

The Evaluation of Cash Transfer Schemes in Africa

Meeting Report
September 2009
Washington D.C.

Candace Miller, ScD
Assistant Professor

**Center for Global Health and Development
Boston University School of Public Health**

This publication was made possible through support provided by the Office of HIV-AIDS, Bureau of Global Health, U.S. Agency for International Development, under the terms of Award No. GHSA-00-03-00020-00. The opinions expressed herein are those of the author and do not necessarily reflect the views of the U.S. Agency for International Development.

Thanks to all meeting participants for generously spending their time to participate in thoughtful discussions. The author takes full responsibility for the information as it is presented; however the content of this report is based on the combined contributions of meeting attendees.

Meeting Participants:

Michelle Adato, Harold Alderman, Charles Shapiro, Gretchen Bachman, Sarah Baird, Diana Chamrad, David Croome, Alan Debrauw, Rena Eichler, Dan Gilligan, Margaret Grosh, Lena Heron, John Hoddinott, Priya Kumar, Esther Lwanga, Amanda Makulec, Maury Mendenhall, Candace Miller, Morris Moono, Allyn Moushey, Harry Mwamlima, Patrick Premand, Laura Rawlings, Gelson Temblego, Thabo Thulo, Fabio Veras, Anna Williams, John Williamson, Rachel Yates

Contents

Contents	3
Executive Summary	4
Overview	6
Lessons from Latin America and beyond	6
Nature of evaluation	7
Program components and design features	9
Methods	11
Impact	12
Operations / Implementation	13
Targeting	15
Economic evaluation	15
Sharing evaluation findings	16
Conclusion	16
References	18
Appendix 1. Meeting Participants	19

Executive Summary

Countries throughout Africa are implementing and scaling up cash transfer schemes. The *Meeting on the Evaluation of Cash Transfer Schemes in Africa*, hosted by the Center for Global Health and Development at Boston University and supported by USAID, brought together donor agency representatives, policymakers and evaluators engaged in the implementation and evaluation of cash transfer schemes. The purpose of the Meeting was to review the evaluations of existing schemes and begin the conversation on 1) how to improve the rigor of cash transfer evaluations in Africa and 2) better connect evaluation results to policymakers and donor agencies, in order that evaluation results assist in evidence-based decision-making on the design and improvement of social protection policies.

Cash transfers schemes are implemented in over 30 countries as the cornerstone of poverty reduction and social protection efforts. Solid study methodology, rigorous analysis, and emphasis on high quality evaluations have enabled Latin American schemes to demonstrate outcomes, identify challenges in implementation, and document the need for linking programs to other safety nets. It is evident that some, but not all, lessons from the first generation programs are transferable to the African context. Still, by employing similar evaluation methodologies, African programs can develop their own evidence base on cash transfers as a social protection strategy.

The following recommendations emerged from the meeting discussions and are offered in order to stimulate high quality evaluation that improves programming. Policymakers and other stakeholders should:

- Choose design features (e.g. transfer size, conditions, linkages to other programs or services) in line with the scheme's goals and evaluate the impact of these components.
- Embed evaluation into the program design phase, in order to pair program goals with evaluation questions and maximize the utility of findings to policymakers.
- Select specific evaluation questions based upon the scheme's objectives, the political context, and the interests of policymakers, donor agencies, and other stakeholders.
- Utilize rigorous methods including combined qualitative and quantitative methods.
- Design evaluations with intervention and comparison groups that minimize selection bias by using randomization, regression discontinuity designs or matching methods.
- Utilize the difference in differences method for analysis to show program impacts among groups over time.
- Obtain adequate sample sizes and a selection of indicators in order to quantify impacts among sub-groups and answer nuanced questions about what works for which group.
- Acknowledge that some outcomes require longer term evaluations, in contrast to the one to three year interventions that have occurred.
- Conduct operational research to evaluate implementation, but with a focus on how to make the program viable, regardless of its current state.
- Given that many African countries have no vital registries or tax information for the poor, adapt targeting processes for the context, but examine these deficiencies to contribute to improved public management for all social protection and development efforts.
- Use economic evaluations to cost schemes and to answer questions about whether transfers yield better outcomes for fewer inputs than other programs.

For the larger community of policymakers, donor partners and international agencies:

- Initiate or expand efforts aimed at building evaluation capacity, including the ability to discern between high and low quality evaluations and how to understand results.

- Also, initiate or expand efforts at building a culture of evaluation within government and among other stakeholders so that there is support for evaluation practices.
- Develop a community of practice to help summarize and communicate evaluation findings, and ensure quality control in evaluation.
- Support government ministries in a systematic, harmonized manner in order to build administrative and management capacity to implement social protection strategies.

While it may be premature to draw the meta lessons on cash transfers across Africa, given the limited data that currently exists, attendees agreed that the evidence suggests that governments should implement cash schemes and in some cases, should extend benefits or scale up programs in order to confront poverty. Policymakers and other stakeholders have important questions and require high quality information in order to design and implement programs. Consequently, meeting attendees called for improved and consistent linkages between policymakers and evaluators and rigorous evaluation methods that yield the data necessary to build quality social protection system across Africa.

Overview

Countries throughout Africa are implementing and scaling up cash transfer schemes. The *Meeting on the Evaluation of Cash Transfer Schemes in Africa*, hosted by the Center for [Global Health and Development at Boston University](#) and supported by [USAID](#), brought together donor agency representatives, policymakers and evaluators engaged in the implementation and evaluation of cash transfer schemes. The purpose of the Meeting was to review the evaluations of existing schemes and begin the conversation on 1) how to improve the rigor of cash transfer evaluations in Africa and 2) better connect evaluation results to policymakers and donor agencies, in order that evaluation results assist in evidence-based decision-making on the design and improvement of social protection policies.

African countries seeking to implement social protection strategies face both an enormous burden of poverty, which requires that funds be used effectively; as well as weak institutions, limited infrastructure, and a chronic lack of resources, which are barriers to effective program implementation. However, evaluations can yield valuable data for policymakers as they design and implement quality cash transfer schemes. Ideally, evaluation findings can help programmers overcome institutional constraints and use limited resources effectively.

At the meeting, attendees reviewed social cash transfer schemes, the politics of schemes, and evaluation designs in Lesotho, Ethiopia, Malawi, Mozambique, Tanzania, South Africa, and Zambia. Attendees also reviewed past experiences with conditional cash transfer schemes in Latin America, where an important feature has been the utilization of evaluations that employ quantitative and qualitative methods to measure impact, understand operations, and examine various targeting methods. From these discussions, meeting participants were able to draw parallels from Latin America to Africa and discuss critical evaluation questions and methods, as well as how results fed into the policy process. Given that cash schemes have only recently been implemented in Africa, with varying degrees of rigor in their evaluations, there are many unanswered questions about cash transfers in the African context.

This report is not a literature review of findings from cash transfer evaluations, but rather serves as a document to assist policymakers, donor organizations, evaluators and other stakeholders as they think through the goals and programmatic components of cash schemes, and the evidence needed to make informed decisions. (*Evaluation meeting materials, including papers and reports from African schemes are available online at: <http://childresearchpolicy.org>*).

Lessons from Latin America and beyond

Conditional Cash Transfers (CCTs) are now present in over 30 countries and are often the cornerstone of poverty reduction and social protection efforts. Programs vary, but share the basic premise that cash transfers are targeted towards poor households. Progressa, Mexico's CCT, is one of the forerunners, initiated in the early 1990's. Since then, CCTs have become increasingly popular. In Latin America, cash schemes are conditional, so that recipients receive cash if they undertake specific actions or behaviors. The rationale is to achieve the dual objectives of:

- Reducing income poverty by providing immediate cash to families, and
- Facilitating human capital development

When cash is conditioned upon school attendance, then cash compensates wages from labor lost for children not working and provides a strong nudge to ensure that children attend school.

Throughout Latin America, evaluations of cash transfer schemes have been a regular component in CCTs and have demonstrated positive impacts on beneficiaries. Solid study methodology, rigorous analysis, and emphasis on high quality evaluations have enabled Latin American schemes to demonstrate outcomes, identify challenges in implementation, and document the need for linking programs to other safety nets (Fiszbein and Schady 2009).

Impact evaluations from CCT programs in Latin America and elsewhere have shown the following:

- Substantial reduction in consumption poverty, with the greatest impact in households where the transfer size is a larger proportion of the household budget.
- Substantial increases in utilization of health and education services, especially in places where there was a low baseline for service utilization, such as in Cambodia, Columbia, Pakistan, and Mexico.

It is evident that some, but not all lessons from the first generation CCT programs are transferable to the African context. While first generation cash programs occurred in rural areas, in low-income countries, they were implemented in a specific time and cultural context. Their evaluation results were derived from programs with reasonable administrative capacities, and where populations had access to existing health and education services. When the context changes, some programmatic aspects may be replicable, while others not. Nevertheless, by employing similar evaluation methodologies, African programs can develop their own evidence base for the use of cash transfers as a social protection strategy.

Nature of evaluation

Evaluation questions

Evaluations require political will and resources. The thrust of an evaluation depends upon program objectives and is contingent upon the political context of the cash scheme, which varies by country depending on goals, competing demands, and the interests of policymakers, donor agencies, and other stakeholders. For example, in Ethiopia, cash transfers were combined with public works. In 2005, the scheme was launched country-wide with approximately one million beneficiaries, making operational and implementation questions particularly relevant, as well as questions about the added benefit in combining cash and public works. In Malawi, the cash scheme grew from debate within the National Social Protection Technical Committee in a political process with frequent dialogue about who to target with limited resources. Documenting the scheme's impact was important to stakeholders as they solicited donor partners and sought approval from government and civil society. In Mozambique, understanding costs, targeting mechanisms, and performance in urban settings were of key interest to policymakers. In Zambia, policymakers were interested in the operations and impacts of different types of schemes, so pilot programs were evaluated to compare universal age-based targeting to community based targeting.

Governments and agencies requesting evaluations and evaluators must be clear and specific about which questions to answer, acknowledging that current evaluations typically focus on about five out of 40 questions. Evaluation topics include, but are not limited to, questions around impact, operations, targeting, costs, and feasibility; and programmatic components such as transfer size, conditions, and graduation. Political economy questions include: Should resources be targeted in order to obtain the largest impact, to reach the most vulnerable groups, or for ease of implementation? Policymakers may need evidence to better target the resource pot or grow the

pot with larger social protection budgets. Economic evaluations may ask how cash schemes fit with and compare to other existing programs in terms of impact, effectiveness, and cost given that they are single programs in larger dynamic systems with multiple components.

In determining evaluation questions, stakeholders should consider both the need for building a solid evidence base and selectivity, particularly in contexts with limited resources. An evidence base with external validity requires that evaluations ask similar questions in various settings. However, limited resources require parsimony in deciding whether the same type of evidence is needed in different countries. While context matters, there may be general evaluation questions that do not require replication. That said, the evidence base in Africa is still at an early stage, suggesting that evaluators and policymakers should use caution in generalizing findings across settings. A community of practice that exchanges information and considers these questions together could help stakeholders find the necessary balance.

Stakeholder capacity

Selecting evaluation topics, agreeing on methodology, and interpreting results is challenging, particularly when the evaluation capacity of stakeholders is limited. Moreover, evaluations are inherently political, with resources and reputations at stake. Powerful networks that want specific results may yield influence regardless of a study's rigor or the accuracy in the interpretation of results. Thus, quality control of evaluation is necessary to ensure rigorous methods and presentation of results. Building evaluation capacity within government and other stakeholders, particularly the ability to discern between high and low quality evaluations, and understand evaluation results, is important. International and donor agencies should consider investing in or expanding efforts to build evaluation capacity within governments and other partners, as well as their own agencies where necessary.

Who should evaluate?

Who evaluates is another important consideration, with different implications depending upon who conducts studies. There are both strengths and drawbacks in utilizing external, non-resident evaluators. Independent and external evaluators may bring greater capacity, and be less amenable to internal pressures than in-country residents, but they are likely to cost more and may be less knowledgeable about context and nuances. Realistically, in many countries, local research institutions require external partners to provide the technical capabilities needed for longitudinal field work in evaluation. In this case, partnerships between resident and non-resident evaluators may make the most sense, maximizing evaluation expertise and local knowledge while minimizing costs. These partnerships usually help to build in-country capacity for evaluation, which is an added benefit. Still, evaluation stakeholders must be open to partnerships, perhaps laying aside perceived political differences between the government and academic elite.

Evaluation Culture

Throughout the world, in comparison to other poverty and development programs with little evaluation, there has been a culture shift towards an expectation that cash schemes will be evaluated. However, the assessment of meeting attendees that have worked in multiple parts of the world is that this shift is occurring at a slower pace in sub-Saharan Africa, which is still behind other resource poor countries in its emphasis on good evaluation practices. For example, in Mexico, there was a political climate supportive of rigorous evaluation, such that the Progressa evaluation had a randomization component supported by the President. In contrast, in one African country, the president reserved the right to repress evaluation results depending upon findings.

This is illustrative of the absence of an evaluation culture in some governments, which may also extend to various stakeholders. It highlights the need to cultivate a culture of evaluation at very high levels, in agencies and government alike.

Oftentimes, governments agree to evaluation at the request of donors or as a donor requirement, rather than demanding evaluation results to inform internal decision making, frank policy discussions, and priority setting. The challenge is: How do stakeholders build greater demand for evaluation and a culture of evaluation among decision and policymakers in Africa? Government, UN and donor agencies, and universities must each play a role. The goal is to move towards the Latin American experience, whereby policymakers expect the assessment of programs, using short-term indicators, along with long-term monitoring of human development gains and losses.

Program components and design features

Policymakers must articulate their scheme's objectives and consider the programmatic components and design features that will yield the desired outcomes. Are benefits provided to fulfill the human right to social welfare? For immediate poverty alleviation or improved food security? Improved human development? These objectives require different programmatic components. For example, the size of the transfer, whether conditions are utilized, whether recipients graduate from schemes, and whether cash is linked to other programs all depend upon the scheme's goals. Different objectives yield distinct evaluation questions as well. A scheme designed to fulfill the right to social welfare might have an evaluation where the main study question is: Did the most vulnerable groups receive the transfer? The main evaluation question of a scheme focused on reducing hunger may ask: Was food insecurity reduced and does the impact depend upon who receives the cash? Determining whether programs impact human development require evaluation study designs with longer timelines than the standard one to three years, and the collection of short- and longer-term indicators of human development outcomes.

Programmatic components of cash schemes to consider include the following:

Size of transfer: Scheme designers must make critical choices about the size of the transfer per beneficiary versus the number of beneficiaries that can be reached. Regarding size, transfers could be set to pull households above a poverty line, for optimum performance in human development, or to break the cycle of intergenerational poverty, each of which has a different cost. Compared to cash transfers in other parts of the world, where the transfer size is below 35% of pre-transfer consumption, in some African countries, transfers valuing \$15 may be as much as 90% of pre-transfer consumption, given the stark poverty among recipients who were previously begging, foraging for food, and living separate from the cash economy. The tradeoff for a higher transfer—which yields impressive impacts—is that only the most destitute, labor constrained households receive benefits. In Malawi and Zambia, 10% of households receive transfers, but more than half the population lives below the poverty line.

Frequency of delivery: What should the schedule of transfer delivery be (e.g. monthly, bimonthly, or annually) and will varying delivery schedules yield similar impacts? Less frequent delivery reduces costs, but may also reduce impacts. Still, some expenses such as school costs are not uniform across the year, so varying the transfer size and delivery frequency might be helpful, particularly in large countries such as Ethiopia, where timely payments have been challenging.

Optimum conditionality: Should conditions be attached or do transfers fulfill a constitutional or human right? If conditions are used, what is the optimum conditionality? Should conditions be

soft or hard, with formal or informal monitoring? In African schemes, scarce resources may lead programmers to rely upon community monitoring or the “IRS” approach (i.e. living in fear of a random audit). Conditions could also be framed as incentives to elicit behavior with bonuses, rather than penalize non-adherents through the loss of benefits. Either way, schemes must balance conditions versus cost as they require a system to monitor adherence. What are stakeholders willing to pay to monitor conditions? It is also important to consider that some conditions, such as school attendance, have implications for other students. If schools are further overcrowded, increased attendance imposes a cost on the children already in school who now have a larger class size. Conditions on attendance may need to be paired with supply side interventions for schools.

Cash-plus: Another programmatic component for scheme designers to consider is: What is the optimum package, cash-plus what? The ‘plus’ may include supply side interventions or linkages to other programs. The first generation CCT programs determined that cash-plus services yielded benefits, but evaluations were unable to determine the source of impacts. In Bangladesh, cash and supply side interventions yielded increased school enrolment, while there was no impact where there were only supply side improvements.

Still, despite increases in service utilization (i.e. health centers and schools), these first generation conditional cash schemes have had mixed success in terms of improving final outcomes in education and health. For example, while cash transfer impacts include increased enrolment, there have been limited improvements in test scores. Realistically, cash transfers may influence utilization, but school performance outcomes may be confounded by issues such as poor school quality and unfavorable teacher-to-pupil ratios. It may take years for supply side interventions aimed at improving school quality to take effect.

Beyond the cash-plus programs in Latin America and elsewhere, in the African context, linking HIV treatment and cash schemes may yield positive benefits on ARV adherence. Cash schemes plus training in income generating activities and business skills may support long term economic strengthening of affected households. Alternatively, it may simply not be feasible for the poorest households in Tanzania or Zambia, to manage income generating activities. Depending upon the goal of the scheme (i.e. to fulfill a right or as an investment in human development and long term economic strengthening) policymakers should have different expectations for cash-plus programs.

As policymakers consider ‘cash-plus’ they must also address the availability of existing services, specifically health and education, given supply side shortages. They may decide to partner or contract with non-traditional service providers or existing NGOs to improve health centers or schools. Nevertheless, as cash plus programs are developed, they require evaluations that can disentangle sources of impact. In some cases, this requires intervention arms where cash households are compared to cash-plus households and other non-intervention households.

Graduation: Should households graduate? In Zambia, some households graduate with a lump sum designed to foster the investments needed to avoid destitution in the future. In Lesotho, the pension is lifelong after age 70. In contrast, in Malawi, whether graduation makes sense is still unclear. Given that program designers are looking for recipients to make long term investments in human development, agriculture, and business, then graduating may prevent this.

Cash as a component of Social Protection Systems: Throughout Africa, while policymakers and evaluators might try to frame cash schemes as programs, for the most part, they are still projects that are not yet embedded in any ministry or the structure of the state. Governments are only beginning to think in terms of a social protection systems. Thus, there are important questions around how programs fit together and complement or conflict, plug gaps or replicate services.

Methods

Ideally, policymakers and evaluators should work together from the onset, so that evaluators join the design team prior to implementation in order to pair program goals with evaluation questions and studies. This early collaboration will yield the greatest benefit to policymakers given their need for data on operational performance and impacts. The evaluation study design depends upon the prioritized evaluation questions, which should be articulated in the program's logical framework. A stakeholder analysis preceding the study can help clarify evaluation questions.

In Latin America, evaluations answered multiple questions, engaging policymakers at different levels. These impact evaluations yielded high quality data because they employed rigorous methods, mainly randomized, longitudinal studies that *causally* linked programs with impacts. These designs should also be applied to the evaluation of cash schemes in Africa, so that data is collected at baseline and other intervals using mixed quantitative and qualitative methods to estimate and describe impacts in intervention versus comparison households.

In *quantitative evaluations*, the comparison of similar intervention and control/comparison groups over time allows an estimation of the counterfactual or the quantification of what would have happened without the intervention compared to what did happen in intervention households. Evaluators seek to reduce *selection bias*, so that results are not biased by the mechanism used to select study groups. Identifying comparable groups is difficult because of operational limitations and ethical choices around who gets benefits when. Nevertheless, comparing similar groups (where one group has an intervention and the other does not) over time is the basis for high quality evaluation. Of note, in many countries, cash schemes are implemented incrementally, which allows evaluators to compare households in communities where schemes are implemented to households in communities where the scheme will be implemented in the future once the program rolls out. The different methods used to reduce selection bias should be based on the scheme, existing and available data, and resource limitations. Methods include the following:

Randomization: In the evaluation of *PROGRESA*, as part of the process of phasing in the program, localities were randomly assigned to receive program benefits starting in 1998 or 2000. Baseline and follow up data was collected from all households. Intervention households that received the transfer in 1998 were compared to control households that did not receive the transfer until 2000. This design takes advantage of the natural rolling out of the scheme and provided two years of impact data to assess. Ultimately, all study households received benefits.

Regression Discontinuity Design: The CCT in eastern Turkey was evaluated using this method. Regression discontinuity design works on the basis that there is a well-defined cutoff that separates those who receive benefits from those who do not. The method requires a hard distinction between groups. For example, an age cutoff allows the comparison of households headed by 40 year old beneficiaries to similar households that are headed by 41 year olds.

Matching methods: The Productive Safety Nets Programme in Ethiopia was evaluated using matching methods. Beneficiary households are "matched" to non-beneficiaries with similar observable characteristics. This method works best when there is some degree of program rationing or discontinuity design that allows people to be excluded, which yields similar households with groups that receive the intervention and groups that do not.

Once comparable groups are identified and baseline and follow up data is collected, the recommended method of analysis is the *Difference in Differences* or the *Double difference*

methodology: The double difference method estimates the differential outcome between the treatment and comparison group over time, removing the effect of observed or unobserved differences between groups at baseline. The double difference is the estimated program impact.

While resource limitations are palpable, the importance of obtaining adequate *sample sizes* to yield sufficient information cannot be overstated. In Ethiopia, large sample sizes allowed analysts to study nuances and tease out impacts on various sub-groups. In Malawi, a sample of 400 intervention and 400 comparison households was adequate to analyze a range of household level outcomes, but was too small to draw definitive conclusions about impacts on child growth among the sub-group of under five year olds, despite promising findings.

Qualitative analysis allows evaluators to confirm and explain quantitative findings on the scheme's impacts, operations, targeting, and other issues. Evaluators can examine, in detail, household decisions around the use of cash, implementer's perceptions of barriers to successful transfer delivery and other key questions. In Turkey, in-depth interviews exposed issues around school performance and safety. In Malawi, focus groups allowed children to express how transfers impacted their lives, and direct observation, record reviews, and key informant interviews allowed the government to make real time improvements in the scheme's operations.

Impact Evaluation

In both Latin American and African schemes, evaluators have assessed the impact of cash on a menu of 'global' individual and household level indicators across programs among various groups including, but not limited to women, children, orphans and vulnerable children, the elderly, the disabled, people living with HIV/AIDS and other chronic illnesses. Indicators included, but are not limited to the following:

Individual level indicators	Household level indicators
Child growth, health, vaccination	Food security, diversity, nutrition
Biological markers (anemia, etc.)	Asset accumulation
Early childhood development (cognitive)	Labor and productivity
Psychosocial wellbeing	Economic empowerment
School enrolment, attendance, performance	Inward and outward migration
Adult and child labor, productivity	Agricultural production
Use of health services	Response to household shocks
HIV risk, ARV adherence	Violence in household
Community level indicators	
Impact on non recipients	
Impact on businesses	
Impact on prices / inflation	
Multiplier effects	

Additionally, evidence to examine unintended consequences may be important for gaining political will and buy-in from civil society, donors and other stakeholders. Past evaluations used data for myth-busting to combat claims that cash transfers encourage pregnancy, violence, or the excessive purchase of alcohol and prostitution.

The indicators chosen to measure impact depend upon program goals and the needs of policymakers. Studies with a selection of indicators and adequate sample sizes allow analysts to examine heterogeneity and quantify differential impacts between sub-groups, to answer the questions: What works for which group? Where, how and why? From a policymaker and

programmatic perspective, these nuances are important for planning and linking services. The evaluation of first generation CCT programs yielded information on average impacts, but did not permit analysts to tease apart impacts from cash, conditions, or the combination, and whether impacts depended upon who got the income.

Recently, Latin American evaluations have yielded information on the differential impacts based on the transfer recipient's gender. In the African context, mothers or grandmothers may receive the transfer, with implications for how it is spent and the associated impact. While evaluations have documented that cash recipients tend to purchase food when they receive transfers, a more nuanced question is whether there are different average expenditures depending upon conditions, targeting, and the size of the transfer.

Additional evaluation questions for beneficiaries with policy and program relevant implications include the following:

- What is the household dynamic once money is received?
- How is money used differently over time?
- How are long-term reliable transfers and transfers that are viewed as short-term used?
- Do people spend transfer income differently than other types of income?
- Are there threshold effects?
- Are there varying impacts based on complex household structures?
- What happens when/if the predictable cash flow stops?

Stakeholders may also have questions related to the impact at the community level or impacts on the local economy. In Malawi, a wider evaluation of the impact of cash on economic development identified and quantified impacts on indirect beneficiaries and local businesses.

It is important to acknowledge that some outcomes require longer term evaluations, in contrast to the one to three year interventions that have occurred so far. For example, quantifying human development gains, the ability of households to withstand seasonal or other shocks or the impact of cash transfers on local markets may require evaluations with a timeline of five to ten years.

In most cases, existing evaluations have examined impacts in order to determine whether the cash scheme is a good program. Moving forward, in the second generation programs, policymakers have a fundamental interest in operations research and evaluation in order to determine how to make the program good, regardless of its current state.

Operations / Implementation Evaluation

Compared to Latin American systems, the government structures and ministerial capacity in many African governments is weak, even before the implementation of cash schemes. Still, despite limitations, African governments are moving to implement cash schemes. Thus, government ministries require systematic, harmonized support from donor partners and agencies in order to build administrative and management capacity. Questions for past and future operations evaluations include the following:

Quality of Management: Is there an organized system, with accountability, quality and back up plans? Is the personnel and human capacity adequate (e.g. staff size, skill matrix)? Where are skill and training deficits? Is staff in log frame devoting appropriate time to scheme?

Institutional roles: How should various institutions be coordinated for optimum implementation? What are the roles of Government, NGOs, UN agencies, donor partners, foundations?

Technology/Equipment: What is required for scheme management and communication? What was purchased? What is/is not working? What are the deficits?

Appropriateness of Costs: What are the start-up costs vs. the costs of running the scheme? How do costs compare to the quality of management (systems, staffing, technology, transparency, internal monitoring and control, communication, targeting accuracy)? What is the priority of the activity vs. budget allocation? What is the appropriateness of expenditures versus returns? What are the planned versus actual costs? Is the monitoring of costs (accuracy, follow up) appropriate?

Level of transparency of process and accounting: What is the level of transparency vis-à-vis the overall management of scheme? Is the financial system (at all levels throughout entire cash flow process) transparent? Are cash transfer files transparent and files accessible and available for review by external groups? Are files accurate and complete? Are deviations explained? What is the system to prevent, measure, and capture fraud, errors and corruption?

Adequacy of internal monitoring and control: To what extent is there monitoring of overall implementation and processes including the targeting and pay processes, removing beneficiaries, and re-targeting? Are systems of monitoring and control adequate to prevent corruption within key processes (i.e. financial management, targeting and approval, vehicle usage)?

Level of cooperation at multiple levels: Is cooperation adequate between and among international, national, and community levels? How well are various activities supported? Is there a clear understanding of others' roles? Are critical processes involving various groups conducted in a cooperative manner (i.e. proposal development, cash flow, and report review)?

Quality of internal and external communication and communication strategies: Is communication adequate for drafting budgets, acquiring resources, problem solving, training of personnel, creating and maintaining linkages, and so on? What is the best way to frame and communicate the scheme's purpose, targeting, expectations, and logistics?

Linkages to existing services and ministries: Are international, national, district and community level operations appropriately linked? Are NGOs, civil society linked to the scheme? Where are the additional areas of linkages? Have they been identified? Is there a plan? Are beneficiaries appropriately included in identifying possible linkages?

Cash management: What are the bottlenecks as cash moves from government or donors to beneficiaries? What are the systems in place to prevent bottlenecks and leakage and do they work? What delivery mechanisms are feasible and cost effective? Should governments partner with banks, post offices, schools or others to deliver cash?

Indexation: Who bears the price risk (the government, donor or recipient) when inflation erodes the value of the cash? How often should indexation occur so that transfers are adjusted in order to maintain their purchasing power? What if indexation allows beneficiaries to purchase goods and services but leaves other poor non-beneficiaries unable to purchase items?

Sustainability: Are the mechanisms in place to sustain the scheme, including training in social protection for civil servants, the inclusion of program-relevant-indicators in regularly collected national data sources, ongoing financial planning for implementation and transfers, and so on?

Fragile states: Most evidence comes from low income, but not fragile countries. However, there is predictable poverty and demonstrable need in places like the DRC and Zimbabwe. With few examples implemented, there is little evidence to determine whether transfers work in contexts where administrative infrastructure is severely weakened and communities are fragile.

Targeting Evaluation

Depending upon fiscal space and priorities, many of the poorest countries such as Mozambique and Zambia, may lack the budget to reach all of the poor and vulnerable groups that really need assistance. Limited resources vis-à-vis widespread need requires a mechanism to target cash. Policymakers and stakeholders must determine who to prioritize with existing resources.

From Latin America to Africa, stakeholders must adapt targeting strategies to settings with lower income and administrative capacity, and higher poverty, because it may not be feasible to use classic means testing to select recipients. In comparison to Latin America, where birth and death registries, payroll and tax systems are systematically kept, in many African countries, there are no vital registries or tax information for the poor, and lists of beneficiaries are kept by hand. Necessary adaptations may include targeting entire vulnerable communities or using community rankings within villages. Moving forward, consistent evaluation, which explains and offers solutions to these deficiencies, can improve public management with the systematic collection and maintenance of information, which benefits all social protection and development efforts. This phenomena, where high quality evaluation leads to improved public management, has occurred throughout Latin America.

The targeting evaluation questions that are important to policymakers and other stakeholders include the following:

- Who is targeted? Is there an accurate understanding of poverty or the target group(s)? How were coverage parameters decided? How accurate was the target group estimation?
- How are targeting concepts operationalized? Do programmatic concepts capture the intended target? Are program definitions of poverty the same as community definitions?
- Are poverty proxies appropriate?
- How well is the targeting or retargeting process implemented?
- What are the outcomes? How well is the targeting process implemented? What are the inclusion and exclusion errors?
- Are there unintended consequences? For example, do transfers that are targeted to households with children cause beneficiaries to shuffle children around? Are there changes in social cohesion in communities that target cash to certain households?
- What are the benefits and drawbacks of strategies?
- What are barriers to access for the most vulnerable?
- How (specifically) can community targeting mechanisms be improved?
- How can programs effectively target in an AIDS-sensitive, but not exclusive manner?

Evaluations can be set up to test different programmatic options, as was done in Zambia where universal age-based targeting was compared to community based targeting.

Economic evaluation

The evidence base on the cost of cash transfer schemes and other economic evaluation questions are still in the earliest stages. While cash transfers may be able to yield better outcomes for fewer inputs than other programs, there is insufficient data to prove the claim. Important evaluation questions include the costing of programs and determining cost effectiveness. Relevant questions including the following:

- What are the costs to beneficiaries and to implementers (local and national government, donor agencies)
- What are the costs of program administration, including targeting, delivery, transfers, salaries, facilities, and opportunity costs? What programmatic alternatives should also be costed and considered?
- What is the optimum balance of costs?
- How much do various impacts (health, access to services, food security) cost?
- How do costs compare to other interventions?
- What is time period of costs? Are benefits ‘for life’?
- Who are various costs paid by? Government, donors, recipients?

Cost effectiveness analysis is used to compare alternative interventions and to identify least-cost strategies to achieve a specified outcome. Indeed, most cash transfer evaluations to date have examined whether the program achieves a particular outcome, such as poverty reduction. Moving forward, asking ‘What is the most cost-effective way to reach a particular outcome?’ is a different evaluation question that involves costing cash transfers and other programs, determining whether programs reached various outcomes (such as improved child growth and health, school attendance, and household food security), and then comparing the costs of the interventions.

A caveat to cost effectiveness analysis is that it is useful to know the figures, but it is difficult to obtain accurate information because other programs may not have been evaluated or costed.

Sharing evaluation findings

Throughout governments and donor agencies, there is much demand for information on cash schemes, including information on promising practices. As schemes scale up, it is essential for evaluators and policymakers throughout Africa to link communities to learn from each other and ensure that current evidence feeds into practice. Attendees agreed that a community of practice is needed to help summarize and communicate information on schemes and the components that demonstrate consistent outcomes. A community of practice can also help build evaluation capacity and ensure quality control in evaluation. In Latin America, a forum was created whereby program managers can discuss operational and strategic issues using a low cost teleconferencing model. A similar structure could be adapted so that African countries can connect and learn from each other.

Additionally, another key gap is in the knowledge management of information that we do have, such as a data archival system where evaluation data can be stored and made publicly accessible. Evaluators agreed that study budgets generally do not allow for the time necessary to make data publicly available. However, open access to evaluation data increases transparency and may allow analysts to answer more questions about the scheme’s impact, particularly as schemes scale up and cash transfers become one tool in a larger social protection system.

Conclusion

There are many gaps in the evidence base on cash transfers in Africa, as well as variation in the robustness of existing evaluation studies. However, compared to other areas of interventions throughout Africa, the cash transfer schemes in Ethiopia, Kenya, Lesotho, Malawi, Mozambique, South Africa, Tanzania, and Zambia have been or are in the process of being evaluated, with studies yielding persuasive information. Still, at the evaluation meeting, policymakers expressed the challenge of knowing when the evidence is sufficient to scale up. In fact, it may be premature for one to draw the meta lessons on cash transfers in Africa, given the limited body of existing evidence. Still, attendees agreed that the evidence suggests that governments should implement cash schemes and in some cases, should extend or scale up programs in order to confront poverty and destitution. Moreover, these schemes need to build robust monitoring and evaluation plans into their schemes in order to yield the data and information necessary to make programmatic changes and determine how cash schemes can be used as viable tools of social protection.

Finally, policymakers and other stakeholders have many questions and require high quality information in order to design and implement programs and measure impact. Meeting attendees called for improved and consistent linkages between policymakers and evaluators and rigorous evaluation methods that yield the data necessary to build quality social protection system across Africa.

References

1. Schady, N. and Fiszbein, A. (2009). Conditional Cash Transfers: Reducing Present and Future Poverty. World Bank, Washington D.C.

Appendix 1. Meeting Participants

<u>Participant</u>	<u>Organization</u>
Michelle Adato	IFPRI, Evaluator South Africa
Harold Alderman	World Bank
Charles Shapiro	Bureau of Western Hemisphere Affairs, US Department of State
Gretchen Bachman	USAID
Sarah Baird	World Bank, Evaluator
Diana Chamrad	Office of HIV/AIDS Global Health, USAID
David Croome	Lesotho, Evaluator
Alan Debrauw	IFPRI
Rena Eichler	HS 20/20, Legal Action Center
Dan Gilligan	IFPRI
Margaret Grosh	World Bank
Lena Heron	PEPFAR
John Hoddinott	IFPRI, Evaluator, Ethiopia
Priya Kumar	Boston University
Esther Lwanga	USAID
Amanda Makulec	Boston University
Maury Mendenhall	PEPFAR
Candace Miller	Boston University, Evaluator, Malawi
Morris Moono	Provincial Social Welfare Officer, Zambia
Allyn Moushy	Poverty Analysis & Social Safety Nets, USAID/EGAT/PR
Harry Mwamlima	Ministry of Economic Planning and Development, Malawi
Patrick Premand	World Bank
Laura Rawlings	World Bank
Gelson Temblego	University of Zambia
Thabo Thulo	Lesotho, Ministry of Finance
Fabio Veras	UNDP, Mozambique, Evaluator
Anna Williams	PEPFAR
John Williamson	USAID, Displaced Children and Orphans Fund
Rachel Yates	UNICEF, Children and HIV/AIDS