

Number 51/ February 2004

# CDD Impact Assessments Study: Optimizing Evaluation Design Under Constraints

Paul Wassenich and Katherine Whiteside

# SOCIAL DEVELOPMENT PAPERS

---

Community Driven Development

Number 51/February 2004

## CDD Impact Assessments Study: Optimizing Evaluation Design Under Constraints

Paul Wassenich and Katherine Whiteside

This publication was developed and produced by the Social Development Family of the World Bank. The Environment, the Social Development, and the Agriculture and Rural Development Families are part of the Environmentally and Socially Sustainable Development (ESSD) Network. The Social Development Family is made up of World Bank staff working on social issues.

Documents in the Social Development Working Paper series are not formal publications of the World Bank. They are published informally and circulated to encourage discussion and comment within the development community. The findings, interpretations, judgments, and conclusions expressed in this paper are those of the author(s) and should not be attributed to the World Bank, to its affiliated organizations, or to members of the Board of Executive Directors or the governments they represent.

For additional copies of this paper, please contact:

Social Development  
The World Bank  
1818 H Street, NW  
Washington, DC 20433

Fax: 202-522-3247

E-mail: [sdpublications@worldbank.org](mailto:sdpublications@worldbank.org)



Printed on Recycled Paper



# CDD IMPACT ASSESSMENTS STUDY: OPTIMIZING EVALUATION DESIGN UNDER CONSTRAINTS

## TABLE OF CONTENTS

Executive Summary	
I. Introduction .....	<b>Error! Bookmark not defined.</b>
II. Study Overview .....	<b>Error! Bookmark not defined.</b>
Sample .....	<b>Error! Bookmark not defined.</b>
Sources .....	<b>Error! Bookmark not defined.</b>
Limitations & Biases .....	<b>Error! Bookmark not defined.</b>
III. Topics .....	4
Hypotheses .....	5
Synthesis of Topical Coverage .....	6
Primary Welfare.....	6
Infrastructure.....	7
Social Dynamics .....	7
Governance .....	7
Targeting.....	7
Comparative Effectiveness .....	8
Research Coordination.....	8
IV. Methods .....	9
Gold Standard Evaluation .....	9
The Constraints.....	10
General Constraints .....	10
Specific CDD Constraints .....	10
Fundamentals of Impact Evaluation .....	11
Tenet 1: Comparison Groups .....	11
Tenet 2: Baseline Data .....	17
Tenet 3: Mixed Methods.....	18
Scaling Down Sound Designs.....	20
V. Concluding Recommendations .....	21
Annex 1: Methods Inventory .....	24
Annex 2: Review of Findings .....	40
I. Primary Welfare .....	40
I.A. Household income and consumption/expenditure proxies .....	41
I.B. Education outcomes.....	43
I.C. Health outcomes .....	47
II. Infrastructure.....	50
II.A. Access to Infrastructure.....	50
II.B. Quality of Infrastructure .....	54
II.C. Quality of Service Delivery .....	56
II.D. Maintenance & Sustainability.....	57
II.E. Complementary Inputs.....	59
II.F. Utilization.....	61
III. Social Dynamics / Social Capital.....	62
III.A. Social Capital, Collective Action, Trust .....	63
III.B. Participation, Voice, Inclusion, Representation, Ownership .....	67
III.C. Gender .....	71
III.D. Empowerment .....	72
III.E. Conflict, Crime .....	73

III.F. Leadership.....	74
IV. Governance .....	75
IV.A. Improved Local Governance.....	75
IV.B. Decentralization.....	77
IV.C. Corruption .....	78
V. Targeting .....	79
V.A. Poverty Targeting .....	79
V.B. Preference Targeting .....	84
VI. Comparative Effectiveness.....	85
VI.A. CDD versus non-CDD.....	85
VI.B. Effectiveness of different delivery mechanisms within CDD .....	87
VI.C. Cost-effectiveness, Economic returns.....	88
Annex 3: Evaluation funding arrangements .....	90
Annex 4: Tallies .....	92
Annex 5: Interview List.....	93
References .....	94

**Acknowledgements:** This paper greatly benefits from the comments of three peer reviewers, Vijayendra Rao, Susan Wong and Michael Woolcock; as well as from the valuable feedback of Sarah Adam, Robert Chase, M. Louise Fox, Camilla Holmemo, Keith MacLean, Kimberly Maynard, Daniel Murphy, Daniel Owen, Jeffrey Ramin, David Warren and Mark Woodward. We would like to thank the task team leaders and research managers who provided information about their evaluations, and everyone we interviewed.



## Executive Summary

The original intent of this study was to conduct a meta-analysis of evidence regarding community driven development (CDD) project impacts, in response to concerns that “evidence on the actual record of CBD/CDD initiatives still lags considerably behind the speed at which such projects are being implemented and ‘scaled up’” (Mansuri & Rao 2003, 3). However, confronting the limited methods and scale of many existing evaluations, the study shifted in focus to a review that identifies knowledge gaps in existing research and emphasizes fundamental principles of sound impact evaluation. The analysis covers 43 evaluations of Bank-sponsored CDD projects, gathered through self-reports and informal snowball sampling. The sample focuses specifically on impact assessment as one tool within monitoring and evaluation systems, and is weighted toward more statistically rigorous studies.

In the spirit of an experimental, learning-by-doing approach to development, we argue first for immediate improvements in the rigor of evaluation methods. Project research teams should conduct the most robust assessment possible given the constraints they face, following three basic evaluation principles: 1) include comparison groups; 2) collect baseline data; and 3) incorporate mixed qualitative and quantitative methods. Second, we recommend longer term efforts to mobilize greater financial and technical resources to improve the number and quality of impact evaluations throughout the Bank.

To promote the evaluation agenda for CDD and the broader Bank community, we recommend:

- Honest stocktaking exercises in all operational areas to assess the quality of evaluation methods, identify hypotheses that require empirical examination, and help coordinate research across projects, sectors and regions.
- Greater CDD research, including impacts on gender, conflict and leadership; project effects on, and interaction with, decentralization processes and public sector management; the relative performance of CDD compared to more centralized approaches in achieving desired outcomes; the impact of various components and processes within a CDD project, to help “unpack the black box” of evaluation work; long-term sustainability of welfare impacts and infrastructure outcomes; and preference targeting, or the alignment of projects with ex ante beneficiary priorities.
- Greater advance planning to integrate evaluation into the design and implementation of projects.
- More direct evaluation financing by the Bank to help alleviate the tension between a) the global public good value and direct benefits to the Bank of evaluation research, and b) individual project or country government reporting needs and priorities. We suggest matching grants.
- Creation of a more effective architecture for sharing: information about evaluations; tools, including those for synthesizing qualitative data; survey instruments, with clear notation of context; and negative or inconclusive results.
- Establishing a peer review board of evaluation experts to help improve evaluation designs, ease concerns about research team independence, and enhance the reliability of findings.

# I. Introduction

This study arises in the context of what some refer to as a “paradigm shift” in the strategy of the World Bank, moving from a “top-down” to more of a “bottom-up” approach to development.<sup>1</sup> Drawing on a rich history of participatory approaches inside and outside the Bank, this shift involves taking advantage of local knowledge and giving the poor greater voice and control over development decisions. The growing popularity of community driven development (CDD) within the Bank is evidenced by its rapidly increasing share of the lending portfolio: conservative estimates place the growth from \$325 million in 1996 to \$2 billion in 2003.<sup>2</sup> Furthermore, Bank President James Wolfensohn emphasizes CDD as one of the key approaches that can be scaled up and leveraged to meet the Millennium Development Goals.<sup>3</sup>

Yet at the same time, CDD faces criticism on both theoretical grounds and concerns about practical application.<sup>4</sup> In this context, Mansuri and Rao (2003, 3) conducted a literature review seeking reliable evidence in response to concerns that “evidence on the actual record of CBD/CDD initiatives still lags considerably behind the speed at which such projects are being implemented and ‘scaled up.’” The confirmed deficit of empirical research in several key areas constrains the potential for an experimental, learning-by-doing approach to development.

The slow progress poor countries have made over recent decades and the limited knowledge about what development interventions work best in various contexts makes evaluation and learning by doing critical. Baker (2000) suggests that, “Evaluating impact is particularly critical in developing countries where resources are scarce and every dollar spent should aim to maximize its impact on poverty reduction. If programs are poorly designed, do not reach their intended beneficiaries, or are wasteful, with the right information they can be redesigned, improved, or eliminated if deemed necessary” (p. vi). Viewing development interventions as experiments, we need to be more systematic about using the work we do as learning opportunities. This implies that careful evaluation should not be an endpoint, but rather an ongoing tool to increase our understanding of the successes and failures we encounter. In this light, impact evaluation should not label projects as successes or failures, but rather provide insights on how to improve project performance. As a complement to experiential knowledge, evaluations document information in a way that can be easily shared with a wide audience.

In response to Mansuri and Rao, this study originally intended to conduct a meta-analysis of World Bank-supported CDD projects. By relaxing constraints on evaluator independence and rigor of methods, the study aimed to compile further evidence, mainly from unpublished studies, on the array of CDD evaluations conducted and the impacts they uncover.<sup>5</sup> However, in the process of reviewing over 60 evaluations, it became clear that too few studies within this gray literature were of sufficient scale and rigor to provide reliable data for a meta-analysis. Moreover, nearly all the studies that generated evidence that was statistically representative of the overall program under investigation have already been

---

<sup>1</sup> Vijayendra Rao interview, 20 August 2003.

<sup>2</sup> CDD Portfolio Matrix maintained by the CDD Anchor.

<sup>3</sup> Daniel Owen email, 11 July 2003.

<sup>4</sup> See Mansuri & Rao (2003) for discussion of arguments for and against CDD.

<sup>5</sup> Mansuri & Rao (2003) include studies of any project with “community participation as a crucial element of its design,” so long as the evaluation was not conducted by those directly in charge of implementing the project in question, except when the study had been published and subject to peer review (p.4). They also review theoretical literature relating to the CDD approach.

comprehensively documented in *Evaluating Social Funds: A Cross-Country Analysis of Community Investments* (Rawlings et al. 2004).

Two main reasons explain why most of the evaluations reviewed for this study did not reach the level of generalizable evidence: insufficient methods and limited scale. First, many of the studies reviewed did not estimate the counterfactual state of participants through the construction of an appropriate comparison group, and thus the resulting findings could not be reliably attributed to the program under investigation. Second, many small-scale studies lacked the sample size necessary to generate statistically significant results. While several generate useful insights, they remain subject to lingering doubt that the findings are not necessarily representative of the overall program.

The problem of infrequent or low quality evaluation is not unique to CDD. Rigorous impact evaluations of World Bank projects are rare in all contexts. Ezemenari et al. (2000) found that only a small percentage (5%–10%) of Bank projects had sound evaluation plans, including defined impact indicators and a reasonable comparison group.<sup>6</sup> In fact, the evidence compiled in this report indicates that the CDD community devotes substantial effort to impact evaluation, likely at least as much as other sectors. As a result, the Bank would benefit from similar stocktaking exercises for all areas of operations to gauge the overall extent, quality and results of evaluation work.

With these ideas and discoveries in mind, the study shifted substantially in focus. While Annex 2 discusses research findings by topic, the rest of the paper moves away from a compilation of what we know about CDD impacts to discussion of evaluation methods and gaps in the research agenda. The study now aims to review completed and planned evaluations to first identify topical areas that are relatively neglected by impact assessments to help steer the CDD research agenda moving forward; and second, capture lessons about how we can more consistently apply sound methodologies to explore impacts in a world of finite resources. We will argue for both immediate improvements in the rigor of methods used within existing constraints, and for longer term efforts to mobilize greater resources for more, higher quality impact evaluation.

The rest of the paper is structured as follows. Section II provides an overview of the study sample, sources, limitations and biases. Section III reviews the thematic areas explored by completed and planned evaluations to identify gaps in the existing knowledge of CDD impacts. Section IV focuses on methods by working through a constrained optimization problem of impact evaluation based on examples from existing CDD studies. Section V concludes with recommendations on how to pursue the evaluation agenda within the Bank. Critical analysis within the body of the paper draws on two detailed annexes: Annex 1 reviews the methods used in each CDD evaluation within the sample, and Annex 2 discusses thematic topics that evaluations have or will explore and research findings to date. The companion website to this paper presents similar material organized by project.<sup>7</sup>

## **II. Study Overview**

### **SAMPLE**

Background research for this review began under broad definitions of both CDD and impact evaluation. First, regarding what constitutes CDD, this study includes World Bank projects that involve substantial

---

<sup>6</sup> The exact figures are 5.4% of projects in FY98, 7.8% in FY99, and 10.2% in FY2000. Counting by lending amounts rather than by project, the figures are 6.2% in FY98, 6.1% in FY99, and 9.0% in FY00.

<sup>7</sup> See <http://lnweb18.worldbank.org/ESSD/sdvext.nsf/09ByDocName/Evaluations>.

participation by community groups in decisions about planning, implementation and management of projects, and that devolve some control over resources to local communities.<sup>8</sup> However, since there is no clear, widespread consensus within the Bank regarding exactly what does and does not count as CDD, we relied on self-reports and did not arbitrate whether reported projects fit the CDD classification.

Second, impact evaluations measure effects on beneficiaries and attribute them to a specific intervention by estimating the counterfactual state of participants.<sup>9</sup> Relaxing this definition, this review considers assessments of CDD projects that address specific questions regarding the performance of a project and its impact on beneficiaries. As such, many of the studies included are not, in fact, impact evaluations because they do not estimate the counterfactual.<sup>10</sup> These studies are included either because they provide interesting insights from a case study approach; or to illustrate how evaluations with weak methodologies compromise the reliability of their findings.

The compilation process began with the ideal evaluation in mind, searching for CDD impact evaluations that came closest to the gold standard, and then explored studies conducted under a series of compromises and constraints. Research focused specifically on impact evaluation, as only one component of overall monitoring and evaluation (M&E) portfolios. We focus on evaluations of specific projects, and do not attempt to re-summarize meta-analyses already conducted.<sup>11</sup> Included CDD projects and evaluations were largely self-reported by those involved, and informal snowball sampling extended the list. Requirements that studies clearly present the methods used and include a minimal attempt at a comparison group means that some studies with weaker methods were reviewed but not included in the paper. The overall process results in a selective sample weighted towards more statistically rigorous evaluations that includes most, if not all, Bank CDD evaluations with findings that are statistically representative to the scale of the project. The implication is that the degree to which the study approximates an inventory tapers off as the methods become less rigorous and studies more frequently excluded.

**Table 1: Evaluations Included (43 total)**

<b>Completed</b>	
Armenia Social Fund (Social Funds 2000)	Pakistan Aga Khan Rural Support Project (AKRSP)
Benin Social Fund gender study	– Khwaja study
Bolivia Social Fund (Social Funds 2000)	Pakistan Aga Khan Rural Support Project (AKRSP)
El Salvador EDUCO	– OED study
Honduras Social Fund (FHIS) (Social Funds 2000)	Panama Rural Poverty and Natural Resources Project
India Dairy: Operation Flood	Peru Social Fund (FONCODES)
Jamaica Social Fund	Zambia Social Fund (ZamSIF) (Social Funds 2000)
Nicaragua Social Fund (FISE) (Social Funds 2000)	
<b>Ongoing</b>	
Albania Second Community Works Project	Indonesia KDP & Conflict study
Armenia Second Social Fund	Indonesia Second Urban Poverty Project (UPP2)
Benin Social Fund	Malawi Social Action Fund (MASAF)
Brazil Northeast Rural Poverty Reduction Program	Moldova Social Investment Fund

<sup>8</sup> For further details on CDD, see Chapter 9 of the PRSP Sourcebook (Dongier et al. 2002), or [www.worldbank.org/cdd](http://www.worldbank.org/cdd).

<sup>9</sup> Two excellent guides to impact evaluation are Ravallion 1999 and Baker 2000.

<sup>10</sup> Many of these also do not claim to be strict impact evaluations.

<sup>11</sup> In particular, the Rawlings et al. (2004) study of six social funds includes many more details than we compile here from the individual project evaluation reports. Similarly, the OED overarching study of social funds is not included here, see World Bank 2002b.

(PCPR) Burkina Faso Second National Agricultural Services Development Project (PNDSA2) Cambodia Rural Investment and Local Governance (RILG) India Andhra Pradesh District Poverty Initiatives Program (APDPIP) – DEC Study Indonesia Kecamatan Development Project (KDP1, 2 &3)	Pakistan National Rural Support Program (NRSP) – DEC study Philippines Comprehensive and Integrated Delivery of Social Services Project (KALAHI-CIDSS) Senegal National Rural Infrastructure Project (PNIR) Senegal Agricultural Services and Producer Organizations Project (PSAOP) St. Lucia Poverty Reduction Fund (PRF) Yemen Social Fund for Development (SFD)
<b>Planned</b> Comoros Social Fund (FADC) Ecuador Social Fund – DEC targeting study Ethiopia Women’s Development Initiative (WDI) Indonesia Support to Conflict Ridden Areas Project (SCRAP) Madagascar Social Fund (FID4) Mongolia Sustainable Livelihoods Project	Nepal Rural Water Supply & Sanitation (RWSS) Nicaragua Social Fund – DEC targeting study Nicaragua Social Fund (FISE 3) Philippines Autonomous Republic of Muslim Mindanao (ARMM) Social Fund (ASFP) Thailand Social Fund Vietnam Northern Mountains Poverty Reduction

## SOURCES

We gathered information on completed evaluations from World Bank project documents (both official and informal drafts), consultants’ final reports and journal articles. Information on ongoing and planned evaluations draws on working documents regarding evaluation design, interim reports summarizing findings of baseline studies, and interviews with relevant World Bank staff (see Annex 5 for interview list).

## LIMITATIONS & BIASES

This report was not conducted as a systematic portfolio review. The authors did not review all Project Appraisal Documents (PADs) and Implementation Completion Reports (ICRs) within the entire Bank CDD portfolio to identify those that mention impact evaluation work. Other CDD impact evaluations may exist that are not included here. As mentioned earlier, the paper focuses on impact evaluation as a specific tool within M&E systems. In sum, the analysis is therefore not representative of the totality of M&E work for Bank CDD projects. As a result, included analyses should be interpreted as snapshots of the selective sample, not of a comprehensive evaluation inventory.

## III. Topics

This section surveys areas where CDD effects are relatively well-documented and those that would particularly benefit from further research. To provide some theoretical background, we first review some hypotheses regarding how CDD interventions work and their expected impacts on beneficiaries. We then turn to the evaluations, presenting a synthesis of thematic areas that evaluations have investigated or plan to explore, which we detail in Annex 2.

## HYPOTHESES

To situate this study in context, it may be useful to briefly highlight a few key points regarding the conceptual underpinnings and historical evolution of the community-driven approach.<sup>12</sup> Reviewing what theory tells us about how CDD aims to address problems inherent in other types of interventions and the mechanism by which it operates will help identify hypotheses for impact evaluations to test empirically.

Drawing on a rich and diverse history of participatory approaches from different countries, organizations, theorists and practitioners, CDD is part of a paradigm shift that aims to address systemic failures in dominant styles of intervention by reconfiguring top-down approaches to development. Lindauer and Pritchett (2002, 5) discuss how immediate historical evidence available to policymakers in the early 1960's suggested that governments with strong planning ministries were the engine of industrialization. Thus the best approach to development assistance for the World Bank was "investable foreign exchange made available to governments for development projects." Two decades later, the failures of central planning "were becoming obvious." Sector by sector, Pritchett and Woolcock (2002, 17-21) explore how responding to the need for essential services with public provision through a centralized civil service created conditions for its own failure. Problems with the "needs-bureaucratic supply" solution include how top-down, standardized models fail to take advantage of (often tacit) local knowledge leading to technological mistakes and inappropriate supply; how "free" provision and inadequate demand assessment lead to low community ownership, utilization and maintenance of services; and how accountability and information systems that flow only vertically upwards and not down to citizens leave room for abuse of discretion and corruption.

Thus with the apparent failures of "Big Development," combined with rapidly degrading common pool resources, Mansuri & Rao (2003, 6) point to a "reawakened interest in the notion of local management of resources and decisions."<sup>13</sup> In this context, CDD had "the explicit objective of reversing existing power relations in a manner that creates agency and voice for the poor, while allowing the poor to have more control over development assistance" (p. 2). They summarize the claims of how CDD restructures development operations by reducing information problems, expanding resources available to the poor, and strengthening the civil capacities of communities (p. 2).

Here follow some central theoretical arguments for how CDD operates as an effective development intervention, drawn from the CDD Chapter within the Poverty Reduction Strategy Sourcebook by Dongier et al. (2002). We organize the hypotheses by theme to trace them to our subsequent review of empirical work.

1. Primary welfare: "using CDD approaches ...can support immediate poverty reduction by efficiently building human and physical assets at the local level." "Community management and accountability can improve education outcomes."
2. Infrastructure: "CDD can make services responsive to demand expressed by poor men and women and as a result can enhance sustainability." "Community management of development investments usually results in lower costs and more productively employed assets." "Community-developed facilities such as health centers, schools, and water supply systems tend to have higher

---

<sup>12</sup> This is clearly a sizeable topic, beyond the scope of this paper. See Mansuri & Rao (2003) for more on the history of participatory approaches and the move toward CDD/CBD.

<sup>13</sup> They identify the present as the second wave of interest in CDD/CBD fully absorbed by mainstream forms of development, citing a first wave in the 1950's via the work of USAID. They also discuss other practical and theoretical origins of participatory approaches.

utilization rates and are better maintained than when investment decisions are made by actors outside the community.”<sup>14</sup>

3. Social Dynamics: “Control over decisions and resources can also give communities the opportunity to build social capital...by extending the depth and range of their networks.” “Targeted community-driven approaches devolve control and decisionmaking to poor women and men, which empowers them immediately and directly.”
4. Local Governance: “strengthening local associations that are inclusive can increase poor people’s voice in local political processes and governance.”
5. Preference targeting, or alignment of projects with community priorities: “demand is better articulated when communities contribute to investment costs and control investment choices.”
6. Comparative Effectiveness: “CDD has the potential to make poverty reduction efforts more responsive to demands, more inclusive, more sustainable, and more cost-effective than traditional centrally led programs.” “The potential for CDD is greatest for goods and services that are small in scale and not complex and that require local cooperation, such as common pool goods”

Because of the heterogeneity in CDD intervention models and implementation contexts, amassing empirical evidence to definitively test these hypotheses about CDD *as an overall approach* may be impossible, if not irrelevant. That said, in the spirit of learning-by-doing the next section highlights a few areas that stand out as having received less attention than others by completed and planned evaluation work for specific project settings.

## **SYNTHESIS OF TOPICAL COVERAGE**

The following findings are based on the selective study sample, which may not include all impact evaluations that address the given topics.<sup>15</sup> Annex 2 provides further details on the topics individual evaluations explore and discusses research findings.

### **Primary Welfare**

Overall, some of the most reliable and comprehensive evidence concerns CDD impacts on primary welfare measures, which cover sub-project impacts on household income, and education and health outcomes. Nearly all of the evaluations in this area with findings that are statistically representative of the overall project in question contributed to the cross-country study of social funds (Rawlings et al. 2004). Project impacts on poverty as measured by household income or relevant proxies is the only area in this category that has not received substantial attention from completed evaluations, however fourteen upcoming studies plan to address this question. Further research would benefit from studies that explore the relative performance of CDD projects in achieving desired welfare outcomes as compared to other types of interventions, and the long-term sustainability of welfare impacts.

---

<sup>14</sup> For infrastructure, it may be instructive to separate hypotheses for how community involvement in the – (a) planning and design stage, and (b) operation and maintenance stage – affects outcomes (M. Louise Fox comments 11 December 2003).

<sup>15</sup> Rawlings et al. (2004) contains more details than we compiled here from the individual project evaluation reports. It contains comprehensive exploration of primary welfare outcomes, as well as issues of infrastructure performance, targeting, community preference, participation, social capital and cost, for six social funds in Armenia, Bolivia, Honduras, Nicaragua, Peru and Zambia. Similarly, the assessment of social funds by the Operations Evaluation Department contains details on social funds in Jamaica, Malawi, Nicaragua, Zambia, Argentina, Bolivia and Eritrea not included here (World Bank 2002b).

## **Infrastructure**

Several evaluations have explored issues regarding the performance of infrastructure implemented by CDD projects, including topics of improved access, quality of infrastructure and services, maintenance and sustainability, utilization, and provision of recurrent inputs. All of these areas are covered by at least five completed or planned studies within the sample, with impacts on access receiving the most attention, from a total of fourteen studies. Two evaluations of maintenance have taken a particularly interesting approach, examining sites where more than one project has been implemented to isolate project-specific factors that affect maintenance performance from differences among communities (see discussions of Khwaja (2002) on Pakistan AKRSP and the DEC Pakistan NRSP study in the annexes).

As a general point, while technical evaluations of CDD project infrastructure by engineers are fairly common, they rarely make a systematic attempt to assess facilities in appropriate comparison areas or those implemented by other agencies. As a result, further insights could be gained by simply shifting the coverage of existing technical fieldwork to include comparison areas. In addition, further research that highlights particular arrangements within CDD programs that effectively provide complementary inputs would be useful, as would long-term utilization and maintenance studies.

## **Social Dynamics**

For this paper, social dynamics serve as a broad category that covers topics of social capital, collective action, norms and trust; participation, voice, inclusion, representation and ownership; gender; empowerment; conflict and crime; and leadership. Several studies have explored these topics, however many have been on a case study scale, and thus few have generated results that are statistically representative of overall programs. Areas of social capital, participation and empowerment will all be further explored by several rigorous upcoming studies. However, fewer planned studies indicate that they will investigate issues of gender, conflict and leadership, and these topics could benefit from further research. Similarly, there is little evidence that establishes a direct causal link between greater community participation more generally and improved project outcomes, and thus further insights would be beneficial.<sup>16</sup>

## **Governance**

The category of governance includes issues of local governance, decentralization, corruption, and public sector reform. Despite being a topic of much discussion, CDD impacts on governance have received the least attention from evaluations in this sample. Ten upcoming studies plan to explore whether CDD projects improve local governance and four will investigate impacts on corruption. However further research is needed on CDD project interactions with decentralization and impacts on public sector reform. No completed and only two planned impact evaluations in the sample explicitly explore decentralization, and no completed nor planned studies address public sector impacts.<sup>17</sup>

## **Targeting**

Two kinds of targeting are often examined as part of evaluations of CDD programs: poverty targeting, where analyses measure the extent to which resources reach poorer segments of the population; and preference targeting, which describes the extent to which the implemented subprojects match the ex ante preferences of beneficiaries. Poverty targeting has been reasonably well studied by six completed studies, plus nine upcoming evaluations will expand knowledge in this area. In contrast, while several completed

---

<sup>16</sup> Emphasized by Vijayendra Rao in BBL presentation, 21 September 2003.

<sup>17</sup> Two papers (not impact evaluations) do, however, provide insights on social funds and institutions, see Bhatia (forthcoming) and Parker & Serrano (2000).

and planned studies explore preference targeting with retrospective questioning, none have collected ex ante baseline data on beneficiaries' priorities. Considering that studies of social funds in Jamaica and Nicaragua indicate that ex post beneficiary satisfaction is not necessarily a reliable indicator of ex ante preferences, further research would benefit from evaluations of preference targeting that incorporate baseline data on beneficiary priorities.<sup>18</sup>

## **Comparative Effectiveness**

The section on comparative effectiveness refers to the interrelated questions of CDD performance relative to alternate interventions, and the comparative performance of different delivery mechanisms within a CDD program. As earlier sections mention, this review discovered little evidence that directly addresses these topics, and only a few rigorous upcoming evaluations that plan to tackle these questions directly. A central problem here is that even high quality comparative evaluations often do not include details on delivery mechanisms used by alternative providers, especially in how they differ from the CDD approach. The DEC Pakistan NRSP evaluation plans to explore both areas and the Brazil PCPR plans to look at the latter, however additional studies in this area would be useful, especially to inform resource allocation decisions across different development approaches.

Lastly, this review did not encounter many comparative cost-effectiveness studies. To the extent that impact evaluations and cost studies are not highly correlated activities, the study sample is not the best source for this information (see Rawlings et al. (2004) for cost details on six social funds and comparators). However, reflecting back on the central hypotheses regarding CDD, more studies that compare the costs and benefits of CDD interventions to that of other providers would be informative.

## **RESEARCH COORDINATION**

From the view that evaluation results are an international public good, filling existing gaps in knowledge about CDD impacts is clearly important. However, the reporting needs of a particular project or country government may rest within areas that have been relatively well-covered, for instance a demonstration that projects improve basic indicators of welfare and living standards. While it is tempting to suggest that project evaluations with limited resources should focus on one new area of research and conduct the most rigorous evaluation possible, this may not satisfy individual project needs. As a result, an effective strategy to round out knowledge of where CDD is most effective and how it could be improved will require additional resources provided directly by the Bank to client governments to support evaluation work.

In Annex 3 we discuss various options currently available for financing evaluations, list some inherent tradeoffs for each alternative, and provide examples of budgets and funding sources from the sample of evaluations. While each approach can be valuable, as a group they are insufficient to address the public good value of research and its direct benefits for the Bank. Therefore, we suggest that the Bank offer matching grants on, say, a one-for-one basis, to double what governments are able to spend on impact evaluations. This would enable project teams to improve the scope and rigor of methods, allow the international research community to test central hypotheses to expand knowledge of CDD processes and impacts, and reduce externalities by ensuring that main beneficiary parties shoulder some of the evaluation costs.

---

<sup>18</sup> See discussion of Rao & Ibáñez 2003 on JSIF and World Bank 2000 on FISE in Annex 2, Section V.B.

## IV. Methods

Having identified several research areas that merit further attention, this section emphasizes some basic principles to keep in mind when planning evaluation designs. We explore impact assessment under a framework of constrained optimization: how to conduct the most rigorous and useful evaluation under finite resources and some preset design parameters. From discussion of the ideal program evaluation and constraints confronting careful evaluation work, we extract three basic tenets that substantially increase the reliability of impact assessment on any scale. We explore each concept in turn, drawing examples from the methods inventory of Annex 1 to ground the conceptual framework presented within the body of existing CDD evaluations.<sup>19</sup>

### GOLD STANDARD EVALUATION

In terms of impact assessment, the highest standard of proof involves a methodologically rigorous and comprehensive approach, which requires substantial resources and evaluation planning concurrent with project design. We outline components of this gold standard evaluation here.

To start, the evaluation team gains an in-depth understanding of the CDD intervention and context in which it takes place. This includes familiarity with a conceptual framework that maps the causal chain by which project inputs and processes intend to produce desired outputs and outcomes. Project and research teams use this framework to identify the links within the causal chain that the evaluation aims to test, to help “unpack the black box” of program evaluation.<sup>20</sup> Project staff randomly assign program participation amongst a large pool of eligible individuals or communities. The research team collects baseline data before participation begins, covering a sample that includes future treatment and control communities, as well as areas with similar projects implemented by alternative delivery mechanisms.<sup>21</sup> Fieldstaff collect follow-up data in the same three areas, contacting the same respondents to establish panel data. Data collection instruments combine detailed quantitative household surveys and in-depth qualitative focus groups, interviews and participant observation. Evaluation work spans a sufficiently extended time horizon to tackle issues of program and impact sustainability; all on a scale that is statistically representative of the overall program. In the context of the KALAH program in the Philippines, one estimate predicts that such an evaluation would cost around six million dollars over five years.<sup>22</sup> Such an evaluation has not yet been conducted for a Bank-sponsored CDD project (or perhaps any Bank project).

Even this model is an imperfect response to constraints. The U.S. Food and Drug Administration (FDA) likely would not allow a new toothpaste onto the market without double-blind clinical trials, where both the participants and evaluators do not know who receives the treatment and who the placebo. Even if such an approach were possible for social programs, CDD interventions and evaluations cannot escape the importance of context. Among several limitations to randomized trials, Ravallion (2003) notes external validity concerns that “if you allow properly for contextual factors it can be hard to make meaningful generalizations for scaling up and replication from trials” (p. 5). While conducting several evaluations in diverse contexts partially accounts for this problem, the point remains that “the institutional context of an intervention may well be hugely important to its impact,” complicating implications for both scaling up an existing project and drawing lessons from one evaluation to another project setting (p. 6).

---

<sup>19</sup> See Baker (2000) and Adam (forthcoming) toolkit for more comprehensive “how-to” guides.

<sup>20</sup> Howard White email, 2 December 2003.

<sup>21</sup> Accounting for other providers should also include calculation of the relative size and importance of project activities in comparison with “parallel activities of government, other donors and private agencies” (Howard White email, 2 December 2003).

<sup>22</sup> Rob Chase email, 23 October 2003.

**Box 1: Elements of the Gold Standard Evaluation**

- Understanding of project and context
- Experimental design (randomization)
- Baseline and follow-up surveys to establish panel data
- Sample covers treatment, comparison and alternative provider areas
- Comprehensive mixed method approach
- Long time horizon for sustainability issues
- Statistical representation on program or national scale

## THE CONSTRAINTS

Several factors help explain why the ideal evaluation has not been conducted for a CDD project, nor perhaps for any other development intervention. While some constraints apply to program evaluation in general, the nature of the CDD approach offers particular challenges of its own.

### General Constraints

To start, evaluation design should ideally take place alongside project design and development. Once a project has begun, many options for making the evaluation more rigorous are no longer available. The fact that initial thinking about evaluation often occurs during or after project implementation remains a serious, although not insurmountable, constraint to the quality of evaluations.

Fundamentally, isolating the impact of any intervention on beneficiaries is difficult because it involves accurately estimating the counterfactual state of participants, which is by definition unobservable. This presents an immediate technical constraint of specialized knowledge often not readily available in the field.

Furthermore, administering large scale surveys quickly becomes expensive, drawing on scarce resources many feel are better spent on the project itself. Determining the appropriate source of evaluation financing is complicated by the public good value of the knowledge generated, and by the fact that the research interests of Bank staff do not always align with those of client governments. While resentment regarding costs may be diffused with a convincing strategy to use study findings to design more effective operations, the time horizon from an evaluation's conception to results may extend beyond the life of the project, or at least the tenure of administrators in charge.

Even if the skills and resources are in place, incentives within a large bureaucracy like the Bank or a country government may not encourage rigorous assessment of project performance. Pritchett (2002) underscores the political economy incentives that support "strategic ignorance" over rigorous program evaluation. If unfavorable results could jeopardize future funding, exact knowledge of program impacts becomes a liability. Facing competition for finances, the program that rigorously evaluates itself becomes vulnerable to as yet unsubstantiated promotional claims from other projects. Thus even well-intentioned actors may avoid evaluation.

### Specific CDD Constraints

Design aspects of typical CDD interventions present additional challenges for evaluating this type of project. First, their demand-driven nature typically involves communities organizing and applying for grants to implement various projects. The unobservable differences between similar communities that do

and do not apply (e.g. degree of social cohesion or capacity for collective action) make ensuring the unbiased comparability of non-participant groups difficult. Similarly, the targeting of CDD programs may apply the program to an entire socio-economic group, eliminating the closest matched comparisons.<sup>23</sup> Furthermore, the lack of control over who participates (and when) may complicate evaluation design, in terms of establishing random program placement or fielding a baseline survey in a sufficient number of future participant and comparison areas.<sup>24</sup>

Finally, some of the main theoretical strengths of CDD involve less tangible topics like social capital, empowerment and participation, which can be difficult to measure. The meaning of these concepts and their interaction with CDD initiatives depends heavily on context, and thus make extracting general lessons or measurement tools from one locale and applying them to another extremely difficult.

## FUNDAMENTALS OF IMPACT EVALUATION

Clearly there are numerous constraints to conducting careful evaluations, and gold standard evaluations are rarely feasible. Yet despite the challenges, impact evaluation remains crucial for increasing our understanding of what types of interventions perform well in various contexts, to improve the effectiveness of development initiatives. Ravallion (2003, 5) writes, “the art of good evaluation is to draw eclectically from the whole menu of methods to find the most cost-effective combination appropriate to each setting.” With that in mind, we argue that first, valuable knowledge can be gained from assessments at lower standards of proof than gold. Thus, research teams should strive to conduct the most robust evaluation possible given the constraints they face, adhering to basic evaluation principles. Second, looking forward, we should aim to improve the quality and quantity of evaluation work overall, by mobilizing greater technical and financial resources for careful evaluation, incorporating evaluation planning into early stages of project design, and setting up a system for quality control.

From the discussion of the gold standard evaluation, we can extract three fundamental components that can substantially improve any effort to assess program impacts:

1. Inclusion of comparison groups;
2. Collection of baseline data; and
3. Incorporation of mixed qualitative and quantitative methods.

Taking each of these ideas in turn, the following sections briefly explain their methodological importance, and then work through examples from Bank CDD evaluations to highlight studies that best employ the techniques as well as those that use innovative solutions under constraints. See Annex 1 for more details on these and other examples.

### Tenet 1: Comparison Groups

Measuring program impacts on beneficiaries requires a strategy to estimate the counterfactual state of participants, or what would have happened had the intervention not taken place. Because the counterfactual is not observable, impact evaluations must include some form of appropriate comparison group. Random assignment of participation is the most rigorous method to construct a comparison group. However, when random assignment of program participation is not feasible or appropriate, there are a

---

<sup>23</sup> For example, if the entire bottom income quartile participates, any comparisons will come from slightly better-off areas. Of course, this type of assignment rule, strictly implemented, presents an opportunity for evaluation based on natural experiment.

<sup>24</sup> However, this does not automatically exclude the potential for experimental design as random assignment of the *offer* to participate can be sufficient.

variety of quasi-experimental techniques that provide useful evidence under a slightly relaxed standard of proof.<sup>25</sup>

## **Experimental Design**

Random assignment to treatment and control groups serves as the “gold standard” design for program evaluation. Newman et al. (1994, 182) emphasize that “whenever a project is of sufficient interest to policymakers to warrant an impact evaluation, program designers ought to consider randomized control design because this methodology yields the most robust results.” Randomization protects internal validity by ensuring that participation is completely exogenous, and thus uncorrelated with other pertinent variables or the error term in a regression. It does so by equalizing the distribution of relevant characteristics across a large sample of eligible participants. Individuals (or communities) randomly excluded from participation represent the counterfactual. Therefore ex post differences between the two groups can be attributed to the intervention under investigation. Even with demand-driven interventions, randomizing the *offer* to participate is sufficient (see Duflo & Kremer 2003).

Common opportunities for experimental design occur when program coverage is phased in, or when program teams wish to test alternative delivery mechanisms. In both cases, random assignment need not deny benefits to eligible participants. Specifically, if a program does not begin with universal coverage, randomizing the order in which individuals or communities participate determines only when, not whether, groups benefit. Similarly, randomly assigning groups to different delivery or implementation mechanisms can still provide all eligible beneficiaries similar end products. Another possibility for demand-driven projects is to use a screening process where potential participants can easily express their interest, but then must take some initiative to actually begin participation in the program (e.g., calling a toll-free number to express interest, and then filling out paperwork to participate, in a U.S. setting). Since many drop out between the first and second steps, a random portion of these can be contacted and encouraged to participate. The evaluation would then track outcomes for the two different groups (actively promoted and not) within the portion that did not take the second step (Zhu 1999).

Randomization can potentially lower the cost of an evaluation by reducing the need to collect data on a large number of non-participants to ensure sufficient coverage of an as-yet unidentified comparison group. For instance, certain matching techniques may require data on 20,000 households to ensure coverage of 2,000 that are actually comparable to beneficiaries. Furthermore, because randomization allows the use of the most straightforward and precise econometric techniques, it can reduce the sample size needed for a given level of statistical precision and the time required to analyze the data.

### ***Random assignment: Bolivia Social Investment Fund Evaluation***

Newman et al. (2002) use a variety of data sources and analytic techniques to evaluate impacts of the Bolivia SIF on education outcomes, health indicators and infrastructure improvements. One section of the evaluation used random assignment of the offer to participate amongst eligible schools to measure education impacts within the Chaco region. Because funding was not sufficient to reach all schools in the region, the SIF used a school quality index to determine needs-based eligibility and then randomized assignment within that group. Specifically, “Only schools with an index below a particular value were considered for SIF interventions, and the worst off were automatically designated for active promotion of SIF education investments. A total of 200 schools were included in the randomization, of which 86 were randomly assigned to be eligible for the intervention” (p. 245). In this way, the SIF was able to target benefits to the worst off (which were not included in the

---

<sup>25</sup> For discussion of the limitations of experimental designs, see Ravallion (2003), and for an accessible overview of the debate between experimental and quasi-experimental designs, see Grossman (1994).

evaluation), and then use random selection to distribute finite resources amongst similarly deserving schools. With baseline and post-intervention data, the authors compared difference-in-difference estimates from the experimental design with those based on propensity score matching. In general, the authors find that while school infrastructure and complementary inputs significantly improved, they demonstrated “little effect on enrollment, attendance, or academic achievement” (p. 248).

***Learning by doing: Pakistan National Rural Support Project – DEC Study***

Mansuri and Rao (2003) emphasize the value of an experimental, learning-by-doing approach to development that greatly benefits from honest monitoring and evaluation efforts. The current study of the Pakistan NRSP conducted by the World Bank Development Economics Research Group (DEC) provides an applied example of these ideas.

The program and research teams were able to agree on a process of random program placement and random assignment of intervention types to explore the efficacy of different NRSP mechanisms. This enables a comparison of outcomes across communities where a specific feature of the program is randomized in order to get a clearer sense of how components of the program impact outcomes of interest. This strategy will allow for an examination of a number of specific hypotheses. The first intervention will examine the impact of bundling microcredit with information on local productive opportunities and enterprise development skills. The intervention design includes 90 randomly selected villages, of which 30 will act as controls; 30 will receive the basic program, which includes credit; and 30 will receive credit complemented with training in entrepreneurship skills and in the identification of local production and marketing opportunities. This sample of 90 villages will also enable research into the role of social mobilization in the take up of public funds, within a context of devolution. With decentralized public funds equally available to all communities, the evaluation will test the hypothesis that areas where NRSP is engaged in the process of social mobilization are better equipped to form an organization, obtain local government funds, and use the funds more effectively (in terms of quality of organization, profile of members, more equitable distribution of project benefits, etc.). Lastly, NRSP management has been testing different models for deepening their social mobilization efforts and improving their targeting within communities, a particular challenge in a context of rapid program expansion to new areas. Thus the study includes a randomized experiment that will allow an assessment of three different credit delivery mechanisms, using another 90 randomly selected villages.<sup>26</sup>

The experimentation inherent in the design of both project expansion and careful evaluation presents a model for programs considering scaling up operations or seeking to test which sub-components or mechanisms are most effective in achieving outcomes of interest. In addition, the evaluation shows how randomization can contribute to an ongoing organizational learning process, especially when limited resources prevent an immediate expansion of operations into every community.<sup>27</sup>

---

<sup>26</sup> For further details and hypotheses regarding the included assessments of microcredit and social mobilization, see Annex 2, Section VI.B: Effectiveness of different delivery mechanisms within CDD.

<sup>27</sup> Ghazala Mansuri interview, 20 August 2003, and email, 12 November 2003.

### **Box 2: Learning by Doing and Operational Impacts**

While at times the feedback loop from evaluations to operations is too slow to significantly effect operations, these examples show that evaluations can have substantial programmatic impacts.

- Financing of sewerage subprojects in the Nicaraguan Social Fund (FISE) was suspended for a few years, based on the evaluation finding that the targeting of this type of subproject was pro-rich.
- Positive impact evaluation results have helped to keep programs alive and running across changes in political administrations in Mexico (PROGRESA/ Oportunidades) and Columbia (Familias en Acción).
- Positive impact evaluation results have also helped promote the expansion of successful development interventions in Mexico (PROGRESA/ Oportunidades), Columbia (Familias en Acción), and Nicaragua (Red de Protección Social).

Source: Laura Rawlings HD Learning Week presentation question and answer session, 13 November 2003, and personal communication, 3 December 2003.

## **Quasi-experimental Design**

Quasi-experiments are far more common than experimental designs. Of the more rigorous evaluations included, this review found seven times as many instances of quasi-experiments (including all forms of matching and natural experiment) than randomized designs (28 versus 4) (see Annex 4 for list).<sup>28</sup> The imbalance is not surprising given that there is often resistance to randomly assigning development interventions and practical difficulties in implementing the design. However, it also suggests there may be more scope for project teams to take advantage of opportunities for experimental design as they arise.<sup>29</sup>

The goal of quasi-experimental matching techniques is to identify comparison communities that are identical in every way except the presence of the intervention, to the greatest extent possible. As such, within the wide variety of matching techniques some are more convincing than others. For example, while propensity score matching within pipeline projects controls for both observable and unobservable selection characteristics, geographic matching controls only for location.

Here follow several common techniques, which are best used in combination to verify results. The reliability of matching estimates can be improved with baseline and follow-up data to further control for time-invariant characteristics, as well as field visits to verify the similarity of identified pairs.<sup>30</sup>

### **A. Propensity score matching**

Propensity score matching is a statistical technique to match communities or individuals by the likelihood that they would participate in a particular program (see Rosenbaum & Rubin 1985). With a large household survey, researchers generate the scores by regressing participation on variables used in the

---

<sup>28</sup> This includes planned evaluations as well as completed. Evaluations that use multiple techniques were counted multiple times.

<sup>29</sup> The present imbalance in the lack of experimental compared to quasi-experimental designs within Bank evaluation work was emphasized by Michael Kremer in his presentation at the July 2003 OED Conference on evaluation and development. See Duflo & Kremer 2003.

<sup>30</sup> The KDP & Conflict study emphasizes the importance of verifying the accuracy of propensity-score matched comparison groups with field visits.

participation selection process. They then match treatment communities with nonparticipating comparisons that have the same predicted probability of participating, i.e. by the closeness of their propensity scores.

### ***Propensity score matching with baseline data: Bolivia Social Investment Fund***

In areas where random program assignment was not possible, Newman et al. (2002) use matching techniques with baseline and post-intervention data to calculate difference-in-difference estimates of program impacts. Using observed characteristics within a baseline dataset, propensity score matching reduces pre-project differences between treatment and comparison groups. Difference-in-difference estimates then further control for time-invariant observed and unobserved characteristics between the two groups. The combination goes a long way to correct for selection effects.

## **B. Pipeline matching**

An innovative pipeline matching technique (using approved, but not yet implemented, projects) first used in the evaluation of the Honduran social fund presents another method to approximate the counterfactual (see Walker et al. 1999). The approach uses pipeline projects as the comparison group, and compares relevant indicators to those of completed projects. The mean difference in outcome variables across the two groups is then interpreted as the average treatment effect. Using pipeline projects corrects for selection bias, or the unobservable dynamics that enable some communities to organize and successfully apply for funding, which becomes important for demand-driven initiatives.

### ***Pipeline matching: Honduras Social Investment Fund (FHIS2)***

The evaluation of the social fund in Honduras (FHIS2) fielded a quantitative household survey in 48 project and 48 pipeline communities, complemented with qualitative fieldwork in a reduced sample of both areas. While there was no further matching beyond pipeline status, the research team used multivariate regression to control for observed factors for which data had been collected.

## **C. Natural Experiment**

Sometimes policies or program selection procedures use inclusion rules that create a natural experiment, or a relatively arbitrary distinction between two very similar groups that approximates a randomized experiment. This allows researchers to investigate program impacts by comparing relevant outcomes of participant groups to those of otherwise comparable non-beneficiary groups.

### ***Natural Experiment: Indonesia Urban Poverty Project (UPP2)***

Participation in UPP2 was based on the relative performance of kecamatan (sub-districts) within their kabupaten (district). Specifically, a composite score calculated with PODES (*Potensi desa*, annual village survey) data was calculated for each kecamatan, and the richest 20% of kecamatans within each kabupaten were excluded from UPP2 participation. This creates a natural experiment. First, two kecamatans with similar scores in different kabupaten could be treated differently because of a difference in their position relative to other kecamatans within their respective kabupaten. As an illustration, suppose two middle income villages have exactly the same score, however one is located within a rich district and the other in a poor district. The first falls within the poorest 80% of villages *within its rich district* and qualifies for UPP2, while the second is relatively better off than 80% of the other villages *within its poor district* and is excluded from UPP2. Thus the villages have the same composite score but only one participates in the program. Second, two similar kecamatans within a single kabupaten near the cut-off score could be treated differently. In this case, the two middle income villages are located in the same district, however the first village's score places it in the 80<sup>th</sup> percentile for the district and it qualifies for UPP2, while the second village's score ranks it in the 81<sup>st</sup>

percentile and renders it ineligible. As before, while the two villages are very similar, only one participates. Thus in both cases, the participation rule creates an opportunity to select otherwise comparable communities from treatment and non-participant groups. With the help of econometric tools, researchers can then estimate project impacts by comparing relevant outcomes across the two groups.<sup>31</sup>

#### **D. Instrumental variables**

Less commonly, evaluations use instrumental variables (IV) to sort out causality when participation is correlated with outcomes of interest. To illustrate, perhaps an evaluation aims to measure the impact of a CDD project on community social capital. However, since the project uses a demand-driven application process, a community's capacity to organize may make it more likely to participate in the program *and* be correlated with its level of social capital. Organizational capacity may be hard to measure, and if the researcher omits it as a factor in the analysis, the results may well be biased: perhaps the apparent positive project impacts on social capital are due to the fact that organizational capacity motivated some communities to participate in the project and to form associations, and not that participation in the CDD project caused more associations to form within the community. Two stage regression with instrumental variables can solve this problem of an omitted variable (organizational capacity) biasing our estimates by first predicting participation, and then examining how the outcome variable (social capital) varies with those estimates. To serve as a valid instrument, the sought-after variable needs to be correlated with program participation, but unrelated to outcome variables beyond its effect on participation. Sometimes good instrumental variables can be found in the context of natural experiments, discussed above, where an exogenous factor influences participation but is not associated with outcomes.<sup>32</sup> Two drawbacks are that first, it is very difficult to find good instruments; and second, IV requires large survey samples to compensate for large standard errors from the two-stage regression.

#### ***Direction of causality & instrumental variables: Peru Social Fund (FONCODES)***

The most convincing quasi-experimental studies combine various techniques to clarify the direction of causality. With quantitative household survey datasets from 1993 and 1996, Paxson and Schady (2002) use instrumental variable analysis to verify the results from nonparametric regression and to clarify the direction of causality.

Regarding the short-run effects of FONCODES on school attendance, the study found that worse-off districts demonstrated large gains in school attendance while better-off did not, and that among poorer districts, those that experienced larger attendance gains tended to receive more FONCODES funding. However, researchers wondered whether an unobserved characteristic, like changes in district preferences for education, might explain the positive association between FONCODES spending and increased attendance: perhaps “a poor district in which people begin to care about education may have larger increases in school attendance rates and generate more proposals for FONCODES school funding than an equally poor district in which preferences for education do not change” (p. 313). If true, then the positive relationship between FONCODES spending and attendance does not reflect the causal impact of FONCODES, but rather the changes in education preferences that increased both program participation and school attendance. So, to test the robustness of the initial positive results and to confirm the direction of causality, the authors then used districts' percentage of votes for Fujimori, the incumbent president, as an instrumental variable for FONCODES participation, as greater FONCODES funds were directed to districts with low Fujimori support, presumably to build

---

<sup>31</sup> Vivi Alatas email, 29 October 2003.

<sup>32</sup> See Angrist & Krueger (2001) for an explanation and list of study examples that use instrumental variables to analyze data from natural and randomized experiments.

popularity.<sup>33</sup> Assuming that “political preferences of districts were not correlated with changes in preferences for education,” this confirmed the finding that districts with higher FONCODES spending experienced greater gains in attendance rates (p. 313).

## **E. Cluster Analysis**

Used in a variety of research fields, cluster analysis presents another set of statistical techniques that can be used to identify similar non-participant groups. It relies on multivariate regression to organize “information to produce relatively homogenous groups,” or classify observations so that those within each cluster are similar to one another in some sense, and dissimilar to those in other clusters (Asia Pacific Policy Institute 2003, A2-1). Biologists use a kind of cluster analysis to classify disparate organisms into species, genus, family, through to kingdom.<sup>34</sup>

### ***Cluster Analysis: Philippines Comprehensive & Integrated Delivery of Social Services Project (KALAHI-CIDSS)***

The Philippines KALAHI baseline study matched future participant and non-beneficiary groups with cluster analysis along several variables within the following categories: human capital, housing and amenities, accessibility in terms of roads and distance to trade centers, and municipality population and land area. Researchers followed an “agglomerative hierarchical procedure,” which begins with as many clusters as observations, and systematically merges clusters with those closest in distance along the various dimensions specified, until one large cluster remains. The ongoing Brazil PCPR evaluation also uses cluster analysis as a matching technique.

## **Triangulation**

Effective triangulation involves verifying study findings with complementary sources of data and/or alternative analytic methods, which can take many forms. Earlier discussion already described how both the Bolivia SIF and Peru FONCODES evaluations used different analytical techniques on similar datasets. Particularly notable in this area, Newman et al. (2002) used three separate analytic techniques on quantitative data to estimate the Bolivia social fund’s impact on child mortality: propensity score matching with recall data, life table estimates with panel data, and estimation of a Cox proportional hazard function (pp. 255-6). Further examples of triangulation with multiple sources and multiple methods are contained in the section on mixed methods.

## **Tenet 2: Baseline Data**

Reference data on the pre-intervention state of beneficiaries is extremely useful in gauging program impacts. Combined with subsequent follow-up data, at minimum baseline data enables reflexive comparison. Even better, if baseline and follow-up data is also collected for comparison groups, it enables a difference-in-difference estimation of program effects.

Overall, about half (18 of 35) of the included evaluations use baseline or panel data.<sup>35</sup> However, the distribution over time reveals a trend toward more frequent collection of baseline data in recent projects: relatively few completed studies used baseline data (3 of 10), whereas the majority of planned or ongoing

---

<sup>33</sup> Paxson & Schady also use the composite FONCODES index used to prioritize investments based on unmet needs prior to interventions as another instrument that determines participation and is plausibly uncorrelated with subsequent changes in education preferences (p. 313).

<sup>34</sup> Manchester Metropolitan University describes cluster analysis on the web at <http://149.170.199.144/multivar/ca.htm>

<sup>35</sup> For some of these evaluations baseline data is not collected before the program begins, only before the program is introduced in some areas.

studies (18 of 25) incorporate baseline or panel data (see Annex 4 for list). This suggests that program managers are responding to concerns voiced about the lack of baseline data collection.

Yet in practice, plans to evaluate often take hold after programs have begun, eliminating the possibility for a true baseline survey. Even where baseline data exists, it may be of questionable quality. In such cases, a nice feature of matching techniques is that they can be used to evaluate impacts where only ex post survey data is available. In addition, retrospective questioning can approximate baseline information, although the reliability of recall data is a subject of concern.

### ***Matching without baseline data: Armenia & Zambia Social Investment Funds***

Facing a lack of baseline data, evaluations of social funds in Armenia and Zambia used similar methods, combining two matching techniques to construct comparison groups. Chase and Sherburne-Benz (2001) and Chase (2002) both use quantitative ex-post data from nationally representative surveys with oversampling in participant communities. They compare post-project outcomes in beneficiary communities with those in two separate comparison groups: pipeline and propensity score-matched communities. Several other rigorous evaluations also use matching techniques where no baseline data was available, including studies of Albania ADF, Benin AGeFIB, and Nicaragua FISE.

## **Tenet 3: Mixed Methods**

Rao and Woolcock (2003) emphasize the value of combining quantitative and qualitative approaches to provide a more comprehensive and nuanced understanding of program impacts than either approach can generate alone. They review the potentially complementary strengths of each, including how quantitative data and analysis can permit generalizations based on statistical representation, provide evidence for causal relationships, and rely on objective indicators; while qualitative data and analysis provide information on context, delve into issues of process, and explore concepts that are difficult to quantify (p. 167). The authors suggest practical integration strategies for program evaluation, enabling both measurement of outcomes and understanding of processes.<sup>36</sup>

Exploiting different types of data and collection techniques can be an effective technique to triangulate findings and explain initially puzzling results. While many evaluations reviewed in this study demonstrate a concerted effort to cross-check results with other sources of data or analytic methods, some missed opportunities to triangulate findings when multiple sources existed or other methods would also have been appropriate for the available data. A typical problem that prevents mixed methods studies from taking advantage of opportunities to triangulate is the challenge of aggregating and synthesizing qualitative data. This problem arose in multiple interviews, in the context of several different projects.<sup>37</sup> Without effective synthesis, the wealth of valuable insights from qualitative work is lost. To explain the problem's cause, two interviewees cited the shortage of individuals with the unique set of skills to effectively identify recurrent themes within large qualitative transcripts. Another explanation involves the

---

<sup>36</sup> Rao and Woolcock refer to a useful framework in Hentschel (1999) that classifies data by how qualitative versus quantitative it is and collection methods by how more or less contextual they are.

<sup>37</sup> Synthesis problems were mentioned regarding beneficiary assessments of Madagascar FID3 and Malawi MASAF, and the qualitative component of the Yemen SFD. The evaluation of St. Lucia PRF discarded earlier qualitative work that was of unsatisfactory quality.

lack of access to or familiarity with analytic methods and data processing techniques to synthesize qualitative data.<sup>38</sup>

The selective sample reviewed here suggests a recent shift in evaluations towards more mixed-methods approaches: a higher proportion of planned or ongoing evaluations (16 of 22) use a mixed methods design as compared to that of completed evaluations (6 of 13) (see Annex 4 for list). Given the importance of mixed method approaches and common practical problems in implementation, here follow a few examples of particularly strong integrated designs and effective frameworks to triangulate findings.

### **Iterative with feedback**

An iterative approach implies that preliminary results from one type of data collection are explicitly intended to inform the design of a further round of data collection. Frequently early qualitative work is used to understand project processes, explore research topics and develop survey instruments; while quantitative work then tests the generality of hypotheses that emerge from the qualitative exploration. Yet to support ongoing exchange between research approaches, iteration involves “regularly returning to the field to clarify questions and resolve apparent anomalies” (Rao & Woolcock 2003, 175).

#### ***Indonesia Urban Poverty Project (UPP2)***

The evaluation of the Indonesia UPP2 uses a comprehensive mixed method design (Rao & Woolcock 2003). To start, initial insights gained from process-focused field visits in UPP1 project areas informed the design of subsequent quantitative investigation. As one example, the interdisciplinary team was surprised by the “key role that facilitators played in the success or failure of a project on the local level,” and thus included an in-depth structured questionnaire specifically for local facilitators in the larger quantitative survey methodology (p. 176). In total, the quantitative baseline will involve five survey instruments in each village, including questionnaires for households, community secondary data, village leader, local activists, and neighborhood activists. To supplement, the research team fielded a separate qualitative baseline that included comparative case studies of twelve communities, using focus groups and in-depth interviews. An interesting aspect here is their use of two snowball samples: one radiating from the municipal office to explore the formal government network, and the other from a local mosque or activist group to investigate informal networks and associations. Together, the qualitative work provides in-depth insights into the roles and processes of local decision-making, while the quantitative work tests the generality of hypotheses emerging from the qualitative exploration.

#### ***Nepal Rural Water Supply and Sanitation (RWSS)***

The technical proposal for the planned Nepal RWSS evaluation emphasizes that the qualitative and quantitative “data collected, methodologies used, focus of the research questions and the mode of analysis need to be developed in unison” (CECI-Nepal 2003, 20). Pairing anthropologists and statisticians, the research team will conduct quantitative household surveys and qualitative process studies that include 12 month stays within selected communities, focus groups, interviews and PRA techniques. The quantitative analysis will “identify the components of initiatives that are most successful and therefore deserve greater investment and it will reveal the correspondence between specific activities and effective indicators to measure their intended progress;” while the qualitative work will focus on “understanding and describing the processes that occur during empowerment and social inclusion, barriers to these processes and the nature of relations between different variables” (pp. 24-27).

---

<sup>38</sup> However Rao and Woolcock (2003) chapter and the Barron et al. (2003) KDP & Conflict field guides are good resources.

## **Merged methods**

A still more integrated merged approach involves combining data types and methods.<sup>39</sup> The merging idea is particularly useful when constructing smaller scale qualitative studies. Oftentimes ethnographic or process-focused studies conduct in-depth exploration of participant groups only. However, especially regarding project impacts, the degree of representation attributed to findings can be greatly enhanced by taking a more quantitative approach to sample selection.<sup>40</sup> This shift involves extending the qualitative investigation into similar non-beneficiary areas.

### ***Indonesia KDP & Conflict Study***

The KDP & Conflict study also presents an example of a comprehensive iterative design, however here we focus specifically on how the research team used quantitative methods to select research sites for extended qualitative fieldwork. At the broadest level, researchers identified two provinces that were demographically very different from one another. Within each province, they then selected one district with high capacity to manage conflict and one with low capacity. Finally, they used propensity score matching to identify villages within each district that were otherwise comparable, except for the fact that one had participated in KDP and the other had not. The central advantage of such purposive sampling is that it provides confidence that “any common themes emerging from across either the program or nonprogram sites are not wholly a product of idiosyncratic regional or institutional capacity factors” (Rao & Woolcock 2003, 178).

### ***Thailand Social Fund***

Similarly, the upcoming Thailand SIF social capital study plans to use quantitative methods to select the sample of participant and non-beneficiary matches for qualitative study. Specifically, 200 of the villages covered by the 30,000 household national socioeconomic survey later participated in the social investment fund. The research team plans to use propensity score matching within that dataset to identify similar non-participant villages for comparison. Qualitative research will then be conducted in both areas, using extended, guided interviews.<sup>41</sup>

## **Scaling Down Sound Designs**

To conclude this section, when faced with severe resource constraints, the most effective approach is to scale down sound evaluation designs. This may include narrowing the range of topics studied or reducing the scope of fieldwork. Maintaining robust sampling and analytic designs on even the smallest scale can improve the reliability and usefulness of findings.

### ***Scaled down evaluation design: Jamaica Social Investment Fund***

For comparison groups, Rao and Ibáñez (2003) matched five randomly selected communities targeted by the project to five non-participant comparison communities through a combination of poverty scores and field observation. Without baseline data, they used retrospective questioning to approximate pre-intervention information. Regarding mixed methods, fieldwork included qualitative focus groups and in-depth key informant interviews as well as a quantitative household survey covering 50 randomly selected households from each community. As a result, while the findings do not necessarily generalize to the JSIF program as a whole, they shed light on how JSIF projects both depend on and potentially effect social capital.

---

<sup>39</sup> For further discussion of merged methods, see Marsland et al. (2000).

<sup>40</sup> Woolcock (2001) explains “thinking quantitatively, acting qualitatively” with an example from St. Lucia.

<sup>41</sup> Rob Chase email, 17 October 2003.

## **V. Concluding Recommendations**

Looking at the aggregate picture of CDD evaluations, this review confirms that while many high quality studies provide insight into CDD project impacts and performance, room remains to improve both the rigor of evaluation designs and topical coverage of investigation. A few general recommendations about how to pursue the evaluation agenda within the Bank seem relevant.

### **Examine evaluation performance across all Bank operations**

Since the need for more, better quality impact assessment is clearly not unique to CDD, the Bank should initiate honest stocktaking exercises in all realms of operations. These studies should aim to evaluate the overall quality of methods, identify key topics for further empirical investigation, and identify opportunities for research coordination across projects, regions and sectors.

### **Explore priority research areas**

With regard to topics of investigation, the CDD community should round out current empirical and experiential knowledge about CDD project impacts with further rigorous evaluation of key hypotheses. Central areas with relatively less research coverage include: CDD impacts on social dynamic issues of gender, conflict and leadership; project impacts on, and interaction with, decentralization processes and public sector management; the comparative performance of CDD versus more centralized approaches in achieving desired outcomes; the impact of various components within a project, to help “unpack the black box” of evaluation work;<sup>42</sup> long-term sustainability of infrastructure outcomes and welfare impacts; and preference targeting. Exploring some of these topics will require the development of new instruments (which partially explains why they have been relatively neglected thus far). Researchers should investigate borrowing relevant tools from other sectors, for example indicators to measure the extent of decentralization.

### **Uphold basic evaluation principles**

For those considering implementing an evaluation, several straightforward steps would address common methodological constraints. Broadly, evaluation designs should incorporate baseline and follow-up data collection for participant and comparison groups, and exploit opportunities for randomized trials. Combining information from quantitative and qualitative sources provides a more comprehensive view of CDD impacts and processes. A sound sampling framework and evaluation methods can improve the reliability of even very small scale assessments facing tight resource constraints.

### **Increase early evaluation planning**

Designing evaluations while designing projects (rather than during or after project implementation) makes it easier to implement basic evaluation principles and provides several other advantages. Concurrent planning can help sharpen and articulate specific project goals; allow collection of true baseline data, including time to develop and field test survey instruments; and enable an experimental design through randomized program placement. As such, early planning can relax constraints on the quality of an evaluation without increasing the budget. Integrated planning processes and early, sustained interaction between evaluation and project teams can further encourage greater buy-in by governments and project management, and ensure that the evaluation will provide operationally relevant details.

---

<sup>42</sup> Emphasized by Howard White in early review meeting for this study.

### **Increase funding from the center**

The tension between the global public good of evaluation results and the information needs and resource constraints of projects is not easily resolved. However, if the CDD and broader development communities are serious about filling existing knowledge gaps with reliable empirical evidence, we must mobilize greater financial and technical resources to support high quality evaluation work. This will clearly involve greater direct evaluation funding from the Bank.

As one idea, the Bank could provide matching funds to client governments specifically for evaluation work. That is to say, the Bank could have a standing policy to contribute funds equal to what governments are able to spend themselves on a given impact evaluation. This would provide an incentive to do evaluations more often, and to do them better. This could help increase the scale of given evaluations and the rigor of methods, as well as finance exploration of topics that may not be high operational priorities. This would also help reduce externalities by getting more of the parties who benefit from impact evaluation findings to contribute resources to the generation of that information.

### **Enhance information sharing**

An effective architecture for information sharing can help facilitate improvements in the quality and scope of CDD evaluations. In particular, knowledge of CDD performance would increase with greater exchange of information on evaluation work, including negative or inconclusive results and planning documents. This should extend to tools, including the sharing of survey instruments and designs, with clear notation of the context in which they were developed; and dissemination of tools for aggregating and synthesizing qualitative data. The website that accompanies this study<sup>43</sup> joins other initiatives to facilitate information sharing processes.<sup>44</sup>

More effectively linking existing resources (internal and external) into a more centralized clearinghouse would ease sharing across sectors and regions. As a start, there should be clearer links between the PovertyNet impact evaluation website, the Bank's internal impact evaluation site, web pages that various anchors and regions maintain on evaluations (including ours), and evaluation pages from external organizations. Given the frequency that topics like public sector reform, decentralization and social capital occur across many different kinds of projects, a single location that provides information on approaches to evaluate them would be a valuable resource. Lastly, a public repository of information on ongoing evaluations would make it more difficult to bury evaluation results, decreasing publication bias.

### **Establish systems for quality assurance**

Setting up quality control mechanisms for impact evaluation within the Bank would help improve the quality of evaluation designs, ease concerns about the independence of research teams, and enhance the reliability of findings.<sup>45</sup> This is especially useful given concerns about the quality of evaluation work, the uneven distribution of familiarity with evaluation techniques, and the fact that most evaluations of Bank projects are not published or subjected to a formal review process. A suggestion here would be to assemble a peer review board of impact evaluation experts that are readily available to critique evaluation designs and provide constructive feedback. Incorporating peer review from the beginning of the

---

<sup>43</sup> <http://Inweb18.worldbank.org/ESSD/sdvext.nsf/09ByDocName/Evaluations>

<sup>44</sup>Notably the East Asia and Pacific Region is working to improve CDD evaluation and assemble and disseminate evaluation tools. (Susan Wong interview, 2 July 2003)

<sup>45</sup> We encountered a number of concerns about the independence of evaluation teams and the disincentives to reporting critical assessments. These range from involved Bank staff burying negative findings, to "independent" local consultants with no incentive to criticize projects run by members of their community. While absolute independence may not be a practical reality, increasing the rigor of evaluation methods and external review would make it harder to manipulate findings.

evaluation process would be particularly valuable, commenting on terms of reference and planning documents while there is still time to establish a sound evaluation design.

## Annex 1: Methods Inventory

This annex is intended as a reference section to catalog the methods used in each evaluation reviewed in this report. Evaluations are presented in alphabetical order by country. Evaluations in very early planning stages are presented separately at the end, since limited information about them was available at the time of writing this report.

### *Albania Development Fund (ADF) [ongoing]*

This quantitative study uses ex-post household data from the 2002 LSMS, covering 3600 households and 288 communities, to compare outcomes in villages that received an ADF subproject over the last 5-10 years and propensity score-matched comparison communities (Carletto et al. 2003). Main topics of investigation include subproject impacts on access to basic infrastructure and community social capital. Several different matching techniques were used to estimate the mean effect of treatment on the treated. For these estimates, the sample size of treatment groups by type of subproject are as follows: 1463 households (92 communities) for roads, 784 households (51 communities) for water projects, and 152 households for schools; plus 122 communities for social capital effects. The various matching techniques drew different numbers of comparison units from available pools of: 2136 households (196 communities) for roads, 2815 households (237 communities) for water projects, and 3447 households for schools. Overall, initial drafts of the study found no significant impacts from any of the infrastructure types on access or social capital. As of December 2003, the research team is using nonparametric regression analysis to explore potential heterogeneous effects along the propensity score distribution.

### *Armenia Social Investment Fund 1 (ASIF1) [completed]*

Using quantitative data from a nationally representative household survey with oversampling in ASIF areas, the analysis examined household-level outcomes in ASIF school and water project areas relative to two comparison groups: propensity score-matched communities and pipeline communities (Chase 2002). In addition to examining overall impacts, the sample was stratified by conflict and earthquake zones to investigate heterogeneous impacts in different areas. Main topics of investigation are poverty targeting, education and health outcomes, and social capital. Sample sizes varied for different parts of the study, as summarized in the table below.

**Table: Sample sizes for ASIF1 evaluation**

	Treatment			Propensity Matched			Pipeline Matched			Unmatched		
	O	E	C	O	E	C	O	E	C	O	E	C
Schools: Household	232	113	20	646	150	87	148	80	5	1298	247	208
Water: Household	340	160	60	1740	420	240	380	80	120	3600	700	580
Schools: Social Capital	620			1740			439			3600		
Water: Social Capital	340			1740			380			3600		

O = group's overall sample size. E = group's earthquake zone subsample size. C = group's conflict zone subsample size.

To determine overall targeting performance, as well as targeting within urban and rural areas, the targeting analysis: (a) compared mean expenditure data for households in ASIF communities with that of a random selection of communities, noting the statistical significance of differences; and (b) constructed concentration curves to compare the distribution of household per capita expenditure among ASIF households and among all households.

### *Armenia Social Investment Fund 2 (ASIF2) [ongoing]*

This evaluation is using beneficiary and institutional assessments in matched case studies of 6 treatment and 8 comparison communities to examine ASIF2 subproject effects on social capital. Early planning

documents propose two selection strategies to identify participant and comparison communities for investigation: one uses propensity score matching based on 2001 Integrated Living Conditions Survey (ILCS) national household survey data; and the other identifies approved projects that were not funded due to resource constraints. Qualitative fieldwork will use interviews and focus group discussions to explore topics of participation, partnership formation and capacity building. An earlier beneficiary assessment fielded in 2001 will also be analyzed. The institutional assessment will focus on macro-level issues of the political environment including the status of decentralization, governance, and community-level social relations.

***Benin Social Fund (AGeFIB) gender study [completed]***

This study used a qualitative anthropological approach, primarily interviews and focus groups with stakeholders (Walker 2002). The field work, covering 31 projects (11 infrastructure, 18 micro-credit, 2 non-AGeFIB micro-credit) was conducted during two one-month visits covering all five regions of the country. As such, while the results shed interesting light on gender dynamics in the context of this project, they are not necessarily representative of the project as a whole. The involvement of local AGeFIB staff in fielding the study was viewed positively as a means of capacity building and information sharing across regions.

***Benin Social Fund (AGeFIB) [ongoing]***

The current impact assessment of the Benin social fund, AGeFIB, is using a survey based on the Bank's CWIQ (Core Welfare Indicators Questionnaire) with additional modules on household consumption and cotton. To be able to assess the impact of AGeFIB's education subprojects (about two thirds of the total), 600 extra households were added to the CWIQ sample and a separate module on schools was included for both participant and control schools, modeled on the survey developed for the Zambia SIF 2000 impact assessment. While acknowledging the loss of seasonal variation in a one-time expenditure survey, the research team was able to improve comparability by fielding the survey in all households within a single season.<sup>46</sup> A representative sample of 6,000 households (2000 urban and 4000 rural) was surveyed. Ex-post propensity score matching will be used to examine sub-project impacts. A draft (green cover) Poverty Assessment report was just completed in October 2003.<sup>47</sup>

***Bolivia Social Investment Fund (SIF) [completed]***

Newman et al. (2002) used a variety of data sources and analytic techniques to evaluate impacts of the Bolivia SIF on education outcomes, health indicators and infrastructure improvements. Five data collection instruments were used: household surveys from 1993 and 1997; school and health center facilities surveys; community surveys; water quality samples; and student achievement tests. Research also took advantage of available data from the 1994 & 1998 Demographic and Health Survey (DHS). With baseline and post-intervention data, they used a combination of random assignment and propensity score matching for difference-in-difference estimation. Specifically, education impacts within the Chaco region were separately estimated using random assignment of the offer to participate amongst eligible schools, and propensity score matching to construct comparison groups. The results were triangulated for verification. Education impacts within the Resto Rural region and the impacts of health centers were estimated using propensity score matching. One section on child mortality impacts uses three separate estimation methods to triangulate results: propensity score matching with recall data, life table estimates for the change in mortality with panel data, and estimation of a Cox proportional hazard function. Reflexive comparison, based on household- and community-level surveys in beneficiary areas, was used

---

<sup>46</sup> John Elder interview, 8 July 2003.

<sup>47</sup> World Bank 2003b.

to measure impacts of the 18 water supply subprojects.<sup>48</sup> Separately, water quality tests were conducted for both the old water sources and new SIF-constructed ones.

***Brazil Rural Poverty Reduction Projects (PCPR) [ongoing]***

This evaluation is intended to measure the performance of PCPR in three states<sup>49</sup> in northeast Brazil—Bahia, Ceará, Pernambuco—along a number of criteria: household welfare impacts, social capital and governance impacts, poverty targeting, and relative cost-effectiveness. It is using a panel data setup for difference-in-difference estimation of program impacts. Household surveys are being administered in PCPR participant communities and in a control group with no project.<sup>50</sup> The overall sampling framework is also stratified by geographic criteria and project type to ensure statistically representative coverage. Fifty case studies are being conducted to collect more qualitative data for analysis of social capital and governance impacts. The baseline round is being collected at the time of project initiation, and the follow-up two years later. Data collection is first pilot-tested for feasibility. Impacts in PCPR communities will be measured at three levels of aggregation: (a) overall; (b) comparing across three levels of decentralized implementation within PCPR; and (c) by subproject type. (FECAMP 2001)

For primary welfare impacts, the technical proposal did not specify sample sizes, but did discuss investigating the feasibility of using pre-existing data to reduce the scope of primary data collection needed. It discussed statistical power calculations to determine the sample sizes necessary to estimate welfare impacts with desired precision, while also taking the cost of data collection into consideration. (FECAMP 2001, 40) This is expected to yield a sample size of about 3500 households, which will be used to collect data on indicators of welfare and social capital.<sup>51</sup> In addition to using quantitative household surveys, the evaluation includes case studies of 50 PCPR municipalities, to collect qualitative data. Analysis of welfare and social capital impacts, based on the household survey, will include the comparison group of areas with no project. Poverty targeting analysis will rely on other secondary data to see where PCPR areas fit in the rural poverty distribution. Cost-effectiveness analysis and case studies will not cover control areas beyond PCPR communities. (FECAMP 2001, 15-17) Municipalities will be matched for comparison using cluster analysis, with data from three sources: Technical Units of each state; the Agricultural Census of 1995-1996; the MIS for PCPR. (FECAMP 2001, 44-49)

***Burkina Faso National Agricultural Services Development Project 2 (PNDSA2) & Senegal Agricultural Services and Producer Organizations Project (PSAOP) [ongoing]***

The baseline surveys for these two sister evaluations have been fielded and follow-up surveys are planned for 2005, but not yet funded. The sample covers 300 villages in each country: 200 participant communities (100 with RPOs that submitted and implemented subprojects; 100 with RPOs that did not submit a subproject proposal, but nonetheless benefited from one); and 100 matched controls with no subproject, to establish data for difference-in-difference analysis. Case studies among some of the sampled villages will provide in-depth information to sharpen hypotheses to be tested in the full sample. The studies will also collect qualitative information on the nature of rural producer organizations and their internal dynamics. (de Janvry et al. 2001)

---

<sup>48</sup> The evaluation design section mentions that matching techniques were used to construct a comparison group from data gathered for the larger health subsample (p. 247), however the main presentation of findings for the impact of water projects (table 10) appears to be based on reflexive comparison for beneficiaries only (p. 261).

<sup>49</sup> A fourth state, Piauí, was initially slated for inclusion, but was dropped due to delays in implementation.

<sup>50</sup> The initial evaluation design included a second comparison group of areas with centralized project delivery, to compare outcomes from PCPR implementation with those under more traditional centralized government implementation. This was dropped due to budgetary constraints.

<sup>51</sup> Claudia Romano email, 6 November 2003.

***Cambodia Rural Investment and Local Governance (RILG) [ongoing]***

The planned impact evaluation of the Cambodia RILG project includes a quantitative socioeconomic baseline survey to be fielded in 2003 to a total sample of 1,920 households in 60 villages within 30 communes, half with access to RILG infrastructure and half without. No further matching of control communities to treatment communities beyond geographic proximity is mentioned. Quantitative fieldwork will be complemented by key-informant interviews in each village to gather data on community characteristics. At least one follow-up survey is planned, although funding is not yet in place. Four case studies will be conducted in one province. Beneficiaries' perceptions will be solicited through semi-structured interviews. (Helmert & Wallgren 2003)

A Knowledge, Attitudes, Practices and Beliefs (KAPB) baseline study has also been done. It includes a quantitative survey of perceptions of 1493 government officials, along with key informant interviews with high-level officials, and focus groups and 523 semi-structured interviews with community members not working in the public sector. Results were triangulated for confirmation. A follow-up study on governance plans to use the KAPB as a baseline.

***El Salvador Community-Managed Schools Program (EDUCO) [completed]***

Jimenez and Sawada (1999) investigate the impacts of EDUCO on education outcomes. They use the percentage of EDUCO schools in each municipality, as well as the formula for targeting EDUCO schools, as instrumental variables to control for student characteristics and selection bias in their regressions to compare achievement and attendance of rural students in EDUCO versus traditional schools. The single cross-section data set was collected by the Ministry of Education and contains observations on 596 ACE (Community Education Association) committee members and 1555 students in 311 schools, covering 162 municipalities out of the country's 262. Many schools in El Salvador are mixed, where EDUCO and traditional management structures share a building. To ensure robust results, the authors use pure schools only, which left a sample of 605 students in 30 EDUCO and 101 traditional schools. Survey data was compiled from five questionnaires: student, parent, school director, teacher, and parents association.

***Ethiopia Women's Development Initiative (WDI) [planned]***

The evaluation of the Ethiopia WDI is exploring the possibility of using random assignment of participation among interested potential beneficiaries to isolate program impacts on empowerment and economic outcomes. An interesting aspect of this study is a series of individual comparisons to explore the separate impacts of participating in program promotional and organizational work versus access to technical assistance and financing. Planning a mixed methods approach, early qualitative participatory techniques to explore empowerment in the Ethiopian context will inform the design of the quantitative survey. Three different comparisons are planned: (1) beneficiary vs. participating non-borrower, to gauge the effect of program financing beyond the effect of participating in the promotion activity only; (2) beneficiary vs. non-beneficiary borrower, to separate promotion and group formation effects beyond the access to financing; and (3) participating non-beneficiary vs. non-participating non-beneficiary, to separate out the effect of self-selection into the promotional activity without program effect. These individual-level comparisons will be complemented by a community-level comparison. The study remains in the planning stage, and further details of the evaluation design, including sample size, are being developed. (Christiaensen & Legovini 2003)

***Honduras Social Investment Fund 2 (FHIS2) [completed]***

Walker et al. (1999) use a combination of quantitative and qualitative methods to estimate subproject impacts on education, health and social capital, as well as assess targeting performance and community participation. A quantitative household survey was administered in 48 project communities and in 48 pipeline communities, covering 2600 households in total. Qualitative fieldwork included 62 key

informant interviews with project staff and local officials, plus focus group discussions in 15 project and 15 pipeline communities. The research team also used specialized survey modules for each type of infrastructure, covering the 48 completed and 48 pipeline projects divided over five infrastructure types, to examine issues of service quality, utilization and sustainability. An engineer assessed the quality of infrastructure provided by FHS and other providers on a three-value scale.

***India Dairy: Operation Flood [completed]***

The OED study of the India Operation Flood program takes a macro approach, as the program was part of a larger policy shift which redefined the dairy market in India (Candler & Kumar 1998). Since the evaluation examines the outcomes of the overall dairy market policy, of which Operation Flood was a relatively small part, it was impossible to isolate the impact of the Operation Flood investments from the effects of overall dairy market reform. The evaluation builds on earlier audits and studies from 1987 and 1996, in addition to fielding a participatory study of farmers and a resurvey of villages studied in 1983. The two village-level participatory studies covered 9 villages (5 treatment, 2 contaminated, 2 control) in the Mysore district of Karnataka state; and 12 villages (5 treatment, 5 contaminated, 2 control) in three milksheds in Madhya Pradesh from the 1983 survey, supported by a survey of secondary data on 200 villages in the same three milksheds. Research topics include impacts on income, women's empowerment, education and targeting performance.

***India Andhra Pradesh District Poverty Initiatives Project (APDPIP) & Rural Poverty Reduction Program (APRPRP) [ongoing]***

Contributing to the DEC CDD research initiative, this evaluation is assessing the initial impacts and targeting performance of APDPIP as well as establishing baseline data for a second related project, APRPRP, to enable eventual comparison between the two programs.<sup>52</sup> The six districts where the APDPIP is active were phased in over the three program cycles and the APRPRP will be implemented in all remaining districts, save the most advanced. Therefore, the sampling strategy for the APDPIP impact assessment incorporates DPIIP participants from each program phase and will use the APRPRP areas as a “without project” comparison group. Fieldwork in the DPIIP areas will revisit 72 villages that were covered in an earlier baseline study, but because of concerns over the quality of that data, the current survey will also include retrospective questioning. Within APRPRP districts, the baseline study will cover both future RPRP-participant and non-participant sub-districts (mandals), selected for similarity along key observable characteristics. Ten households will be visited per village. To ensure a sufficient number of potential and actual participants, household selection will not be entirely random, but weighted toward target groups, including scheduled caste members and households in the poorer end of the poverty distribution (as identified by an earlier participatory identification of the poor (PIP) exercise which covers 12 million households). While the exact distribution across villages and treatment areas remains to be determined, approximately 4,500 households will be included.

Several questionnaires will be administered in each village: 10 household, 1 village, and 1-2 habitation; plus short surveys on self help groups (SHG), village organizations (VO) and Mandal Samakhyas (MS); and a local government finances sheet. There are separate household questionnaires for males and females, to be conducted simultaneously with an enumerator of the same sex (if possible) to facilitate honest responses to sensitive topics, like domestic violence. Since both projects contain pilots of innovative sub-components (for example, support to the disabled), the questionnaires will collect information on the demand for such specialized services “to facilitate the use of data collected as a comparison group for other studies (either qualitative or quantitative).” Main topics for exploration regarding APDPIP impacts include “social attitudes and cohesion, productive activities, investment by the

---

<sup>52</sup> Klaus Deininger email attachment, 10 October 2003, “Aide Memoire,” [July 2003], 2. The research team consists of Klaus Deininger, Hanan Jacoby and Ghazala Mansuri, all from DECRG.

poor, school attendance (especially by girls), intra-household decision-making and allocation of resources.”<sup>53</sup> Another hypothesis to test is whether the rapid expansion of the program worsens targeting and project performance. Lastly, data will provide an opportunity to test the correspondence between qualitative measures of poverty from the PIP with quantitative measures from the household survey. Fieldwork began in late 2003, and the final report is anticipated in August 2004.

In addition, a substantial land titling and land purchase assistance program is envisioned in the second phase of the APDPIP. The land titling program is being done in conjunction with the government of Andhra Pradesh. DPIP community organizations will assist individuals in resolving disputes related to land and identify households who require assistance with land purchase. The baseline for this is planned for early 2004. Program placement is intended to be random. The research team has designed the questionnaire for the baseline (which will take place in December) to obtain information at the plot level on all land in the possession of sample households. In this case, they hope to follow plots over time to look at the productivity and equity impacts of land titling and purchase under this program.<sup>54</sup>

### ***Indonesia Kecamatan Development Project (KDP1, KDP2, KDP3) [ongoing]***

The evaluation of KDP takes an integrated approach that spans the three phases of the project over a ten-year period. Structures of the evaluation, both quantitative and qualitative, were put in place during the first phase of the project, collected de facto baseline data for phases 2 and 3, and are being further developed and expanded during the current phases.

The preliminary study of Indonesia KDP1, conducted by the Demographic Institute at the University of Indonesia, compares quantitative responses in surveys of villagers’ and community leaders’ perceptions of various process and impact outcomes. The sample included 74 pairs of KDP and non-KDP communities. Regarding sample selection, within each kecamatan participating in KDP, half the villages were randomly selected for eligibility to participate in the first year, and the other half were eligible to participate in the second year. The sample thus included first year participants as treatment groups, and used the other villages within the same kecamatan for the comparison group (Wong 2003). Since the two rounds of data collection were only one year apart, this data is intended to give a basic idea of short-term outcomes from KDP1, and to serve as de facto baseline data for KDP2 and KDP3. In each community 15 household and 15 community leader respondents were interviewed, and differences in responses across KDP and non-KDP areas were compared with the statistical significance of differences noted (one-sided hypothesis tests). The complete dataset included panel data from 2000 and 2001 on 2039 household respondents and 2054 community leader respondents, as well as 524 economic loan recipients (UOI 2002, 7). The report also includes qualitative findings based on direct observation of participation in KDP meetings.

Additional reports look at microcredit and infrastructure components. The microcredit report is based on direct observation from a three-week field visit and a review of secondary sources. While it does not measure impacts, it provides a good summary of reasons given for the poor performance of the microcredit portion of KDP, mostly compiled from local consultants’ reports; and it gives guidelines for designing a more effective microcredit component for KDP2 (Holloh 2001). An infrastructure study focuses on quality of infrastructure, cost effectiveness, and economic rates of return evaluated 235 KDP1 projects, however data problems resulted in only 41 being included in the cost-benefit analyses (Dent 2001, 5).

---

<sup>53</sup> Klaus Deininger email attachment, 10 October 2003, “Aide Memoire,” [July 2003], 2.

<sup>54</sup> Ghazala Mansuri email, 12 November 2003.

For the current and upcoming phases of KDP2 & KDP3, a feasibility and sample size study is currently underway. The quantitative impact evaluation work will most likely include a 4,600-household panel survey building on data collected for KDP1. Given that the KDP1 dataset uses a pipeline comparison group and KDP2 & 3 will expand to half the country, the communities included there are all treatment areas for KDP2 & 3. Methods for constructing a new comparison group are currently being developed. The upcoming evaluation will also include a study using Susenas national household survey data to examine impacts of KDP on poverty welfare indicators. In addition, evaluation work for KDP2 & 3 will include further evaluation of the microfinance portfolio; qualitative case studies; infrastructure audits; a study on the economic internal rate of return; a study on knowledge, attitudes, practices, and beliefs of government officials; and an evaluation of capacity-building.

Key performance indicators across all phases of the project include impacts on poverty reduction, technical quality of infrastructure, social capital, community participation, governance (responsiveness, transparency), access to microcredit, and cost effectiveness of KDP investments.<sup>55</sup>

### ***Indonesia KDP & Conflict study [ongoing]***

The KDP & Conflict study uses a combined qualitative-quantitative approach to examine factors that contribute to conflict and conflict resolution and how participation in KDP projects may impact conflict management (Barron et al. 2003a). The qualitative fieldwork incorporates three research tools: semi-structured interviews, focus group discussion and participant observation. Qualitative fieldwork aims to collect both case data on specific instances of conflict and general socioeconomic data on the contexts studied. The quantitative survey instrument is being designed as the qualitative fieldwork takes place, explicitly to incorporate and test hypotheses that arise from the qualitative research. The study intends to include a short conflict module in an upcoming national decentralization survey. This household survey will further collect quantitative data on the nature of conflict and broader demographic and institutional characteristics, including retrospective questioning to approximate baseline data.

Regarding sampling, the qualitative work will cover eight kecamatan (sub-district) and the quantitative will cover twenty-four kecamatan. The broad sampling framework covers two very different provinces, within which one district with high capacity to manage conflict and one with low capacity were selected, and then within those one kecamatan that had a KDP1 project for three years and one propensity-matched that is scheduled to have a KDP2 project in future. The quantitative work includes two additional kecamatan pairs per district. Within the sub-districts, the sampling strategy further aims to identify a series of matched conflict cases. These include similar situations of tension where one turned violent and the other was resolved peacefully; similar conflict situations in KDP versus non-KDP areas; peaceful resolutions of conflict in generally violent areas; and violent cases of conflict in typically peaceful areas. In total, 56 case studies are being developed.

### ***Indonesia Support to Conflict-Ridden Areas Project (SCRAP) [very early planning]***

The project is currently in the pipeline, and scheduled to go to the Board in the fourth quarter of Fiscal Year 2003. Planning documents for SCRAP indicate that evaluation work will include baseline and follow-up surveys in participant and comparison areas to establish data for difference-in-difference estimates of project impact. Both qualitative and quantitative methods will be used. A pre-project situation analysis will be conducted in each district and a selection of villages. The situation analysis and additional qualitative fieldwork including focus group discussions and in-depth interviews will inform the design of a quantitative household survey. Evaluation work will further include in-depth qualitative case studies, community participatory monitoring and evaluation, a pilot of community-based health and

---

<sup>55</sup> Susan Wong interview, 12 November 2003.

education monitoring, and a module to assess the impact of training activities that uses a series of follow-up questionnaires and direct observation. It is worth noting that work being done for the Indonesia KDP & Conflict study (see above) will feed into the M&E and survey designs for SCRAP. Key research topics include household welfare, service access and quality, private sector development, village governance, social capital, and conflict resolution.<sup>56</sup>

### ***Indonesia Urban Poverty Project 2 (UPP2) [ongoing]***

An extensive mixed-methods evaluation of UPP2, involving a quantitative survey and in-depth case studies, is in the early stages of implementation. Main research topics include UPP2 impacts on collective action, community well-being and poverty alleviation. Baseline, mid-term and follow-up surveys are planned for 157 UPP2 Kelurahan (villages) and 98 comparisons to construct a panel dataset for difference-in-difference analysis. The government is in the process of contracting the quantitative baseline, which will involve implementing five survey instruments in each Kelurahan: 30 household questionnaires, one community secondary data questionnaire, one village leader questionnaire, one group interview with 4 to 8 local activists, and multiple neighborhood activist questionnaires.

Two identification strategies for selecting matched controls make use of the natural experiment created by the arbitrary eligibility cut-off established for UPP2 participation.<sup>57</sup> Based on a composite score calculated with PODES (*Potensi desa*, annual village survey) data, the richest 20% of kecamatan (sub-districts) within each kabupaten (district) were excluded from UPP2. Therefore, similar kecamatan in different districts are treated differently (included/excluded) because of their score relative to other kecamatan within their respective districts, allowing for matched comparisons to be used as the first identification strategy. The second identification strategy uses regression discontinuity design, since similar kecamatan within the same province that have scores just below/above the cut-off mark are included/excluded. In both cases, excluded kecamatan similar to those that were included serve as good control groups.

A qualitative baseline study has already been fielded in a sub-sample of locations, drawing on methodologies developed for the KDP & Conflict study. Qualitative data was collected through focus groups, interviews and direct observation. The case study sampling frame also involves matched pairs of control and participant districts, selecting the most urbanized and most rural villages and then poorest and wealthiest neighborhood within each. Field team members compiled an inventory of collective action cases from which three were chosen for in-depth study. Further details will be developed when the full research plan is drafted in December 2003.<sup>58</sup>

### ***Jamaica Social Investment Fund (JSIF) [completed]***

Rao and Ibáñez (2003) used in-depth case studies of five matched pairs of communities combining qualitative and quantitative methods to examine social capital, preference targeting and community capacity for collective action. Five randomly selected communities targeted by the project were matched through a combination of poverty scores and field observation to five non-participant comparison communities. Fieldwork included qualitative focus groups and in-depth key informant interviews as well as a quantitative household survey covering 50 randomly selected households from each community. Retrospective questioning was used to approximate baseline data including ex-ante preferences. In the household level, responses were compared across households in the participating communities and the

---

<sup>56</sup> Susan Wong email attachment, 2 July 2003. "Monitoring and Evaluation System Concept Note For SCRAP" [Draft 2, 9 April 2003].

<sup>57</sup> Vivi Alatas email, 29 October 2003.

<sup>58</sup> Menno Pradhan email, 20 October 2003.

average of five “nearest neighbor” propensity-score matched households from the matched “control” community. These comparisons estimated the average treatment effect, with bootstrapped standard errors to test for statistical significance. The findings shed light on how JSIF projects both depend on and potentially affect social capital, but the constraint of the small sample size means they do not necessarily generalize to the JSIF program as a whole.

The main topics of the evaluation focus on social dynamics (specifically social capital, participation, and conflict), on leadership, and on preference targeting.

#### ***Malawi Social Action Fund (MASAF) [ongoing]***

The poverty alleviation and sustainability review is a two-part study of MASAF (Centre et al. 2002). The first inception phase is largely a desk study complemented by stakeholder interviews. Poverty targeting is measured using correlation coefficients between MASAF expenditures per capita and poverty headcount index. Political targeting is measured using correlation coefficients between expenditures and political alignment, as measured by the vote share for the political party currently in power. The second phase calls for a case-study analysis using focus group discussions of factors associated with positive outcomes in poverty reduction; complemented by an analysis of existing datasets, including IFPRI income maps, to determine MASAF impacts on poverty. It is unclear if the number and selection of case studies will be sufficient to be broadly representative. Implementation of this study was delayed by drought and food shortages during 2002/2003. As of mid-November, the final draft report is with reviewers and will be available as soon as comments from reviewers have been incorporated into the final draft.<sup>59</sup>

#### ***Moldova Social Investment Fund (MSIF) [ongoing]***

The terms of reference (TOR)<sup>60</sup> for the recently contracted baseline study include a general outline for the MSIF impact evaluation. The document calls for retrospective questioning regarding participant status three years prior to explore impacts of MSIF1 projects, while establishing baseline data on MSIF2 participants. The evaluation should include a combination of quantitative household survey and qualitative focus groups, sampling in participant and non-beneficiary communities, to a maximum of 2000 respondents divided across project types. Potential areas of investigation are MSIF transparency, beneficiary satisfaction, empowerment, participation, targeting, social capital, institutional development, sustainability, specific sub-project outcomes, development of private contractor markets, and community perceptions of MSIF performance. The TOR further stipulates that an independent consulting firm will carry out the evaluation. Field work is beginning in November 2003, and a report should be completed by February 2004.

#### ***Nepal Rural Water Supply and Sanitation (RWSS) [planned]***

The evaluation design indicates a mixed methods approach, combining quantitative individual surveys with a process study involving several months of observation in selected communities, focus groups, interviews and PRA techniques. The evaluation will compare relevant outcomes across participant, pipeline, and non-participant communities. The first round of analysis will compare RWSS1 participant communities with RWSS2 pipeline communities for an ex-post impact assessment, while further collecting baseline data on non-participants. Contingent on obtaining funding for a potential second round, follow-up surveys will build on this data to establish difference-in-difference estimates across the three groups once RWSS2 projects have been implemented. By including RWSS1 participants in the follow-up survey, the research team hopes to attain information on longer term project impacts.

---

<sup>59</sup> Hope Phillips emails, 3 July 2003, 12 November 2003.

<sup>60</sup> Anush Bezhanyan email attachment, 9 July 2003, “TOR”.

The sampling frame will be too small to be statistically representative at the community level, but will reflect regional, ethnic and project implementation diversity. It includes a total of 60 communities: 20 RWSS1, 20 RWSS2, and 20 controls. Twenty-two households will be randomly selected per community, interviewing one man and one woman in each, yielding a total of 2640 individuals, sufficient for statistical representation in terms of individual effects. Control communities will be drawn from communities that NGOs have identified as potential communities, but which have not yet become involved in the program.<sup>61</sup>

The main topics covered in the evaluation are household income, health outcomes, and access to infrastructure, and the interaction of these outcomes to empowerment and inclusion, particularly for women and Dalits.

### ***Nicaragua First Emergency Social Investment Fund (FISE1) [completed]***

Pradhan and Rawlings (2002) used ex-post household survey data from FISE records and the 1998 LSMS to construct two comparison groups, one matched by facility characteristics, the other by propensity score, to measure FISE impacts. For the analysis based on facility matching, sample size ranges were 85 – 358 for education variables and 25 – 1196 for health; while for the propensity score matched section, sample size ranges were 77 – 341 for education variables, 36 – 1169 for health, and 30 – 451 for water/sanitation. In addition, two levels of targeting were analyzed: (a) community-level using all potential beneficiaries, and (b) household-level targeting using direct beneficiaries only. For each group, the benefit incidence, concentration coefficient and share of FISE funds to those below the poverty and extreme poverty line were calculated; as well as separate figures for education, health, water, sewerage and latrine projects.

Case studies were also conducted within the ex-post evaluation sample comparing 20 pairs of FISE and non-FISE health centers; and 24 pairs of FISE and non-FISE schools (both run by the Ministry of Education); plus technical audits for 10 FISE latrines and 10 FISE water systems (World Bank 2000). This report also included a FISE household survey, compared against nationwide data from the 1993 LSMS, 1992 and 1998 FISE poverty maps, and a Beneficiary Assessment (BA). The survey data was used to calculate benefit incidence relative to poverty, overall and for each sector of interventions. The BA served to gauge beneficiaries' feelings on participation. The sample consisted of 43 projects in 22 municipalities, with a total of 246 key informant interviews and 24 focus groups, all beneficiaries, with no comparison group for this component of the evaluation.

The main topics of analysis were education and health outcomes, access to water and sanitation infrastructure and poverty and preference targeting.

### ***Pakistan Aga Khan Rural Support Project (AKRSP) – Khwaja study [completed]***

Khwaja (2002) examines community-maintained infrastructure projects to determine the relative effects of community- versus project-specific factors on project maintenance. Data was gathered on 132 infrastructure projects in 99 communities, randomly selected from the population of AKRSP communities. The sample included 33 sites with two infrastructure projects within a single community, allowing Khwaja to isolate the effects of project-specific factors from the effects of community characteristics. Data was collected from four sources, including three questionnaires (group, individual and hereditary leader) and a technical survey by engineers. The combination of community and technical surveys provided data on both participant perceptions and contextual details as well as objective technical assessments of maintenance success. The maintenance indicator used combines the three highly positively

---

<sup>61</sup> CECI-Nepal 2003; Lynn Bennett interview, 17 July 2003.

correlated measures of physical score (proportion in initial physical condition), functional score (percentage of initial project purpose satisfied) and maintenance-work score (percentage of required maintenance needs carried out). Khwaja is currently working on a revised draft, which should be ready in early 2004.

The main focus of the study is on the effects of participation and collective action on infrastructure maintenance. This includes discussion of leadership, social capital, and comparative effectiveness of CDD (NGO project implementation) vs. non-CDD (local government project implementation).

***Pakistan Aga Khan Rural Support Project (AKRSP) – OED study [completed]***

The OED evaluation of AKRSP used individual and group discussions in 23 participant communities and one non-AKRSP village. As the AKRSP was active in nearly all communities in the region and had no baseline data, attributing causality was not possible for this study, so the research team aimed to ensure “consistency and coverage rather than to produce statistically significant analysis.” (World Bank 2002a, 2) They administered four semiformal questionnaires on community organizations, infrastructure, natural resource management and microfinance; fielded an institutional survey to AKRSP staff regarding their perceptions on the organization’s goals, competence and management practices; and estimated the economic rate of return and cost-effectiveness of investments. They used a case study approach to evaluate the overall project and the sub-program areas in terms of relevance, efficacy, efficiency, institutional development impact and sustainability. (World Bank 2002a)

***Pakistan National Rural Support Project (NRSP) – DEC study [ongoing]***

One component of the DEC CDD research initiative will examine the role of community organizations in maintaining community infrastructure in the context of the Pakistan National Rural Support Program.<sup>62</sup> The research team will examine whether community infrastructure created under this program is of better quality, more cost effective and more sustainable (in terms of design, construction, cost, maintenance) than similar (in size and type) infrastructure created by the government. The evaluation strategy is to compare similar projects provided by the NRSP and by the government within the same village. The sample is being drawn from 9 districts all over the country where the NRSP is currently active. Within these districts, 140 villages have some 450 comparable projects, all of which are included in the sample. As there are 6 main project types in the sample, the survey will include detailed on-site assessments of each sample project by a team of engineers. Extending earlier research, the evaluation will test the hypothesis that “building community organizations (including NGO facilitation of groups, etc.) can improve outcomes. The study will also look at the impact of community characteristics on projects outcomes and the extent to which community organizations mitigate the negative effects of community heterogeneity (caste, religion) or inequality on the maintenance of village infrastructure.”<sup>63</sup> Towards this end, field work will include a complete census of all villages both to determine how community characteristics affect project outcomes (including capacity to obtain CDD funds) and to compile a count of direct and indirect beneficiaries. Within-village comparisons of specific projects allow the study to control for both unobserved community characteristics, and project specific interactions between such characteristics and outcomes to allow for a cleaner comparison of outcomes which can be attributed to the implementing agency. In sum, this design will enable a “more rigorous test of the proposition that community assets created through CDD initiatives are better maintained” than those by more

---

<sup>62</sup> The current RSP stems from the Agha Khan Rural Support Program (AKRSP) mentioned earlier. Information was drawn from an interview with Ghazala Mansuri on 20 August 2003, as well as from two informal documents drafted by Mansuri: an early CDD research initiative proposal, “Research Program on Community Based Rural Development (CBRD),” and the response to subsequent comments, “Community Driven Development Component.”

<sup>63</sup> Mansuri, “Research,” 17. (See previous footnote.)

conventional means, particularly by a centralized bureaucracy.<sup>64</sup> This component of the study will be completed by mid-2004.

In addition, the program and research teams were able to agree on a process of random program placement and random assignment of intervention types to explore the efficacy of different NRSP mechanisms. This enables a comparison of outcomes across communities where a specific feature of the program is randomized in order to get a clearer sense of how components of the program impact outcomes of interest. This strategy will allow for an examination of a number of specific hypotheses. The first intervention will examine the impact of bundling microcredit with information on local productive opportunities and enterprise development skills. The intervention design includes 90 randomly selected villages: 30 will act as controls, 30 will get the basic program, which includes credit, and 30 will have credit complemented by entrepreneurship skills training and training in the identification of local production and marketing opportunities. The baseline for this intervention will be completed in March 2004 and the intervention will begin immediately after this. This sample of 90 villages will also enable research into the role of social mobilization in the take up of public funds, within a context of devolution. With decentralized public funds equally available to all communities, the evaluation will test the hypothesis that areas where NRSP is engaged in the process of social mobilization are better equipped to form an organization, obtain local government funds, and use the funds more effectively (in terms of quality of organization, profile of members, more equitable distribution of project benefits, etc.).

Lastly, NRSP management has been testing different models for deepening their social mobilization effort and improving their targeting within communities, a particular challenge in a context of rapid program expansion to new areas. Thus the study includes a randomized experiment that will allow an assessment of three different credit delivery mechanisms. This will be done in another 90 randomly selected villages—approximately one-half of which overlap with the sample for the CPI study.<sup>65</sup>

### ***Panama Rural Poverty and Natural Resources Project (RPNRP) [ongoing]***

The draft report of a cost-effectiveness study involving the Panama RPNRP and four other World Bank projects investigates the association between community participation and infrastructure quality, costs and maintenance. Drawing on data from a technical facility assessment (98 subprojects), cost data from implementing agencies (80 projects), and a multiple-choice social survey of community perceptions (1,435 interviews over 42 subprojects), it conducts chi-squared tests regarding several hypotheses. The study finds significant associations between above-average community participation (across all included projects, using various indices of participation) and: (a) above-average quality (measured by community perception of project quality, engineer's assessment of construction quality, services, and technical assessment of functionality); (b) below-average costs; and (c) above-average project maintenance. While it finds statistical associations between these variables, the report notes that chi-squared tests do "not quantify the effect exerted by the 'treatment' (participation, in the example) on the results," and thus recommends further analysis of the compiled database using other techniques, like regression analysis ("Panama," pp. 47, 59).

There is also a separate paper that explores the impacts of the Panama RPNRP on social capital, summarizing two earlier studies (Santucci 2003). Qualitative fieldwork using rapid participatory rural appraisal (RPRA) techniques was conducted in 25 program areas and 5 non-program areas. Twenty people were interviewed per site, including the president of the local project council (committee of sustainable development), six formal leaders, two individuals from the private sector, and 12 individuals

---

<sup>64</sup> Mansuri, "Community," 3. (See footnote above.)

<sup>65</sup> For further details and hypotheses regarding the included assessments of microcredit and social mobilization, see Annex 2, Section VI.B: Effectiveness of different delivery mechanisms within CDD.

“casually met” during field visits. Separate questionnaires were used for the different respondent groups, and included retrospective questioning to approximate the situation in 1998. Qualitative responses were then assigned quantitative codes and weights; total points were summed across questions; and a final score was calculated as a percentage of the maximum points possible. Index values were compared over time and used in regression analysis.

### ***Peru Social Fund (FONCODES) [completed]***

The main topics of investigation are education outcomes (namely attendance rates), poverty targeting, and comparative effectiveness of FONCODES relative to similar non-CDD programs. While the overall sample sizes are presented for the different data sets that are used, the exact sample sizes of treatment and comparison communities used in the analysis are not explicitly stated.

To estimate the short-run effects of FONCODES on school attendance, Paxson and Schady (2002) use both nonparametric and instrumental variable regressions with national survey data from 1993 and 1996. The sample size of 349 districts draws on data from the 1993 census, 1996 INEI (national statistics office) household survey (more than 18,000 households) and the FONCODES index (used to determine funding priorities, based on data from 1993 Population and Housing census). To verify their results and clarify the direction of causality, they used vote share for the party in power as an instrument for FONCODES participation (as greater funds were directed to districts where Fujimori was less popular, presumably to build political support) (Paxson & Schady 2002, 313).

For poverty targeting, they use data from the INEI and FONCODES index to analyze district-level poverty targeting by regressing FONCODES expenditures on district per capita income, and compare that with the same regression using expenditures under INFES (Instituto Nacional de Infraestructura Educativa y de Salud, a centralized government program). Household-level targeting is analyzed based on 1996 INEI household survey data using three separate techniques to triangulate results. The first runs logit regressions of the probability of having benefited from spending by a particular program (FONCODES, INFES, or parent committees) on the log of per capita household income. The second uses nonparametric regressions on the same variables, to allow for nonlinear relationships. The third also uses nonparametric regression, graphing the probability of benefiting from one of the three programs against the standard deviation of a household’s income above or below its district mean. (Paxson & Schady 2002, 304-309)

### ***Philippines Comprehensive and Integrated Delivery of Social Services Project (KALAHI-CIDSS) [ongoing]***

Two baseline surveys, of households and of local leaders, are currently being fielded in 60 participant and 60 comparison communities, matched through cluster analysis. The beneficiary communities will participate in Phase III of the program, which begins in January 2004. The overall sample size is 2400 households. Financed by the Trust Fund for Environmentally and Socially Sustainable Development, the baseline was designed and is being implemented by a team from the University of the Philippines. While funding and timing have not yet been finalized, the team will conduct mid-term and follow-up surveys to construct a panel data set. Responding to the KALAHI project team's operational needs, the baseline survey collects data about three categories of potential effects: (a) project outputs (access to facilities, particularly those most likely to be affected by KALAHI sub-projects, such as health, schooling, water and sanitation, roads); (b) poverty indicators (agricultural assets, housing, amenities, consumption and expenditure, subjective poverty measures); and (c) empowerment and governance (collective action, social cohesion and inclusion, trust and solidarity, groups and networks, voice, governance).<sup>66</sup>

---

<sup>66</sup> Camilla Holmemo emails, October 2003.

***Philippines Autonomous Republic of Muslim Mindanao (ARMM) Social Fund (ASFP)***  
*[planned]*

An evaluation of this social fund project is planned and will likely follow a similar research design to the Philippines KALAHI-CIDSS evaluation, with additional focus on conflict.

***Senegal National Rural Infrastructure Program (PNIR)*** *[ongoing]*

The ongoing impact evaluation of Senegal PNIR uses national household survey (ESAM) data on all 320 communautés rurales (CRs, the lowest level of elected government) and propensity score matching to identify 19 pairs of treated and matched non-participant CRs for analysis.<sup>67</sup> The approximate numbers for the sample size are 800 households overall; 400 controls that will not participate in PNIR during its four years of implementation; and 400 treatment, of which 280 will participate in PNIR in years 1 & 2, 80 in year 3, and 80 in year 4. The 2001 national household survey (ESAM2) will be used as the baseline, with abridged follow-up surveys every six months in the treatment and comparison areas only. PNIR impacts will then be estimated using difference-in-difference analysis on the constructed panel data set. The evaluation also includes technical audits of infrastructure: quality of physical structure and services, maintenance (technical norms by type of infrastructure, fraction of project budget devoted to maintenance), sustainability and costs. While this exercise is separate from the household surveys, it appears that it will be conducted for infrastructure in both PNIR and comparison villages.

To test political economy theories of poverty targeting and elite capture, household-level distribution of impacts within villages will be regressed against three indicators and controls to test if (a) “median voter,” (b) most disadvantaged, or (c) local elites best capture project benefits. The respective indicators would be a negative coefficient on (a) the squared difference between household *i*'s welfare and the welfare of the median household for that village, (b) the difference between household *i*'s welfare and the welfare of the poorest (or worst off) household, or (c) the difference between household *i*'s welfare and the welfare of the richest (or best off) household.<sup>68</sup>

***Senegal Agricultural Services and Producer Organizations Project (PSAOP)*** *[ongoing]*

See *Burkina Faso PNDSA 2 & Senegal PSAOP*, above.

***St. Lucia Poverty Reduction Fund (PRF)*** *[ongoing]*

Fieldwork will be conducted in 48 communities—36 intervention and 12 control—with a designed sample of 24 households per site for 1152 households overall. The control group was selected from among project pipeline communities, refined through nearest neighbor matching (based on project type, district, and poverty index) to get 36 potential control communities, from which the 12 final controls were randomly selected.<sup>69</sup> Survey instruments include a household survey, a project survey, and qualitative study using focus groups and in-depth interviews. Main topics for investigation include project impacts on access to infrastructure and social capital, cost efficiency, infrastructure sustainability, poverty targeting and alignment of projects with community priorities.<sup>70</sup>

***Yemen Social Fund for Development (SFD)*** *[ongoing, near completion]*

In a two-stage evaluation, the first stage used 1999 national survey data to compare household-level outcomes in participant (74 projects, 1352 households) and pipeline communities (109 projects, 1996 households), using multivariate regression to control for some systematic differences across groups. In the

---

<sup>67</sup> Adama Touré email attachment, 17 June 2003, “Impact.”

<sup>68</sup> Adama Touré email attachment, 17 June 2003, “Evaluation,” 11.

<sup>69</sup> Ian Walker email, 30 October 2003.

<sup>70</sup> Ian Walker email, 21 September 2003. Laura Rawlings email, 25 September 2003.

second stage, “variables were revisited in the 2003 impact evaluation study, which generated ‘ex-post’ data for the projects and communities that were in the pipeline dataset in 1999. This made possible a before-and-after comparison. There is no non-intervention control group in this study methodology. Multivariate analysis is used to control for sampling differences and trend changes in other independent variables (apart from SFD intervention) that might have explained changes in the impact variable in the period 1999-2003. For variables not covered in the 1999 baseline dataset, the 2003 survey includes retrospective questions to approximate baseline data.”<sup>71</sup> Qualitative work in 32 project areas (split between urban and rural areas) consisting of focus groups and in-depth interviews was used to understand processes and specific causes of success or failure.

Research topics include impacts on education, health, and infrastructure outcomes as well as targeting performance. The draft final report,<sup>72</sup> dated September 2003, contains more details on the methodology.

### ***Zambia Social Recovery Project / Social Investment Fund (ZAMSIF) [completed]***

Chase and Sherburne-Benz (2001) use ex-post data from the 1998 Zambia Living Conditions Monitoring Survey (similar to the LSMS, covering approximately 13,500 households) with oversampling in 99 ZAMSIF communities (an additional 2,950 households). They compare post-project outcomes in ZAMSIF communities with two separate comparison groups: pipeline and propensity score-matched communities. The analysis focuses on household-level impacts on education and health, and community-level participation and social capital outcomes. They also examine poverty targeting in three ways: (a) based on a number of different poverty measures, namely headcount, poverty gap, and poverty gap squared; (b) comparing adult equivalent per capita household expenditures; and (c) using concentration curves of ZAMSIF funding across income deciles.

Additionally, the evaluation also includes a 1998 facilities survey of 68 schools (43 ZAMSIF and 25 comparisons) and 30 health centers (16 ZAMSIF and 14 comparison), with data on utilization rates, physical resources, and staffing. (Chase and Sherburne-Benz 2000)

## **Very Early Planning Stages**

The following planned evaluations are still being designed, and may change considerably before implementation.

### ***Comoros Social Fund (FADC)***

The Comoros FADC is in the process of planning an impact evaluation. Details have not yet been determined.<sup>73</sup>

### ***DEC targeting study on social funds in Nicaragua and Ecuador***

As part of the DEC research initiative on CDD projects, one study plans to examine the effectiveness of social fund targeting in Nicaragua and Ecuador. Early planning documents include as research topics exploration of characteristics of communities that receive social fund investment (including the role of factors like extreme poverty, distance from the center, mayoral political affiliation, and NGO facilitators), and how the use of poverty maps could improve social fund targeting. In addition, the evaluation plans to

---

<sup>71</sup> Ian Walker email, 30 October 2003.

<sup>72</sup> Because it contains no comparison group, this study (ESA Consultores 2003d) falls outside the scope of this report, and has not been reviewed for this paper. Because findings differ considerably between the two Yemen evaluations (ESA Consultores 2002 & 2003d), they are not reported in Annex 2.

<sup>73</sup> Maryanne Sharp email, 30 September 2003.

examine whether projects within participating communities are targeted towards the poor, including issues of internal inequality and the potential association between high poverty and low capacity for decentralized decision-making. The description of the analysis calls for the use of existing datasets including national poverty maps, social fund MIS data and survey data developed in the evaluation of Nicaragua FISE.<sup>74</sup>

***Madagascar Social Fund 4 (FID4) [not yet contracted]***

While not yet contracted, a possible evaluation of FID 4 would include impacts on local capacity and decentralization as priority research areas.<sup>75</sup>

***Mongolia Sustainable Livelihoods Project [planned]***

This project is in the early stages of implementation, and is making arrangements for impact evaluation.<sup>76</sup> No further details are currently available.

***Nicaragua Third Emergency Social Investment Fund (FISE3) [not yet contracted]***

The call for expressions of interest for the FISE3 evaluation indicates that the evaluation aims to compare outcomes among FISE communities with different subproject implementation structures: centralized FISE implementation, decentralized municipal government management, and community-based management.<sup>77</sup> Topics for investigation include service utilization, beneficiary satisfaction, social capital, local capacity development, infrastructure sustainability, cost-efficiency and the contribution of FISE to broader sectoral goals. Further details will be determined in the terms of reference anticipated in 2004.

***Thailand Social Investment Fund***

This mixed-methods evaluation plans to use national household survey data (Socio-Economic Survey, 30,000 households) from 1998 and 2000, and subsequent qualitative investigation, to explore social fund impacts on community social capital. Specifically, the study will take advantage of over 200 villages within the SES samples that later participated in the social investment fund, and use propensity score matching to construct appropriate comparison groups to match participant villages. Qualitative research will then be conducted in both areas, using extended guided interviews.<sup>78</sup>

***Vietnam Northern Mountains Poverty Reduction Project***

This project is in the early stages of implementation, and is making arrangements for impact evaluation.<sup>79</sup> No further details are currently available.

---

<sup>74</sup> Berk Ozler email, 5 September 2003.

<sup>75</sup> Frank-Borge Wietzke email, 11 September 2003.

<sup>76</sup> Robin Mearns email, 13 August 2003.

<sup>77</sup> "Evaluación" 2003.

<sup>78</sup> Robert S. Chase email, 17 October 2003.

<sup>79</sup> Robin Mearns email, 13 August 2003.

## **Annex 2: Review of Findings**

Evaluations of CDD projects cover a broad range of topics. This annex organizes them into six themes: primary welfare, infrastructure, social dynamics, governance, targeting, and comparative effectiveness. Within each theme, several topics are included. Primary welfare includes income, education, and health impacts. Infrastructure covers issues of access, quality, maintenance and sustainability, complementary inputs, and utilization. Social dynamics include impacts on social capital, participation, gender, empowerment, conflict, and leadership. Governance topics include local governance, decentralization, corruption, and public sector reform. Targeting performance refers to both poverty- and preference-based targeting. Comparative effectiveness includes CDD versus non-CDD approaches, various delivery mechanisms within a CDD framework, participation effects, and cost-effectiveness or economic returns.

For each topic, we attempt to answer three questions: What has been done? What do we know about impacts? What will we know in the near future?<sup>80</sup> For the first question, we list all evaluations that address a given topic, beginning with the more rigorous research designs. We limit discussion of impacts to evaluations with robust results that generalize to the universe of the particular CDD program. Details provided for planned and ongoing evaluations reflect currently available information, however the evaluations may not be implemented as described.

Where there are several studies with results that generalize to the universe of a given project, a frequency table of reliable findings is included. The frequency tables summarize findings across a number of studies, presenting loose summations of positive, insignificant and negative findings. The tabulations are “loose” in the sense that specific indicators and methods are typically not consistent across evaluations. As such, the tables intend to provide a quick overview of the direction of findings; not a rigorous meta-analysis of the magnitude of impacts or specific causal relationships. The usual problems with publication bias certainly apply here.

More specifically, results for indicators are only reported in the frequency tables if they demonstrate significance at the ten percent level, at minimum. In calculating the frequency of findings, each evaluation is counted as a single instance. Therefore, if the evaluation used different methods and found different results, each finding for a specific indicator is represented as a fraction of the total findings. For example, if an evaluation used two separate comparison groups and found positive beneficiary impacts in relation to one comparator and no significant difference from the other, half a point each is recorded under positive and insignificant findings. Here, positive refers to the direction of beneficial impact, i.e. a positive finding for school drop-out rates would indicate a significant reduction. The tables record information from the most aggregated sample, while the paragraphs that describe the methodologies and findings delve into more specific subsets, for example differential impacts in urban versus rural areas. Publication bias (within studies) applies: tested variables not reported in the original research documents are clearly also omitted in the presentation of findings here.

### **I. PRIMARY WELFARE**

Overall, the most reliable and extensive group of evidence concerns CDD impacts on primary welfare measures, particularly education and health outcomes. Much of this is documented in the Rawlings et al. (2004) cross-country study of social funds, which we do not (re-)review in this report. For education, eight rigorous studies documented statistically significant impacts on various outcome measures that, in aggregate, present a positive view of CDD project impacts on education. The distribution of evaluation

---

<sup>80</sup> Evaluations covering several topics appear under each relevant section.

results across nine indicators totaled 10.0 positive, 10.5 insignificant and 0.5 negative findings (see Table 1). Health impacts have also been reasonably well documented, with six rigorous studies producing generalizable results. However, the overall impact is somewhat less clear because there is a large proportion of insignificant findings and several studies combine results across different infrastructure types (See Table 2). Since virtually all the generalizable evidence is from social funds, it may not necessarily apply to other interventions. In addition, there are no large scale studies that compare the performance or efficiency of CDD interventions in achieving these outcomes in relation to other types of interventions. Finally, while there is little evidence of CDD impacts on poverty as measured by household income or relevant proxies, several rigorous planned studies should expand knowledge in this area.

## **I.A. Household income and consumption/expenditure proxies**

Currently, there is little evidence documenting CDD impacts on household income. However, fourteen ongoing and planned evaluations will explore this topic using a variety of evaluation techniques.

### **What has been done? What do we know about impacts?**

One completed study has examined project impacts on income, consumption, expenditure, poverty headcount or savings.

#### ***Pakistan Aga Khan Rural Support Project (AKRSP) – OED study***

The OED study conducted ex-post interviews in 23 AKRSP and one non-AKRSP village (World Bank 2002a, 2). While the study reports that average farm incomes more than doubled between 1991 and 1997, it acknowledges that the lack of comparative or baseline data prevents any clear attribution to AKRSP interventions (World Bank 2002a, 7).

### **What will we know?**

Several ongoing and planned studies intend to address income impacts using a variety of methodologies. Nearly all of the studies here include clear plans to approximate the counterfactual state of participants. As such, strong research over the next few years should fill the current gap in knowledge about CDD poverty impacts.

#### ***India Andhra Pradesh District Poverty Initiatives Project (APDPIP)***

The female respondent household questionnaire contains questions about income shocks (crop failure, loss of livestock, family member illness or death, loss or damage to assets, and marriage expenses), coping mechanisms (mortgage or sell assets, use savings, withdraw children from school, reduce consumption, borrow, etc.) and food insecurity (number of days when two square meals unavailable over 12 months). It further includes detailed expense and consumption profiles for the previous 30 days. The male respondent household questionnaire includes questions about land ownership, use, related investments, crop and livestock production; housing characteristics and non-land assets; employment and wage profiles for each family member; and savings, access to and demand for credit. Two key hypotheses to test in this category involve income diversification and gender roles in intra-household expenditure allocation: a) “By allowing access to income generating options that were not previously open to the household, the project is expected to help reduce the reliance on wage labor and migration, implying an increase in income, a reduction in its variability, and the concentration of larger amounts of income in the hands of women,” and b) “If women and men have different preferences over types of household expenditures, greater economic power of the woman will result in more being spent on children’s

nutrition, education, and health, and less on liqueur and tobacco. This will reflect itself (over time) in improved nutrition and educational performance of children of program households.”<sup>81</sup>

#### ***Indonesia Urban Poverty Project (UPP2)***

Currently in the early stages of implementation, the UPP2 evaluation will be fielded in 157 participant Kelurahan (villages) and 98 comparisons. Baseline, mid-term and follow-up household surveys will include questions on economic status (assets and consumption), loans, savings and perceptions of poverty.<sup>82</sup>

#### ***Senegal National Rural Infrastructure Project (PNIR)***

Using difference-in-difference analysis of 19 treated and 19 matched non-participant localities, the impacts of Senegal PNIR infrastructure on well-being will use indicators from the national household survey relating to expenditures, including food expenditure or food consumption, as well as tobacco and alcohol expenditures (to measure bias against children in household consumption patterns).<sup>83</sup>

#### ***Burkina Faso National Agricultural Services Development Project (PNDSA2) & Senegal Agricultural Services and Producer Organizations Project (PSAOP)***

Sister evaluations of Rural Producer Organizations (RPO) in Burkina Faso and Senegal plan to use a difference-in-difference design based on baseline and follow-up surveys in participant communities and matched controls. The study plans to use income proxies, including the ability to eat more than one meal per day and consumption of durable goods, to examine the impact of membership in RPOs on quality of life.

#### ***Nepal Rural Water Supply and Sanitation (RWSS)***

Using a mixed methods approach, the research design for the RWSS evaluation plans to compare relevant outcomes across participant, pipeline and non-participant communities for eventual difference-in-difference estimates. Although specific indicators are still being developed, one of the hypotheses that the evaluation seeks to explore is the positive impact of RWSS participation on family income, which in turn leads to greater empowerment of women and Dalits.<sup>84</sup>

#### ***Benin Social Fund (AGeFIB)***

Using data on a representative sample of 6,000 households, ex-post propensity score matching will be used to examine sub-project effects. The survey includes an expenditure module to measure impacts on standards of living. A draft (green cover) Poverty Assessment report contains estimates of household consumption and poverty.<sup>85</sup>

#### ***Brazil Rural Poverty Reduction Projects (PCPR)***

The research design calls for panel data fielded two years apart in PCPR communities and those with no project. The household survey will collect data on consumption proxies, including type of food consumed, durable goods, and non-durable goods, as well as data on physical capital and savings. (FECAMP 2001, 84-85)

#### ***Ethiopia Women’s Development Initiatives Project (WDI)***

The evaluation of the Ethiopia WDI is planning to use random assignment of participation among interested potential beneficiaries to isolate program impacts on empowerment and economic outcomes. Early planning documents mentioned financial capital and economic resources as elements of individual agency. (Christiaensen & Legovini 2003)

---

<sup>81</sup> Klaus Deininger email, 10 October 2003.

<sup>82</sup> “UPP-2 Evaluation Household Questionnaires.” (2003).

<sup>83</sup> Adama Touré email attachments, 17 June 2003.

<sup>84</sup> Lynn Bennett interview, 17 July 2003.

<sup>85</sup> World Bank 2003b.

### ***Philippines Comprehensive & Integrated Delivery of Social Services Project (KALAHI-CIDSS)***

The household survey includes poverty indicators regarding housing characteristics, economic activity, assets, three-day diet recall, one-month expenditure recall and self-rated poverty status. The baseline round is currently being fielded in participant and comparison communities, matched using cluster analysis, to be followed-up with mid-term and end-of-project rounds for a panel data set. (APPC 2003a, 12-18)

### ***Indonesia Kecamatan Development Project (KDPI, KDP2, KDP3)***

For the preliminary study, the perception survey of villagers and community leaders contained six questions on impacts of infrastructure projects. These explored the impact on household income, access to markets, school participation (for females as well as males) and access to health facilities. The proportions of positive responses among KDP participants exceeded those of non-KDP for all questions, however there were substantial problems in several districts. (UOI, 2002, p. xiii and p. 35) The response of community leaders was also more favorable than household respondents. However, both KDP household respondents as well as community leaders felt the impact of the project on the income of the poor was significant. (UOI, 2002, p. 24) Given that these findings are based only upon a one-year treatment period, they should be considered preliminary only. These topics will be further explored in upcoming phases of impact research.

### ***Cambodia Rural Investment and Local Governance Project (RILG)***

Fielded in participant and nearby nonparticipant villages, the planned quantitative socioeconomic baseline study includes questions on household expenditure, savings and asset profile. The baseline will potentially be followed up with subsequent fieldwork to establish difference-in-difference comparisons.

### ***Indonesia Support for Conflict-Ridden Areas Project (SCRAP)***

The draft design for the SCRAP household survey includes questions on household characteristics and consumption modules.

### ***Malawi Social Action Fund (MASAF)***

The second phase of the MASAF poverty and sustainability review will use focus group discussions and structured interviews, complemented by an analysis of existing datasets, to determine MASAF impacts on poverty (Centre et al. 2002).

## **I.B. Education outcomes**

Eight rigorous completed and ongoing studies document CDD project impacts on measures of educational outcomes. As a group, they document an equal number of positive and insignificant effects, with nearly no negative impacts. The regional distribution of these studies is five in Latin America, three in Eastern Europe and one in Africa, however planned evaluations should contribute to available information on CDD impacts in East Asia and Africa. In general, CDD interventions seem to positively influence education outcomes.

### **What has been done?**

Seven completed evaluations used sufficiently rigorous methods to produce reliable results, summarized below. It is important to note that six of these studies were of social funds. While these six provide a fairly clear idea of education impacts from social funds, these findings may not necessarily generalize to other CDD mechanisms.

### ***Bolivia Social Investment Fund (SIF)***

To assess the impact of the Bolivia social fund on education outcomes, researchers used a combination of random assignment of eligibility and propensity score matching with baseline and post-intervention data (Newman et al. 2002). Authors find that while school infrastructure and complementary inputs significantly improved, they demonstrated “little effect on enrollment, attendance, or academic achievement” (p. 248). Regarding indicators of student-level outcomes, results from randomized eligibility show that only a reduction in the drop-out rate was statistically significant. No significant impacts were found for repetition rate, enrollment ratio or attendance. For the difference-in-difference estimates with the matched comparison group, there appears to have been a significant decrease in the drop out rate and increases in school enrollment and attendance for beneficiary schools. Again, no impact found on the repetition rate.

### ***El Salvador Community-Managed Schools Program (EDUCO)***

Jimenez and Sawada (1999) used the percentage of EDUCO schools in each municipality, as well as the EDUCO targeting formula as instrumental variables to control for selection bias into the program and school characteristics. They estimated the effect of EDUCO on third grade achievement test scores in mathematics and language, and on days of school missed due to a teacher absence. They found that EDUCO participants showed no significant difference in achievement test scores, compared to similar traditional school students. However, teacher attendance in EDUCO schools was better than in traditional schools, once differences in school characteristics and selection bias into the EDUCO program were controlled for.

### ***Peru Social Fund (FONCODES)***

The study used both nonparametric and instrumental variable regressions on panel data to estimate the short-run effects of the FONCODES on school attendance. They found that worse-off districts demonstrated large gains in school attendance while better-off did not, and that among poorer districts, those that experienced larger gains tended to receive more FONCODES funding. To test the robustness of these positive results, and to confirm the direction of causality, they then used districts’ percentage of votes for Fujimori, the incumbent president, as an instrumental variable for FONCODES participation, as greater funds were directed to districts with low Fujimori support, presumably to build popularity. Assuming that “political preferences of districts were not correlated with changes in preferences for education,” this confirmed the finding that districts with higher FONCODES spending experienced greater gains in attendance rates. (Paxson & Schady 2002, 311-313)

### ***Armenia Social Investment Fund (ASIF1)***

Chase (2002) found mixed evidence regarding educational impacts. Using data from a national household survey, oversampled in ASIF areas, the analysis examined education indicators for households in ASIF project areas relative to those in both propensity score matched communities and pipeline communities. In earthquake zones, households in ASIF areas spent more on education and sent a larger proportion of their children to school than those in the pipeline comparison group. Yet in conflict zones, households in beneficiary communities spent significantly less on education and demonstrated no significant difference in enrollment compared to both pipeline and matched groups. The mixed results suggest different social fund impacts in the two areas (Chase 2002, 230).

### ***Nicaragua Emergency Social Investment Fund (FISE1)***

Using two comparison groups, matched by facility characteristics and by propensity scores, Pradhan and Rawlings (2002) found positive and significant effects of FISE projects on primary school enrollment, education gap, and age starting first grade. Furthermore, the enrollment effect was stronger for girls than

for boys, and impacts on education gap and age-for-grade more pronounced for poorer quintiles. However, they found no significant impacts on age-for-grade overall, nor on attendance. (Pradhan & Rawlings 2002, 288-289)

***Zambia Social Recovery Project / Social Investment Fund (ZAMSIF)***

Chase and Sherburne-Benz (2001) compared household-level education outcomes in communities with a ZAMSIF school to those in pipeline and propensity score-matched communities, and found the following statistically significant results. ZAMSIF areas had higher attendance rates than both comparison groups in urban areas, but no significant difference in rural areas. Compared to the pipeline communities, ZAMSIF areas showed higher attendance in age-appropriate grade, in both urban and rural areas, and a higher share of household expenditures on education in urban areas. Compared to propensity score-matched communities, the difference age-for-grade attendance was insignificant, and ZAMSIF areas showed a greater household expenditure share on education in rural areas. (Chase & Sherburne-Benz 2001, Table 3)

***Honduras Social Investment Fund 2 (FHIS2)***

The research team used multivariate regression to compare household-level education outcomes in FHIS communities to those in pipeline communities. While no significant impacts were found on school enrollment rates, positive and significant impacts were discovered for age-for-grade performance (Walker et al. 1999, 58).

***Albania Development Fund (ADF)***

Using propensity score matching on a single ex-post cross section, Carletto et al. (2003) have so far found no significant effects of ADF in terms of education spending or average years of schooling. They are exploring potential heterogeneous effects along the propensity score distribution, using nonparametric regression.

***India Dairy: Operation Flood***

This OED evaluation examines the outcomes of India’s dairy market policy, of which Operation Flood was a relatively small part. The small sample size and difficulty isolating impacts of the program prevent direct attribution of impacts. Anecdotal evidence from the selected cases suggested that milk income from the dairy cooperatives helped families afford children’s educational expenses (Candler & Kumar 1998, 56).

**What do we know about impacts?**

Seven studies above followed sufficiently rigorous methodologies to generate education results that are representative to the scale of the overall project:

- Bolivia SIF
- Nicaragua FISE
- Honduras FHIS
- Albania ADF
- Zambia ZamSIF
- Armenia ASIF- Peru FONCODES
- El Salvador EDUCO

The table below shows that on balance, the various studies of CDD interventions tend to find many more positive, or at least neutral impacts, than negative ones.

**Table 1: Impacts on Education**

Education	Positive findings	Insignificant	Negative findings
-----------	-------------------	---------------	-------------------

Enrollment	Bolivia, Nicaragua	Bolivia, Honduras	
Drop out rate	Bolivia		
Age-for-grade	Honduras, Zambia	Nicaragua, Zambia	
Age in first grade	Nicaragua		
Education gap	Nicaragua		
Attendance	Bolivia, Armenia, Peru, Zambia, El Salvador	Bolivia, Armenia, Zambia, Nicaragua	Nicaragua
Repetition rate		Bolivia	
Education spending	Zambia	Zambia, Armenia, Albania	
Test scores		El Salvador	
Years of schooling		Albania	
<b>Total Indicators</b>	<b>10</b>	<b>10.5</b>	<b>0.5</b>

### What will we know?

In addition to the numerous completed studies, at least six planned studies will explore CDD program impacts on education measures. All of these intend to compare education outcomes against an estimated counterfactual.

#### ***Philippines Comprehensive & Integrated Delivery of Social Services Project (KALAHI-CIDSS)***

The household questionnaire for the evaluation includes questions on highest education level attained, attendance and grade. (APPC 2003a, 4) The baseline round is currently being fielded in participant and comparison communities, matched using cluster analysis.

#### ***India Andhra Pradesh District Poverty Initiatives Project (APDPIP)***

The female respondent household questionnaire contains questions about school enrollment, fees, school meals, drop out, and enrollment at bridge school.<sup>86</sup>

#### ***Benin Social Fund (AGeFIB)***

The study includes a separate school module focused on quality of infrastructure that will also record enrollment details. (Coulombe 2002, 4)

#### ***Brazil Rural Poverty Reduction Projects (PCPR)***

The household survey is to collect data on illiteracy and years of schooling, and the evaluation also plans to use achievement test scores for children. PCPR results will be compared against those from areas with no projects, using panel data for difference-in-difference estimation. (FECAMP 2001, 89)

#### ***Indonesia Support for Conflict-Ridden Areas Project (SCRAP)***

The draft design of the SCRAP household survey indicates that education outcomes will be included.

#### ***Moldova Social Investment Fund (MSIF)***

A baseline survey and impact analysis for MSIF has recently been contracted. The terms of reference call for retrospective questioning regarding participant status to explore impacts of MSIF1 projects, while establishing baseline data on MSIF2 participants. The study will include education variables on attendance, enrollment and school results.

<sup>86</sup> Klaus Deininger email, 10 October 2003.

## **I.C. Health outcomes**

Overall, the more rigorous evaluations point towards positive impacts of CDD projects on health outcomes. Particularly strong positive evidence comes from the Bolivia SIF assessment of impacts on child mortality rates. The evaluation of Zambia documented positive impacts on child vaccination rates. The least clear evidence regards impacts on the frequency of seeking medical treatment, which is surprising given that it should be one of the most immediate impacts of health interventions.

### **What has been done?**

Five completed and ongoing evaluations have explored CDD impacts on health outcomes, and their results are discussed below.

#### ***Bolivia Social Investment Fund (SIF)***

For the health portion of the Bolivian SIF evaluation, Newman et al. (2002) used difference-in-difference estimation involving participant and propensity score-matched comparison health centers. Their analysis found significant increases for the beneficiary group in the proportion of women receiving prenatal care, yet significant decreases in cases of cough treated. However, they found no significant impacts on use of public health facilities, fraction of attended births, fraction of cases of diarrhea that were treated, incidence of diarrhea nor incidence of cough. To estimate impacts on under-age-five mortality, they used three separate methods: propensity score matching with recall data, life table estimates for the change in mortality with panel data, and estimation of a Cox proportional hazard function. Strikingly, all three methods found a significant reduction in child mortality in the participant groups and no change for the comparison groups. After checking for other relevant interventions, the authors conclude that the findings strongly suggest positive impacts of SIF health centers on under five mortality rates.

#### ***Armenia Social Investment Fund (ASIF1)***

With few relevant indicators in the dataset, the Armenia ASIF impact evaluation examined the impact of water projects on health outcomes using ex post matching techniques to both pipeline and propensity score matched comparisons. Chase (2002) found limited significant evidence to suggest that a lower proportion of treatment households reported illness and inactivity due to illness over the prior twelve month period. No significant impacts on proportion of households reporting ill children were found. (Chase 2002, 233)

#### ***Zambia Social Recovery Project / Social Investment Fund (ZAMSIF)***

Using household survey data, Chase and Sherburne-Benz (2001) found the following statistically significant differences in health outcomes between ZAMSIF areas and pipeline and propensity score-matched comparison areas. ZAMSIF households had significantly higher self-reported rates of sickness, lower likelihood of seeking treatment when sick, and were less likely to go a hospital, but equally likely to go to a health center. No significant differences were found in incidence of diarrhea or share of household expenditures on health. (Chase & Sherburne-Benz 2001, Table 4) ZAMSIF areas also had significantly higher child vaccination rates for DPT, but roughly equal rates of child vaccinations against BCG, polio, or measles (Chase & Sherburne-Benz 2001, Table 5).

#### ***Honduras Social Investment Fund (FHIS2)***

The study found significant impacts of health centers on increasing the probability that individuals sought professional attention for a health problem. Limited data prevented an estimation of impacts on vaccination rates, and no significant results were found for the impact of water projects or sewerage systems on the incidence of diarrhea (Walker et al. 1999, 61-62).

### ***Nicaragua Emergency Social Investment Fund (FISE1)***

The ex post evaluation of FISE1 found inconclusive results regarding health impacts. Regarding contact rates for households and children with diarrhea specifically, and indicators of acute malnutrition (wasting, prevalence of underweight), Pradhan and Rawlings (2002) find significant impacts when comparing the treatment group to one of two comparison groups (propensity score matched or matched by facility characteristics), but no findings are supported by both comparators. They find insignificant differences for other indicators measured, including those regarding incidence of cough, prenatal care, assisted births, DPT and polio vaccination rates, and stunting. Furthermore, they find no significant impacts from water, latrine or sewerage investments on the incidence of diarrhea, wasting, stunting or underweight prevalence. (Pradhan & Rawlings 2002, 288-292)

### **What do we know about impacts?**

The following table loosely summarizes the findings from the five reliable studies regarding health impacts listed above. In addition to a variety of variables and methods across evaluations, the tabulations below include health impacts arising from different types of infrastructure: health posts, latrines and water systems.

**Table 2: Impacts on Health**

<b>Health indicators</b>	<b>Positive findings</b>	<b>Insignificant</b>	<b>Negative findings</b>
Prenatal care	Bolivia	Nicaragua	
Assisted births (skilled staff)		Nicaragua, Bolivia	
Institutional births	Nicaragua	Nicaragua	
Child BCG vaccination	Zambia	Zambia	
DPT vaccination	Zambia	Nicaragua	
Polio vaccination		Nicaragua, Zambia	
Measles vaccination		Zambia	
Complete vaccine schedule (BCG, polio, measles)			
Complete vaccine schedule on time			
Treatment all hhs		Zambia, Bolivia	
Sought treatment when ill (contact rate)	Nicaragua Honduras	Nicaragua Bolivia	Zambia
Contact rate under age 6	Nicaragua	Nicaragua	
Contact rate over age 5		Nicaragua	Nicaragua
If child diarrhea, sought treatment	Nicaragua	Nicaragua	
Fraction diarrhea cases treated		Bolivia	
Fraction cough cases treated			Bolivia
If treated, went to hospital (+=less hospital)	Zambia		
If treated, went to health center (+ more health center)	Zambia		
Went to hospital all HHS (= less hospital)	Zambia		
Went to health center all HHS (+ more health center)		Zambia	
Incidence of diarrhea		Zambia, Nicaragua (health, water, sewerage, latrine), Bolivia	

Diarrhea if sick (+=less diarrhea)	Zambia		
Incidence of cough		Nicaragua, Bolivia	
Child mortality	Bolivia		
HHS member reported sick (+ more sickness reported) <sup>87</sup>	Zambia		Armenia
Inactivity due to illness (less inactivity)	Armenia		
Ill children		Armenia	
Diarrhea if sick (+ less diarrhea)	Zambia		
Health expenditures		Zambia	
Health share of total expenditures		Zambia	
Wasting prevalence	Nicaragua (health facility)	Nicaragua (health propensity, water, sewerage, latrine)	
Stunting prevalence		Nicaragua (health, water, sewerage, latrine)	
Underweight prevalence	Nicaragua (health facility)	Nicaragua (health propensity, water, sewerage, latrine)	
<b>Total Indicators</b>	<b>15.5</b>	<b>37.0</b>	<b>4.0</b>

Similar to the spread of educational impacts, there is a larger number of positive and insignificant findings compared to negative ones. This suggests that in general, evaluations of CDD interventions tend to find somewhat positive impacts. However, the large portion of insignificant results merits further investigation.

## What will we know?

In addition to the completed evaluations above, eight planned and ongoing studies intend to explore CDD project impacts on health outcomes.

### ***Brazil Rural Poverty Reduction Project (PCPR)***

The household survey will collect anthropometric data, namely body-mass index for women of child-bearing age and of children under age five, as indicators of nutritional status. The household survey will also ask how many family members (a) have physical difficulties due to health problems, (b) were sick over the past twelve months, (c) received medical treatment over the past twelve months, (d) missed work for health reasons over the past two weeks, and (e) wanted but could not get treatment over the past two weeks. Panel data will be collected in PCPR areas, for comparison against areas with no projects. (FECAMP 2001, 86-88)

### ***Nepal Rural Water Supply and Sanitation Project (RWSS)***

While specific variables remain to be decided, one hypothesis to test involves improved knowledge about health as an intermediate step in increasing empowerment.<sup>88</sup>

### ***Philippines Comprehensive & Integrated Delivery of Social Services Project (KALAHI-CIDSS)***

The household questionnaire for the baseline study includes health indicators of occurrence and type of illness during the previous six months, and visits to a health practitioner. (APPC 2003a, 3)

<sup>87</sup> Although Armenia and Zambia found opposite results, both interpreted the results in a positive light (one as raised awareness, versus the other as lower incidence).

<sup>88</sup> Lynn Bennett interview, 17 July 2003.

The baseline round is currently being fielded in participant and comparison communities, matched using cluster analysis.

#### ***India Andhra Pradesh District Poverty Initiatives Project (APDPIP)***

The female respondent household questionnaire contains questions on health status, treatment, expenditures, prenatal care, child immunizations, information about health issues, family planning, and contraception.<sup>89</sup>

#### ***Moldova Social Investment Fund (MSIF)***

The terms of reference for the Moldova SIF evaluation mention health related impacts of gas supply systems (reduction of chill cases and chronic diseases) and of water systems (reduction in cases of infection).<sup>90</sup>

#### ***Benin Social Fund (AGeFIB)***

The household survey for the Benin AGeFIB impact evaluation includes health measures of prenatal care, incidence of illness, and usage of medical services.

#### ***Senegal National Rural Infrastructure Project (PNIR)***

Household surveys will collect data on health indicators of days of work or school missed because of illness, and child anthropometric measures of height, weight, and body mass index.<sup>91</sup> The surveys will be administered several times in 19 pairs of PNIR and matched comparison communities to construct a panel data set.

#### ***Indonesia Support for Conflict-Ridden Areas Project (SCRAP)***

Early planning documents note that health outcomes will be measured in the SCRAP household survey, including child vaccination rates.

## **II. INFRASTRUCTURE**

This section focuses on CDD impacts on (A) access to basic socioeconomic infrastructure, (B) quality of infrastructure, (C) quality of service delivery, (D) sustainability and maintenance of facilities and services, (E) the provision of complementary inputs necessary for infrastructure to function, and (F) utilization rates.

### **II.A. Access to Infrastructure**

While nearly all Bank projects track the progress of infrastructure implementation, few verify impacts on access by attempting to estimate the counterfactual state of beneficiaries, taking into account other providers.<sup>92</sup> Three completed studies collected generalizable evidence on access to basic infrastructure, and they present a favorable picture of CDD impacts. Table 3 (on page 52) shows 9.0 positive, 8.5 insignificant and 0.5 negative findings across sixteen indicators. Ten ongoing studies should generate further reliable evidence, as most of them are planning to collect baseline and follow-up data in participant and matched comparison areas. Incorporating data from relevant line ministries into project MIS databases could facilitate periodic evaluation of changes in access, and further improve targeting.

---

<sup>89</sup> Klaus Deininger email, 10 October 2003.

<sup>90</sup> Anush Bezhanyan email attachment, 9 July 2003, "TOR" p. 4.

<sup>91</sup> Adama Touré email attachment, 17 June 2003, "Evaluation," 7.

<sup>92</sup> While it may seem impossible for a CDD infrastructure project to *decrease* access, it is important to note that many other agencies actively construct schools, health centers and water/sewerage systems. If other agencies are more effective than a particular CDD intervention and avoid communities that have a CDD project, one could potentially find negative CDD impacts on access, as communities would have been better off without the CDD program as they would have received facilities from another, more effective, provider.

## **What has been done?**

### ***Armenia Social Investment Fund (ASIF1)***

Using both propensity-matched and pipeline comparison groups, the impact evaluation explores impacts of potable water projects on household access to water. Chase (2002) finds positive, significant impacts on the proportion of households with central water systems, cold running water and indoor water tap;<sup>93</sup> and mixed impacts on proportion with a flush toilet. Unlike the educational impacts discussed earlier, the impacts of water projects are more consistent across conflict areas and earthquake zones. (Chase 2002, 232-234)

### ***Albania Development Fund (ADF)***

Using propensity score matching, the ongoing Albania ADF analysis has thus far found no significant impacts from road projects on distance to primary school, ambulatory or doctor, and bus station; nor from water systems on the probability of having a toilet inside dwelling, running water inside dwelling or having water continuously (Carletto et al., tables 11-2). Regarding distance to schools, only two of seven estimates showed significant positive impacts; and for distance to bus, one of seven showed a negative impact with significance. Further research currently focuses on heterogeneous effects along the propensity score distribution.

### ***Nicaragua Emergency Social Investment Fund (FISE1)***

The evaluation found positive, significant FISE1 impacts on access to water and sanitation infrastructure, using recall information from 1993 (before FISE investments) collected in both treatment and propensity score matched communities. Pradhan and Rawlings (2002, 290-293) find greater increases in the share of households with piped water and flush toilets, and greater decreases in the distance to the nearest water source, in FISE1 areas versus matched comparison areas.

### **Bolivia Social Investment Fund (SIF)**

To measure impacts of water supply projects, the evaluation of Bolivia SIF used reflexive comparison in beneficiary areas (Newman et al., 2002). Baseline and follow-up surveys were administered for all 18 water projects constructed in 1992, on both the household and community level, although no facility survey was fielded. Authors reported that the main change was a reduction in the distance to the nearest water source.

## **What do we know about impacts?**

Three evaluations included a convincing approximation of the counterfactual state of participants to estimate the impact of CDD projects on access to infrastructure:

- Albania ADF                      - Nicaragua FISE                      - Armenia ASIF

Although based on a limited number of evaluations, the table below presents a favorable picture of CDD impacts on access. Together, the three studies generated nearly equal numbers of positive and insignificant findings, and nearly no findings of negative impacts on access.

---

<sup>93</sup> For water tap, one comparison group demonstrated significant positive impacts, the other no impacts (Chase 2002, 232).

**Table 3: Impacts on Access to Infrastructure tend to be positive.**

	Positive findings	Insignificant	Negative findings
Distance to primary school		Albania (roads)	
Distance to doctor/health center		Albania (roads)	
Distance to bus station		Albania (roads)	
Distance to water source ex post		Nicaragua (water)	
Change in distance to water source (pre-post)		Nicaragua (water)	
Probability of/share of HHS with toilet inside dwelling (ex post)	Nicaragua (sewerage)	Albania (water)	
Change in share of HHS with toilet	Nicaragua (sewerage)		
Share of HHS with no toilet (ex post)	Nicaragua (latrine)		
Change in share HHS no toilet	Nicaragua (latrine)		
Probability of running water inside dwelling, share piped water	Nicaragua (water)	Albania (water)	
Change in share of piped water	Nicaragua (water)		
Probability of having water continuously		Albania (water)	
Proportion of HHS with indoor water tap	Armenia (water)	Armenia (water)	
Proportion of HHS with central water system	Armenia (water)		
Proportion of HHS with flush toilet	Armenia (water)		Armenia (water)
Proportion of HHS with cold running water	Armenia (water)		
<b>Total Indicators</b>	<b>9.0</b>	<b>8.5</b>	<b>0.5</b>

## What will we know?

### *Indonesia Urban Poverty Project (UPP2)*

The Indonesia UPP2 evaluation design indicates that surveys will be fielded in 157 participant and 98 matched villages, to collect baseline, midterm and follow-up data. The local leader questionnaire contains questions on access (distance and time) to nearest market, school, hospital, bank and other financial institutions; and the household survey examines access to water, sewerage, and electricity infrastructure.<sup>94</sup>

### *Brazil Rural Poverty Reduction Projects (PCPR)*

The household survey asks questions related to access to health facilities, particularly in terms of reasons for inability to get treatment, and about access to education, in terms of distance and time to facilities. Panel data from PCPR areas will be compared with areas with no projects. (FECAMP 2001, 88-89)

### *Nepal Rural Water Supply & Sanitation Project (RWSS)*

Using quantitative and qualitative methods in participant, pipeline and non-beneficiary comparison communities, the Nepal RWSS project plans to explore the connection between project impacts on increasing equitable access to water and improving health knowledge; and subsequent poverty reduction and greater empowerment.<sup>95</sup>

<sup>94</sup> “UPP-2 Evaluation: Lurah and PKK (Family Welfare Movement) Chairperson Questionnaires” (2003). “UPP-2 Evaluation: Household Questionnaires” (2003).

<sup>95</sup> Lynn Bennett interview, 17 July 2003.

### ***Indonesia Kecamatan Development Project (KDP1, KDP2, KDP3)***

The preliminary impact study for KDP1 included perceptions on access to services in surveys of households and community leaders in KDP and non-KDP matched areas. Both household respondents and community leaders indicated higher satisfaction regarding access to health facilities in KDP than non-KDP subprojects, by small but significant margins of about 3-5 points. Regarding access to markets, household respondents were slightly happier with this result under KDP subprojects than non-KDP, but community leaders were equally satisfied with KDP and alternatives regarding this outcome. (UOI 2002, 22) These topics will be revisited in subsequent phases of impact research.

### ***Senegal National Rural Infrastructure Project (PNIR)***

Establishing difference-in-difference estimates every six months in 19 participant and 19 matched areas, the ongoing Senegal PNIR evaluation considers indicators of access to basic social and economic services, measured by distance (including to water source, school, health center, market and road).<sup>96</sup> It further includes a focus on the access of women and other vulnerable groups to services.<sup>97</sup>

### ***Philippines Comprehensive & Integrated Delivery of Social Services Project (KALAHI-CIDSS)***

The household and community leader surveys include sections regarding access to infrastructure. Specifically, these include questions of time to health facility, school, post office, and telephone; road networks; difficulties fetching water; and retrospective questions about being better or worse off as compared to one year previous regarding access to water, electricity, health facility, school, work, market, source of financing and technology. (APPC 2003a & 2003b) The baseline round is being fielded in October 2003 in participant and comparison areas matched by cluster analysis.

### ***Cambodia Rural Investment & Local Governance Project (RILG)***

To eventually assess the cost-benefit ratio and economic rates of return on Cambodia RILG infrastructure, the commissioned baseline study includes a section on savings in rural people's time as part of the household level benefits generated by the project. The terms of reference mention developing proxies to value "changes in access to important facilities and services" based on frequency of access as well as cash and time required per contact (Helmert & Wallgren 2003, 10).

### ***Malawi Social Action Fund (MASAF)***

The initial desk review of the MASAF poverty alleviation and sustainability review includes a plan to explore existing datasets to determine MASAF impacts on poverty reduction in the second stage of the evaluation. The report mentions access to potable water, education and health services as aspects of poverty to examine using maps developed by the ministries of education and health (Centre et al. 2002, 56). The second analysis stage is primarily using a case study approach.

### ***St. Lucia Poverty Reduction Fund (PRF)***

The ex post evaluation of the St. Lucia social fund will attempt to approximate the pre-intervention state of beneficiaries in relation to access and service quality with existing secondary sources and recall questions, and compare those values with ex post data from project and household surveys, in PRF and control communities.<sup>98</sup> Exact methodology remains to be determined.

### ***India Andhra Pradesh District Poverty Initiatives Project (APDPIP)***

The main village questionnaire a series of questions regarding access to basic infrastructure.

---

<sup>96</sup> Adama Touré email attachment, 17 June 2003, "Evaluation," 2.

<sup>97</sup> Adama Touré email attachment, 17 June 2003, "Impact," 3.

<sup>98</sup> Ian Walker email, 30 October 2003.

## II.B. Quality of Infrastructure

Among the evaluations reviewed for this study, only the Bolivia SIF evaluation produced generalizable results on CDD impacts on quality of infrastructure. Although only for one region, it found quite positive results. A few rigorous upcoming studies intend to explore these effects and should enhance knowledge in this area, including examination of the relative performance of projects implemented by CDD programs in comparison to that by other providers. In general, while technical evaluations of CDD project infrastructure by engineers are not uncommon, they rarely make a systematic attempt to assess facilities in appropriate comparison areas or those implemented by other agencies. Much could be gained by shifting the coverage of existing technical assessments to include both intervention and comparison areas.

### What has been done?

#### *Bolivia Social Investment Fund (SIF)*

The randomized eligibility section of the Bolivia SIF analysis found positive, significant impacts on the quality of school infrastructure in terms of classrooms being in good condition, however no significant impacts on students per classroom (Newman et al. 2002, 249). Results from the difference-in-difference estimate confirmed the increase in classroom condition, and found further impacts on reducing the number of students per classroom and increasing the space per student within classrooms.

#### *Honduras Social Investment Fund 2 (FHIS2)*

Regarding quality of construction, the Honduras FHIS2 evaluation used an engineer to assess the quality of infrastructure provided by FHIS and other providers on a three value scale, “good, regular, bad.” The report indicates the following case study-based insights. While school projects exhibited a similar pattern of quality across providers, FHIS health centers received much more favorable ratings. Key problems identified for FHIS schools stemmed from inflexibility in the design that reduced community involvement and led to poor adaptation to local context. For water and sewerage projects, while FHIS received generally positive ratings for original construction, they were not as high as for other providers. FHIS latrines received much higher positive marks than other providers. Lastly, data from the household survey revealed a much higher percentage of “good” quality assessments by beneficiaries in completed project as compared to pipeline project areas. Amongst FHIS beneficiaries only, water projects received relatively low ratings and higher frequency of problems. (Walker et al. 1999, 37-39).

#### *Panama Rural Poverty & National Resources Project (RPNRP)*

Drawing on data from a technical facility assessment and a multiple choice social survey of community perceptions, the Panama cost-effectiveness study uses a chi-squared statistic to test the significance of the association between above-average community participation (across all included projects, using various indices of participation) and above-average quality (measured by community perception of project quality, engineer’s assessment of construction quality, services, and technical assessment of functionality). The report finds a positive association between participation and infrastructure quality. (“Panama,” 58)

### What do we know about impacts?

Only the Bolivia SIF evaluation provides generalizable results. The study found predominantly positive impacts of interventions improving the quality of (mainly school) infrastructure.

**Table 4: Impacts on Infrastructure Quality**

Quality of Infrastructure	Positive findings	Insignificant	Negative findings
Classrooms in good condition	Bolivia (random)		
Fraction of classrooms in good	Bolivia (random)		

condition	Bolivia (match)		
Students per classroom	Bolivia (match)	Bolivia (random)	
Square meters per student	Bolivia (match)		
Number of patient rooms	Bolivia (match)		
<b>Total</b>	<b>4.5</b>	<b>0.5</b>	<b>0</b>

## What will we know?

### ***Pakistan National Rural Support Program (NRSP) – DEC study***

Part of the DEC study of the NRSP will include 140 villages with some 450 infrastructure schemes to explore whether community infrastructure created under this program is of better quality than similar (in size and type) infrastructure created by the government. The evaluation will test the hypothesis that “building community organizations (including NGO facilitation of groups etc) can improve outcomes.”<sup>99</sup> Fieldwork will include a complete census of all villages and technical engineer assessments of six types of infrastructure, implemented by the NRSP and by the local government.

### ***Senegal National Rural Infrastructure Project (PNIR)***

The Senegal PNIR evaluation includes an examination of physical characteristics of infrastructure (including quality of facility and services), maintenance (technical norms by type of infrastructure, fraction of project budget devoted to maintenance), sustainability and costs. While this exercise is separate from the household surveys, it appears that it will be conducted for infrastructure in PNIR and comparison villages.<sup>100</sup>

### ***Benin Social Fund (AGeFIB)***

A separate module on schools will be included in the Benin AGeFIB fieldwork, based on the survey developed for the Zambia SIF 2000 impact assessment. To be fielded in both participant and control schools, it includes questions on physical characteristics of school facilities, level of equipment, enrollment, quality of teachers, etc.

### ***Cambodia Rural Investment & Local Governance Project (RILG)***

The terms of reference for the Cambodia RILG evaluation mentions baseline information on quality of infrastructure, but it is not clear how details will be used in future.

### ***India Andhra Pradesh District Poverty Initiatives Project (APDPIP)***

The main village questionnaire contains questions regarding quality of school infrastructure.

### ***Indonesia Support for Conflict-Ridden Areas Project (SCRAP)***

The SCRAP household survey will include questions on service quality, utilization and beneficiary satisfaction.

### ***Malawi Social Action Fund (MASAF)***

The second stage of the poverty and sustainability review plans to assess the performance of MASAF “in comparison with other school construction initiatives,” potentially using boreholes, as well as in terms of staffing trends, supervisory capacity, benefits and costs, and infrastructure quality (Centre et al. 2002, 55). It is unclear what exact methodology will be used and whether this component will be included in the final review.

<sup>99</sup> Mansuri, “Research,” 17. Ghazala Mansuri email communication 11/12/03.

<sup>100</sup> Adama Touré email attachment, 17 June 2003, “Impact,” 7.

## II.C. Quality of Service Delivery

Regarding CDD effects on the quality of service delivery, only two studies reported reliable evidence, and those for only a few indicators. While several rigorous upcoming studies indicate an intention to explore this area, few details are available on how impacts on service quality will be measured. As noted in the previous section, while technical evaluations of CDD project infrastructure by engineers are fairly common, they rarely make a systematic attempt to assess facilities in appropriate comparison areas or those implemented by other agencies. As a result, much could be gained by simply shifting the coverage of existing technical audits to include both intervention and comparison areas.

### What has been done?

#### *Armenia Social Investment Fund (ASIF1)*

Armenia ASIF found contradictory results regarding household reporting of school service improvement, with significant impacts of opposite sign depending on whether the propensity score-matched or pipeline group was used as comparator. Additionally, there was a significant increase in households that reported water service improvements, but no significant impact on household perception of sanitation improvements. (Chase 2002, 231-233)

#### *Bolivia Social Investment Fund (SIF)*

The random eligibility section of the Bolivia SIF evaluation found no significant impacts on the proportion of teachers with a professional degree (Newman et al., 249). Separately, water quality tests were conducted for both the old water sources and new SIF-constructed ones. Little change was found between the quality of water from old and new SIF sources, which was attributed to a number of problems that had since been solved. Similar tests for subsequent projects yielded more favorable results.

### What do we know about impacts?

Little evidence is currently available regarding CDD impacts on quality of service delivery. Only two studies generated reliable evidence, the Bolivia SIF and Armenia ASIF evaluations, and they reported impacts for only a few indicators.

**Table 5: Impacts on Quality of Service Delivery**

Services	Positive findings	Insignificant	Negative findings
HHS report school service improved in prior 12 months	Armenia		Armenia
HHS report water service improvement	Armenia		
HHS report sanitation improvement		Armenia	
Fraction of teachers with professional degree		Bolivia (random)	
<b>Total</b>	<b>1.5</b>	<b>2</b>	<b>0.5</b>

### What will we know?

While a few upcoming studies intend to use rigorous methods, planning documents provide only cursory details regarding how impacts on quality of service delivery will be measured. These include:

#### *Benin Social Fund (AGeFIB)*

The school module includes questions on the number of trained versus untrained teachers, and years of teaching experience (Coulombe 2002, 27).

### ***Burkina Faso National Agricultural Services Development Project (PNDSA2) & Senegal Agricultural Services & Producer Organizations Project (PSAOP)***

The joint study of the RPOs intends to explore the impacts of the project on producer organizations in terms of quality of service to members, effectiveness of representation of members' interests, and delivery of public goods to members. This approach is slightly different from the other evaluations mentioned, as it explores project impacts on the quality of service producer organizations provide to their members, not the quality of service the project delivers to participating POs.

### ***Senegal National Rural Infrastructure Project (PNIR)***

The research plan for the Senegal PNIR study mentions that quality of services will be included in the infrastructure audits, however provides no further details.<sup>101</sup>

## **II.D. Maintenance & Sustainability**

Several interesting studies have provided insights on maintenance and sustainability outcomes using more of a case-based approach. While the results so far do not necessarily generalize to the overall projects in question, several larger-scale studies plan to explore the maintenance performance of CDD projects, including in relation to other types of interventions.

### **What has been done?**

#### ***Pakistan Aga Khan Rural Support Project (AKRSP) – Khwaja study***

Using a dataset that includes 33 sites where two infrastructure projects were implemented within a single community, Khwaja used ordinary least squares (OLS) and instrumental variables (IV) regressions to isolate project-specific factors that affect maintenance performance from differences among communities (Khwaja 2002, 11). Overall, he finds that project specific factors have a greater impact on successful maintenance than community characteristics, concluding that “(good) projects can succeed (even) in bad communities” (Khwaja 2002, i). Specifically, he finds the following statistically significant relationships: regarding community characteristics, land inequality has a U-shaped effect on maintenance, social heterogeneity has a negative impact, and the presence of a project leader has a positive effect. For project-specific factors, project complexity has a negative effect, observed inequality in distribution of project returns has a U-shaped effect, and continuation projects are better maintained than new projects (Khwaja 2002, 17-23). Furthermore, the analysis indicates that community participation in non-technical decisions positively impacts maintenance, while participation in technical decisions has a negative effect, controlling for overall project complexity (Khwaja 2002, 22).

#### ***Nicaragua Emergency Social Investment Fund (FISE1)***

While not generalizable to the universe of Nicaragua FISE projects, case studies within the ex post evaluation compared 20 FISE to 20 non-FISE health centers and 24 FISE to 24 non-FISE schools (both run by the Ministry of Education); and technically evaluated 10 FISE latrines and 10 FISE water systems. The case studies explored utilization, operational sustainability and physical sustainability of infrastructure investments. Regarding operational sustainability, researchers found a greater (and increasing) number of staff, and larger endowments of equipment and medical materials in FISE clinics, although comparable problems with inadequate medicinal supplies for both. In comparison to non-FISE schools, the FISE schools demonstrated a larger increase in the number of teachers and staff; better access to piped water, electricity and bathrooms; an increase in the number of classrooms; more parents reporting improvements in the condition of their school; yet no difference in the supply of textbooks.

---

<sup>101</sup> Adama Touré email attachment, 17 June 2003, “Impact,” 7.

FISE sewerage systems reportedly received regular maintenance work, and system administrators associated them with improvements in roads, presence of insects and incidence of sickness. For physical sustainability, they found comparably low levels of electricity and piped water in both, yet FISE clinics had more frequent maintenance work though less sufficient maintenance funds, and greater community contributions. For the ten FISE water systems studied, a divide became apparent between successful subterranean water systems in wealthier areas and unsuccessful gravity-based surface systems in poorer areas. (World Bank 2000)

### ***Honduras Social Investment Fund 2 (FHIS2)***

The engineer case study assessments in the Honduras FHIS2 evaluation found favorable maintenance performance of FHIS water systems, sewerage projects and latrines compared to other providers. However, a high proportion of FHIS beneficiaries in the household survey and focus groups felt supervision systems were weak, and attributed many existing problems to insufficient supervision (Walker et al. 1999, 38-40).

### ***Panama Rural Poverty & National Resources Project (RPNRP)***

Using chi-squared tests on variables across several projects in Panama, the cost-effectiveness study found a positive association between above-average community participation and above-average infrastructure maintenance (“Panama,” 50).

### ***Pakistan Aga Khan Rural Support Project (AKRSP) – OED study***

The OED evaluation of AKRSP discusses institutional and financial sustainability. Anecdotal evidence indicates that many of the village organizations had “matured” and survived “independently of their progenitor;” several other programs have replicated components of the AKRSP model in other contexts; however, the broader institutional setting of the region depends heavily on the AKRSP and thus if it were to withdraw abruptly it would leave a “large institutional gap.” The organization is not financially self-sustaining and a task force that examined sustainability prior to the OED report noted that the AKRSP may be facing “donor fatigue” and a “very tight” financial future if changes were not enacted to “strengthen financial management and increase cost recovery.” (World Bank 2002a, 14-17) Sustainability of the program was assessed as “highly likely,” however there was not sufficient budget to survey maintenance conditions for many project sites (World Bank 2002a, 111).

## **What do we know about impacts?**

None of the completed studies produced broadly applicable findings. While the Khwaja (2002) study is the most rigorous, the sample that supports many of the most interesting results covers only a small number of sites. The collection of case studies in the several studies above generates useful insights, it does not provide an empirical basis of results to generalize beyond the particular contexts.

## **What will we know?**

### ***Pakistan National Rural Support Program (NRSP) – DEC study***

Part of the DEC study of the NRSP will include 140 villages with some 450 infrastructure schemes to explore whether community infrastructure created under this program is more sustainable (design, construction, cost, maintenance) than similar (in size and type) infrastructure created by the government. The evaluation will test the hypothesis that “building community organizations (including NGO facilitation of groups etc) can improve outcomes. The study will also look at the impact of community characteristics on projects outcomes and the extent to which

community organizations mitigate the negative effects of community heterogeneity (caste, religion) or inequality on the maintenance of village infrastructure.”<sup>102</sup> Fieldwork will include a complete census of all villages and technical engineer assessments of six types of infrastructure, implemented by the NRSP and by the local government. This study design will enable a “more rigorous test of the proposition that community assets created through CDD initiatives are better maintained” than those by more conventional means, particularly by a centralized bureaucracy.<sup>103</sup>

#### ***Indonesia Kecamatan Development Project (KDPI, KDP2, KDP3)***

In the first phase of research, perception surveys fielded in participant and matched communities indicated that household members and community leaders felt the community’s ability to maintain the project in future was higher in KDP subprojects than in non-KDP subprojects (UOI 2002, 22). Maintenance issues will be further studied in upcoming phases of the impact evaluation.

#### ***Senegal National Rural Infrastructure Project (PNIR)***

The PNIR evaluation includes a technical examination of infrastructure maintenance, assessing adherence to technical norms by type of infrastructure, fraction of project budget devoted to maintenance, and unspecified sustainability details.<sup>104</sup>

#### ***Moldova Social Investment Fund (MSIF)***

The evaluation will investigate MSIF’s “impact on institutional capacity and whether the facilities in which SIF funds were invested are properly maintained.”<sup>105</sup>

#### ***St. Lucia Poverty Reduction Fund (PRF)***

The evaluation plans include questions on sustainability, maintenance, and financing arrangements from PRF infrastructure. Data on these topics will be collected from a project survey, household survey and qualitative work, however details on the methodology remain to be decided.<sup>106</sup>

#### ***Malawi Social Action Fund (MASAF)***

The first desk study phase of the MASAF poverty alleviation and sustainability review notes a lack of comparative data regarding maintenance arrangements for other interventions, and encountered confusion regarding maintenance responsibilities during limited field visits to MASAF sites (Centre et al. 2002, 24). It is unclear whether and how these issues will be covered in subsequent case study research.

## **II.E. Complementary Inputs**

As a contributing factor toward project maintenance, continuous provision of complementary inputs helps ensure that infrastructure investments effectively provide services. Only one completed study provides reliable evidence on the provision of recurrent inputs for CDD sub-projects. Research plans for two upcoming studies include specific indicators in this area. Further research that highlights particular arrangements within CDD programs that effectively prevent supply shortages over time would be informative.

---

<sup>102</sup> Mansuri, “Research,” 17. Ghazala Mansuri email communication 11/12/03.

<sup>103</sup> Ghazala Mansuri email attachment, 20 August 2003, “Community,” 3.

<sup>104</sup> Adama Touré email attachment, 17 June 2003, “Impact,” 7.

<sup>105</sup> Anush Bezhanyan email, 6 November 2003.

<sup>106</sup> ESA Consultores 2003c, 8.

## What has been done?

### *Bolivia Social Investment Fund (SIF)*

Evaluation of Bolivia SIF schools based on random assignment of eligibility found positive, significant impacts on the provision of blackboards, desks, teachers' tables and sanitation facilities (Newman et al. 2002, 249). However, no significant impacts were associated with the fraction of schools with electricity or textbooks. The difference-in-difference estimates confirmed the positive findings regarding sanitation facilities and lack of impact on electricity; found no evidence for the impact of desks per student; and found further positive impacts on supply of textbooks and a reduction in the number of students per teacher. For health centers, the difference-in-difference analysis found positive, significant impacts on number of beds, fraction of clinics with sanitation facilities and availability of medical supplies; while no impact on fraction of clinics with electricity or on an index measure of availability of medical equipment in good condition.

### *Honduras Social Investment Fund 2 (FHIS2)*

Matched infrastructure case studies examined issues of service quality, utilization and sustainability. This included some discussion of complementary inputs, where health post staffing was found to be generally deficient. For schools, operating patterns were found to be generally similar across the two groups, including for textbook provision, however FHIS schools had higher availability of desks. The level of dissatisfaction was found to be higher for the FHIS group, including complaints about lack of materials and personnel (Walker et al. 1999, 47-50).

### *Nicaragua Emergency Social Investment Fund (FISE1)*

The Nicaragua FISE comparative infrastructure case studies contain details on complementary inputs. These are discussed in the previous section with regards to maintenance and operational sustainability.

## What do we know about impacts?

The only generalizable evidence on CDD project impacts on the provision of complementary inputs comes from the Bolivia SIF evaluation, which reports an equal number of positive and insignificant findings.

**Table 6: Impacts on the Provision of Complementary Inputs**

<b>Complementary Inputs</b>	Positive findings	Insignificant	Negative findings
Blackboards	Bolivia (random)		
Blackboards per classroom	Bolivia (random)		
Desks		Bolivia (random)	
Desks per student	Bolivia (random)	Bolivia (diff)	
Teachers' tables		Bolivia (random)	
Teachers' tables per classroom	Bolivia (random)		
Fraction of schools with sanitation facilities	Bolivia (random) Bolivia (diff)		
Fraction of schools with electricity		Bolivia (random)	

		Bolivia (diff)	
Textbooks		Bolivia (random)	
Textbooks per student	Bolivia (diff)	Bolivia (random)	
Students per teacher (+ fewer)	Bolivia (diff)		
Beds per health clinic	Bolivia (diff)		
Fraction of clinics with electricity		Bolivia (diff)	
Fraction of clinics with sanitation facilities	Bolivia (diff)		
Fraction of clinics with water		Bolivia (diff)	
Availability of medical equipment in good condition		Bolivia (diff)	
Availability of medical supplies	Bolivia (diff)		
<b>Total Indicators</b>	<b>8</b>	<b>8</b>	<b>0</b>

### What will we know?

Only two upcoming evaluations include a study of complementary inputs in the research design. However, it is possible that other planned infrastructure studies will include recurrent input indicators without stating so explicitly in the research design.

#### ***Benin Social Fund (AGeFIB)***

The supplementary school module contains questions on complementary inputs like latrines, water, electricity, telephone, and desks.

#### ***India Andhra Pradesh District Poverty Initiatives Project (APDPIP)***

Regarding health centers, the main village questionnaire contains questions of number of doctors and availability of medicines.

## II.F. Utilization

While utilization rates overlap with the analysis of Primary Welfare impacts, especially school enrollment and health clinic contact rates discussed above, they merit separate treatment given the CDD emphasis on aligning projects with demand for services. Here, the evidence from five evaluations regarding education is on balance positive, while that for health is more mixed.

### What has been done?

#### ***Nicaragua Emergency Social Investment Fund (FISE1)***

The case studies included subproject impacts on utilization rates. For the 20 pairs of health posts, utilization was higher for FISE than non-FISE posts and increasing over time, although rates for both remained low. In the 24 pairs of primary schools, enrollment had increased in FISE schools while declined in non-FISE schools. The study of ten FISE sewerage systems found low connection rates, while rates were high overall for FISE water systems. (World Bank 2000)

#### ***Honduras Social Investment Fund 2 (FHIS2)***

Regarding the connection between community participation and utilization rates, the Honduras FHIS2 evaluation compared the percentage of households that use the infrastructure among those that were consulted on the project versus those that were not. The report found a higher proportion of users amongst households that were consulted for all types of infrastructure. A separate analysis using the FHIS and Honduran Health Ministry (MSP) databases, compared the average number of “attentions” (consultations, including vaccinations) for FHIS posts to the overall average for all posts in the MSP database (however only 60% of FHIS clinics could be identified in the dataset). In this comparison, FHIS urban posts and

restored rural posts were found to be above average, though new rural posts were below average (Walker et al. 1999, 27, 48).

### **What do we know about impacts?**

Four studies discussed in Sections I.B. & I.C. documented generalizable evidence on the utilization of school and health projects:

-Bolivia SIF                      - Nicaragua FISE                      -Honduras FHIS                      - Zambia ZAMSIF

Overall, the four studies find positive impacts on education utilization measures and more mixed evidence for health centers. Specifically, as discussed earlier under impacts on education, of the four studies with clear evidence on enrollment rates, evaluations of Bolivia SIF (matching), and Nicaragua FISE found positive significant impacts while Bolivia (random assignment) and Honduras FHIS found no significant impacts. No studies included here found negative impacts.

Regarding health centers, the health impact frequency table (Table 2) recorded ten utilization indicators in the form of contact rates, or treatment sought.<sup>107</sup> Within those, the distribution of findings regarding utilization shows 5.5 positive, 5 insignificant and 3.5 negative impacts from evaluations of Zambia ZAMSIF, Nicaragua FISE, Honduras FHIS, and Bolivia SIF.

### **What will we know?**

Three upcoming studies plan to explore utilization issues:

#### ***Indonesia Urban Poverty Project (UPP2)***

The household survey for the Indonesia UPP2 evaluation contains questions regarding use of health and social safety net facilities.<sup>108</sup>

#### ***Philippines Comprehensive & Integrated Delivery of Social Services Project (KALAHI-CIDSS)***

The household survey for the KALAHI-CIDSS contains questions on visits to health practitioner by type of facility, school attendance and frequency of travel. (APPC 2003a, 3-6) The baseline round is being fielded in October 2003 in participant and comparison areas matched by cluster analysis.

#### ***Indonesia Support for Conflict-Ridden Areas Project (SCRAP)***

The SCRAP household survey will include questions on service quality, utilization and beneficiary satisfaction.

## **III. SOCIAL DYNAMICS / SOCIAL CAPITAL**

This section presents findings and highlights upcoming studies in five categories: (A) Social Capital, Norms, Trust, Collective Action, Initiative and Resource Mobilization; (B) Participation, Voice, Inclusion, Representation, Ownership; (C) Gender; (D) Empowerment; (E) Conflict and Crime; and (F) Leadership.

---

<sup>107</sup> Those health utilization indicators are treatment all households, treatment when ill, contact rate under 6, contact rate over 5, child diarrhea sought treatment, fraction diarrhea cases treated, fraction cough cases treated, if treated went to hospital, if treated went to clinic, went to hospital all households, and went to health center all households.

<sup>108</sup> "UPP-2 Evaluation: Household Questionnaires" (2003).

While many CDD impact evaluations examine issues of social dynamics and social capital, few do so rigorously. Only three completed evaluations—of social funds in Armenia and Zambia, and of the AKRSP in Pakistan—have documented social dynamics impacts at a scale representative of the entire project. They do not constitute, by themselves, a sufficient body of evidence to make any broad claims about the interactions of CDD interventions and social dynamics. In fact, the two social fund evaluations seem to have found somewhat contradictory results: one finds evidence that CDD projects build up social capital, while the other suggests that CDD projects may drain existing social capital. The study from Pakistan found that community participation in non-technical subproject decisions enhanced maintenance outcomes, but that community participation in technical decisions was detrimental (Khwaja 2002).

However, several upcoming evaluations will carefully examine social capital and related questions in the context of CDD projects. For those with information available, most designs include rigorous methods to estimate the counterfactual, and many are of a sufficiently large scale to generate results representative of the overall project. These should add considerable knowledge to the literature of CDD and social capital in the coming years.

### **III.A. Social Capital, Collective Action, Trust**

Evaluations looking at structural and/or cognitive social capital have tended to focus more on collective action, initiative and resource mobilization than on norms, values and feelings of trust and social cohesion. Five such studies have been completed, and two have generalizable findings. One, on Zambia ZamSIF, indicates that a CDD approach augments cognitive and structural social capital. A similar study on Armenia SIF, finds that CDD projects draw on existing stocks of social capital, perhaps crowding out other community initiatives.

#### **What has been done?**

Five completed evaluations have addressed issues related to social capital. Two studies (Armenia SIF 1; Zambia ZamSIF) have findings which are reliably representative of the project as a whole.

##### ***Armenia Social Investment Fund (ASIF1)***

Using a nationally representative survey, with both propensity score-matched and pipeline comparison groups, Chase (2002) found that communities who got ASIF1 subprojects were initially better endowed with social capital, but that they expended it on the ASIF1 subprojects. This is indicated by a generally lower probability of undertaking other similar initiatives in communities that had completed ASIF1 subprojects, but a generally higher probability of independent community initiatives in ASIF1 pipeline communities (p. 237).

##### ***Zambia Social Recovery Project / Social Investment Fund (ZAMSIF)***

Chase & Sherburne-Benz (2001) constructed two separate comparison groups, propensity score-matched and pipeline matched, from a nationally representative household survey. Econometric analysis showed that participating in ZAMSIF impacted structural and cognitive indicators of social capital differently in rural and urban areas. In rural areas, community members were more likely to participate in ZAMSIF subprojects than in similar interventions in comparison communities. This led to positive effects on subjective feelings of community cohesion, as well as increasing the likelihood of the community undertaking other local development initiatives. In urban areas, ZAMSIF subprojects actually had lower community participation than similar subprojects in comparison communities, and were less likely than matched urban communities to undertake other local development initiatives.

### ***Pakistan Aga Khan Rural Support Project (AKRSP) – Khwaja study***

Khwaja (2002) investigates the effects of social capital and collective action on project maintenance in his study of rural support projects in northeast Pakistan (cf sections III.B and III.F on participation and leadership in this annex). Primarily based on a subsample of 33 AKRSP communities that implemented two subprojects, he finds that project-specific factors tend to have economically and statistically larger effects than community-specific factors. He then suggests that social capital factors are better viewed as part of the enabling environment than an outcome to be targeted directly by interventions, and that elements of project design are the policy tools that can respond to these constraints (p. 26). As such, “(good) projects can succeed (even) in bad communities” (p. i).

As noted in a footnote, he also found some statistically significant associations between project maintenance and various other indicators of community characteristics/social capital that fell beyond the main focus of the study. High community-reported unity, no report of land disputes or of problems in raising cash for collective work, and higher membership in community organizations were all associated with higher project maintenance. However, other indicators showed no significant correlation with maintenance outcomes: trust, problems in raising labor contributions, or number of community organizations. (p. 20)

### ***Jamaica Social Investment Fund (JSIF)***

The Jamaica study was not a formal impact evaluation, but rather an assessment of how CDD projects and social capital interact, based on five pairs of matched case studies of social fund communities, and matched comparison communities. The evidence did suggest that leadership, drawn from elites, was crucial in mobilizing the community to access social fund resources. Many of these case studies highlight how the process of mobilizing the community around a CDD project can enhance social cohesion, but also show how these may not be sustainable. (Rao & Ibáñez 2003)

### ***Panama Rural Poverty & National Resources Project (RPNRP)***

Based on case studies of 25 RPNRP areas and 5 non-program areas, the study notes positive effects on enhancing social capital, including improving the role of women and youth, stronger relationships between community-level and formal local institutions / ministries, and beneficiaries’ opinion about community self-management (Santucci 2003).

## **What do we know about impacts?**

The Armenia and Zambia studies find mixed results in terms of how a CDD project enhances or detracts from social capital. The other completed studies are either based on samples too small to draw statistical inference, or have potential sources of bias that cloud the results.

## **What will we know?**

Many ongoing and planned evaluations are examining questions related to social capital. As a group, they represent a step forward in terms of using sound methodologies to evaluate social capital impacts of CDD projects.

Ten studies have rigorous evaluation designs of a scale that could generate reliable results, representative of the overall program. Their evaluation designs all include comparison groups of communities to isolate the effects of the program being evaluated.

### ***Pakistan National Rural Support Program (NRSP) – DEC study***

Within the DEC study of the NRSP, a sample of 90 villages will enable research into the role of social mobilization in the take up of public funds, within a context of devolution. The Pakistani government has decentralized, making public funds equally available to all communities. The evaluation will test the hypothesis that areas where NRSP is engaged in the process of social mobilization are better equipped to form an organization, obtain local government funds, and use the funds more effectively (in terms of quality of organization, profile of members, more equitable distribution of project benefits, etc.). Such findings would contribute to a clearer understanding of how important the social mobilization aspect of the program is in empowering poor communities and facilitating collective action.

In addition, NRSP management has been testing different models for deepening their social mobilization effort and improving their targeting within communities. This is a particular challenge in a context where the program is rapidly expanding to new areas. Credit is a key program component and the program has thus far used community members to identify potential credit recipients and to collect loans. This has worked well in some cases but has often led to poor targeting, excessive reliance on community activists who have used the credit program to expand their influence locally and in some cases has even led to substantial default. Due to this social organizers have often had to short change social mobilization efforts to focus on credit delivery and recovery. NRSP is thus attempting to redesign its microcredit delivery system by separating credit operations from their social mobilization work, and creating village credit branches with fulltime field workers who will take over many of the day to day credit related tasks and report to a credit officer rather than the social organizer. This provides an opportunity to examine whether an increase in social mobilization efforts lead to higher levels of collective action capacity in the target community. The study examines this issue with a randomized experiment that will allow an assessment of three different credit delivery mechanisms. This will be done in another 90 randomly selected villages—approximately one-half of which overlap with the sample for the CPI (Community Physical Infrastructure) study above.

### ***Indonesia Urban Poverty Project (UPP2)***

Household surveys, village head surveys and case studies are looking at social capital, including social networks, community organization, and social cohesion. Early planning documents identify impacts on collective action as a central research topic. The quantitative part of this evaluation is using a natural experiment panel data design based on cut-off criteria for UPP eligibility. Qualitative case studies are selected from within the quantitative sample to capture maximum variety of village characteristics in each area.

### ***Burkina Faso National Agricultural Services Development Project (PNDSA2) & Senegal Agricultural Services & Producer Organizations Project (PSAOP)***

These evaluations are investigating the nature of existing rural producer organizations and their interaction with development efforts in rural areas. (Collion 2002, 3-4; de Janvry & Sadoulet 2002, 8-12) The research design includes in-depth case studies, and baseline and follow-up surveys in a total sample of 300 program and comparison villages for panel data. The surveys will assess the capacity of rural producer organizations (RPOs) to mobilize public or private services.

### ***Brazil Rural Poverty Reduction Projects (PCPR)***

Using a household survey, the evaluation will look at PCPR's effects on social capital, and gather information on perceptions of civic engagement and patterns of associations and relations with government. This will be complemented by case studies in 50 PCPR municipalities, sampled to cover all three states in the study, and to include 10-15 with pre-existing municipal councils from the first project phase, 20-25 that will form municipal councils during the current phase, and 10-15 that will not have municipal councils during the current phase. These case studies will involve

180 focus groups, interviews with local politicians, and participant observations of PCPR implementation. The social capital analysis is strongly linked to analysis of its interaction effects vis-à-vis local governance. (FECAMP 2001, 147-186)

***Philippines Comprehensive & Integrated Delivery of Social Services Project (KALAHI-CIDSS)***

The household survey covers the following topics: household participation in community collective action, groups and networks; social cohesion, trust and solidarity within the local community; transparency and accountability of local institutions; participation in community decision-making and activities. (APPC 2003a, 18-23) The community leader questionnaire contains a section on people's organizations. (APPC 2003b, 4) The research team began fielding the baseline survey in October 2003 in 60 pairs of treatment and matched comparison communities.

***Albania Development Fund (ADF)***

Using the 2002 LSMS household survey, this ongoing study examines the extent to which ADF communities independently engage in local development initiatives, as compared to matched non-ADF communities. Thus far, the analysis has found no significant impacts on measures of whether the community has an organization to discuss important issues, how often community members meet for community activities, how often the community makes common proposals for assistance from politicians or institutions, and whether there is always someone in the community willing to help. The analysis will look for heterogeneous effects along the propensity score distribution.

***Indonesia Kecamatan Development Project (KDP1, KDP2, KDP3)***

In terms of resource mobilization, the preliminary KDP1 evaluation found that "villagers contribute an average of 17% of project costs despite the relative poverty of the villages and the fact that there is no mandatory contribution amount" (GOI 2002, 10). This was not compared against non-KDP subprojects. However, surveys of community members and leaders indicated a higher probability of participation in KDP subprojects than in non-KDP ones (UOI 2002, 21). The surveys also indicated that community members and leaders were more aware of the role of community organizations in the KDP subprojects than in non-KDP ones (UOI 2002, 23). These topics will be revisited in upcoming phases of the impact evaluation.

***India Andhra Pradesh District Poverty Initiatives Project (APDPIP)***

The female respondent household questionnaire contains questions on degree of trust and solidarity with people within and outside the same village, by caste and religious groups, and with officials, police, and project staff.

***Indonesia Support for Conflict-Ridden Areas Project (SCRAP)***

The draft design for the SCRAP household survey includes social capital issues of group dynamics, trust, solidarity, collective action, social cohesion and inclusion.

Two further studies use rigorous approaches, albeit on a smaller scale, that include comparison groups to estimate the counterfactual state of beneficiaries.

***Indonesia KDP & Conflict***

This evaluation is looking at the influence of participating in KDP on community organization for problem solving and conflict resolution, both through changing institutional structures and through cognitive changes in norms and trust (Barron et al. 2003a).

***Armenia Social Investment Fund (ASIF2)***

Comparing 6 ASIF and 8 control communities, this evaluation will examine ASIF's effect on building social capital, focusing on (a) social norms, (b) social networks and partnerships, and (c) institutional capacity.

Two studies are still in relatively early planning stages, with few confirmed details on methodology.

#### ***Moldova Social Investment Fund (MSIF)***

The terms of reference state that this evaluation will look at issues of social capital, participation and empowerment. It calls for a mix of qualitative and quantitative methods with baseline and follow-up surveys, and stipulates that an independent consulting firm is to be contracted to carry out the evaluation, but that the “dissemination of results is the prerogative of MSIF team.”<sup>109</sup>

#### ***St. Lucia Poverty Reduction Fund (PRF)***

A small comparison group of non-PRF communities will be studied to explore issues of PRF impacts on social capital. A quantitative household survey and qualitative fieldwork will be used.<sup>110</sup>

### **III.B. Participation, Voice, Inclusion, Representation, Ownership**

Many studies have looked at participation and related issues. Only two—El Salvador EDUCO and Pakistan AKRSP—have generated reliable results of how participation affects other outcomes. In Pakistan, Khwaja (2002) found that community participation in non-technical subproject decisions enhanced positive outcomes, but that community participation in technical decisions was detrimental. In Ecuador, Jimenez and Sawada (1999) found that more visits by parents associations to schools was correlated with higher test scores. The two completed large-scale studies focusing on the level of participation in social fund projects, in Zambia and Nicaragua, found no evidence that the social funds achieved better participation than similar non-CDD projects.

Many ongoing and planned studies also cover these issues, most of them with rigorous evaluation designs that estimate the counterfactual state of beneficiaries.

#### **What has been done?**

Eight completed evaluations have addressed participation, voice, inclusion, representation, or ownership. We also mention a paper (Isham et al. 1995) that lies outside the scope of this study, but uses interesting methods and finds positive results regarding higher participation and project effectiveness.

#### ***El Salvador Community-Managed Schools Program (EDUCO)***

Jimenez and Sawada (1999) found that more participation by parents in schools, as measured by the number of visits by a parent association to classrooms, was significantly and positively correlated with higher test scores in math and language, but had no significant effect on teacher attendance. However, “the number of days students miss is negatively related to intangible EDUCO effects” (Jimenez & Sawada 1999, 440), suggesting that community-based school management does improve student, and perhaps teacher, attendance compared to traditional, more centralized schools.

#### ***Pakistan Aga Khan Rural Support Project (AKRSP) – Khwaja study***

Primarily based on a subsample of 33 AKRSP communities that implemented two subprojects, Khwaja (2002) found that community participation in non-technical decisions had positive effects on project outcomes, but that participation in technically complex decisions negatively affected outcomes.

---

<sup>109</sup> Anush Bezhanyan email attachment, 9 July 2003, “TOR” p. 7.

<sup>110</sup> Ian Walker email, 30 October 2003.

### ***Zambia Social Recovery Project / Social Investment Fund (ZAMSIF)***

Chase & Sherburne-Benz (2001) found that urban areas demonstrated less community involvement in ZAMSIF schools than in schools in comparison areas (maybe due to other, non-demand driven government programs implemented by the Fund in urban areas), used their inputs less, showed less impact on community cohesion and were less likely to undertake further initiatives. In contrast, in rural areas, more community involvement (although not statistically significant) increased cohesion and led to a greater likelihood to undertake further initiatives (although not uniformly).

### ***Nicaragua Emergency Social Investment Fund (FISE1)***

Focus groups generally indicated that community participation was high in prioritization of needs and project identification, but low in terms of design, construction and supervision (World Bank 2000, 37). Broken down by subproject type, participation is nearly universally lower in FISE subprojects than in non-FISE projects. FISE health centers showed lower community participation in design, but more in labor contribution to construction, than non-FISE health posts (p. 47). Participation was lower for the design and implementation of FISE schools, compared with non-FISE school projects (p. 52). Community participation was minimal for water and sewerage subprojects (pp. 59 & 64).

### ***Jamaica Social Investment Fund (JSIF)***

The five matched pairs of case studies gave insights into who participated in and really drove applications for funding, showing that the process was typically led by a minority within the community, and a leader drawn from the educated and wealthy elite (Rao & Ibáñez 2003). This is further discussed in the section on Preference Targeting later in this report.

### ***Honduras Social Investment Fund 2 (FHIS2)***

Based on focus groups in 15 pairs of FHIS participant and pipeline communities, the study found that more in-depth participation by the community at each stage of project design, implementation, and administration led to a greater sense of ownership. Looking at complementary inputs as short-term indicators of sustainability, FHIS subprojects fared no better than non-FHIS comparisons. None of the cases studies showed fully satisfactory results in subproject maintenance (Walker et al. 1999, 55).

### ***Benin Social Fund (AGeFIB) – gender study***

The gender study of Benin AGeFIB, while not an impact evaluation, highlights several important issues within the local context. Through interviews and focus groups, Walker (2002) found that AGeFIB had difficulty communicating with communities that were disorganized or experiencing conflict. The role of the village facilitator was key, but the M&E system did not take it into account. While the project promoted gender equity, most of the project personnel were male, most of the village committees had only one female representative, which was required. The low female literacy rate may have limited women's participation on subproject management committees; or, this may be a rationalization of their low participation, since many committee functions did not require literacy.<sup>111</sup>

### ***Panama Rural Poverty & National Resources Project (RPNRP)***

This cost-effectiveness study finds correlations between participation and positive project outcomes. Chi-squared tests show that above-average community participation is associated with lower costs (per beneficiary and per infrastructure unit), higher quality (in terms of community perceptions and technical assessments of infrastructure and service quality), and better infrastructure maintenance ("Panama," 51).

---

<sup>111</sup> John Elder email, 24 November 2003.

### ***Rural Water Projects***

While the participation study by Isham et al. (1995) lies beyond the scope of this review as it covers many projects beyond World Bank CDD initiatives, it is worth noting here both for the assessment of the relationship between greater participation and project performance, and for the methods used to synthesize qualitative data. The study uses systematic case review by two independent coders to transform qualitative case studies of 121 rural water supply projects into data that can be statistically analyzed. Using a variety of methods, they report statistically significant and empirically substantial evidence to support a causal relationship between participation and project effectiveness.

### **What do we know about impacts?**

Khwaja (2002) and Jimenez & Sawada (1999) are the only ones within our sample that measured the effect of participation on further project outcomes. Khwaja found that community participation in non-technical decisions had positive effects on project outcomes, but that participation in technically complex decisions negatively affected outcomes. Jimenez and Sawada found that parent participation in schools positively correlated with higher test scores in math and language.

The two large-scale studies of participation (of social funds in Zambia and Nicaragua) both documented that social fund communities did not necessarily demonstrate more or better community participation than other communities with similar projects. In Zambia, urban ZamSIF projects involved less community participation than similar comparison projects, but rural ZamSIF projects did seem to have more community participation than similar projects in other rural communities. In Nicaragua, FISE communities generally had higher participation in initial project selection, but less in implementation, than comparison communities.

Walker's (2002) gender study on Benin AGeFIB, while not an impact evaluation, highlights several important issues within the local context. These included the difficulty of communication with disorganized communities, the key role of the facilitator, and low literacy as a possible impediment to increasing women's participation.

### **What will we know?**

Ten ongoing and planned evaluations consider issues of participation, voice, inclusion, representation, and/or ownership.

One evaluation stands out in its plan to randomize the offer of participation, to compare the relative efficacy of different interventions, as well as measure their interaction effects.

#### ***Ethiopia Women's Development Initiatives Project (WDI)***

This evaluation is planning to use random assignment of participation among interested potential beneficiaries to isolate program impacts on empowerment and economic outcomes. An interesting aspect of this study is a series of individual comparisons to explore the separate impacts of participating in program promotional and organizational work versus access to technical assistance and financing. Further details are being worked out. (Christiansen & Legovini 2003)

Five studies look at participation and related questions in the context of CDD projects, with appropriately selected comparisons, and at a scale that may permit statistical inference.

***Burkina Faso National Agricultural Services Development Project (PNDSA2) & Senegal Agricultural Services & Producer Organizations Project (PSAOP)***

These two sister evaluations are looking at issues surrounding participation in and leadership of rural producers organizations (RPOs). Initial case studies will help sharpen hypotheses to be tested in quantitative surveys in 300 villages. (de Janvry et al. 2001)

***India Andhra Pradesh District Poverty Initiatives Project (APDPIP)***

Both the female and male respondent household questionnaires contain questions about participation in political and social groups. These include community actions to address problems with basic infrastructure, social problems and common property resources; individual participation and willingness to participate in such actions; and voicing issues in meetings.

***Indonesia Urban Poverty Project (UPP2)***

The household survey and case studies include modules or questions on participation in the program.

***Nepal Rural Water Supply & Sanitation Project (RWSS)***

This study will examine changes in empowerment and inclusion during the course of RWSS activity; the impacts of particular program components on empowerment and inclusion; and the relationship between RWSS investments in empowerment, inclusion and other development outcomes. Particular emphasis is being placed on inclusion and empowerment of Dalits and women. The study will use mixed qualitative and quantitative methods, and a sample designed to capture variation across regions, ethnicities, and program implementation.

***Indonesia Kecamatan Development Project (KDPI, KDP2, KDP3)***

The preliminary KDP study found that participation rates for community members and leaders in KDP areas were generally favorable and positive compared to non-KDP areas. Also KDP area respondents were more aware of the role of the poor and of community organization in KDP projects than non-KDP areas (UOI 2002, 21 & 23). This topic will be further researched in subsequent phases of the impact evaluation.

***Philippines Comprehensive & Integrated Delivery of Social Services Project (KALAHI-CIDSS)***

The household questionnaire includes sections on participation in institutions of local governance and in community decision-making and activities (APPC 2003a, 18-23). The research team will investigate the relation of participation with empowerment and other project outcomes.

One study is using a rigorous sampling framework, including propensity score matching of comparison communities, but at a smaller scale. As such, the results, while broadly representative of the overall project, may not permit statistical inference of significant impacts.

***Senegal National Rural Infrastructure Project (PNIR)***

Regarding local governance and participation, topics include representation in decision-making, especially of marginalized groups; capacity of rural councils to obtain and manage resources (including tax collection and mobilization of other financial resources); the credibility of rural councils; capacity of rural councils to defend local interests against the central state; conflict resolution capacity; participation in elections; access to information on local budgets and plans; and access of women to services, information and decision-making.<sup>112</sup>

***Moldova Social Investment Fund (MSIF)***

The terms of reference state that this evaluation will look at participation issues.

---

<sup>112</sup> Adama Touré email attachment, 17 June 2003, "Impact," 3.

### **III.C. Gender**

Not many studies have looked specifically at gender issues in the context of CDD projects. A few describe women's participation, without delving into factors that might enhance or constrain it. The notable exception is the Benin AGeFIB gender study, which took a more in-depth qualitative approach to explore the reasons behind the observed low levels of women's participation.

#### **What has been done?**

##### ***Benin Social Fund (AGeFIB) – gender study***

Based on qualitative fieldwork, this study found that women were often precluded from taking a more active role in local institutions, perhaps by their low literacy levels (Walker 2002). Many groups had only the minimum required number of female representatives (namely, one), typically the treasurer since it was felt that women were more trustworthy. The study generated interesting findings, but the sampling was not random, so it is unclear whether the findings are entirely representative. The results are better seen as generating hypotheses to test, or raising awareness about potential issues in future implementation. The study was particularly useful in highlighting issues which, from an operational standpoint, AGeFIB needed to work on.

##### ***Indonesia Kecamatan Development Project (KDP1, KDP2, KDP3)***

The government's summary report regarding the preliminary KDP1 study describes women's participation in KDP meetings: "Women generally comprise some 26 – 45% of all participants in KDP village and kecamatan meetings...That being said, the Program needs to do more to engage villagers, especially the poor and women. In general, women and the poorer members did not actively participate in the meetings" (GOI 2002, 10). The Demographic Institute report discusses stakeholders' perceptions on women's empowerment. Household members' and community leaders' perceptions regarding subproject impacts on women's empowerment were the same in KDP and non-KDP subprojects (UOI 2002, 22). The report further notes "the KDP areas appear to be involving women and the poor in infrastructure projects more effectively than do the non-KDP areas, but there are still major obstacles in the way of full equality in either setting..."(p. 34). Furthermore, community members and especially leaders felt that the impact of the KDP infrastructure project on women's empowerment was significant compared to the non-KDP areas (p. 24). These issues will be further explored in subsequent impact research.

##### ***India Dairy: Operation Flood***

This report includes description of women-only dairy cooperative societies and their relation with women's empowerment (Candler & Kumar 1998). Anecdotal evidence is mentioned of women reporting reduced domestic tension if they had an independent, if modest, source of income.

##### ***Panama Rural Poverty & National Resources Project (RPNRP)***

Retrospective questioning in 25 participant and 5 comparison areas suggests that women's role in civil society increased during the course of the project (Santucci 2003).

#### **What do we know about impacts?**

The Indonesia KDP1 evaluation is the only one that attempts to measure gender impacts in a statistically rigorous way. The main insight produced is that while KDP made significant progress increasing participation rates of women, room remains to enhance the quality of their participation.

The Benin AGeFIB gender study also generated interesting findings, but the sampling framework for the villages that were included in the study was not random, so the results are better seen as generating hypotheses to test, or raising awareness about potential issues in future implementation.

### **What will we know?**

Only two upcoming studies are looking at gender issues.

#### ***India Andhra Pradesh District Poverty Initiatives Project (APDPIP)***

The female respondent household survey contains questions on participation in household decision-making, autonomy, treatment by husband, respect from family and identity outside the home. Including indicators from the participation section, the study aims to test the hypothesis that: “Given the program’s focus on organizing women, many of whom have limited education and low social status, participation in project activities will increase the economic (income) as well as social (self-esteem) status of women and their ability to articulate themselves.”<sup>113</sup> Questions under the income section further explore gender differences in intra-household resource allocation.

#### ***Nepal Rural Water Supply & Sanitation Project (RWSS)***

Gender is included in the empowerment and inclusion study, women being considered one of the disadvantaged groups in Nepal.

## **III.D. Empowerment**

Although several studies include empowerment as an area of investigation, the term is rarely adequately defined, especially in terms of what indicators will be used and how they can be measured. However, the five country Measuring Empowerment study should provide guidance in these areas, aiming to both assess impacts and develop relevant indicators and instruments.

### **What has been done?**

#### ***India Dairy: Operation Flood***

This report includes description of women-only dairy cooperative societies and their relation with women’s empowerment. Anecdotal evidence is mentioned of women reporting reduced domestic tension if they had an independent, if modest, source of income. (Candler & Kumar 1998)

### **What will we know?**

Notably, a five country Measuring Empowerment study has been launched, with evaluation results anticipated in early 2005. The study objectives are: “(a) to assess the effect of investments in empowerment on both empowerment outcomes and development outcomes, and (b) to develop a set of indicators and instruments which can be used for tracking and evaluating these investments and their outcomes.”<sup>114</sup> Included projects are the Bolivia Poverty Reduction Strategy, Ethiopia Women's Development Initiatives Project, Nepal Rural Water Supply and Sanitation Program, Indonesia Kecamatan Development Project, and Brazil Participatory Budget Initiative.

---

<sup>113</sup> Klaus Deininger email attachment, 10 October 2003, “February 2003 Aide Memoire,” 6.

<sup>114</sup>For further information, see:

<http://Inweb18.worldbank.org/ESSD/sdvext.nsf/68ByDocName/CurrentInitiativesMeasuringEmpowermentStudy>

Including three that contribute to the empowerment study, this review includes six ongoing and planned evaluations of individual CDD projects that include empowerment as a topic for study. Most of these also include participation and related issues, and are discussed in section III.B above as well.

***Ethiopia Women's Development Initiatives Project (WDI)***

Based on random assignment of participation, this study plan to explore the empowerment and economic impacts from participating in program promotional and organizational work, versus those from access to technical assistance and financing. Early qualitative participatory techniques exploring empowerment in the Ethiopian context will inform the design of the quantitative survey. Still in the planning stage, further details of the evaluation design are being developed.

***Burkina Faso National Agricultural Services Development Project (PNDSA2) & Senegal Agricultural Services & Producer Organizations Project (PSAOP)***

In terms of empowerment, these two evaluations look at how improving rural producers organizations (RPOs) can lead to better services for farmers, both from the government and from community-based organizations to which farmers belong. The study aims to find where empowerment breaks down, or is insufficient. It will address the question of why, in spite of the richness of civil society, development outcomes remain so poor in rural Senegal and Burkina Faso, by testing several hypotheses that will help pinpoint the constraints restraining outcomes.

***Philippines Comprehensive & Integrated Delivery of Social Services Project (KALAHI-CIDSS)***

Empowerment in this evaluation largely overlaps with concepts of participation, social capital, and governance, as discussed in those sections of this annex. A particular interest of the research team is the interaction effects of empowerment and governance.

***Nepal Rural Water Supply & Sanitation Project (RWSS)***

This study will examine changes in empowerment and inclusion during the course of RWSS activity; the impacts of particular program components on empowerment and inclusion; and the relationship between RWSS investments in empowerment, inclusion and other development outcomes. Particular emphasis is being placed on inclusion and empowerment of Dalits and women. This study will use mixed qualitative and quantitative methods, and a sample designed to capture variation across regions, ethnicities, and program implementation.

***Malawi Social Action Fund (MASAF)***

This study may conduct focus groups with beneficiaries to assess perceptions of empowerment.

***Moldova Social Investment Fund (MSIF)***

The terms of reference state that this evaluation will look at issues of empowerment.

### **III.E. Conflict, Crime**

Few studies examine how CDD projects can reduce or augment conflict within and between communities. Only four upcoming studies, three in Indonesia and one in Senegal, are planning to investigate the influence of CDD projects on conflict and violent crime.

#### **What has been done? What do we know about impacts?**

No completed evaluations have looked at CDD and conflict.

## **What will we know?**

### ***Indonesia KDP & Conflict***

Among the several evaluations of KDP, this one is focusing specifically on how KDP influences conflict. It seeks to answer when KDP reduces or increases conflict, what external enabling conditions seem necessary in either case, and what specific elements of KDP affect communities' ability to manage conflict. The quantitative survey instrument is being designed as the qualitative fieldwork takes place, explicitly to incorporate and test hypotheses that arise from the qualitative findings. Case studies will seek to compare similar situations of tension where one turned violent and the other was resolved peacefully; and similar situations of tension in KDP versus non-KDP areas. Others will cover peaceful resolutions of conflict in generally violent areas, and violent cases of conflict in typically peaceful areas. Overall, a total of 56 case studies will be developed.

### ***Indonesia Urban Poverty Project (UPP2)***

This evaluation design includes components to look at issues of dispute resolution and village-level problem solving.

### ***Indonesia Support for Conflict-Ridden Areas Project (SCRAP)***

Early planning documents indicate that the SCRAP evaluation system will focus on factors that affect the root causes and levels of conflict, reconciliation and non-violent resolution of problems within communities, attitudes toward peace and violence, conflict resolution mechanisms, the incidence of violence and crime, and access to legal services. Potential case study topics further include alternative dispute resolution mechanisms, young men in conflict situations, the role of women in conflict, and fora for cooperation and reconciliation.

### ***Senegal National Rural Infrastructure Project (PNIR)***

The PNIR evaluation plans to look at its effect on local councils' abilities to anticipate and resolve conflicts (religious, inter-village, ethnic, professional, etc.), using a panel data set covering 19 pairs of PNIR and matched comparison communities.<sup>115</sup>

## **III.F. Leadership**

Two studies have discussed the role of leadership as a key factor in the success of CDD projects. While leadership is not really an outcome of interest, these studies do highlight the fact that it is often a critical input in the CDD process, and that studies which fail to account for it may be missing an important piece of the process that determines other impacts. This could lead to problems of bias and misplaced attribution. Only one upcoming study has explicitly included leadership as a factor to take into account. While not always appropriate for inclusion in every research design, leadership merits more attention in future CDD evaluations.

## **What has been done?**

### ***Pakistan Aga Khan Rural Support Project (AKRSP) – Khwaja study***

Khwaja (2002, 39-41) shows that projects are better maintained in communities with the presence of a project leader, as instrumented by attributes of the local hereditary-leader household. Maintenance is better still with a higher quality leader. Moreover, the leadership effect is stronger for more complex projects.

---

<sup>115</sup> Adama Touré email attachment, 17 June 2003, "Impact," 7.

### ***Jamaica Social Investment Fund (JSIF)***

Rao and Ibáñez (2003) highlight the role of influential and well-connected individuals or groups within communities in determining the choice of projects to be implemented and in accessing social fund resources for the project. All five case studies of Social Fund projects found evidence that leaders, in the form of established organizations or influential individuals, were able to get their choice of project implemented, often in spite of a preference for a different project by a majority of the community.

### **What do we know about impacts?**

These two studies highlight the importance of leadership in influencing other outcomes in CDD projects, yet do not document exactly how or how much leadership affects other impacts.

### **What will we know?**

#### ***Indonesia Urban Poverty Project (UPP2)***

This evaluation is including local leadership as a factor to be accounted for. The household survey will gather data on local leadership, including characteristics of leadership and scores for local leaders and activists.

## **IV. GOVERNANCE**

Evaluations investigating governance are presented below in the following categories:

(A) improving local governance, (B) decentralization, and (C) corruption.

Very few completed evaluations of CDD projects have considered questions of governance.<sup>116</sup> To our knowledge, no reliable evidence has as yet been generated regarding CDD impacts on governance. This is surprising, given that this is a topic of much discussion. Several evaluations in the pipeline will investigate CDD impacts on local governance and interaction effects with decentralization. As a result, some reliable evidence should be available over the next few years. A few studies are investigating whether CDD projects have an impact on corruption and transparency of local-level institutions and government. While these should produce some evidence, this topic would benefit from further attention.

Lastly, research so far has largely neglected the question of CDD program influences on reform in line ministries and higher levels of government. While many people discuss the influence that CDD projects may have on governance, we have found no studies that have investigated how CDD projects impact public sector reform at higher levels, encourage or hinder changes in line ministries, or effect or delay reform at any level of government above the local. No studies provide any empirical evidence to support or refute one side or the other in this debate, nor do any seem to be planned. Some studies investigating local governance include marginal components on public sector management, but the focus is more on how local level institutions can better demand what they need from higher levels of government than on how supply of centralized services is improved.

### **IV.A. Improved Local Governance**

While no rigorous studies have produced evidence of CDD impacts on local governance, a number of ongoing and planned evaluations seek to answer related questions. On the whole, these upcoming

---

<sup>116</sup> However, two desk studies provide insights on social funds and institutions. See Bhatia (2003) and Parker & Serrano (2000).

evaluations appear to have good research designs, and should generate some interesting evidence over the next few years.

### **What has been done? What do we know about impacts?**

One completed study has looked at local governance in the context of CDD projects, but the evidence is limited to the descriptive and anecdotal.

#### ***Panama Rural Poverty & National Resources Project (RPNRP)***

Based on retrospective questioning in 25 participant and 5 comparison areas, study findings suggest that this project may have improved local government in terms of more interaction with community-level institutions and better awareness of citizens' needs (Santucci 2003).

### **What will we know?**

Twelve ongoing or planned studies will look at whether and how CDD projects affect local governance. One of these is still in very early planning stages, and is listed separately below. Nearly all of the others have reasonably rigorous evaluation designs, but they vary in their focus on governance as a topic for evaluation versus governance as part of the descriptive context to be considered in evaluating other outcomes. They are listed roughly in order of their apparent interest in local governance as an outcome for evaluation.

#### ***Indonesia Urban Poverty Project (UPP2)***

A village-head questionnaire includes questions on local government organization and participation in higher-level government meetings. A household survey includes questions on satisfaction with local government, contact with government officials and participation in village-level government. An activist interview includes questions about village-level oligarchy, village-level government organization, and government or community organizations at the super-village level that community members can participate in, along with case studies of the largest community project initiated by government and by the community. Qualitative case studies in a subset of villages include the following aspects of local governance: trust in and problem solving by kelurahan (village) government, participation in decision making at levels of neighborhood and community, leadership, access to information on government decisions, predictability of decision-making, accountability, representation, and new institutions.

#### ***Philippines Comprehensive & Integrated Delivery of Social Services Project (KALAHI-CIDSS)***

A primary focus of this study is on empowerment and governance, and the household survey includes the following topics: household participation in community collective action; transparency and accountability of local institutions; participation in community decision-making and activities (APPC 2003a, 18-23). In addition, the local leader interview collects information on local government organization and interaction between different levels of government (APPC 2003b).

#### ***Senegal National Rural Infrastructure Project (PNIR)***

PNIR works at the lowest level of elected government. Regarding local governance, topics for research include capacity of rural councils to obtain and manage resources (including tax collection and mobilization of other financial resources); the credibility of rural councils; and the capacity of rural councils to defend local interests against the central state.<sup>117</sup>

---

<sup>117</sup> Adama Touré email attachment, 17 June 2003, "Impact," 3.

***Burkina Faso National Agricultural Services Development Project (PNDSA2) & Senegal Agricultural Services & Producer Organizations Project (PSAOP)***

These two sister evaluations are looking at the role of rural producers organizations (RPOs), and how these influence local governance, in the sense of making public services more responsive to client demand.

***Brazil Rural Poverty Reduction Projects (PCPR)***

This study will look at how PCPR affects social capital, and its interaction effects vis-à-vis local governance, as detailed in the social capital section above. This evaluation includes some questions addressing communities' relations with higher levels of government and bureaucracy. (FECAMP 2001)

***Cambodia Rural Investment & Local Governance Project (RILG)***

The KAPB (Knowledge, Attitudes, Practices and Beliefs) study documented views on accountability, participation, predictability, and transparency. As it stands the KAPB is a very thorough description of how citizens and government officials think about governance in Cambodia, but not an impact evaluation as such. A follow-up study on governance plans to use the KAPB as a baseline, which could generate interesting results on how perceptions have changed, though this may not be directly attributable to RILG.

***Armenia Second Social Investment Fund (ASIF2)***

This evaluation is using matched case studies to examine ASIF's effect on institutional quality and governance at the community level; and document the status of decentralization.

***Indonesia KDP & Conflict***

The KDP conflict evaluation will include local governance institutions as part of its investigation of the environmental context.

***Indonesia Support for Conflict-Ridden Areas Project (SCRAP)***

The draft design for the SCRAP household survey indicates that the evaluation will explore issues of village-level decision-making, leadership selection, accountability and transparency.

***Malawi Social Action Fund (MASAF)***

This project plans to evaluate MASAF's influence on public sector management, using case studies and key informant interviews as data collection tools.

This study is still in early planning stages, and the evaluation topics and designs are not firm:

***Moldova Social Investment Fund (MSIF)***

This evaluation aims to assess long-term improvement in accountability and management by local government.

## **IV.B. Decentralization**

Though popular topics of discussion, the interaction effects of decentralization and CDD projects are not well evaluated. No completed evaluations have studied it, and the planned ones are few. This is an area that could benefit from more attention in future research.

### **What has been done? What do we know about impacts?**

No completed evaluations have looked at decentralization in relation to CDD projects.

## **What will we know?**

Two ongoing evaluations are looking at the interaction of CDD projects and decentralization.

### ***Pakistan National Rural Support Program (NRSP) – DEC study:***

Within the DEC study of the NRSP, a sample of 90 villages will enable research into the role of social mobilization in the take up of public funds, within a context of devolution. The Pakistani government has decentralized, making public funds equally available to all communities. The evaluation will test the hypothesis that areas where NRSP is engaged in the process of social mobilization are better equipped to form an organization, obtain local government funds, and use the funds more effectively (in terms of quality of organization, profile of members, more equitable distribution of project benefits, etc.). Such findings would contribute to a clearer understanding of how important the social mobilization aspect of the program is in empowering poor communities and facilitating collective action.

### ***Senegal National Rural Infrastructure Project (PNIR)***

In the context of decentralization in Senegal, PNIR works at the lowest level of elected government, the communauté rural (CR). This evaluation aims to study governance and participation in relation to decentralization. Topics for research at both the CR and village level include representation in decision-making, especially of marginalized groups; capacity of rural councils to obtain and manage resources (including tax collection and mobilization of other financial resources); the credibility of rural councils; capacity of rural councils to defend local interests against the central state; conflict resolution capacity; participation in elections; access to information on local budgets and plans; and access of women to services, information and decision-making.<sup>118</sup>

## **IV.C. Corruption**

As with decentralization, many claims are made about the merits of CDD in terms of enhancing transparency, but few evaluations directly investigate the impact of CDD projects on corruption. This is another area for greater research.

## **What has been done?**

### ***Indonesia Kecamatan Development Project (KDPI, KDP2, KDP3)***

Household members and community leaders were more often satisfied with financial transparency in KDP subprojects than in non-KDP subprojects, by margins of 7 and 9 points, respectively, both of which are significant at the one-percent level (UOI 2002, 22). Subsequent phases of research should generate further insights.

## **What do we know about impacts?**

So far, no studies have investigated the impacts of CDD projects on corruption outside of the project. The Indonesia KDP study focused on a comparison of stakeholder perceptions of transparency within the project.

---

<sup>118</sup> Adama Touré email attachment, 17 June 2003, “Impact,” 3.

## **What will we know?**

Four upcoming evaluations include topics related to corruption.

### ***Indonesia Urban Poverty Project (UPP2)***

A household survey, key informant interviews and case studies will all look at transparency of local government.

### ***Philippines Comprehensive & Integrated Delivery of Social Services Project (KALAHI-CIDSS)***

The household survey includes questions on transparency and accountability of local institutions and activities (APPC 2003a, 22).

### ***Cambodia Rural Investment & Local Governance Project (RILG)***

The KAPB (Knowledge, Attitudes, Practices and Beliefs) baseline study includes stakeholders perceptions on governance, decentralization and corruption.

### ***Moldova Social Investment Fund (MSIF)***

This evaluation is looking at transparency within the social fund, but not necessarily at its effect on transparency in broader government at any level.

## **V. TARGETING**

Two kinds of targeting are often examined as part of evaluations of CDD programs. One is poverty targeting, where the analysis measures the extent to which resources are channeled to the poorer segments of the population. The other is preference targeting, where the analysis describes the extent to which the implemented subprojects match the ex-ante preferences of the beneficiaries.

### **V.A. Poverty Targeting**

The question of pro-poor targeting has been reasonably well studied so far, yet the results from different studies are not consistent. Moreover, poverty targeting of CDD programs is rarely compared against that of other delivery mechanisms. This is curious, given the emphasis placed on CDD's ability to reach the poorer segments of the population, and highlights it as an area for further research. Further studies should continue to disaggregate targeting analysis by sector and by rural-urban or other relevant geographic stratifications. As more studies add to the literature, it will be possible to see if any patterns emerge. It is also worth noting that different measures of poverty can generate different results.

The four reliable, completed studies generate mixed findings on the effectiveness of poverty targeting in CDD programs. The results suggest that community-level targeting tends to be progressive, while household-level targeting is less often so. However, results often vary considerably between rural and urban areas, and among different sectors. One illustrative example of this comes from the Zambia ZamSIF study, which found that the targeting of rural education and urban health investments was pro-poor while the targeting of urban education and rural health investments was pro-rich. This same study also highlights how the results of poverty targeting analyses depend on the definition of poverty used. Headcount versus poverty gap measures can yield different results, so it is important to include as many as are feasible from the data available.

## **What has been done?**

Seven completed studies have looked at poverty targeting.

**Nicaragua Emergency Social Investment Fund (FISE1)**

At the municipal level, FISE1 targeting was progressive overall. At the community level, education investments were neutral and health investments were progressive. Sewerage subprojects were regressive, possibly because they were only practicable in urban areas. At the household level, latrine projects were very pro-poor. Sewerage projects were very pro-rich. Water investments were neutral. Both education and health were “slightly less progressive at the household level than the community level” (World Bank 2000, iv).

Targeting the poorest of the poor (poorest 17% living in extreme poverty) was limited at both community and household levels. Primary education and latrines were progressively distributed among this segment of the population, but only latrines remain progressive for the poorest 10%.

**Table 7: Nicaragua FISE poverty targeting was progressive overall, but variation among sectors was high**

	Municipal		Community		Household	
	Extreme Poverty	High Poverty	Extreme Poverty	High Poverty	Extreme Poverty	High Poverty
<b>Overall</b>	Progressive	Progressive				
<b>Education</b>	Progressive	Progressive	Slightly Progressive	Slightly Progressive	Slightly Progressive	Progressive
<b>Municipal Infrastructure</b>	Regressive	Slightly Progressive				
<b>Health</b>	Progressive	Progressive	Neutral	Very Progressive	Regressive	Progressive
<b>Latrines</b>	Neutral	Neutral	n.a.	n.a.	Very Progressive	Very Progressive
<b>Sewerage</b>	Slightly Progressive	Slightly Progressive	Very Regressive	Very Regressive	Very Regressive	Very Regressive
<b>Water</b>	Very Progressive	Progressive	n.a.	n.a.	Regressive	Neutral
<b>Social Assistance</b>	Very Regressive	Regressive				
<b>Reforestation</b>	Regressive	Progressive				

n.a. = not applicable

**Zambia Social Recovery Project / Social Investment Fund (ZAMSIF)**

ZAMSIF relies on its menu of subprojects to encourage self-targeting of poorer communities. Chase and Sherburne-Benz (2001) used data from 3,185 ZAMSIF households and 13,484 comparison group households, roughly half in urban and half in rural areas to examine the household-level poverty targeting performance of ZAMSIF. They compared targeting based on different poverty definitions, namely headcount, poverty gap, and poverty gap squared, although the statistical significance of differences for these comparisons was not reported. They also compared household expenditures per adult equivalent, where statistical significance was reported, and they graphed concentration curves of ZAMSIF funding across income deciles. They also examine results based on the poverty line (an income K46,286 per adult

equivalent, enough to cover basic food needs, plus 30% to cover basic non-food needs), and based on the extreme poverty line (an income of K32,232 per adult equivalent, enough to cover only basic food needs).

Based on poverty and extreme poverty headcounts, targeting was more or less neutral overall, although somewhat progressive below the extreme poverty line in rural areas. Using poverty gap or poverty gap squared measures, based on poverty or extreme poverty lines, showed that targeting was progressive overall, particularly in rural areas. In urban areas, these measures showed neutral targeting below the poverty line, but progressive targeting below the extreme poverty line. Comparing adult equivalent per capita expenditure showed that targeting was progressive overall (beneficiaries spent 8% less compared to the overall population), and progressive in rural areas but regressive in urban areas, particularly Lusaka. Looking at education and health subprojects separately found somewhat odd results. Education investments were progressive in rural areas and regressive in urban; whereas health investments were regressive in rural and progressive in urban. Concentration curves were neutral to mildly progressive, including for the poorest two deciles, so the poorest of the poor were targeted reasonably well, although urban spending was again shown to be slightly regressive. (Chase & Sherburne-Benz 2001, Tables 1 & 2 & Figure 3) The authors point out that part or all of ZAMSIF's success in reaching the poor can be explained by Zambia's high overall poverty incidence. Nationwide, 72% are below the poverty line, 56% are below the extreme poverty line (p. 9).

#### ***Honduras Social Investment Fund 2 (FHIS2)***

This evaluation found that FHIS2 targeting of beneficiaries was quite progressive. Of the population in areas where FHIS2 implemented subprojects, 19% were in the poorest decile and 13% were in the second poorest decile. In terms of actual users of services and infrastructure provided by FHIS2, 22% were from the poorest decile and 15% were from the second poorest. A full 29.5% of the resources spent on FHIS2 subprojects directly benefited those in the poorest decile (Walker et al. 1999, 22). These findings are noteworthy in that household-level targeting was actually more progressive than community-level targeting, in contrast to most other studies' findings.

#### ***Armenia Social Investment Fund (ASIF1)***

Although ASIF targeted areas with poor infrastructure, rather than poor areas per se, community-level targeting was mildly progressive (Chase 2002). People in ASIF communities were poorer on average than other Armenians. However on the household level, targeting in all areas was nearly income-neutral, being slightly progressive in urban areas (mostly Yerevan, the capital), and slightly regressive in rural areas. The requirement of a 10% local contribution to the subproject may have excluded very poor rural communities.

#### ***Peru Social Fund (FONCODES)***

Paxson and Schady (2002, 304-309) regress FONCODES education expenditures on average per capita income for each district. The same is done for a centralized government program, INFES. FONCODES is effective in targeting poor districts, whereas INFES targeting is neutral or even slightly regressive. The authors point out that this is to be expected, given that FONCODES focuses on rural primary schools, and INFES focuses on urban secondary schools.

Household targeting was progressive overall, but slightly regressive within districts. Logit regressions of benefit incidence on household per capita income showed that FONCODES' targeting was progressive, even more so than parents' committees, while INFES was regressive. Nonparametric regressions confirmed these results. However, nonparametric regressions of benefit incidence on the standard deviations of a household's per capita income from the district average showed that within districts FONCODES and INFES were both slightly regressive, while parents' committees performed slightly better.

### ***Indonesia Kecamatan Development Project (KDPI, KDP2, KDP3)***

Poverty targeting was evaluated only for microfinance activities, which constituted 14% of the overall KDPI resource distribution. Economic Loan Schemes were pro-rich in their targeting, while non-KDP loans were neutral in targeting poor or rich borrowers. Qualitative data indicated that this was because of widespread perceptions that the poor would not be able to repay the loans, or related to this, would not use the money for productive purposes with a short-term return, as required (UOI 2002, 28-29). Further qualitative findings indicated that KDP loan schemes were too complicated and not well targeted to those who really wanted them (p. 30). Targeting will be further studied in later phases of the impact evaluation.

### ***India Dairy: Operation Flood***

While Operation Flood was not designed as a poverty project and a formal evaluation of poverty targeting was not part of this report, the study notes that small milk producers, the majority of whom were poor, were the main beneficiaries. It also mentions that the extremely poor could not be beneficiaries, since one had to possess or obtain at least one milch animal to participate (Candler & Kumar 1998, 6).

## **What do we know about impacts?**

Four studies have produced reliable results on poverty targeting of CDD projects:

- Nicaragua FISE 1      - Honduras FHIS2      - Peru FONCODES
- Armenia SIF            - Zambia ZamSIF

Based on the results of these four social fund studies, poverty targeting has a very mixed record in CDD projects. The only clear pattern that emerges is that targeting poorer communities is generally easier than targeting poorer households. But results can vary considerably across sectors, and across urban and rural areas. Results can also depend on the measure of poverty used.

## **What will we know?**

Ten ongoing and planned evaluations will explore poverty targeting.

### ***Benin Social Fund (AGeFIB)***

This study will make use of a national population census and conduct a survey of 6000 households. This information will be used to look at community- and household-level poverty targeting (Coulombe 2002).

### ***Brazil Rural Poverty Reduction Projects (PCPR)***

This evaluation will look at the effectiveness of poverty targeting using secondary data in conjunction with the household survey data collected in PCPR areas to see where PCPR communities fit within the distribution of rural poverty (FECAMP 2001, 99-128).

### ***Burkina Faso National Agricultural Services Development Project (PNDSA2) & Senegal Agricultural Services & Producer Organizations Project (PSAOP)***

These two sister studies will collect qualitative information on who participates in and who is excluded from rural producers organizations, and who self-selects to be in (de Janvry et al. 2001). The studies aim to find out whether or not rural producer organizations (RPOs) in Senegal and Burkina Faso are efficient vehicles for reaching the rural poor.

### ***Cambodia Rural Investment & Local Governance Project (RILG)***

This study includes some qualitative work asking about the appropriateness of beneficiaries and how well subprojects address their needs.

### ***Malawi Social Action Fund (MASAF)***

The targeting analysis calculates correlation coefficients between MASAF expenditures per capita and poverty headcount index for both district and traditional authority levels. It also examines correlation between expenditures and political alignment measured by the vote share for the political party currently in power. Although the evaluation is still ongoing, some findings have already been generated for the targeting analysis. It appears that MASAF funding is more or less neutral both in terms of poverty and politics. No correlations between expenditures and poverty headcount or political alignment were found to be statistically significant. Smaller districts seem to be favored with higher per capita expenditures.

### ***Moldova Social Investment Fund (MSIF)***

This study will use a mix of qualitative and quantitative data. The terms of reference call for an evaluation of whether the poorest villages were targeted and how well the proxy poverty targeting criteria used by MSIF to identify poor communities worked.

### ***Pakistan National Rural Support Program (NRSP) – DEC study***

The DEC NRSP study will use a randomized experiment in 90 randomly selected villages to assess three different credit delivery mechanisms. NRSP management has been testing different models for deepening their social mobilization effort and improving their targeting within communities. This is a particular challenge in a context where the program is rapidly expanding to new areas. Credit is a key program component and the program has thus far used community members to identify potential credit recipients and to collect loans. This has worked well in some cases but has often led to poor targeting, excessive reliance on community activists who have used the credit program to expand their influence locally and in some cases has even led to substantial default. Due to this social organizers have often had to short change social mobilization efforts to focus on credit delivery and recovery. NRSP is thus attempting to redesign its microcredit delivery system by separating credit operations from their social mobilization work, and creating village credit branches with full time field workers who will take over many of the day to day credit related tasks and report to a credit officer rather than the social organizer. This provides an opportunity to examine which types of incentive schemes for field workers are likely to produce the greatest impact on credit targeting, outreach and delivery costs.

### ***Senegal National Rural Infrastructure Project (PNIR)***

This ongoing evaluation ties the assessment of poverty targeting and elite capture to political economy theories of intra-village allocation of project benefits. To do so, household-level distribution of impacts within villages will be regressed against three indicators and controls to test if (i) “median voter”, (ii) most disadvantaged, or (iii) local elites best capture project benefits. These correspond, respectively, to (i) Median Voter Theorem, (ii) Rawls’ theory of social justice, and (iii) elite capture models. The respective indicators would be a negative coefficient on (i) the squared difference between household  $i$ ’s welfare and the welfare of the median household for that village, (ii) the difference between household  $i$ ’s welfare and the welfare of the poorest (or worst off) household, or (iii) the difference between household  $i$ ’s welfare and the welfare of the richest (or best off) household.<sup>119</sup>

### ***St. Lucia Poverty Reduction Fund (PRF)***

For poverty targeting, the St. Lucia PRF evaluation plans to look at household and community level targeting. Data will be drawn from a number of sources, including the 1991 and 2001 censuses, a household survey, administrative records and qualitative fieldwork.

---

<sup>119</sup> Adama Touré email attachment, 17 June 2003, “Evaluation,” 11.

## **V.B. Preference Targeting**

The most rigorous way to analyze preference targeting performance is to collect baseline data on what people think the greatest needs in their community are, and then compare this to what is actually implemented. None of the completed or upcoming studies do this. Some use retrospective questioning to get at ex-ante priorities; others simply ask about how satisfied beneficiaries are with subprojects. These are second-best options, especially since the Jamaica JSIF and Nicaragua FISE1 studies emphasize that ex post beneficiary satisfaction is not necessarily a reliable indicator of ex ante preferences. Studies need to be carefully designed so that it will be apparent when beneficiaries think, “We did not get what we wanted, but at least we got something,” versus “This project is what we needed most.” Given that effective preference targeting is often cited as an advantage of CDD projects, this should be emphasized more in the research agenda moving forward.

### **What has been done? What do we know about impacts?**

None of the completed evaluations have actually collected baseline data on beneficiary preferences or priorities, so no reliable evidence has been documented regarding CDD effectiveness in terms of preference targeting.

#### ***Jamaica Social Investment Fund (JSIF)***

Examining five pairs of matched case studies, Rao and Ibáñez (2003) found that sub-projects were quite poorly aligned with communities’ ex-ante preferences, based on retrospective questioning. Nevertheless, most beneficiaries expressed satisfaction with them ex-post.

#### ***Indonesia Kecamatan Development Project (KDPI, KDP2, KDP3)***

Based on surveys in 74 matched pairs of KDP and non-KDP communities, both community leaders and household members felt that subprojects were the highest priority for raising household incomes more often in KDP than in non-KDP subprojects, by margins of 3.4 percentage points for household respondents and 6.0 points for community leaders, both of which are significant at the 1% level (UOI 2002, 22). Subsequent phases of the impact evaluation should provide further details.

#### ***Nicaragua Emergency Social Investment Fund (FISE1)***

A separate beneficiary assessment (BA) was conducted, but it seems limited to questions about participation in the process and satisfaction with the outcome. It was not possible to evaluate ex-ante subproject preferences and whether these were matched by the subprojects actually implemented, since no baseline data existed on community development preferences. Rather than asking retrospective questions about priorities to see if they matched FISE investments, the BA asked FISE communities to name the most beneficial project. They often named FISE projects, but the evaluation does not claim this is conclusive evidence that projects were so well aligned with ex-ante priorities. It is also notable that there is a gap between the percentage expressing satisfaction with the FISE experience, at 82%, and the percentage naming FISE projects as the most beneficial in their community, at 71% (World Bank 2000).

#### ***Honduras Social Investment Fund 2 (FHIS2)***

Based on retrospective questioning, most sub-projects were in line with community priorities, although sewerage and latrine sub-projects often were not. Many expressed interest in road and electricity projects, which were not available on the menu of sub-projects in FHIS (Walker et al. 1999).

## What will we know?

Four ongoing or planned evaluations are looking at preference targeting. None of them appear to be using a baseline to measure beneficiaries' ex ante preferences or priorities for development projects.

### ***Benin Social Fund (AGeFIB)***

This study intends to look at how well subprojects respond to the needs of targeted communities.

### ***Burkina Faso National Agricultural Services Development Project (PNDSA2) & Senegal Agricultural Services & Producer Organizations Project (PSAOP)***

These two sister evaluations are assessing how effectively the rural producer organizations represent members' interests. The overall research design calls for a panel data design with surveys of representatives of the community and the producers organization, as well as some case studies.

### ***St. Lucia Poverty Reduction Fund (PRF)***

The evaluation will look at the correspondence between community priorities and implemented infrastructure. Data instruments will include recall data in household surveys and qualitative discussions regarding project selection and decision-making.<sup>120</sup>

## VI. COMPARATIVE EFFECTIVENESS

This section reviews evaluations that explore the interrelated questions of (A) CDD effectiveness as compared to alternate interventions; (B) the relative performance of different delivery mechanisms within a CDD program; and (C) differences in the cost-effectiveness and economic returns of initiatives.

Several studies compare various performance measures of CDD projects to other providers of unclear origin. Without clearly explaining differences in implementation mechanisms across the two, they do not provide clear evidence of the relative merits of CDD versus non-CDD implementation of projects. Overall, this review discovered little evidence that directly address these topics, and only a few rigorous upcoming evaluations that plan to tackle these questions directly.

### VI.A. CDD versus non-CDD

A review of available evidence confirms the findings of Mansuri and Rao (2003, 19) regarding a lack of robust evaluations that compare CDD interventions to similar projects implemented by less participatory, and perhaps more centralized or top-down, development mechanisms. A central problem here is that even high quality comparative evaluations often fail to include details on the delivery mechanism used by the alternative provider, especially in how it differs from the CDD approach. Two upcoming evaluations plan a rigorous treatment of these concepts, however additional studies in this area would inform resource allocation decisions across different development approaches.

## What has been done?

### ***Pakistan Aga Khan Rural Support Project (AKRSP) – Khwaja study***

A distinguishing aspect of Khwaja's study is the inclusion of sites where two infrastructure projects were implemented within a single community. This enables him to isolate project-specific factors from differences among communities, relating to both project complexity and external implementing agency (Khwaja 2002, 11). Controlling for indicators of participation, he finds that in comparison to similar

---

<sup>120</sup> Ian Walker email, 30 October 2003.

infrastructure projects implemented by local government, the Local Bodies and Rural Development (LB&RD) section, AKRSP projects were better maintained, regardless of project type, complexity of project and current physical condition (p. 23). He notes that this calls for further investigation, to see what the NGO is doing beyond involving communities in project decisions to distinguish its performance from that of local government bodies (p. 27).

### ***Peru Social Fund (FONCODES)***

On the district level, Paxson and Schady (2002) compare the targeting performance of the Peru FONCODES program to a more centralized government program, INFES. At the household level, FONCODES' poverty targeting performance is compared against both INFES and parents' committees. However, the paper does not provide details regarding differences in implementation mechanisms among the three groups.

### ***Indonesia Kecamatan Development Project (KDP1, KDP2, KDP3)***

The initial study conducted by the University of Indonesia of KDP1 impacts on community organization compared household and community leader responses in 74 matched pairs of KDP and non-KDP areas. The report compares difference in means across the two and provides the significance level for hypothesis tests. However, while the KDP projects receive higher participation scores, the paper does not explicitly provide details on the other implementing mechanisms: "In KDP areas, the selection was from all KDP infrastructure projects, while in non-KDP regions, the criteria was simply to choose the best among any infrastructure projects likely to have the same investment characteristics as those supported by the KDP program – they were to be economic investments, not social investments" (UOI 2002, 20). These topics will be further explored in subsequent phases of research.

### ***Nicaragua Emergency Social Investment Fund (FISE1)***

Case studies compared FISE1 health and education facilities and water and sanitation infrastructure to those implemented by other providers to examine the relative performance in terms of utilization and sustainability. However, details on the implementation and management process supported by the other providers are unclear (World Bank 2000, 40-64).

### ***Honduras Social Investment Fund 2 (FHIS2)***

The evaluation of Honduras FHIS2 used an engineer to technically assess the quality of infrastructure implemented by FHIS in comparison to other providers. Again, the implementation mechanism used by the other providers is not clear from the report (Walker et al. 1999, 37).

Other comparative studies focusing on cost-effectiveness and participation are discussed below.

## **What do we know about impacts**

Thus far there is little evidence to document the relative performance of CDD versus non-CDD initiatives. Further information could be assembled by researching the implementation mechanisms used by the other providers included in some of the completed studies above.

## **What will we know?**

### ***Pakistan National Rural Support Program (NRSP) – DEC study***

Part of the DEC Pakistan NRSP study will examine whether community infrastructure created under this program is of better quality, more cost effective and more sustainable (design,

construction, cost, maintenance) than similar (in size and type) infrastructure created by the government. To do so, the research team will compare similar projects provided by the RSPs and by the government within the same village, relying on a sample of 140 villages with some 450 comparable projects. This design will enable a “more rigorous test of the proposition that community assets created through CDD initiatives are better maintained” than those by more conventional means, particularly by a centralized bureaucracy.<sup>121</sup>

## **VI.B. Effectiveness of different delivery mechanisms within CDD**

There is little evidence on the relative performance of different delivery schemes within a CDD program. Three upcoming studies plan to evaluate different mechanisms within their programs, one with a rigorous design already in place, and two with only tentative evaluation plans at present.

### **What has been done? What we know about impacts?**

At present there appears to be insufficient information on the performance of different delivery mechanisms within CDD programs.

### **What we will know?**

Two upcoming studies plan to explore the relative performance of different service delivery mechanisms within a CDD program.

#### ***Pakistan National Rural Support Program (NRSP) – DEC study***

The NRSP program and research teams were able to agree on a process of random program placement and random assignment of intervention types to explore the efficacy of different NRSP mechanisms. This enables a comparison of outcomes across communities where a specific feature of the program is randomized in order to get a clearer sense of how components of the program impact outcomes of interest. This strategy will allow for an examination of a number of specific hypotheses.

The first intervention will examine the impact of bundling microcredit with information on local productive opportunities and enterprise development skills. The provision of credit to the poor for microenterprise development usually assumes that liquidity is the main constraint on productive activity. If information on available productive opportunities and market conditions is asymmetrically distributed, however, then a specific intervention to deal with this market failure should improve both credit uptake among poorer households and the establishment of profitable microenterprises within the target group. The intervention design includes 90 randomly selected villages: a third of which will act as controls, another third which will get the basic program, which includes credit, and a final third where credit will be complemented with entrepreneurship skills training and training in the identification of local production and marketing opportunities. The baseline for this intervention will be completed in March 2004 and the intervention will begin immediately after this.

Second, NRSP management has been testing different models for deepening their social mobilization effort and improving their targeting within communities, a particular challenge since the program is rapidly expanding to new areas. Credit is a key program component and the program has thus far used community members to identify potential credit recipients and to collect loans. This has worked well in some cases but has often led to poor targeting, excessive

---

<sup>121</sup> Ghazala Mansuri email attachment, 20 August 2003, “Community,” 3.

reliance on community activists who have used the credit program to expand their influence locally, and in some cases, has even led to substantial default. Due to this, social organizers have often had to short change social mobilization efforts to focus on credit delivery and recovery. NRSP is thus attempting to redesign its microcredit delivery system by separating credit operations from their social mobilization work, and creating village credit branches with fulltime field workers who will take over many of the day to day credit related tasks and report to a credit officer rather than the social organizer. This provides an opportunity to examine which types of incentive schemes for field workers are likely to produce the greatest impact on credit targeting, outreach and delivery costs. To examine these issues, the study is designing a randomized experiment that will allow an assessment of three different credit delivery mechanisms. This will be done in another 90 randomly selected villages—approximately one-half of which overlap with the sample for the CPI (Community Physical Infrastructure) study.

### ***Brazil Rural Poverty Reduction Projects (PCPR)***

The ongoing evaluation by the University of Campinas (FECAMP 2001) is comparing effectiveness across three resource allocation mechanisms within PCPR, each with different degrees of decentralization: (A) PAC projects (25% total costs) where community associations submit an application to the technical unit, (B) FUMAC (56%) where municipal councils act as intermediaries between community associations and the technical unit, and (C) FUMAC-P (9%) where municipal councils manage the annual budget and approve community association proposals directly and the technical unit delegates budget to the municipal councils.

## **VI.C. Cost-effectiveness, Economic returns**

Few studies investigate relative cost-effectiveness or economic rates of return of CDD projects versus other alternatives. From the few here, no clear picture emerges of the comparative cost performance of CDD operations. However, for cost details on six social funds and comparators, see Rawlings et al. (2004).

### **What has been done?**

#### ***Pakistan Aga Khan Rural Support Project (AKRSP) – OED study***

Regarding cost effectiveness, the OED study tabulated costs for both AKRSP projects and comparators, including the Pakistan National Rural Support Program, India AKRSP, and two programs in Pakistan implemented by the International Fund for Agricultural Development (IFAD) (World Bank 2002a, 3). Since comparisons are drawn to other “similar organizations,” this appears to be more of a statement regarding the cost performance of the AKRSP program itself, and not the CDD approach (World Bank 2002a, 11). Cost calculations place the Pakistan program at operating costs of \$42 per household and total costs of \$385 per household (over the previous five years; for total program life of 17 years, the latter figure is \$890). These are both substantially higher than those for India AKRSP (\$19 operating and \$210 total costs per household) and Pakistan NRSP (\$9 operating, \$11 total), but lower total costs per household than IFAD’s Pakistan Barani (\$705) and IFAD’s Pakistan Northern Areas Development (\$1,255) (World Bank 2002a, 51). The report concludes that Pakistan’s AKRSP is at the “top end of the range” (p. 11). Furthermore, the report states that the high per household costs are justified by a high economic rate of return on investments of 16 – 24 percent from 1982 to 1999 (pp. 8 & 67). No ERR calculations were performed for the comparator providers.

Similarly, based on AKRSP calculations, the OED study also compares the performance of its microfinance division with the average performance of other microfinance providers in the region. The report found that AKRSP had lower operating costs as both a percentage of the average loan portfolio

(12% versus 20% regional mean) and as cost per borrower (\$16 versus \$19) (World Bank 2002a, 104-105).

***Panama Rural Poverty & National Resources Project (RPNRP)***

This study used chi-squared tests and found a positive association between above-average participation and below-average costs across six projects.

**What we will know?**

***Pakistan National Rural Support Program (NRSP) – DEC study***

Part of the DEC study of the NRSP will include 140 villages with some 450 infrastructure schemes to explore whether community infrastructure created under this program is more cost effective than similar (in size and type) infrastructure created by the government.

***Cambodia Rural Investment & Local Governance Project (RILG)***

The inception report for the Cambodia RILG baseline study plans to collect data that can be “used in future” to conduct cost-benefit analyses and ERR calculations (Helmert & Wallgren 2003, 10).

***Brazil Rural Poverty Reduction Projects (PCPR)***

The technical proposal indicates that cost-benefit and cost-effectiveness analysis will be included in the evaluation, with details to be worked out in the early stages of implementation (FECAMP 2001, 128-146).

***Indonesia Kecamatan Development Project (KDPI, KDP2, KDP3)***

Upcoming phases of research will explore the cost-effectiveness of KDP projects.

***Indonesia Support for Conflict-Ridden Areas Project (SCRAP)***

The SCRAP project plans to include studies of cost-effectiveness, economic internal rates of return and cost-benefit.

***Malawi Social Action Fund (MASAF)***

The first stage of the desk review discusses earlier cost effectiveness studies of MASAF and other agencies, but it discovered serious problems with the data and methodologies that prevented any systematic conclusions (Centre et al. 2002, 23). It is unclear whether this topic will be revisited in the second stage of the study.

## Annex 3: Evaluation funding arrangements

Here follow some examples of how much evaluations cost, what sources of funding are available to finance them, and inherent advantages and disadvantages of various funding options, as reported by Bank staff members.

**A. Government finances through project loan:** Particularly for large-scale evaluations that require significant primary data collection, and hence large budgets, funding from the government through project loans typically covers a significant part, if not all, of the evaluation budget. This option gives control over the evaluation contracting process to governments. On the positive side, this may increase government buy-in and ownership of the evaluation process, making it more likely that the research conducted will be relevant to government policy needs and concerns. Conversely, using entirely external funding with Bank staff running the show poses the risk that the government does not buy into or care about the evaluation results, precluding any programmatic or policy impacts. From an operational perspective, 2% of a loan seems like a small price to pay to explore whether the remaining 98% is working well and meeting project objectives. In addition, since governments must meet multiple conditions and manage many contracts with each project, it is not immediately clear why impact evaluation should be treated differently. Yet on the negative side, there are concerns that the government will not hire the highest quality consultants, that the Bank loses direct control over the survey process, and as a result, data quality may suffer. Similarly, there may be little in-country demand for sophisticated methodologies, which are viewed as expensive and not substantially more useful for operations than “quick and dirty” assessment techniques. Lastly, many projects simply do not include a realistic monitoring and evaluation budget from the start, which also makes it difficult to organize sufficient funding later.

**B. Project Budgets:** can cover the time of Bank staff to prepare terms of reference and consult with clients, as well as related travel expenses.

**C. Analytic and Advisory Services (AAA):** This can also be used to help fund Bank staff time to provide technical analytical and advisory services, as the name implies.

**D. Research Support Budget (RSB):** Like other Bank budget sources, RSB can cover the time and travel of consultants or Bank staff, often researchers from DEC, to do technical analytic work, such as data analysis. While flexible and advantageous, RSB is usually not enough to cover the full cost of the evaluation, is competitive and thus difficult to get, and involves high transaction costs.

**E. Trust Funds:** are “financial and administrative agreements between the Bank and external donors, under which the donor entrusts funds to the Bank to finance a specific development-related activity.”<sup>122</sup> Trust funds are useful because some of the donors are also interested in impact evaluations. However, they come with restrictions and involve high transaction costs.

### Examples

**Indonesia KDP** The entire monitoring and evaluation (M&E) system for KDP is financed by the government using part of the project loan. The total M&E budget included \$2.0 million (0.73% of total

---

<sup>122</sup> <http://tf.worldbank.org/>

\$273 million project loan) for KDP1, \$2.9 million (0.69% of total \$422 million project loan) for KDP2, and \$3.1 million (0.83% of total \$375 million project loan) for KDP3. (It is worth noting that the total M&E budget for KDP is not broken down into components in the Project Appraisal Documents, so these figures are not strictly comparable to those for other projects below, where we list only the impact evaluation budget.)

**Brazil PCPR** The \$500,000 evaluation budget is entirely funded by the three participating state governments, using money from World Bank loans.

**Burkina Faso & Senegal RPOs** The sister evaluations of rural producers organizations in Burkina Faso and Senegal are using four different sources of resource contributions for the first round of data collection: (1) \$165,000 from the Trust Fund for Environmentally and Socially Sustainable Development (TFESSD) for universities' operating expenses across both Burkina Faso and Senegal project evaluations; (2) \$70,000 from each project's budget for national survey expenditures; (3) World Bank administrative budget for staff time and travel; (4) universities at Berkeley (California) and Clermont-Ferrand (France) for professors' time.

**Pakistan NRSP** The DEC NRSP evaluation totals \$274,000, drawn from Pakistan Poverty Alleviation Fund (\$200,000); DEC Research Support Grant (\$43,000); and GENFUND (\$31,000) (approximate figures).

**Benin AGeFIB** The total cost of the AGeFIB impact evaluation was \$225,000, funded entirely from a Japan Policy and Human Resources Development Fund (PHRD) grant. The earlier gender study cost \$48,000, funded by a Norwegian trust fund.

**Philippines KALAH-CIDSS** The baseline survey for the KALAH project totaled \$99,500 and was funded entirely by the Trust Fund for Environmentally and Socially Sustainable Development. The costs break down as follows: survey instrument design and field testing (\$27,000); data collection for 2400 households (\$55,000); data input, cleaning, and report writing (\$17,500).

**Nepal RWSS** Two trust funds support this study. The Trust Fund for Environmentally and Socially Sustainable Development (TFESSD) has contributed \$150,000 and GENFUND has contributed \$50,000. An additional \$20,000 study was co-funded by the Social Development and Gender anchors. The project team is currently exploring options for funding a second round of data collection to construct a panel data set for difference-in-difference estimation.

Sources: Lynn Bennett, John Elder, Camilla Holmemo, Ghazala Mansuri, Vijayendra Rao, Laura Rawlings, Claudia Romano, Susan Wong

## Annex 4: Tallies

	✓: included in evaluation design --: information not available	Experimental	Quasi- experimental	Mixed methods	Baseline/ panel data
<b>Completed</b>					
Armenia Social Fund (Social Funds 2000)			✓		
Benin Social Fund gender study					
Bolivia Social Fund (Social Funds 2000)	✓	✓	✓		✓
El Salvador EDUCO			✓		
Honduras Social Fund (FHIS) (Social Funds 2000)			✓	✓	
India Dairy: Operation Flood				✓	✓
Jamaica Social Investment Fund (JSIF)			✓	✓	
Nicaragua Social Fund (FISE) (Social Fund 2000)			✓	✓	
Pakistan Aga Khan Rural Support Project – Khwaja			✓		
Pakistan AKRSP – OED				✓	
Panama Rural Poverty and Natural Resources Project				✓	
Peru Social Fund (FONCODES)			✓		✓
Zambia Social Fund (ZAMSIF) (Social Funds 2000)			✓		
<b>Ongoing</b>					
Albania Second Community Works Project			✓		
Armenia Second Social Fund			✓	✓	
Benin Social Fund			✓		
Brazil Rural Poverty Reduction Program (PCPR)			✓	✓	✓
Burkina Faso National Agricultural Services Dev't Project			✓	✓	✓
Cambodia Rural Investment and Local Governance			✓	✓	✓
India Andhra Pradesh DPIP – DEC	✓	✓	✓		✓
Indonesia Kecamatan Development Project (KDP)			✓	✓	✓
Indonesia KDP & Conflict study			✓	✓	
Indonesia Second Urban Poverty Project (UPP2)			✓	✓	✓
Malawi Social Action Fund (MASAF)			--	✓	
Moldova Social Investment Fund			✓	✓	✓
Pakistan National Rural Support Program – DEC	✓				✓
Philippines KALAHI-CIDSS			✓		✓
Senegal National Rural Infrastructure Project (PNIR)			✓		✓
Senegal Ag. Services & Producers Orgs Proj (PSAOP)			✓	✓	✓
St. Lucia Poverty Reduction Fund (PRF)			✓	✓	
Yemen Social Fund for Development (SFD)			✓	✓	✓
<b>Planned</b>					
Comoros Social Fund (FADC)	--	--	--	--	--
Ecuador Social Fund – DEC targeting study	--	--	--	--	--
Ethiopia Women's Development Initiative (WDI)	✓			✓	✓
Indonesia Support to Conflict Ridden Areas Project			✓	✓	✓
Madagascar Social Fund (FID4)	--	--	--	--	--
Mongolia Sustainable Livelihoods Project	--	--	--	--	--
Nepal Rural Water Supply & Sanitation (RWSS)			✓	✓	✓
Nicaragua Social Fund – DEC targeting study	--	--	--	--	--
Nicaragua Social Fund (FISE 3)	--	--	--	--	--
Philippines ARMM Social Fund (ASFP)	--	--	--	--	--
Thailand Social Fund			✓	✓	
Vietnam Northern Mountains Poverty Reduction	--	--	--	--	--

## **Annex 5: Interview List**

Interviews took place between June and November 2003.

- Mark Austin. 29 August 2003. Panama Rural Poverty and Natural Resources Project & Nicaragua Protierra.
- Lynn Bennett. 17 July 2003. Nepal Rural Water Supply and Sanitation Project.
- Gabriela Boyer. 29 August 2003. Panama Rural Poverty and Natural Resources Project & Nicaragua Protierra.
- Robert S. Chase. Multiple. St. Lucia Poverty Reduction Fund & Thailand Social Fund.
- Marie-Hélène Collion. 31 July 2003. Burkina Faso PNDSA2 & Senegal PSAOP.
- Klaus Deininger. 10 October 2003. India Andhra Pradesh District Poverty Initiative Project & Uganda Local Government Development Project.
- John Elder. 8 July 2003. Benin AGeFIB.
- Hernando Garzón. 29 August 2003. Panama Rural Poverty and Natural Resources Project & Nicaragua Protierra.
- Camilla Holmemo. Multiple. Philippines KALAHI.
- Ghazala Mansuri. 20 August 2003. Pakistan Rural Support Program & DEC CDD evaluation initiative.
- Daniel Murphy. Multiple. São Tomé Social Sector Support Project.
- Lance Morrell. 1 July 2003. Uganda Local Government Development Project.
- Rikke Nording. Multiple. Armenia Social Investment Fund.
- Daniel Owen. 3 July 2003. Angola Social Action Fund, Sierra Leone National Social Action Project & São Tomé Social Sector Support Project.
- Hope Phillips. 15 July 2003. Malawi Social Action Fund & Tanzania Social Action Fund.
- Vijayendra Rao. 20 August 2003. Indonesia Urban Poverty Project 2.
- Laura Rawlings. 30 July 2003. Nicaragua Social Investment Fund 3.
- Pierre Rondot. 31 July 2003. Burkina Faso PNDSA2 & Senegal PSAOP.
- Gloria Rubio. 30 July 2003. Nicaragua Social Investment Fund 3.
- Adama Touré. 29 August 2003. Senegal National Rural Infrastructure Project.
- Andrea Vermehren. 30 July 2003. Nicaragua Social Investment Fund 3.
- Ian Walker. 6 August 2003. Yemen Social Fund, St. Lucia Poverty Reduction Fund, Honduras Social Investment Fund 2, Egypt Social Fund.
- Wendy Walker. 2 July 2003. Benin AGeFIB.
- Howard White. 21 August 2003. OED Social Funds study.
- Susan Wong. 2 July 2003. Indonesia KDP1, KDP2, KDP3 & SCRAP.
- Michael Woolcock. 29 July 2003. Indonesia KDP conflict study.
- Ye Zhang. 27 June 2003. Albania Development Fund.

## References

Many of the references listed here are informal working documents not publicly available. Where possible, we have listed the persons who gave them to us, and their World Bank organizational code. Throughout the text, we have used in-text parenthetical citation for formal or finished documents, and footnotes for informal working drafts and personal communications.

- Alatas, Vivi. 2003. Email, 29 October 2003. Indonesia
- Angrist, Joshua D. and Alan B. Krueger. 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." NBER Working Paper 8456. Cambridge, MA: National Bureau of Economic Research.
- APPC (Asia-Pacific Policy Center). 2003a. "Living Condition Survey: Interview Schedule for Family." [Emailed from Camilla Holmemo, SDV, 27 October 2003] Philippines
- APPC (Asia-Pacific Policy Center). 2003b. "Living Condition Survey: Interview Schedule for Barangay Official." [Emailed from Camilla Holmemo, SDV, 27 October 2003] Philippines
- Arcand, Jean-Louis. 2002. "Producer Organizations in Burkina Faso and Senegal : A Synthesis of Case Studies in Twenty Villages." September 2002. [Emailed from Marie-Hélène Collion, MNSRE, 25 July 2003] Burkina Faso, Senegal
- Baker, Judy. 2002. *Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners*. Washington, DC: The World Bank.
- Barron, Patrick, Rachael Diprose, Yatrin Kaniu, Claire Q. Smith, & Michael Woolcock. 2003a. "Field Research Guide and Training Manuals: KDP & Community Conflict Negotiation Study," Jakarta: World Bank. [available from the authors] Indonesia
- Barron, Patrick, Rachael Diprose, David Madden, Claire Q. Smith, & Michael Woolcock. 2003b. "Do Participatory Development Projects Help Villagers Manage Local Conflicts? A Mixed Methods Approach to Assessing the Kecamatan Development Project, Indonesia." CPR Working Paper No. 9, September 2003, Washington, DC: Social Development Department, The World Bank. Indonesia
- Bennett, Lynn. 2003a. "Proposed Consultancy to Review and Redesign the Women's Technical Support Services WTSS Component of RWSS II." Email attachment, 15 July 2003. Nepal
- Bennett, Lynn. 2003b. Interview, 17 July 2003. Nepal
- Bezghanyan, Anush. 2003. "Terms of Reference: Impact Evaluation of Moldova Social Investment Fund Project." Email attachment, 9 July 2003. Moldova

- Bhatia, Mukhmeet. 2003. "Social Funds: A Review of Public Sector Management and Institutional Issues." June 2003. Washington, DC: Social Protection Team, Human Development Network, World Bank.
- CECI-Nepal (Canadian Centre for International Studies and Cooperation). 2003. "Technical Proposal Presented to the World Bank Nepal for the Project entitled: Measuring and Learning About Empowerment and Social Inclusion." [Emailed from Lynn Bennett, SDV, 15 July 2003.] Nepal
- Candler, Wilfred, & Nalini Kumar. 1998. *India: The Dairy Revolution*. Washington, DC: Operations Evaluation Department, World Bank. India
- Carletto, Calogero, Robert S. Chase, & Ye Zhang. 2003. "Detecting the Performance of Community Driven Development Program: Evidence from Albanian Development Fund." First Draft, July 2003. Washington, DC: World Bank. Albania
- Centre for Social Research, Malawi; Institute for Development Studies, UK; & Institute for Urban and Regional Research, Norway. 2002. "Poverty Alleviation and Sustainability Review of the Malawi Social Action Fund (1996-2001)." [from Hope Phillips, AFTH1] Malawi
- Chase, Robert S. 2003. "BTOR: Thai Social Capital Evaluation." Email, 17 October 2003. Thailand
- Chase, Robert S. 2002. "Supporting Communities in Transition: The Impact of the Armenian Social Investment Fund." *World Bank Economic Review* 16(2): 219-240. Armenia
- Chase, Robert S., and Lynne Sherburne-Benz. 2000. "Zambia Social Recovery Project: Evaluating the Household Impact." Presentation. [Available from the authors.] Zambia
- Chase, Robert S., & Lynn Sherburne-Benz. 2001. "Household Effects of African Community Initiatives: Evaluating the Impact of the Zambia Social Fund." Washington, DC: World Bank. [available from authors] Zambia
- Christiaensen, Luc, & Arianna Legovini. 2003. "Reflections on Evaluating Ethiopia's Women Development's Initiative." PowerPoint presentation at the Advanced Impact Evaluation Workshop for Africa and MENA Region, Hilton, Addis Ababa, Ethiopia, June 9-12, 2003. (available from the authors) Ethiopia
- Coulombe, Harold. 2002. «Evaluation Socio-Economique du Fond Social Béninois: Proposition Méthodologique.» 30 July 2002. [from John Elder, AFTH3] Benin
- Deininger, Klaus. 2003. Email attachments 10 October 2003. "Mid Term Appraisal of District Poverty Initiatives Project (DPIP) of Andhra Pradesh Household Questionnaire Respondent: Female." "Mid Term Appraisal of District Poverty

Initiatives Project (DPIP) of Andhra Pradesh Household Questionnaire Respondent: Male.” “APRPRP Baseline / APDPIP Mid-term evaluation Main Village Questionnaire.” “ APRPRP Baseline / APDPIP Mid-term evaluation SHG Questionnaire.” “APRPRP Baseline / APDPIP Mid-term evaluation VO/MS Questionnaire.” “Andhra Pradesh District Poverty Initiatives Project (APRPR) Monitoring & Impact Assessment and Land Component – Supervision Mission Aide Memoire” [July 2003 mission]. “Andhra Pradesh District Poverty Initiatives Project (APRPR) Land Component – Supervision Mission DRAFT Aide Memoire” [February 2003 mission].

- de Janvry, Alain, Elisabeth Sadoulet, Marie-Hélène Collion, & Pierre Rondot. 2001. “Impact of producer organization capacity-building component in agricultural services projects” (BTOR June 2001). Burkina Faso, Senegal
- de Janvry, Alain, & Elisabeth Sadoulet. 2002. “Some Hypotheses for the Analysis of Producers Organizations in Senegal and Burkina-Faso and Strategies for Empirical Testing.” May 2002. Burkina Faso, Senegal
- Dent, Geoffrey. 2001. “Ex-Post Evaluation of Kecamatan Development Program (KDP1) Infrastructure Projects: Final Report.” November 2001. [from Susan Wong, EASES] Indonesia
- Dongier, Philippe, Julie Van Domelen, Elinor Ostrom, Andrea Rizvi, Wendy Wakeman, Anthony Bebbington, Sabina Alkire, Talib Esmail, and Margaret Polski. 2002. “Chapter 9: Community-Driven Development.” in Poverty Reduction Strategy Sourcebook. Washington, DC: World Bank. [URL: <http://www.worldbank.org/poverty/strategies/sourctoc.htm>, accessed 6 November 2003]
- Duflo, Esther, & Michael Kremer. 2003. “Use of Randomization in the Evaluation of Development Effectiveness.” [presented at Fifth Biennial World Bank Conference on Evaluation and Development] Washington, DC.
- Elder, John. 2003. Interview, 8 July 2003. Benin
- ESA Consultores. 2002. “Yemen Social Fund for Development Impact Evaluation, Phase 2 Report: Findings from the 1999 National Poverty Survey.” [Ian Walker, ESA Consultores, email 5 August 2003] Yemen
- ESA Consultores. 2003a. “Consultancy to provide support on Impact Evaluation for Yemen Social Fund for Development, Report #2 Detailed Design of 2002/2003 Impact evaluation survey.” [Ian Walker, ESA Consultores, email 6 August 2003] Yemen
- ESA Consultores. 2003b. “Yemen Social Development Fund Impact Monitoring and Evaluation Manual.” [Ian Walker, ESA Consultores, email 6 August 2003] Yemen
- ESA Consultores. 2003c. “Consultancy for Impact Evaluation of St Lucia Poverty St. Lucia

Reduction Fund. Report #1: Methodology Report.” September 2003. [from Rob Chase, SDV]

- ESA Consultores. 2003d. “Yemen Social Fund for Development: Impact Evaluation Study: Draft Final Report.” September 2003. [Ian Walker, ESA Consultores, email attachment 30 October 2003] Yemen
- “Evaluación del Impacto del Fondo de Inversión Social de Emergencia en Nicaragua Request for Expressions of Interest.” 2003. [Online posting on UN Development Business (UNDB) 25 August 2003]. Nicaragua
- Ezemenari, Kene, Janet Owens & Gloria Rubio. 2000. “Progress on Impact Evaluation Plans in Bank Projects: Comparison of Fiscal Years 1998-2000.” December 2000. Washington, DC: PRMPO, World Bank.
- FECAMP. 2001. “Proposta Técnica apresentada pela Fundação Economia de Campinas – FECAMP: EDITAL No. 012/01: Estudos de Impacto dos Novos Projetos de Combate à Pobreza Rural dos Estados da Bahia, Ceará, Pernambuco e Piauí.” 31 July 2001. [Mailed from Claudia Romano, LCSER] Brazil
- GOI (Government of Indonesia). 2002. “Kecamatan Development Program 1998-2002: Final Report. June 2002.” Jakarta: Ministry of Home Affairs, Directorate General of Community and Village Empowerment & KDP National Secretariat and National Management Consultants. [from Susan Wong, EASES] Indonesia
- Grossman, Jean Baldwin. 1994. “Evaluating Social Policies: Principles and U.S. Experience.” *World Bank Research Observer* 9(2): 159-180.
- Helmets, Kent, & Pia Wallgren. 2003. “Draft 3 Inception Report: Baseline Survey for Commune Council Investments.” 4 April 2003, Washington, DC: World Bank. [from Louise Scura, EASRD] Cambodia
- Hentschel, Jesko. 1999. “Contextuality and Data Collection Methods: A Framework and Application to Health Service Utilisation.” *Journal of Development Studies* 35(4): 64-94.
- Holloh, Detlev. 2001. “Review of the KDP Microcredit Approach.” September 2001. [from Susan Wong, EASES] Indonesia
- Isham, Jonathan, Deepa Narayan, & Lant Pritchett. 1995. “Does Participation Improve Performance? Establishing Causality with Subjective Data.” *World Bank Economic Review* 9(2): 175-200.
- Jimenez, Emmanuel, and Yasuyuki Sawada. 1999. “Do Community-Managed Schools Work? An Evaluation of El Salvador’s EDUCO Program.” Working Paper Series on Impact Evaluation of Education Reforms, Paper No. 8, February 1998. Washington, DC: Development Research Group, World Bank. El Salvador

- Jimenez, Emmanuel, and Yasuyuki Sawada. 1999. "Do Community-Managed Schools Work? An Evaluation of El Salvador's EDUCO Program." *World Bank Economic Review* 13(3): 415-441. El Salvador
- Khwaja, Asim Ijaz. 2002. "Can good projects succeed in bad communities? Collective Action in Public Good Provision." Cambridge, MA: Kennedy School of Government, Harvard University. [URL: <http://ksghome.harvard.edu/~.akhwaja.academic.ksg/papers/BaltJune02.pdf>, accessed 5 November 2003.] Pakistan
- LIFAD. 2001. « Evaluation d'impacts des approches et actions de l'AGEFIB. » Laboratoire d'Ingénierie de Formation de d'Assistance en Développement Local. [from John Elder, AFTH3] Benin
- Lindauer, David L. and Lant. Pritchett. 2002. "What's the Big Idea? The Third Generation of Policies for Economic Growth." [June 2002 draft for article in *Economia* 3(1), from <http://ksghome.harvard.edu/~.lpritch.academic.ksg/research.htm>, accessed 9 December 2003].
- Mansuri, Ghazala. 2003a. Interview, 20 August 2003. Pakistan
- Mansuri, Ghazala. 2003b. "Research Program on Community Based Rural Development (CBRD)," and the response to subsequent comments, "Community Driven Development Component." Email attachments, 20 August 2003. Pakistan, Nicaragua, Ecuador, India, Indonesia, Burkina Faso, Senegal
- Mansuri, Ghazala, & Vijayendra Rao. 2003. "Evaluating Community-Based and Community-Driven Development: A Critical Review of the Evidence." September 2003. Washington, DC: Development Research Group, World Bank.
- Marsland, N., Wilson, I., Abeyasekera, S., & Kleih, U. 2000. "A Methodological Framework for Combining Quantitative and Qualitative Survey Methods." [URL: <http://www.rdg.ac.uk/ssc/develop/dfid/booklets/qqa.html>, accessed 20 October 2003]
- Mascarell, Caroline. 2003. "Social Investment Fund II Project: Social Capital Assessment." Concept Note, May 2003. [from Rob Chase, SDV] Armenia
- Newman, John, Menno Pradhan, Laura Rawlings, Geert Ridder, Ramiro Coa & José Luís Evia. 2002. "An Impact Evaluation of Education, Health, and Water Supply Investments by the Bolivian Social Investment Fund." *World Bank Economic Review* 16(2): 241-274. Bolivia
- Newman, John, Laura Rawlings, & Paul Gertler. 1994. "Using Randomized Control Designs in Evaluating Social Sector Programs in Developing

Countries.” *World Bank Research Observer* 9(2): 182-201.

- Nording, Rikke. 2003a. Interview, 1 July 2003. Armenia
- Nording, Rikke. 2003b. “Social capital impact assessment 2003—Armenia Social Investments Funds.” July 2003. (available from the author) Armenia
- Owen, Daniel & Julie Van Domelen. 1998. “Getting an Earful: A Review of Beneficiary Assessments of Social Funds.” Washington, DC: World Bank Human Development Network, Social Protection Team.
- Ozler, Berk. 2003. “Research Program on Community Based Rural Development.” Email attachment, 5 September 2003. Nicaragua, Ecuador
- “Panama Cost-Effectiveness and Sustainability of Small Rural Investments in Panama Synthesis Report.” [Draft Report from Mark A. Austin, LCSER, Email 5 August 2003] Panama
- Parker, Andrew, & Rodrigo Serrano. 2000. “Promoting Good Governance through Social Funds and Decentralization.” Social Protection Discussion Paper No. 0022, September 2000. Washington, DC: Social Protection Unit, Human Development Network, World Bank.
- Paxson, Christina, & Norbert R. Schady. 2002. “The Allocation and Impact of Social Funds: Spending on School Infrastructure in Peru.” *World Bank Economic Review* 16(2): 297-319. Peru
- PCPR. 2001. “Carta Convite: Avaliação do Programa de Combate à Pobreza Rural (PCPR) no Nordeste do Brasil.” 21 January 2001. [from Claudia Romano, LCSER] Brazil
- Pradhan, Menno. 2003. “Terms of Reference for Evaluation Consultant.” “Instruments for the evaluation of World Bank financed community development projects in Indonesia.” (V. Alatas & M. Pradhan). “UPP-2 Evaluation Lurah and PKK (Family Welfare Movement) Chairperson Questionnaires.” “UPP-2 Evaluation Village Profile/Monograph Questionnaires (Map, Village Monograph and Profile).” “UPP-2 Evaluation Household Questionnaires.” “UPP-2 Evaluation Main Respondent.” “UPP-2 Evaluation Interview with Village Activist.” “Evaluasi P2KP-2 Dusun/RW Activist.” Email attachments, 20 August 2003. Indonesia
- Pradhan, Menno, Laura Rawlings, & Geert Ridder. 1998. “The Bolivian Social Investment Fund: An Analysis of Baseline Data for Impact Evaluation.” *World Bank Economic Review* 12(3): 457-482. Bolivia
- Pritchett, Lant. 2002. “It Pays to be Ignorant: A Simple Political Economy of Rigorous Program Evaluation.” Cambridge, MA: Harvard University Kennedy School of Government, April 29, 2002. [URL: [http://ksghome.harvard.edu/~lpritch.academic.ksg/ignorance\\_v2\\_r1.pdf](http://ksghome.harvard.edu/~lpritch.academic.ksg/ignorance_v2_r1.pdf),

accessed 3 October 2003]

- Pritchett, Lant and Michael Woolcock. 2002. "Solutions when *the* Solution is the Problem: Arraying the Disarray in Development." [20 March 2002 Draft from <http://www.st-edmunds.cam.ac.uk/vhi/research/pritchet.pdf>, accessed 8 December 2003].
- Rao, Vijayendra, & Michael Woolcock. 2003. "Integrating Qualitative and Quantitative Approaches in Program Evaluation," in Francois Bourgingnon and Luiz Pereira Da Silva (edited) *Tool Kit for Evaluating the Poverty and Distributional Impact of Economic Policies*, World Bank and Oxford University Press. [URL: <http://www.cultureandpublicaction.org/bijupdf/ch08.pdf>, accessed 20 October 2003]
- Rao, Vijayendra, & Ana Maria Ibáñez. 2003. "The Social Impact of Social Funds in Jamaica: A Mixed-Methods Analysis of Participation, Targeting, and Collective Action in Community-Driven Development" World Bank Policy Research Working Paper No. 2970, February 2003. Jamaica
- Ravallion, Martin. 1999. "The Mystery of the Vanishing Benefits: Ms. Speedy Analyst's Introduction to Evaluation." World Bank Working Paper No. 2153, 1 July 1999.
- Ravallion, Martin. 2003. "Randomized Trials of Development Policies and Projects: Some Comments." Unpublished conference paper, World Bank.
- Rawkins, P. M. 2003. "Armenia Social Investment Fund II Project: Institutional Assessment." 19 June 2003. [from Rikke Nording, SDV] Armenia
- Rawlings, Laura. 2003. Email, 25 September 2003. St. Lucia
- Rawlings, Laura, Lynne Sherburne-Benz, & Julie Van Domelen. 2004. *Evaluating Social Funds A Cross-Country Analysis of Community Investments*. Washington, DC: The World Bank. Armenia, Bolivia, Honduras, Nicaragua, Peru, Zambia
- Rosenbaum, Paul R. & Donald B. Rubin. 1985. "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score." *The American Statistician* 39: 33-38.
- Salmen, Lawrence. 1999. "Beneficiary Assessment Manual for Social Funds." Social Protection Discussion Paper Series. Washington, DC: World Bank.
- Santucci, Fabio M. 2003. "Social capital formation: the case of the Rural Poverty and Natural Resources Project in Panama." Very First Draft, June 2003. [from Mark A. Austin, LCSER, Email attachment, 5 August 2003] Panama
- Sharp, Maryanne. 2003. Email, 30 September 2003. Comoros

- Tevoedjre, M. 2001. « Rapport d’audit sur la Fonctionnalité des Infrastructures socio-communautaires. » [from John Elder, AFTH3] Benin
- Touré, Adama. 2003a. “Impact du PNIR sur la réduction de la pauvreté en milieu rural: Proposition de recherche.” “Evaluation d’impact du PNIR: Mission de mai 2003. Pour mémoire.” Email attachments, 17 June 2003. Senegal
- Touré, Adama. 2003a. Telephone interview, 29 August 2003. Senegal
- UOI (University of Indonesia). 2002. “Evaluating of Kecamatan Development Program (KDP) Impacts on Community Organization and Household Welfare (the Second Round Survey): Impact Assessment Study: Final Report.” July 2002. Jakarta: Demographic Institute, Faculty of Economics, University of Indonesia. [from Susan Wong, EASES] Indonesia
- Walker, Ian, Rafael del Cid, Fidel Ordoñez, & Florencia Rodríguez. 1999. “Ex-Post Evaluation of the Honduran Social Investment Fund (FHIS2).” July 1999. Tegucigalpa: ESA Consultores. Honduras
- Walker, Wendy M. 2002. “Qualitative Evaluation of Gender Impacts in the Benin Social Fund.” February 2002. [available from the author] Benin
- Walker, Wendy M. 2003. Interview, 3 July 2003. Benin
- Wietzke, Frank-Borge. 2003. Email communication, 11 September 2003. Madagascar
- Woolcock, Michael. 2001. “Social Assessments and Program Evaluation with Limited Formal Data: Thinking Quantitatively, Acting Qualitatively.” Social Development Briefing Note No. 68. Washington, DC: World Bank. St. Lucia
- Wong, Susan. 2003. “Indonesia Kecamatan Development Program: Building a Monitoring and Evaluation System for a Large-Scale Community-Driven Development Program.” Environment and Social Development Unit East Asia and Pacific Region (EASES) Discussion Paper, May 2003. Indonesia
- World Bank. 1997-1998?. “PHRD Funding Proposal: Implementation Grant. John Elder, Task Team Leader. Benin Social Fund (P035645), Implementing Agency: Agence de Financement des Initiatives de Base (AGeFIB).” Benin
- World Bank. 2000. “Nicaragua: Ex-Post Impact Evaluation of the Emergency Social Investment Fund (FISE).” Report No. 20400-NI, Human Development Sector Management Unit, Latin America and the Caribbean Region, World Bank: Washington, DC. Nicaragua
- World Bank. 2001. *Improving the Monitoring and Evaluation System for KDP2*. July 2001. [from Susan Wong, EASES] Indonesia
- World Bank. 2002a. *The Next Ascent: An Evaluation of the Aga Khan Rural Support Program, Pakistan*. Washington, DC: Operations Evaluation Pakistan

Department (OED), World Bank.

World Bank. 2002b. *Social Funds Assessing Effectiveness*. Washington, DC: Operations Evaluation Department (OED), World Bank.

World Bank. 2003. "Benin: Poverty Assessment DRAFT." 30 September 2003. Benin  
[from Maurizia Tovo, AFTH2]

Zhang, Ye. 2003. Interview, 27 June 2003. Albania

Zhu, Shu-Hong. 1999. "A Method to Obtain a Randomized Control Group Where It Seems Impossible: A Case Study in Program Evaluation." *Evaluation Review* 23(4): 363-377.