice—a belief resulting in part from people's observation of the fever which can be a side effect of some live vaccinations. The qualitative research also offered insights into why people do not necessarily participate in formal health care services, including the check-ups required by the program. First people employ a range of traditional healing practices at home and visit traditional healers, though such practices coexist pragmatically with "biomedical" responses, the latter being more widely applied in cases of "serious" illness. Generally people associated health care services with serious illness, and otherwise are not that interested in going to doctors. This is exacerbated by people's experience of poor treatment and lack of respect from health care professionals, as problems with language, and the issue of shame and body-centered embarrassment, where it is seen as improper or uncomfortable for women or girls to go to male doctors.

6. Conclusion: *Q-Squared or Big Q+small q?* Reflections on the Status of Mixed-Method Research for Social Program Evaluation, and other thoughts

Quantitative and qualitative research methods are being integrated in the large-scale evaluations of social protection programs commissioned by governments and donors. With respect to CCT programs, the evaluations in Mexico, Turkey, Jamaica, Colombia, and, most recently, El Salvador, have required qualitative methods, established in the request for proposals and the contractual deliverables. The evaluation of Brazil's nutrition CCT, the CCT in Honduras, and the first phase of the Nicaragua program did not require them. The fact that many have been integrated signals a recognition that the impacts of social policy will be mediated by social and institutional processes and relationships, and that understanding them will increase the chances of

achieving the desired results. Santiago Levy, former general director of the Mexican Social Security Institute, former Deputy Minister of Finance in Mexico, and central architect of PROGRESA, wrote that "combining quantitative and qualitative methods provides a rich source of information and a positive feedback loop among evaluation, program design, program operation, and program continuity" (Levy 2006).

There is still a long way to go, however, with respect to how mixed method research is appreciated by researchers, donors, and policymakers in social program evaluation. While quantitative methods are a given, qualitative methods are still not. Where they are used they are often underfunded so that they can not achieve the depth that is their strength, and where they turn up important findings these are often overlooked, to the detriment of the programs that could learn from them. In the Third International Conference on Conditional Cash Transfer Programs held by the World Bank in Istanbul in June 2006, out of over 45 presentations on CCTs, only two specifically reported on qualitative research findings (Ahmed et al. 2006b[29]; Adato 2006b), and the session on how to do evaluation did not include qualitative methods (see World Bank 2006).

An example of missed opportunities through the exclusion of qualitative methods can be seen in the evaluation of Brazil's CCT program Bolsa Alimentacão, where IFPRI researchers considered including a qualitative component and discussed it with the government, but it was not prioritized. The quantitative research then found a small *negative* effect on children's weight gain, and the researchers speculated that this was due to a perverse incentive: "there have been anecdotal—and impossible to substantiate—reports of beneficiary mothers deliberately keeping their children malnourished to qualify for the benefits" (Morris et al. 2004:2340). This is a critically important issue, and it could have been substantiated or refuted, with well-designed qualitative research. This issue involves the type of behavior and beliefs that lend themselves well to being studied through the ethnographic methods described in this paper. Papers based on survey data alone sometimes attempt to explain reasons for survey outcomes, offering plausible hypotheses, whereas qualitative research could establish whether the hypotheses are correct or not. The World Bank and the Inter-American Development Bank, which have often been involved in the contracting of evaluations because of their role in financing loans for the programs, have shown an increasing appreciation of qualitative methods for CCT evaluations.[30] Whether or not qualitative research is included depends on the whether their value is understood by government officials and donors. Donors in particular can play an important role in advancing the use of mixed method evaluation.

It is important that qualitative studies are of high quality, demonstrating their value, in order to continue this progress. Superficial or ad-hoc approaches will be counterproductive. These include exercises that are labeled as qualitative research but have no systematic research design, applications of interview guides, or regard for confidentiality, and there is insufficient time in the field to establish rapport. Where good

[29] This one reported only a few results from the First Qualitative Assessment in Turkey. The Second study reported on in this paper had not yet been completed.

[30] For example, the recent Turkey and El Salvador CCT evaluations, in which the World Bank and IDB were involved, respectively, in defining terms of reference, both required qualitative studies.

qualitative and good quantitative research is being carried out, thee two are often not well integrated (see, for example, the separate quantitative and qualitative reports on gender issues in the PROGRESA evaluation in Adato et al. 2000). This is largely due to professional biases in the context of resource constraints—the tendency not to appreciate the '*other*'s methods as much as one's own, and in the context of time and resource constraints, not to prioritize integration (see Adato et al. 2007; Place et al. 2007). Furthermore, the disciplinary compartmentalization embodied in professional peer-reviewed journals provides disincentives for integrated publications. Even the terminology of quantitative and qualitative "components," as we as the researchers often refer to them in practice, signifies separation rather than integration. In the Nicaragua study, findings were integrated in the context of a six-page policy brief (Maluccio et al. 2005), but not in other publications or reports. In the Turkey study, the final evaluation report included quantitative and qualitative research findings by issue, with an effort made to relate the findings (Ahmed et al. 2007), but the study could have gone further in integrating issue identification and data analysis throughout the research and in the final product. In the Nicaragua and Turkey evaluations, the survey results were used to identify and prioritize some of the questions for the qualitative study, but much less the other way around. It was recognized that the qualitative studies had identified important issues with significant policy implications, and the prevalence of these findings could have been established through the survey. The ideal format for integration would be iterative stages of research and analysis, with qualitative and quantitative research each used for identification of issues to be investigated with the other method, and interpretation of the findings of the other method, in several alternating rounds. It is difficult, however, for a second round panel survey to integrate new questions, since by definition it must ask the same questions each time. But retrospective questions can be used for some issues, and additional questions can be added as single round survey questions.

In the CCT evaluations discussed in this paper, the quantitative and qualitative research team worked closely with the government in designing the evaluation, in aspects of its implementation, and in determining the implications of the findings for policy and program redesign. This runs the risk of a collegial relationship developing between the evaluators and evaluated, where it becomes harder to be critical and report negative findings. The quantitative and qualitative results from Nicaragua and Turkey demonstrate that such criticism is made anyway, but it is impossible to deny that it is harder, with a tendency for both parties to want to put things in as positive a light as possible, without violating professional standards and ethics. To make criticism constructive is not necessarily problematic, but there is a line that needs to be watched and sometimes drawn explicitly. The advantage of working closely with the government is that there is a receptive audience for the research findings—in most cases the evaluation is part of a loan contract and thus there is a structural incentive to at least explicitly consider the findings. Recommendations are part of the quantitative and qualitative reports, and both teams of researchers work directly with the government to find policy and program responses to the findings.

Whether the recommendations are followed, however, is highly contingent. It may depend on whether the recommendations are seen as feasible or desirable from a

technical, administrative, financial, or political standpoint. Generally speaking, however, quantitative findings are more systematically considered in the formulation of policy and program responses to evaluation findings. This is mainly because they are seen as 'representative,' whereas qualitative findings are not. If qualitative studies are well-designed, however, and their findings are strong enough and occur across the different sites and households included, there is usually a strong message that should not be ignored, e.g. in the nature of the constraints on girls education identified in Turkey, or the concerns about the social conflict related to the household targeting system in Mexico. Qualitative findings also can resonate with policymakers and program managers.[31]

The quantitative and qualitative studies in Nicaragua both had some influence on program design in later stages—although it is difficult to rigorously attribute impact—policy changes are often the result of the interaction of many factors. The survey results contributed to the program's continuation into a second phase, because the government had to show effectiveness before receiving the second tranche of the IDB loan for funding phase 2. The size of the transfer was somewhat reduced in the second phase, because of the large impact demonstrated in the first phase. The weight gain conditionality was discontinued in the second phase—although it was already seen as problematic, the survey data analysis showed it adversely affected poorer people. Finally, geographic targeting was used when the program expanded into the region of Wiwili. The quantitative analysis had shown that such a large percentage of people were poor that it did not make sense to target by household. Although this decision had largely been made before the qualitative results on targeting came out, a program official also cited the qualitative findings in explaining this decision. Furthermore, in response to the qualitative findings on targeting—particularly that which found that children in some non-beneficiary households experienced shame because they did not have the benefit of new clothing or the backpack of supplies—we recommended that non-beneficiary households also be given the backpack. The government did not adopt that recommendation, but they did give non-beneficiaries the *bono de la oferta*—the cash for the family to give to the teacher—including them at least in this way.[32] It is too early to know what if any recommendations the Government of Turkey will adopt in response to the quantitative or qualitative research in the CCT evaluation.[33] We did, however, work closely with the SRMP office in developing policy responses, for example, around the issue of transportation, improving communications, and better targeting.

[31] The economist Binayak Sen once summarized the complementarity of methods in this way: "numbers give one a feeling of facts; qualitative stories give one a feeling of truth" (Adato et al. 2007).

[32] As of this writing, the new government in Nicaragua had decided not to continue the CCT program, so we will not see further impacts of the evaluation in the foreseeable future.

[33] At the time this paper was written, the Turkey CCT is in a period of transition. Due to political and institutional issues, the old program run by the Social Risk Mitigation Program was closed in 2007 and a new program directly under the Prime Ministry, General Directorate of Social Assistance and Solidarity, is under development. While we had extensive discussions with the SRMP about policy changes to respond to our research results, we do not know how the new program administrators may use them. We do know they have been reviewing the reports..

With more resources, qualitative studies could provide greater benefits. More geographical coverage would increase the number of insights and understanding of regional diversity. They would also enable the use of panel approaches, where baseline research is conducted before the intervention starts. While qualitative methods can be quite good at making comparisons retrospectively because of the time and space available to explore issues in narrative form, probe, and enable recall, observing communities at baseline will provide better information. The budgets for qualitative studies are almost always much smaller than that for the quantitative; while this is justifiable in light of sample sizes needed, there is a cost to making the qualitative study too small. There are additional topics for qualitative research that have not yet been fully explored. Studies of policy processes, political economy, politics, and program operations can provide important insights that help to explain program outcomes. While some operational issues were examined in the Nicaragua and Turkey studies (and more so in Mexico; see Adato et al. 2000a), investigation of a wider range of issues would have been valuable. Study of policy processes in the context of CCT programs has yet to be carried out in the course of program evaluation. This is partly because of budget and time limitations, as feedback is needed as soon as possible. Policy processes are also less likely to be recognized by government as a priority, even if they would have important implications for policy and implementation.

One might be concerned that the use of ethnographic approaches, in particular residential fieldwork, is too expensive, but in fact it is probably not. The monthly cost of hiring BA and MA students in the program countries is normally quite affordable, and daily subsistence for living in communities is also low. More importantly, this approach provides employment and professional training to students from low-income countries, providing opportunities to earn degrees, and building long-term capacity for research. This can be accomplished while providing depth of insight in evaluation findings that can not be gained from other methods.

References

Adato, M., T. Roopnaraine, N. Smith, E. Altinok, N. Çelebioğlu, and S. Cemal. 2007. An evaluation of the conditional cash transfer program in Turkey: Second qualitative and anthropological study. Project Report. Washington, D.C.: International Food Policy Research Institute.

Adato, M., R. Meinzen-Dick, P. Hazell, and L. Haddad. 2007. "Integrating Social and Economic Analyses to Study Impacts on Livelihoods and Poverty: Conceptual Frameworks and Research Methods," in *Agricultural Research, Livelihoods, and Poverty*, M. Adato and R. Meinzen-Dick, eds. Baltimore: Johns Hopkins University Press.

Adato, M. 2006a. "Social protection for Children and the Implications of high HIV and AIDS prevalence for (Re)Structuring Interventions," XVI International AIDS Conference, Toronto, August 13-18, 2006.

Adato, M. 2006b. "Empowerment and Social Cohesion in Conditional Cash Transfer Programs," Presentation at the Third International Conference on Conditional Cash Transfers, Istanbul, June 26-30, 2006.

Adato, M., J. Hoddinott and L. Haddad. 2005. Power, Politics, and Performance: Community Participation in South African Public Works Programs. Washington, D.C.: International Food Policy Research Institute.

Adato, M. A. Ahmed and F. Lund. 2004. Linking safety nets, social protection, and poverty reduction: Directions for Africa. 2020 Africa Conference Brief 12. Washington, D.C.: International Food Policy Research Institute.

Adato, M. and T. Roopnaraine. 2004. *A Social analysis of Red de Protección Social in Nicaragua.* Final report to the Red de Protección Social, Ministry of Mi Familia, Government of Nicaragua. Washington, D.C.: International Food Policy Research Institute.

Adato, M. 2000. The Impact of PROGRESA on Community Social Relationships. Final Report to PROGRESA, Government of Mexico. Washington, D.C.: International Food Policy Research Institute.

Adato, M., D. Coady, and M. Ruel. 2000. An Operations Evaluation of PROGRESA from the Perspective of Beneficiaries, *Promotoras*, School Directors, and Health Staff. Final Report to PROGRESA, Government of Mexico. Washington, D.C.: International Food Policy Research Institute.

Adato, M. and D. Mindek. 2000. "*Progresa* and Women's Empowerment: Evidence from Six Mexican States" in Adato, M., B. de la Brière, D. Mindek, and A. Quisumbing 2000. The Impact of PROGRESA on Women's Status and Intrahousehold Relations. Final Report to PROGRESA, Government of Mexico. Washington, D.C.: International Food Policy Research Institute.

Ahmed, A., M. Adato, A. Kudat, D. Gilligan, T. Roopnaraine, and R. Çolasan. 2007. Impact Evaluation of the Conditional Cash Transfer Program in Turkey: Final Report. Washington, D.C.: International Food Policy Research Institute.

Ahmed, A., A. Kudat, and R. Çolasan. 2006b. Evaluation the Conditional Cash Transfer Program in Turkey. Presentation at the Third International Conference on Conditional Cash Transfers, Istanbul, June 26-30, 2006.

Ahmed, A., D. Gilligan, A. Kudat, R. Çolasan, H. Tatlidil, and B. Ozbilgin. 2006a. Interim impact evaluation of the Conditional Cash Transfer Program in Turkey: A qualitative assessment. Washington, D.C.: International Food Policy Research Institute.

Ayala Consulting (2006). *Country profiles,* report prepared for the Third International Conference on Conditional Cash Transfers, Istanbul June 26-30, 2006.

Behrman, J.R., H. Alderman and J. Hoddinott, 2004. "Hunger and malnutrition", in *Global crisis, global solutions,* B. Lomborg, ed. Cambridge: Cambridge University Press.

Bernard, H.R. 2002. *Research methods in anthropology: Qualitative and quantitative approaches.* Walnut Creek, CA: Altamira Press.

Brewer, J., and A. Hunter. 1989. *Multimethod research: A synthesis of styles*. Newbury Park, CA: Sage.

De la Brière B. and A. R. Quisumbing. 2000. The Impact of PROGRESA on Intrahousehold Decisionmaking and Relative Schooling Achievements of Boys and Girls in The Impact of PROGRESA on Women's Status and Intrahousehold Relations. Final Report to PROGRESA, Government of Mexico. Washington, D.C.: International Food Policy Research Institute.

Burawoy, M. 1998. The extended case method. *Sociological Theory*, 16 (1): 4-33.

Carney, D., ed. 1998. *Sustainable rural livelihoods: What contribution can we make?* Department for International Development: London.

Conway, T. and A. Norton. 2002. "Nets, Ropes, Ladders and Trampolines: The Place of Social Protection within Current Debates on Poverty Reduction," *Development Policy Review* 20 (5): 533-540.

Creswell, J. W. 1995. *Research design: Qualitative and quantitative approaches.* Thousand Oaks, Calif., U.S.A.: Sage.

Davis, B. and S. Handa 2006. The Experience of Conditional Cash Transfers in Latin America and the Caribbean. *Development Policy Review* 24(5): 513-536.

Denzin, N. K. 1978. "The logic of naturalistic inquiry," in *Sociological methods: A sourcebook,* ed. N. K. Denzin. New York: McGraw-Hill.

Devereux, S. and R. Sabates-Wheeler. 2004. Transformative Social Protection. IDS Working Paper 232, Brighton: Institute for Development Studies.

Greene, J. C., V. J. Caracelli, and W. F. Graham. 1989. Towards a conceptual framework for mixed-method evaluation design. *Educational Evaluation and Policy Analysis* 11.

Guhan, S. 1994. Social Security Options for Developing Countries. *International Labor Review* 133(1):35-53.

Hallman, K., D. Lewis and S. Begum. 2007. "Assessing the Impact of Vegetable and Fishpond Technologies on Poverty in Rural Bangladesh, in *Agricultural Research, Livelihoods, and Poverty*, M. Adato and R. Meinzen-Dick, eds. Baltimore: Johns Hopkins University Press.

Hoddinott, J.R. Behrman, J.A. Maluccio, R. Flores, and R. Martorell. 2007. Nutrition in early childhood and income of Guatemalan adults: A follow-up, more than twenty-five years later, of a community-randomized supplementation trial. Unpublished mimeo.

International Food Policy Research Institute. 2002. Sistema de Evaluación de la Red de Protección Social de Nicaragua: Evaluación de Focalización. Washington, D.C.: International Food Policy Research Institute.

Jick, T. D. 1979. Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly* 24.

Kudat, A. 2006. Evaluating the conditional cash transfer program in Turkey: A qualitative assessment. Project Report. Washington, D.C.: International Food Policy Research Institute.

Levy, S. 2006. *Progress Against Poverty: Sustaining Mexico's Progresa-Oportunidades Program*. Washington, D.C.: The Brookings Institution.

Maluccio, J. A., and R. Flores. 2005. *Impact evaluation of a conditional cash transfer program: The Nicaraguan Red de Protección Social*. Research Report 141. Washington, D.C.: International Food Policy Research Institute.

Maluccio, Adato, Flores and Roopnaraine. 2005. "Nicaragua: Red de Proteccion Social—Mi Familia. Breaking the Cycle of Poverty." Washington, D.C.: International Food Policy Research Institute.

Maluccio, J.A., and R. Flores. 2004. Impact evaluation of the pilot phase of the Nicaraguan *Red de Protección Social*, Food Consumption and Nutrition Division Discussion Paper No. 184. Washington, D.C.: International Food Policy Research Institute.

Martorell, R. 1995. Results and Implications of the INCAP Follow-Up Study, *Journal of Nutrition* 125: 1127S-1139S

Mitchell, J.C. 1987. Case studies. In Ellen, R. (ed.) *Ethnographic Research*. London: Academic Press.

Morley, S. and D. Coady. 2003. *From social assistance to social development: Targeted education subsidies in developing countries*. Washington, D.C.: Center for Global Development and International Food Policy Research Institute.

Morris, S., P. Olinto, R. Flores, E.A.F. Nilson, and A.C. Figueiró. 2004. "Conditional Cash Transfers are Associated with a Small Reduction in the Rate of Weight Gain of Preschool Children in Northeast Brazil," *The Journal of Nutrition* 134: 2336-2341.

Norton, A., T. Conway and M. Foster. 2002. Social Protection: Defining the Field of Action and Policy, *Development Policy Review* 20 (5): 541-568.

Place, F., M. Adato, and P. Hebinck. 2007. Understanding Rural Poverty and Investment in Agriculture: An Assessment of Integrated Quantitative and Qualitative Research in Western Kenya. *World Development* 35(2): 312-325.

Ravallion, M. 2001. The mystery of the vanishing benefits: An introduction to impact evaluation. *The World Bank Economic Review* 15(1): 115-140.

Rawlings, L. and G. Rubio. 2005. Evaluating the Impact of Conditional Cash Transfer Programs. *The World Bank Research Observer* 20(1).

Samson, M. 2006. *Are conditionalities necessary for Human Development?* Presentation at the 3rd International Conference on Conditional Cash Transfers: Conference Proceedings, Istanbul, June 26-30, 2006.

Skoufias, E. 2005. *PROGRESA and Its Impacts on the Welfare of Rural Households in Mexico*. IFPRI Research Report No. 139. Washington, D.C.: International Food Policy Research Institute.

Tashakkori, A., and C. Teddlie. 1998. *Mixed methodology: Combining qualitative and quantitative approaches.* Thousand Oaks, Calif., U.S.A.: Sage Publications.

Walker, S. 2007. "Child Development: Risk Factors for Adverse Outcomes in Developing Countries," Vol. 369.

World Bank. 2006. 3rd International Conference on Conditional Cash Transfers: Conference Proceedings. Washington, D.C.: The World Bank. http://web.worldbank.org/WBSITE/EXTERNAL/WBI/WBIPROGRAMS/SPLP/0,,contentMDK:20892674~pagePK:64156158~piPK:64152884~theSitePK:461654,00.html

World Bank. 2001. Project Appraisal Document (PAD) on a Proposed Hybrid Investment/Adjustment Loan in the Amount of $500 Million to the Republic of Turkey for a Social Risk Mitigation Project/Loan. Washington, D.C.: World Bank.