

Is Empowerment Efficient?

A Data Envelopment Analysis of 260 Local Associations in Bangladesh



Aldo Benini

January 2009

Is Empowerment Efficient?

A Data Envelopment Analysis of 260 Local Associations in Bangladesh

Aldo Benini

RDRS Bangladesh

January 2009

Cover photo by RDRS:

Heavy symbolism is on display in this meeting of the executive committee, Khangaon Union Federation, one of 260 associations of poor people in northwestern Bangladesh, whose relative efficiency this study analyzes.

Seven out of the nine committee members in Khangaon are women, a proportion almost as high as in the general membership. Among the member groups active in regular meetings and savings, not one was a men's group in this federation in 2005. Yet the dependency on male expertise and leadership is visible in this image, with the chairman operating a calculator and explaining the accounts. It is typical of many federations in which women form the vast majority of the active members. Active women members and profitable businesses are intermediary steps to fighting domestic violence and aiding members threatened by destitution.

© Aldo Benini and RDRS Bangladesh 2009

Suggested citation:

Benini, Aldo (2009). *Is Empowerment Efficient? A Data Envelopment Analysis of 260 Local Associations in Bangladesh*. Dhaka and Rangpur, RDRS Bangladesh and North Bengal Institute.

Author e-mail address:

abenini@starpower.net



RDRS Bangladesh

Rangpur Dinajpur Rural Service

In association with LWF/World Service, Geneva

House 43, Road 10, Sector 6, Uttara

Dhaka-1230, Bangladesh

Telephone +880 2 8954384-86 Fax +880 2 8954391

E-mail: rdrs@bangla.net, Website: <http://www.rdrsbangla.net>

Published by



RDRS Bangladesh, Jail Road, Engineer Para

Rangpur, Bangladesh

Table of Contents

Preface.....	6
Executive Summary	7
Acknowledgement	9
Introduction.....	10
[Sidebar:] Data Envelopment Analysis and social justice	13
Organization of the paper.....	15
Theoretical model	16
[Sidebar:] An intuitive introduction to DEA	18
DEA applications in microfinance.....	20
The supporting NGO and the 260 local associations.....	22
RDRS Bangladesh	22
The federations.....	23
Structure	23
Relevance.....	26
Effectiveness and efficiency	26
A model of federation efficiency	27
Institutional factors	29
Environmental factors.....	29
Quality of the NGO support.....	30
Results.....	30
The distribution of efficiency.....	30
Efficiency and effectiveness	32
Institutional and environmental determinants.....	34
Conclusions.....	36
Empowerment through second-best institutions?	36
DEA for multi-objective NGOs	38
Appendix.....	39
DEA considerations	39
Data sources and descriptive statistics.....	43
Results of the second-stage regressions	45
Auxiliary model of dispute caseloads.....	46
Bibliography	47
About the author	50

Tables and Figures

Table 1: Effects of institutional and environmental variables	34
Table 2: Descriptive statistics [several sub-tables]	43
Table 3: Second-stage regression results	45
Table 4: Dispute caseload model - descriptive statistics and regression results.....	46
Figure 1: Stylized input-output transformation.....	17
Figure 2: The efficiency frontier in a simple model	18
Figure 3: Technical and scale efficiency.....	19
Figure 4: Overview map of the Federation working area in 2005	24
Figure 5: Basic structure of a Union Federation.....	25
Figure 6: A model of RDRS Federation efficiency	28
Figure 7: Distribution of efficiency scores	31
Figure 8: Efficiency in response to membership	31
Figure 9: Geographical distribution of relevance, effectiveness and efficiency scores.....	33
Figure 10: Extent of bias correction using bootstrapping	42

Preface

“Efficiency” is routinely listed as one of the standards of program evaluations, alongside relevance, effectiveness, impact and sustainability. In actual practice, accurate estimates of efficient resource use have proven difficult. The diverse services that development NGOs produce cannot be simply added up in one figure. Efficiency is all the more evasive a concept when we look at empowerment programs. These emphasize process, capacity and ownership, manifest in a great variety of individual and group actions. Not surprisingly, credible measures of “efficient empowerment” have been offered primarily in micro-finance, with its well defined inputs and outputs, most of which are simply monetary.

RDRS Bangladesh pursues efficiency as a value in its day-to-day management in the way of diligent and responsible use of our budgets and other resources. The same standard applies to those of our programs that we may call “empowerment programs”. As many readers know, the support of federations of poor people is our flagship program in this area. These associations – there were 260 when the data was collected; there are over 350 by now - are fairly homogeneous in their basic setup, and the RDRS support program has many standardized components.

This has resulted in a fair degree of comparability, which this study exploits for a more precise estimate of how efficient the federations are. As such it is an extension of the assessments that for many years RDRS and, increasingly, the federations themselves have conducted of their relative strength. How much do these supposedly empowering organizations achieve (effectiveness), and at what price (efficiency)?

Two points stand out. First, effectiveness and efficiency go less closely hand in hand than one might expect. There are many federations that have not been noted as high performers, but which have put their modest resources to good use. In other words, many are moderately effective, and yet work efficiently. Empowerment may thus come in two flavors – as the demonstrated ability to achieve objectives, and as the best use of what little resources that the poor control through their organizations.

Second, this study breaks new methodological ground. It uses a method that overcomes the problem of “counting apples and oranges”. Admittedly, not all NGO monitoring data will be good enough to support efficiency estimates. But for well defined contexts such as the RDRS federations, the study sets an example of how several inputs and outputs, some without monetary prices, can be meaningfully related to each other. Since empowerment programs yield benefits, but also consume resources, such efficiency estimates are a necessary and helpful complement to our understanding of their successes and limitations.

Kamaluddin Akbar
Executive Director
RDRS Bangladesh

Dhaka, 3 January 2009

Executive Summary

Empowerment, a concept with a successful twenty-century cultural career, has been recognized for its relevance and, increasingly, effectiveness in liberating the poor, both at the individual and local community level. Efforts to create valid measurement tools have advanced, with a focus on causality, thus on effectiveness of empowerment programs. The efficiency of such programs, in other words considerations of optimal resource use, has not been investigated widely, with the exception of microfinance projects. Such programs are sheltered from efficiency pressures by the subsidies of aid chains and by the need to work out, in precarious social environments, organizational arrangements that produce credible empowerment effects in the first place.

I take advantage of a unique dataset on 260 local associations of poor people in northwestern Bangladesh, founded and fostered by the NGO “RDRS Bangladesh”, with a homogeneous basic structure and contiguous working areas. A capacity building grant from the EU in 2002 helped further revitalize these so-called “Union Federations” or, synonymously, “RDRS Federations”. Using Data Envelopment Analysis, a technique suited to relate diverse inputs and outputs without pre-existing weights or prices, I examine how efficiently the federations transformed membership and financial inputs in 2003 into female activism and business profits in 2005. I compare these efficiency estimates with effectiveness estimates from federation self-assessments and from a NGO-defined index. I also relate them to institutional and environmental factors.

The 260 federations are fairly distinctly separated into a large group of inefficient operators, and a smaller group of efficient ones, with a noticeable mid-field gap. For this I lack a cogent explanation. Congestion in 2003 with inactive members and fluctuations in tree sales may account for some of it. More importantly, efficiency and effectiveness are unrelated or, depending on the effectiveness measure, weakly negatively related. Their predictors have opposite effects. For example, loans from RDRS had effects that are positively related to self-assessed effectiveness. The effect on efficiency is negative. The aid chain shelters effectiveness, but not efficiency, from the impacts of higher ambient poverty. Better NGO support does improve federation efficiency whereas the relationship with effectiveness is less clear.

I seek a coherent interpretation of these various findings with the help of Dani Rodrik's allegory of "second-best institutions". The organizational arrangements that prove effective in poor societies are not always those recommended by context-free best-practice thinking. Moreover, aid chains tolerate vast inefficiencies, in order to keep flows predictable and hedge against various contingencies. While empowerment programs struggle with effectiveness, efficiency takes a back seat. This raises a moral dilemma about how to treat partners who use their modest resources efficiently, but, due to accidents of the aid chain or local environment, are still far from being effective.

This article chiefly addresses NGO managers and monitoring staff, for a greater conceptual awareness of effectiveness and, newly, efficiency. Most technicalities have been relegated to the appendix, but a sidebar offers a first introduction to Data Envelopment Analysis. This technique deserves further exploration as a tool to help mitigate reporting and evaluation challenges typical of multi-objective NGOs, an organizational form common to many empowerment programs.

The study is organized in a way that facilitates selective reading. I give a brief introduction to empowerment and its measurement on pages 10 sqq. A notional model of local associations and their efficiency follows on page 16. RDRS Bangladesh and the basic structure of the federations that it supports are described starting on page 22. Some readers may want to leap to the specific model of federation efficiency (page 27), and hence to the estimated distribution of efficiency scores and their determinants (page 30). The results are interpreted in the concluding section (page 36), followed by a technical appendix.

The technically interested reader may note that I estimated a 3-input 4-output variable-returns-to-scale input-oriented DEA model. To safeguard against outlier influence, I re-estimated it using Wilson's R-based bootstrapping algorithm. It returned a bias-corrected Shepard distance, the reciprocal of a Farrell-type efficiency measure. In a second stage, I regressed the efficiencies on institutional and environmental factors. This model and, with the same predictor set, models of effectiveness are truncated regression models.

Acknowledgement

The data used for this article were collected as part of my research into RDRS Federations between 1997 and 2006. Most of it originated in the federation monitoring system based in Rangpur, Bangladesh. My main support groups at the time were the RDRS Bangladesh Social Organization Unit, led until 2007 by Azizul Karim, and the North Bengal Institute, an RDRS affiliate, led until 2008 by Jasim Uddin Ferdous.

Paul von Büнау, Technical University, Berlin, Germany, provided comments on an earlier draft and statistical advice. Maurizio Pisati, University of Milano Bicocca, Italy, helped with formatting the color-graded maps on page 33. Pia Schneider, The World Bank, Washington DC, helped with literature.

I decisively benefited from the guidance on Data Envelopment Analysis that Hong Son Nghiem, University of Queensland, Brisbane, Australia, offered at various stages of this project.

The cover photo, by RDRS, and part of its caption as well as the federation area map on page 24 were used in my 2006 federation study (Benini 2006). The background section on RDRS and on its federations borrowed from an article on the federation monitoring system (Benini 2007).

The views expressed in this article are entirely my personal, and the individuals and organizations mentioned bear no responsibility for any error of fact or interpretation.

Introduction

Empowerment as a development concept

Empowerment is one of the most victorious concepts in twentieth-century global culture. Bröckling (2007) identifies it as one of the main strategies in the development of modern subjectivity. He assigns to it a place side by side with such mainstays of modernity as “creativity”, “quality”, and “projects”. It has numerous dimensions defying universally shared definition; this protean character allows very diverse actors, purposes and interpretations to find shelter under this umbrella concept. In part, empowerment ideas have been co-opted by corporate actors of whom earlier versions were suspicious or even potentially hostile (Bartunek and Spreitzer 2006).

With considerable cultural lag over developed-nation usage, “empowerment” made a forceful entrée into international development in the late eighties, with the World Bank Development Report 2000-01 being a watershed event in its rapid diffusion (World Bank 2001). The report applied the concept to two themes – making state institutions more responsive to the poor (chapter 6), and removing social barriers and building social institutions (chapter 7) – that have since gained prominence in aid policy and programming, in and beyond the Bank. The overall influence of the concept has kept growing; with 21 million documents using “empowerment” that Google returned in October 2007 (Benini 2008: 32), it has outgrown affiliated literatures such as on “rural development / community development” (14 million hits) and is still far ahead of the more recent “rights-based” writing (1.5 million hits). The World Bank has elevated empowerment on an equal footing with economic growth, enshrined in “Growth and Empowerment” by its chief economist (Stern, Dethier et al. 2005). Adoption in the development NGO world has been no less enthusiastic although the gains in programmatic clarity over earlier semantics are not always self-evident.

The stunning growth of empowerment movements has not been matched by a commensurate growth in firm knowledge. The poverty of what we know about empowerment leaves vast claims to its effectiveness virtually unexamined in terms of rigorous scientific inquiry even when most anecdotal evidence strongly argues for it. The dispositional character¹ of “empowerment” makes it hard to measure although this difficulty is more practical than philosophical; Amartya Sen’s capabilities model of development (Sen 2001), in which the empowered are freed to realize their full functions in society, is itself dispositional.

Measuring empowerment

Only in recent years have coordinated efforts to understand and control measurement challenges in developing-society empowerment gained momentum. Again, the World Bank took the lead (Alsop and Heinsohn 2005: ; Narayan-Parker 2006b). The volume edited by Narayan-Parker is particularly helpful as it assembles contributions that shed light on the measurement challenges

¹ “Dispositional” in the sense of Carnap (1936). In a statistical context, one may think of a “latent variable” that cannot be directly observed.

encountered at four societal levels: individual, household, community and local governance, and nation state.

The authors' main common concern is to advance models of empowerment and methods of measuring it that will make causal attribution safer. This is a much needed focus, given the tautological usages of which some of the empowerment jargon is guilty - vide an ingrained project reporting style that weaves together multiple dispositional terms, with empowerment causing better awareness, and awareness in turn further empowering the poor. To safeguard against the loss of testability, and ultimately credibility, some (Khwaja 2006) want us to make a clear distinction between models in which empowerment is an end in itself, and those that treat it as a means to other aspects of poor people's welfare. This may be a suitable perspective for the econometric model-builder, but is not borne out in the everyday lives of the stakeholders. These do not follow the age-old analytic distinction between expressive and instrumental behavior, but rather meld the two continuously.

While Narayan-Parker is about measurement, and only indirectly about policy, it is safe to assume that more rigorous studies will impact the policy discourse chiefly at the point of effectiveness. In the classic sequence of the major hurdles that innovations need to take, first demonstrating their relevance, then effectiveness, and ultimately efficiency and sustainability, empowerment had already proven its relevance. The wide adoption of the concept in the concerned professional communities was sufficient for that, regardless of the volume and quality of the work on the ground, which, of course, has reinforced the vitality of the former.

Beyond effectiveness

This paper goes beyond the effectiveness perspective. It attempts to determine the relative efficiency of participants in an empowerment program. Efficiency, which the OECD DAC evaluation glossary defines as "a measure of how economically resources/inputs (funds, expertise, time, etc.) are converted to results" (OECD DAC 2002: 21) is understood here as a property of the participants' agency, modified by their social environment. Higher efficiency means that given the resources available, the participants produce stronger empowerment effects. It also implies that criteria can be justified on which different actors for whom resource use and outcomes are known can be compared, ideally through a quantitative metric that validly captures the quality of the resources-to-outcomes conversion. The basic elements are the same as in Narayan-Parker's conceptual framework – opportunity, agency, and outcomes (Narayan-Parker 2006a: 7) – but with the addition of a measure of input-output conversion². Whether inputs and outputs are monetized is of secondary importance although it helps if they are measured at least at interval level.

The motivation to zoom in on the efficiency of empowerment programs is manifold. In the course of field research in Bangladesh, Cambodia and Tanzania, I concluded that community empowerment programs can be effective, at least in an initial wave of collective mobilization for

² The difference between outputs, outcomes and impacts is important in evaluation (as is the one between resource availability and actual resource use on the input side), but is of minor concern for the time being. I am aware that some readers who conceive of empowerment primarily in a human dignity perspective may take offense with the use of input-output terminology in this context. It is being introduced here in the build-up to the particular efficiency model on which the empirical work is built.

better household and local community outcomes (Benini and Benini 1997: ; Benini 2006: ; Benini 2008). In addition, such programs can have immensely positive effects on the self-value and dignity of the poor, an achievement that is particularly difficult to express directly in quantitative models. However, while the variation in effectiveness was accessible, notably through the communicated results of the participants' own assessments, virtually nothing could be said about the variability in the effort made to achieve those results. In a rare exception for the NGO world, Lutheran World Federation managers in Cambodia produced financial data that led to a global estimate of US\$ 60,000 per village, or \$ 50 per resident, as the total cost of a typical nine-year stay in its community empowerment program (Benini 2008: op.cit., 144). Although the 56 participant village communities in the sample self-rated their progress, cost allocation practices in the supporting LWF program ruled out individual estimates.

Interest in the variability of empowerment program efficiency is justified on several grounds. Resources for such programs have potential alternative uses; the benefits of these have to be weighted against the claimed empowerment effects. This is particularly the case in programs that are highly subsidized (as opposed to, say, self-financing micro-credit), which, while generously funded, may produce dazzling empowerment results. In the longer term, these may be unsustainable. Since the empowered will have grown in voice and assertiveness, it is likely that some of the most vocal participants will be show-cased, regardless of the disproportionate resources that their empowerment consumed. This will distort the representation of empowerment results with a double bias, once to atypically high achievers, and twice by overlooking the difference between their resource levels and the average or long-term sustainable levels. This leads to the assumption that the most effective participants will not always be the most efficient ones. Efficiency studies can thus inject greater realism, particularly with a view to sustainable outcomes.

Efficiency in the local and wider contexts

Second, just as effective empowerment is context-dependent, one must assume that degrees of efficiency vary systematically with the local environments. Factors of variation may operate on the selection of participants into empowerment programs as well as on their behavior within. For example, in comparing the information effort that the Tanzania and Cambodia programs expended to guide the selection of households, full household census fared poorly against the nimbleness and selective follow-up of the wealth ranking approaches (ibid.: 149). An example of an efficiency challenge for the "behavior within" concerns the relationship between community facilities (e.g., health clinics), which may win over the village elites, and the part directly invested in the poorest households, given limited budgets. How is the input-output rapport among the program participants affected by factors outside their immediate control? How do program history or community poverty levels affect efficient resource use?

Third, empowerment programs take place and produce results within complex institutional arrangements, some of which may be characterized as aid chains (Bebbington 2005) or as a hierarchical contests for aid (Epstein and Gang 2005). Donors need to dispose of budgets while carrying out policy guidelines, and intermediary NGOs try to minimize the unpredictability of grant flows. Epstein and Gang make a theoretical argument that NGO intermediation can lead to low-efficiency traps directing aid away from the poorest. This may be a variant of the budget-maximizing bureaucratic manager thesis (Niskanen 1968), but an empirical study of its

manifestations in empowerment program has, to my knowledge, not yet been undertaken. Selection bias in social assistance and development programs has, of course, been amply studied (often under the topic of targeting), but cogent proof of how this affects the efficiency of the kinds of participants in my focus might be quite demanding.

The efficiency of the empowered

My objectives are more modest. I wish to characterize a set of empowerment program participants in a way that distinguishes between effectiveness and efficiency, and to identify significant correlates of their relative efficiency³. Literature searches did not produce any hints that such studies exist in the field of development NGOs or programs of empowering the poor, with the notable exception of micro-finance (on this, more below). Take, for example, Bangladesh, a country with a dense and relatively well-researched NGO sector. An in-depth study of NGO economics and governance (Zaman and al. 2006) makes only one brief reference to efficiency or cost-effectiveness outside micro-finance, to NGO-run schools. Similarly, an Australian government evaluation of NGO projects in Vietnam concluded, somewhat resignedly, that *“cost-effectiveness is seen as a characteristic of NGOs, but the Team was unable to assess this due to lack of monitoring data”* (AusAid 2000: 51). For a more general map of where the foci of efficiency research rest, I turned to Google Scholar. The search engine returned 23,600 “efficient firm” hits, as compared to 482 for “efficient NGO”⁴. “Efficient empowerment”, with only 15 hits, stayed outside poverty reduction altogether, in slightly bewildering worldviews like “Feminist bureaucracy as hybrid form” (Ashcraft 2001), which may or may not overlap with the gender perspectives offered in the developing-country empowerment literature.

[Sidebar:] Data Envelopment Analysis and social justice

In professional communities imbued with an empowerment philosophy, there is a strong, if sometimes implicit link with social justice demands. Much of this discourse is couched in purely qualitative terms. Where it meets with institutional demands for quantitative data, such as in project monitoring, these are often qualified by calls for qualitative supplements and multi-method triangulation. The intuition that quantitative methods can reinforce social justice reasoning is rare (see, however, e.g. Basso and Funari 2003).

DEA holds potential for that. This is unrecognized because empowerment thinkers and economists speak separate languages. One of the ways DEA contributes to clearer empowerment thinking is by contrasting effectiveness and efficiency where appropriate. For example, William Cooper, one of the pioneers of DEA, relates the story of a critical study of the way Texas rewards excellence in school administration. The State-mandated Excellence Standards for Texas, which have inspired also the “No Child Left Behind” initiative of the George W. Bush presidency,

“represent statements of desired goals and schools are rewarded (or not rewarded) on the basis of their achievements. Nothing is said about the amounts (or varieties) of resources used. Hence it should be no surprise that only one of the excellent-rated schools included in this study were found to be efficient. The other schools rated as excellent by the State of Texas had all expended excessive resources in achieving these goals. On the other hand many schools that failed to achieve excellence were nevertheless found to be efficient in producing desired outputs under very difficult conditions.” (Cooper, Seiford et al. 2007: 66)

³ Correlates of effectiveness were one of the subjects of the earlier studies referenced above.

⁴ As of 11 July 2008.

These included high numbers of minority students, of economically disadvantaged students, and of students with limited English language proficiency. The study team recommended to base state budgetary allocations to school districts not only on “excellence” (effectiveness), but also on efficiency. Cooper notes, perhaps unsurprisingly, that the recommendation was not accepted.

The situation of failing “*to achieve excellence [while nevertheless being] efficient in producing desired outputs under very difficult conditions*” is clearly a frequent experience in empowerment programs. Moreover, it is a situation that challenges their leaders ethically. What to do with those who struggle hard and use their modest resources judiciously, but remain unable to reach program objectives?

A unique dataset and a different method to analyze it

I take advantage of a unique data set on 260 local associations of poor people in northwestern Bangladesh in order to estimate the efficiency of their resource use in achieving certain empowerment outcomes. I studied these federations, as they are called locally, in line with my research interests in organizations of the poor; twenty years earlier, well before they existed, I had been program coordinator of the NGO that was to found and foster them. In the conceptual scheme of “Measuring Empowerment”, the federations are collective assets (Narayan-Parker 2006a: op.cit., 11), the empowerment outcomes considered are of a mixed intrinsic and instrumental character, a mixture that came natural in the views of many of their 130,000 members. The homogeneity of their constitution and support history, their geographical contiguity filling up a poverty crescent within a border region, and the one-to-one matching with the lowest local government entity (the “Union”) make for extraordinary comparability.

Earlier (in Benini 2008) I struggled with the great diversity of small projects that the federations were pursuing and the scarcity of financial data that frustrated the analysis of the income-generating ones. In this paper, I use a technique known as Data Envelopment Analysis (DEA) to admit also non-financial inputs and outputs. Arguably, the most valuable inputs that the federations have received are trained recruits, in the shape of so-called “primary groups”, neighborhood groups of 15 to 25 members that the supporting NGO identified among the poor, organized, trained, supported with loans and extension services, and finally evaluated and credentialed as fit to join their local federations⁵.

Hybrid and “second-best” institutions

Before I outline how the rest of this paper proceeds, I wish to briefly situate this study within other research fields that may not habitually overlap with empowerment concerns. The federations studied here, but also some of the NGOs fielding empowerment programs are hybrid organizations, combining multiple organizational personalities and tied into larger networked alliances. Mainstream economics and economic sociology have in recent decades seen a burgeoning literature on hybrid forms (for an overview, Ménard 2004), with sub-themes such as franchising, the cost of monitoring, and long-term relationships vs. opportunism that apply to empowerment programs as well. Second, the coordination problems for which multi-sectoral rural programs of earlier days were infamous (Ruttan 1984: ; Gsänger 1994) do exist in empowerment programs, but they can nowadays be studied more fruitfully through the lenses of participation research (Khwaja 2001: ; Khwaja 2006: op.cit.) and of the economics of scale and

⁵ The downsides of this recruitment monopoly, part of a managed participation regime, are highlighted in Benini (2008: op.cit.).

scope, epitomized by the Grameen Bank's expansion into the telecom sector⁶. Last, public sector economists remind us that productive efficiency is only one of the necessary performance dimensions although there is no excuse for not trying hard to measure it (Pestieau and Tulkens 1993). Rodrik (2008) has given this idea a new twist, by arguing that in the hazardous socio-economic circumstances of poor countries "second-best institutions" may often be more productive than the theoretically most efficient alternatives recommended by "best practice" reviews. This leads back to one of the ideas of this paper, that the correlation between effectiveness and efficiency may be weak.

Organization of the paper

This study is written for readers with various interests in the subject. Some may not be concerned with the specific setting, theory and techniques. They may fast-forward to the results (page 30) concerning the efficiency of the associations studied and to the conclusions drawn in a wider multi-objective NGO context.

Readers who wish to familiarize with the theoretical model of efficient grassroots associations, concepts of measuring efficiency and with the basic functioning of the associations studied may work through some or all of the following sections:

I first offer a notional model of local organizations of poor people supported by an NGO. These associations transform inputs from their task environment into outputs that affect the welfare of their members and possibly some non-members. The efficiency of this transformation is determined by institutional and social-environment factors, including the quality of the NGO support.

The formal technique that I use to estimate the relative efficiencies of a set of such grassroots organizations is known as "Data Envelopment Analysis" (DEA). In the main body of the article, I give an intuitive introduction to DEA whereas I relegate more technical aspects to an appendix. In the world of development NGOs, efficiency studies have been conducted almost exclusively on microfinance questions. I reference a few studies, particularly such as look into gender and social vs. financial objectives.

My own material is from outside micro-finance (although a loan variable is used to explain differences in effectiveness and efficiency), from the multi-objective NGO "RDRS Bangladesh" and the 260 federations of organized program participants that it founded and supported. A background section on RDRS and the federations discusses why the latter are relevant and how their effectiveness (but not efficiency) was measured in the past. Next I propose a model of federation efficiency with three inputs and four outputs and report on findings from DEA estimates.

I map out the geographical distribution of the federations' relevance (partially captured by member density), effectiveness and efficiency and comment on some of the distinct clusters that appear in the RDRS working area for each of those measures. However, my main focus is on the

⁶ The mobile phone is an important part of their practical solution!

apparent disconnect between effectiveness and efficiency and on their respective institutional and environmental determinants.

I offer two conclusions. Substantively, I interpret key findings in the light of Dani Rodrik's metaphor of "second-best institutions", institutions that are effective in their extremely difficult poverty environments, although they may be inefficient in terms of context-free efficiency standards. Formally, I recommend to further experiment with DEA models in multi-objective NGO contexts, whose monitoring systems struggle with the comparability of unpriced activities.

In the appendix, I describe the initial steps in (but not the full) formal aspects of a DEA model, justify the choice of model orientation, and describe the safeguards that I took in order to mitigate some of the commonly known issues in naive DEA. I show the extent of bias correction. I give data sources, descriptive statistics, and summaries of the second-stage regressions.

Theoretical model

In a general formulation, I posit an NGO that induces poor people to enroll in local associations in which they produce outputs. These, when crossing the associations' boundary, become inputs to the members' and some non-members' welfare. The associations produce the outputs by transforming inputs acquired from their task environment, some of them from the supporting NGO. In this case, inputs for the associations are outputs produced by the NGO, whether under grants from donors favoring associations of the poor, or from other sources. The associations' outputs, given input levels, depend on environmental factors.

Associations with unknown production functions

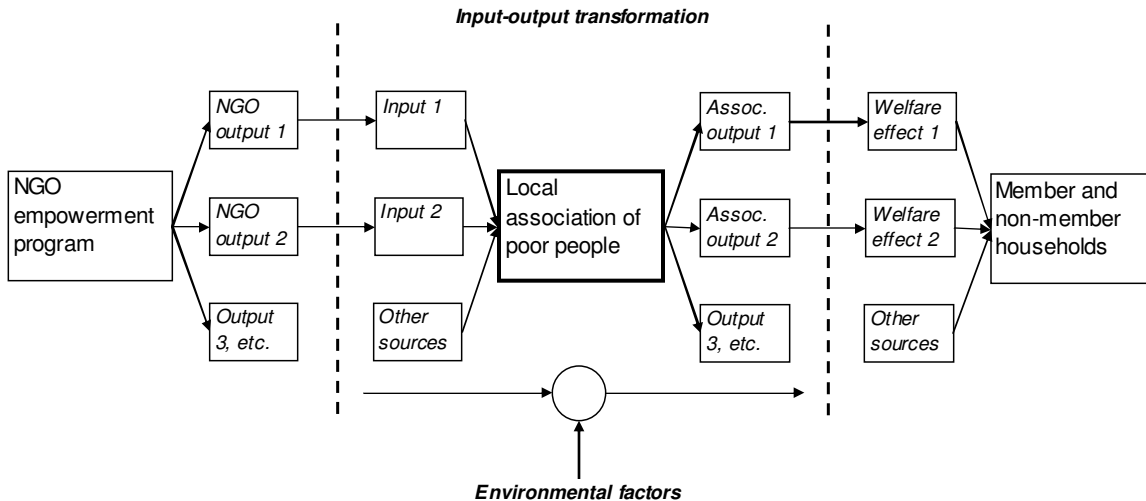
The process is understood and observed incompletely. The associations produce multiple kinds of outputs, some more intuitively associated with empowerment, others less so. The production function is largely unknown although observers may perceive some outputs as mutually reinforcing (efficiencies of scope), and others in conflict. The transformation of associational outputs to welfare effects is known anecdotally only; some outputs of a character that suggests empowerment effects as an end-in-itself (e.g. participation of women who otherwise have no forum outside their households) may be assumed to result in proportionate welfare effects. Likewise, important inputs may go unobserved. Neither for the inputs nor for the outputs is there a pre-established or consensual weighting system as is needed for the denominator and numerator of a single quantitative efficiency ratio. Importantly, the factors causing some poor people to enroll, and others not, are known anecdotally at best, perhaps because detailed baseline information was lost or kept in formats that cannot be assimilated to association data. There is no counterfactual of how empowered the poor would have become had the NGO not founded and fostered the associations of which they are members. Although they and the NGO field workers may have rich knowledge of what factors make for greater or lesser efficiency in associational life, these may be incompletely documented (e.g. access to associations offered by other NGOs in the same locality) or suffer from high measurement errors (e.g., third-party estimates of community poverty levels).

In the light of these challenges, I restrict the model to the transformation of inputs received from the supporting NGO to outputs of the associations, and to the effect of observed environmental factors on the rates of the transformation, the efficiency. In order to mitigate the ambiguity that

certain assets can be legitimately considered inputs as well as outputs⁷, I assume that the associations take time to achieve the transformation, and inputs can be meaningfully observed at time t_1 , and outputs at t_2 . Of course, this does not mitigate the problem that transformation rates may be determined by unobserved factors, including the ways the NGO conducts member recruitment and other supports for particular associations.

Schematically, the model can be diagrammed like this:

Figure 1: Stylized input-output transformation



The empirical model, as indicated, is limited to what happens between the two dashed lines, the input-output transformation and the environmental effects on its efficiency. However, it goes without saying that the selection of specific input and output variables needs to be justified in terms of the supporting NGO’s program and of the plausible welfare effects.

Empowerment, choice and diversity

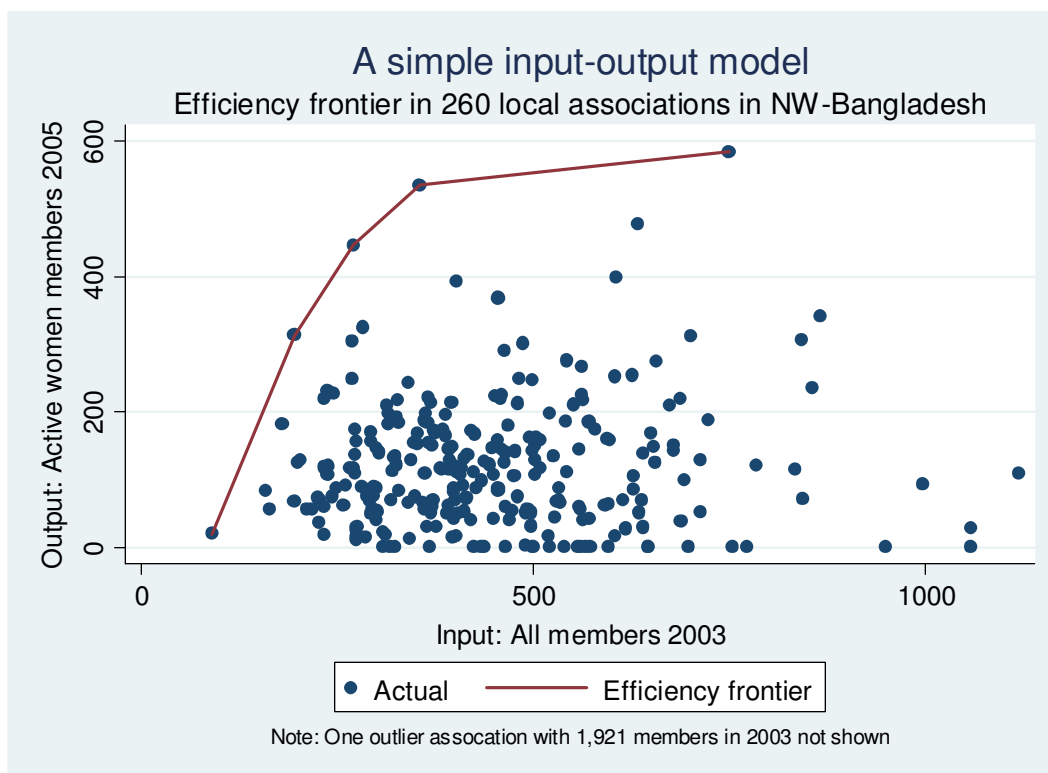
On the formal side, I estimate the efficiency of individual local associations with a technique known as “Data Envelopment Analysis” (DEA). The sidebar below attempts an intuitive introduction for non-technical readers while the formal statement is relegated to the appendix. In the empowerment context, it is important to note that because empowered people enjoy greater choice, we expect to see greater diversity of collective and individual assets among the participants of empowerment programs. DEA has the potential to make monitoring data from such programs, but also from other kinds of multi-sectoral NGO monitoring data accessible to analysis. Such data is oftentimes so heterogeneous that it resists traditional analysis forms.

⁷ For example, when the NGO supplies recruits for the associations, some of these become formal members, and then more or less active members. A member may in short succession be a recruit and an active member. More commonly, micro-finance efficiency models are challenged by input-output efficiency. Deposits, depending on model purpose, are either inputs or outputs.

[Sidebar:] An intuitive introduction to DEA

An artificially simple input-output model offers a first intuitive access to the logic of DEA. Let us consider the 260 local associations of poor people in northwestern Bangladesh from which the data for this study originated. We consider only one input: the number of members that each association had in mid-2003, as a result of members joining from neighborhood groups credentialed by the NGO that had founded, and was supporting, the association. These men and women, guided by a nine-member executive committee, worked together on multiple objectives. One was continued access for women to actively participate in their neighborhood groups and in the association itself, against pressing livelihood demands or cultural resistance. The output selected for this demonstration is the number of female members active two years later, in mid-2005.

Figure 2: The efficiency frontier in a simple model

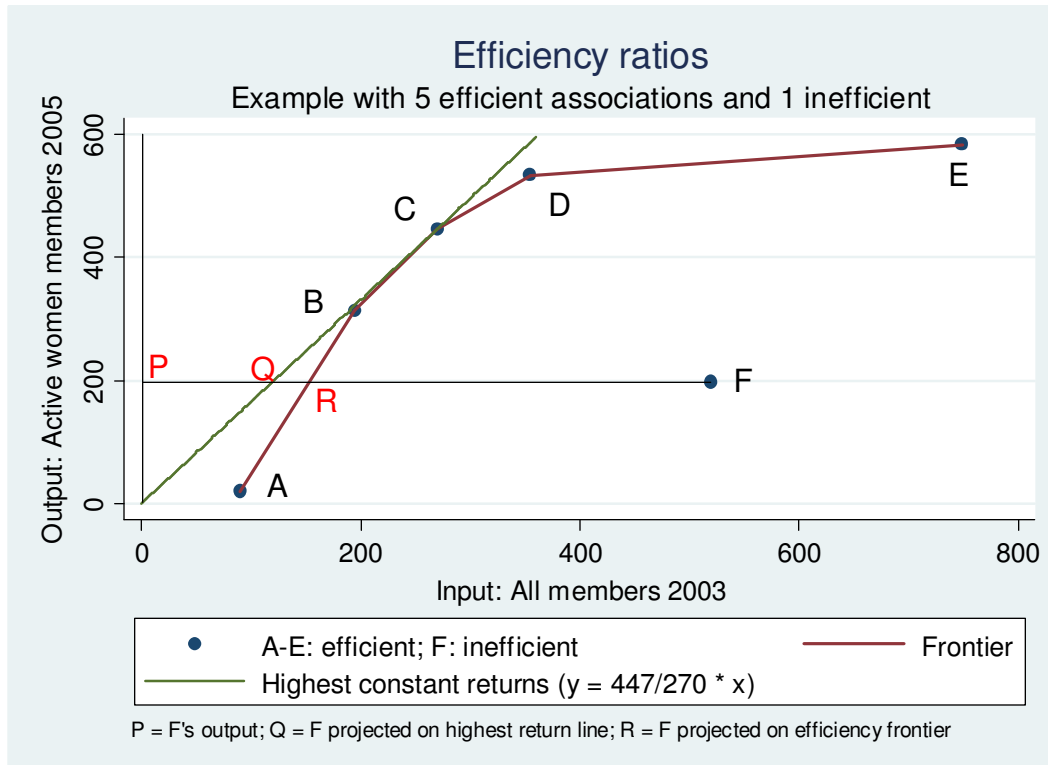


Thus, efficiency is understood as the relationship between the number of active women in 2005 and that of all members in 2003. The assumption is that in the two-year period the quality of decisions, by the committee and by other members, has a significant influence on who among the women enrolled by 2003 would still be active in 2005, and how many of those credentialed in the interval would care to join the association and become active members. But how should the efficiency itself be expressed in a meaningful figure?

A basic idea of DEA is that the most efficient units form a frontier which none of the other, less efficient units crosses. In the above diagram, the efficient associations are the five represented by points connected with a brown line. It is obvious that for each of the five points the following observation holds: All other associations, i.e. points in the diagram, that had at least the same number of members in 2003, had a less favorable ratio of output to input. None of the other 255 points is above the frontier.

At the same time, it is equally obvious that this ratio is not identical for all of the five efficient associations. The largest association on the frontier was almost twice as large as the second largest, but in terms of active women members two years later, the difference was relatively small. Thus, a simple uniform measure will not do all of them justice. The next diagram, which for explanation retains only one association below the frontier (F), facilitates the necessary distinctions.

Figure 3: Technical and scale efficiency



Of the five efficient associations, A – E, C is the one with the most favorable output to input ratio. One can see this by drawing a tangent from the origin (i.e., from 0, 0); it meets the frontier in C (B almost touches it). This green line is technically known as the line of “highest constant returns”. Any point below it has a less favorable ratio. But with the frontier segments BC and CD both very close to it, we might naively speculate that only relatively small associations – between, say, 200 and 350 members – were efficient in creating active participation among its female members. This is not correct when we look back to the full diagram, but it is an example of how DEA can stimulate new thinking.

Now to the one inefficient association in the second graph, F, which had 520 members in 2003. In 2005, 198 of its women members were rated as active members. F is inefficient because numerous smaller associations benefited from higher numbers of active women. But exactly how inefficient is F?

DEA expresses efficiency in relation to the most efficient performers, not to averages in a sample or population. But even this leaves several meaningful options, of which we discuss two here. Both answer the question: How many members would an *efficient* association have needed in 2003 in order to maintain 198 active women members in 2005, as F did? If we expect F to be similarly efficient as the units A and B, then we may select the point where the line for 198 active women, PF, intersects with the line AB. This is the point R. An association represented by R would have had 153.6 members in 2003. However, if our point of comparison is the association with the best ratio, C, then we would compare F to the hypothetical Q, with $198 = 447/270 * x$, or $x = 119.6$ members as the efficient input level.

One of the strengths of this method is that both efficiency comparisons are relevant and are related to each other in a simple way. When we compare F to the best members with similar output, A and B, the relationship between R and F matters. Quantitatively, we express this as the ratio $PR / PF = 153.6 / 520 = 0.295$. If we compare F to the unit with the best ratio, C, the measure to use is $PQ / PF = 119.6 / 520 = 0.23$. So which should be taken? Was F 29.5 percent efficient or only 23 percent?

To answer that, the relationship between points Q and R is of interest. The hypothetical R is, so to speak, the efficient nemesis of F since it is on the frontier. However, R is less efficient than the equally hypothetical Q because Q is a member of the line of highest returns. Thus even efficient associations in the 2003-membership range between A and C are less efficient than Q or C. This is often phrased as “R is producing at an inefficient scale”, and the measure $PQ / PR = 119.6 / 153.6 = 0.779$ is aptly known as “scale efficiency”. This measure is at the heart of the question whether organizations should “scale up” to produce more efficiently, or whether they have become too big in efficiency terms.

For what concerns F directly, the comparison with A and B, $PR / PF = 153.6 / 520 = 0.295$, is, by convention, its “technical efficiency”. In comparing F with the highest technical returns, $PQ / PF = 119.6 / 520 = 0.230$, is its “overall efficiency”⁸. It is easy to see that $PQ / PF = (PR / PF) * (PQ / PR)$, in other words: Overall efficiency = technical efficiency * scale efficiency. This identity always holds, and practical management will likely want to look at both technical efficiency (comparison with the best peers) and scale efficiency (are we too small, too big to be efficient?).

This six-association scenario was deliberately simple, with one input and one output only. The mechanics defy easy graphic representation when the variety of inputs or outputs increases. Also the question then arises how to weight the different inputs, or different outputs for which there is no common denominator. The beauty of DEA is that it does not require pre-defined weights or prices. It creates input and output weights virtually, allowing each unit (each local association in our example) to maximize its output to input ratio (or minimize its input to output ratio), given certain constraints. For example, all inputs and outputs must have some positive worth, or at least not be harmful, and thus must have weights greater than, or equal to, zero.

Another advantage is that DEA does not require knowledge of how the outputs are actually produced. It does not require us to specify, in the example of association F, how 520 members, male and female, active and passive, in 2003 miraculously resulted in 198 active female members two years later. All that is required is the assumption that this type of association tries to maximize active participation among women, and that a greater intake of male and female recruits supports this objective. That this assumption may be questionable is obvious when we look at the right half of the first graph in this sidebar. The efficiency of large associations seems to be very poor, with few exceptions.

Lest this appear overly hyped, some of the downsides of naive DEA have to be mentioned. The method is sensitive to outliers. Looking again at the first diagram, it leaps to the eye that there is a lot of “empty space” between the frontier and the “second-rank” efficient points below. Also, related to that, DEA assumes that no measurement errors occur. This is unrealistic, and all the more so in the non-financial parts of development NGO monitoring systems, or in sample surveys with unreliable data quality. However, new DEA techniques developed in recent years have mitigated these deficiencies. My analysis applies some of them, as reported in the appendix.

DEA applications in microfinance

Before proceeding with my data, it is appropriate to pause for a few remarks concerning DEA applications to microfinance. The microfinance revolution has been hailed as liberating and empowering, by giving access to savings and loan services to poor people who used to be

⁸ This terminology is not universally followed, but I refrain from presenting others, which may confuse the reader.

completely excluded. DEA methods to gauge the efficiency of microfinance providers are attractive for more than one reason. Numerous studies of traditional banking efficiency have used DEA models. With that sector microfinance shares the ready availability of financial, thus interval-level, data.

Outreach and sustainability

Beyond precedent and convenience, DEA offers more to microfinance studies. The usual financial-ratio analysis – there is a confusing number of ratios in use – is not always well suited to take into account issues of scope and scale (Gutiérrez-Nieto, Serrano-Cinca et al. 2007). But these are key to microfinance: Should borrowers receive technical assistance from the loan provider, or not? How are scale and technical efficiency distributed among a group of providers of different size? Increasingly, a double yardstick of outreach and sustainability has been applied to the assessment of such organizations (Yaron 1994). Whether these objectives are mutually reinforcing, or in conflict, is an empirical question; and findings from different studies do not agree. A study of Grameen Bank branch efficiency (Hassan and Tufte 2001) found that female-only branches were more efficient than male-only and mixed branches, an indication that outreach and sustainability are not fundamentally opposed. Looking at numerous microfinance institutions, Lensink et al. (2008), by contrast, found “*a trade-off between outreach to the poor and efficiency.*” Efficiency was higher in those institutions that “*focus less on the poor and/or reduce the percentage of female borrowers*”.

Neither study used DEA⁹. Nghiem (Nghiem, Coelli et al. 2006: ; Nghiem 2007), however, brought DEA to fruition while investigating 44 microfinance schemes in Vietnam. He estimated a mean technical efficiency of 75 percent, and a mean scale efficiency of 29 percent. With 21 percent mean overall efficiency, this immature industry appeared to tolerate numerous very inefficient providers. Earlier, the same researcher (Nghiem 2003) had provided separate estimates of their financial and social efficiency. He classified their corporate personalities in a 2 x 2 table as “beginners” (inefficient on both counts), “charity workers” (socially efficient, financially inefficient), “business persons” (socially inefficient, financially efficient) and “experts” (efficient of both). Only 5 out of 44 earned the “expert” predicate. But the “beginners” are so numerous that one cannot really see a definitive opposition between business and charity approaches.

Nghiem then proceeded to the next stage. He estimated the effects that various environmental and institutional factors had on the technical efficiency. Providers operating in areas that the government classified as “very poor” were less efficient, as one would expect. Other results were counterintuitive. The age of the institution showed a U-shaped effect that started improving only after 13 years of operations. His specific findings are not very relevant for our purpose here, but these second-stage regressions models, while common after the first-stage DEA, are not always straightforward to interpret. Just as the first stage does without traditional price or cost functions, the correct model to account for variable efficiency may not be available. However, the use of DEA in microfinance should be taken as an encouragement for empowerment studies to integrate data on financial activities (such as from village banks) with other empowerment variables that have a meaningful input or output interpretation.

⁹ Both used parametric frontier regression (Kumbhakar and Lovell 2000).

The supporting NGO and the 260 local associations

RDRS Bangladesh

The 260 local associations with whose efficiency I am concerned are locally known as federations, often as “RDRS Federations” or “Union Federations”. They were all founded and supported for many years by the same NGO, RDRS Bangladesh. Their operation and context can only be understood once RDRS has been properly introduced.

RDRS was established in 1971 as a field program of the Geneva-based Lutheran World Federation (LWF) when Bangladesh was an emerging nation and the vast majority of its population lived on the edges of starvation. Its first task was to provide relief and rehabilitation for refugees and those left destitute after the War of Independence. RDRS derives from “Rangpur Dinajpur Rural Service”, named after the Rangpur and Dinajpur region in north-west Bangladesh.

During the period 1976 to 1990, RDRS completed its transformation from a relief agency to a multi-sectoral rural development NGO, retaining its regional identity and focus in the northwestern poverty belt. Its working area nowadays comprises almost 10,000 sq km, spreading across 37 sub-districts with 307 Union Councils. Among an estimated population of 8 million, 1.5 million are involved in the RDRS development programs.

During the late 1980s and through the 1990s, a radical shift took place in RDRS’ philosophy and field activities towards a group-based delivery system, with Union Federations and other community-based organizations emerging as the medium for the message. In this decade, RDRS, like many other Bangladeshi NGOs, built up a large micro-credit program.

A mid-sized NGO among 8 million people

In 1997, after 25 years as a field office with expatriate senior administrators, RDRS became an autonomous, national development NGO, governed by a board of trustees and run by Bangladeshi managers. The supportive relationship with LWF and its partners continues, with aid agencies in Nordic countries and in Holland as its long-term core partners. In 2004, RDRS was working with over 17,500 organized groups, with members drawn from 281,000 households. It had a total staff of 1,958, of whom twenty-nine percent were women, and administered resources worth US\$ 6.6 million¹⁰. Field coordination of the three dozen projects that RDRS administers in a normal year is done through four substantive departments, “Micro-Finance”, “Livelihoods”, “Social Empowerment”, and “Health”. The first three are headquartered in Rangpur, the major city of the working area. In terms of size, RDRS is a mid-field player in the Bangladeshi NGO population, much smaller than the brandname giants BRAC, Proshika and ASA (Grameen Bank, technically, is not an NGO), yet larger than all but 10 – 15 other NGOs.

¹⁰ This information has been compiled from the 2003 and 2004 Annual Reports.

The federations

The RDRS federation movement was initiated by a charismatic staff member, A.H. Bhuiyan, who, in the late 1980s, convinced initially reluctant field coordinators to let him experiment with associations of small groups of landless laborers, marginal farmers and poor women. The demand to federate came from two opposites, from top management and from groups in high-density small-scale irrigation areas adopting tens of thousands of RDRS treadle pumps. These users wanted to leverage their grown incomes and social influence beyond their local hamlets. The experiment was rapidly generalized across the entire RDRS working region. Most federations were set up between 1992 and 1993, with only a few added later, and some reorganized among river island communities. The 260 came to cover a contiguous area (see map overleaf). This is different from the interregional patchwork that the working areas of Bangladeshi NGOs typically form.

Structure

Organizationally, federation support resides, as “Social Organization Unit” alongside the “Women’s Rights Unit” and some education projects, within the “Social Empowerment Department”. RDRS classically exemplifies the multi-sectoral rural development NGO, a type of organization whose effectiveness has been studied extensively, inside and outside empowerment frames. Much less has been written about its efficiency.

In 2002, RDRS won a Euro 2.5 million grant from the European Union (EU) for what has been called the “*Federation Capacity Building and Social Mobilization Project*”. As a result of the capacity building project, the administration of most of the 260 federations was substantially strengthened, dead and dormant groups were revitalized or taken off the books, and neighborhood groups that had not yet joined a local federation were evaluated and encouraged to join. Between 2003 and 2005, membership grew from 118,000 to 129,000, with many federations growing rapidly, and other shedding members in response to internal conflicts or simply updating of old records. As mentioned earlier, I take advantage of this two-year difference in order to achieve a clearer distinction between inputs and outputs in the model while the basic federation structure remained unchanged.

Each of the 260 federation areas is defined by the local Union, the smallest unit in the administrative and local government denomination. The 129,000 members in 2005 related to their federations through 8,000 neighborhood groups, which would send representatives to monthly General Meetings. The typical (median) federation had 460 members. The range was from 120 to 1,200. Sixty percent were women.

The federations have continued to grow beyond the study period. RDRS founded 37 new federations outside its 2005 working area. By mid-2008, the 297 federations counted close to 172,000 members (Karim 2008). Since January 2008, the EU has supported RDRS with a new project, “Empowering the Poor through Federations”.

Figure 4: Overview map of the Federation working area in 2005

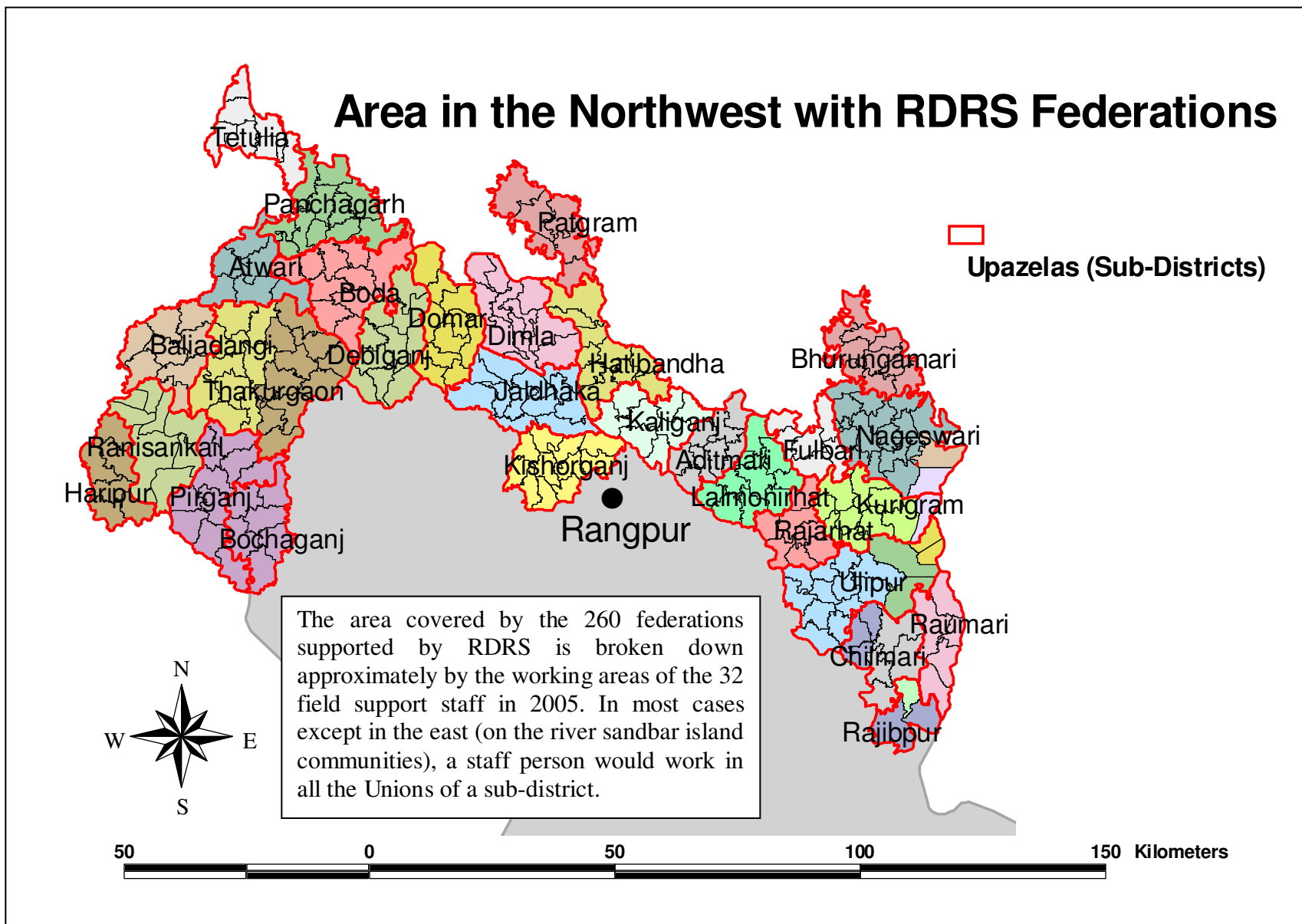
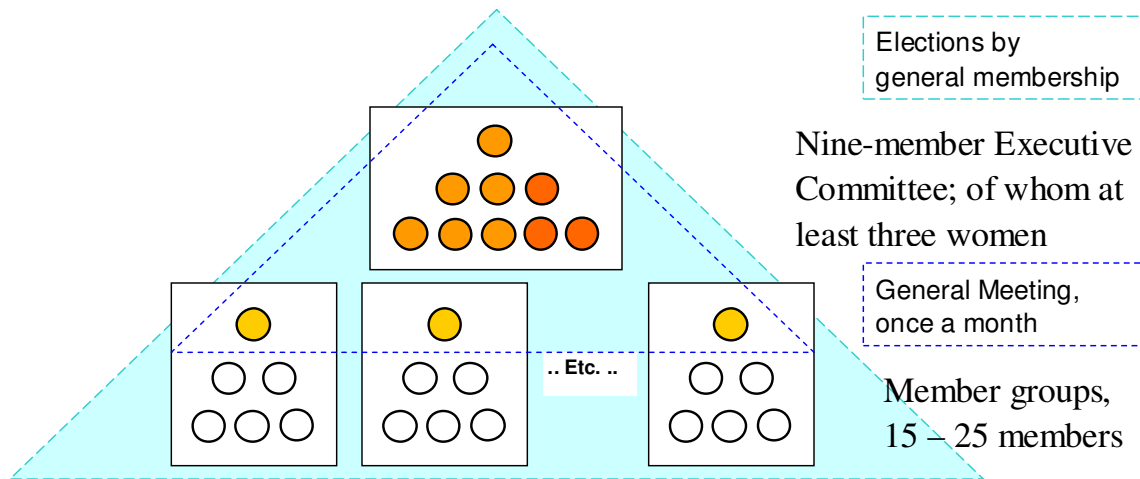


Figure 5: Basic structure of a Union Federation



The basic structure of a Union Federation has two levels. Some of the local terminology is counterintuitive. At the base, we find neighborhood groups called “secondary groups” because they were derived from earlier unfederated “primary” ones. A secondary group has between 15 and 25, either all-male or all-female, members. The second level is made up of the nine-member executive committee. Once a month, secondary groups each send one member to attend a meeting with the executive committee, known as the “General Meeting”. Executive committee members are elected by all individual members. Beyond this diagram, the federations within each sub-district form a coordination committee for mutual help and to plan larger activities.

These figures, from mid-2005, take meaning in historic context. RDRS organized and supported neighborhood groups, including thousands that have not yet joined the federations, throughout the eighties and nineties as well as, in a more focused micro-finance approach, in recent years. In periodic reviews, it “graduated” (a term borrowed from an education perspective) successful groups and asked them to join the federations.

This credentialing has made RDRS the effective intake agent for the federations. It has produced a double definition of membership. *Affiliated* members are those on the books of the Federation. *Active* members are counted as those groups that meet certain behavioral criteria, including regular savings deposited with RDRS. About a third of the members were rated active in 2005, with a strong upward trend since 2003. Among the active ones, the proportion of women was even higher, 66 percent¹¹.

¹¹ After 2004, RDRS abandoned the practice of group graduations. Members of RDRS-organized neighborhood groups, but also other poor people have since been encouraged to enroll in their local federations on an individual basis.

More consequentially, the particular recruitment mechanism has produced a trade-off of inclusiveness for quality. RDRS has made sure that only experienced and (at the time of graduation) active groups would join, and that wealthy and influential community members could not infiltrate. The price of this “managed participation regime” is that the Federation members have remained a minority among the local poor. On the upside, capture by local elites was eluded.

The federations combine multiple personalities, as area-based development associations co-extensive with the local government areas, as special-interest associations of the poor, and as cooperative-like structures that pool resources for common business or social ventures. The range of activities and local initiatives is extremely wide. Large tree plantation projects on the business side and dispute resolution on the social represent activity types that are frequent and semi-standardized. Surrounding them are a host of sundry and temporary initiatives that, despite a fairly good monitoring system (detailed in Benini 2007), may never be fully known to the central support unit in RDRS. Federation finances have been audited yearly, but comparable formats were made available only for one business year, 2004-5.

Relevance

Are the federations relevant for the poor people in the Northwest? This can be answered in two ways, by considering their functions in the societal context, and by the number of people who cared to join and work actively. Clearly, for their normal livelihoods of the poor, the federations are marginal. Most members draw their incomes from activities unaided by the federations. However, the federations have an important quasi-insurance function, by arranging emergency aid when destitution threatens and by deterring violence against women. It is not by accident that women call their federations “shelters for the poor”, and poor men see them as defenders of their dignity against predatory officials and local goons. For many women, their local federation is strictly their only public forum of any consequence. Many Federations have fielded candidates in Union Council elections, with widely variable success from capturing the all-import council chair to increased internal dissension.

Estimates of the member density depend on the choice of enumerator and denominator. An estimated 3.5 percent of *all* households in the 260 Unions were home to an *active* federation member in 2005; this gives a lower bound. Admitting all *affiliated* members, and working with the relatively low Union poverty estimates from the Bangladesh Bureau of Statistics and World Food Program (Bangladesh Bureau of Statistics 2004), one arrives at a more favorable ratio of one member to every four poor households. By all estimates, the federations have enrolled a minority of the poor even where they speak for all of them and render service to non-members as well.

Effectiveness and efficiency

One of the effectiveness indicators is the income situation of the members. Their annual incomes were one third higher than those of a reference population. This was a program effect, not a selection effect, in the sense that members improved skills and incomes *after* coming into the RDRS fold. The differences typically were larger for executive

committee members, most of whom had had longer participant careers with RDRS. Many worked up their way to village middle class positions. This enabled them to give unpaid time and labor, in exchange for social capital used to solidify their fragile ascension.

Other evidence of effective federations is largely anecdotal or relative to each other, without an external standard, not least because there were no competing institutions around. That women who are victims to domestic violence or dowry-related abuse find recourse in the federations is commonly believed, but studies on the longer-term effects of these para-legal dispute resolutions do not exist. Federation leaders seem to have their own development typologies that distinguish among psycho-social, social, financial and institutional aspects, but participatory self-assessments arrived late in the evolution of the federations, relying on imported formats. Earlier measurement attempts combined indicators collected through the routine RDRS monitoring systems. The quality categories based on index ranges – usually three categories – served more the need to direct attention to particularly strong and weak federations than a desire to gauge the degree of goal attainment over all of the 260. I shall nevertheless use some of these measures – a bureaucratic one, and another, self-assessed one – in a bid to compare institutional and environment effects on both effectiveness and efficiency.

The distinction between these two played no conscious role in the everyday lives of the RDRS support staff, as the short-lived effort to make federation accounts comparable attests. RDRS had no particular incentive to address the larger question whether multi-objective associations like its federations were confined within narrow productivity barriers, but it certainly has provided types of support that made federation management more efficient, such as by equipping all chairpersons with a mobile phone.

A model of federation efficiency

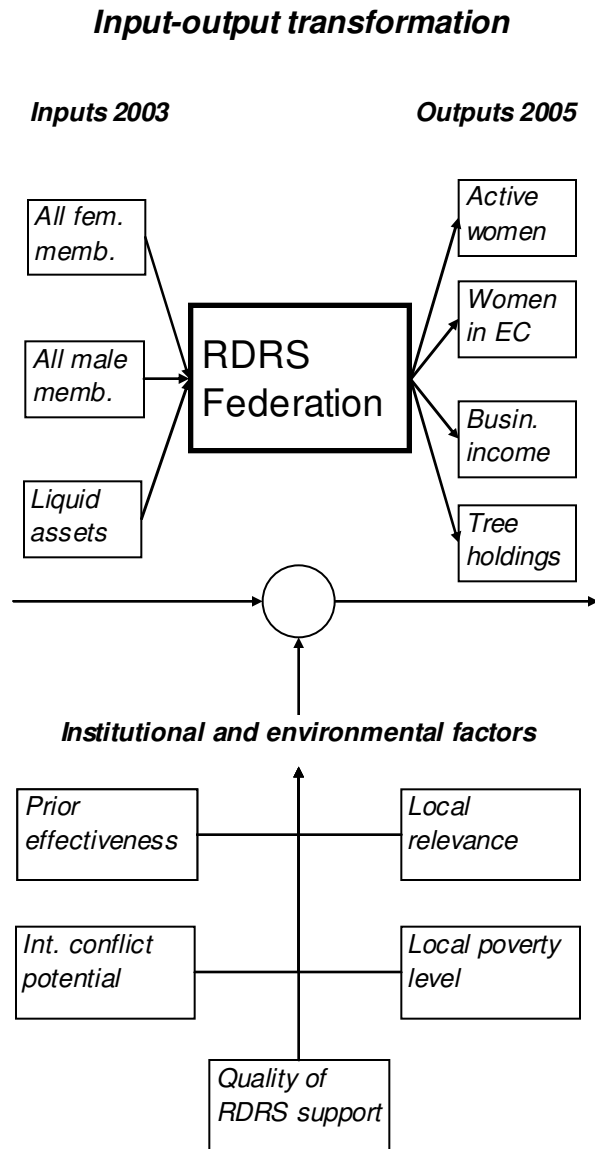
We consider how the 260 federations, given key inputs in 2003, produced outputs in 2005, and how the conversion was modified by institutional and environmental factors. The variety of relevant inputs and outputs included in the estimated model does not cover the entire spectrum of what the federations mobilized and produced, but it does capture essential elements. Similarly, the demarcation between inputs and institutional factors is precarious. I gave input status to variables when they had a plausible positive worth for the production of outputs and the federations had substantial control over their usage between 2003 and 2005. I placed them among the modifiers if the positive worth seemed uncertain, or if the federations had no essential decision making power over them during this period.

Outputs

I selected four outputs that have a plausible rapport with empowerment effects. Active women members are assumed to have greater self-worth and enough power to overcome traditional resistance to participation in a public forum; they also are the ones best placed to extend hands to other women in need. More female executive members tend to give greater voice and impact to women's concern in federation decision-making although their number alone is an incomplete indication, absent information on whether the key positions of chairperson, treasurer and secretary were held by women. Note that elections

had taken place in 2002 already; the point here is the ability of the federation to retain the elected women in the committee.

Figure 6: A model of RDRS Federation efficiency



Business income and future income from tree sales buy the financial capacity of the federations to provide earning opportunities to needy persons (mostly through small loans); they are shorthand for the ability to save members (and occasionally others) from utter destitution. Since trees are planted, raised and sold in lots, income from their sale is periodic, and both variables are needed to express the business stature in a given year.

Inputs

On the input side, male and female members credentialed by RDRS and present in 2003 are included besides the federations' liquid assets. The assumption is that a larger membership in 2003 would go hand in hand with more *active* women in 2005. These would not necessarily be the same individuals as in the 2003 female member pool (they might include new recruits), but their activism would be related to the potential present in the 2003 membership, both male and female. The liquid asset variable is shorthand for estimates, by RDRS monitors, of cash on hand, bank savings, and value of short-term projects¹².

While the seven input and output variables all have a plausible relationship with human and material capabilities, it bears repeating that they are only a conceptual sampling from the full universe of the added agency that the federations provide for the local poor. Similarly, the modifier variables do not span all the dimensions that affect federation efficiency. In the regional context, for example, a measure of NGO micro-lending

¹² The detailed accounts that auditors produced in comparable format for the financial year 2004-5 (used on the output side) were not available for 2002-3.

competition in the sub-district would be valuable because providers discourage borrowers to join or stay with organizations that they perceive as controlled by competitors.

Because the operationalizations are peculiar to the context, the modifier boxes in the above diagram display the theoretical concept, not the indicators themselves. I introduced five concepts two of which chiefly speak to institutional factors, and two to the local social environment. The fifth belongs to both. I discuss them one by one.

Institutional factors

Prior effectiveness: I assume that federations that were more effective in 2003 would translate inputs more efficiently. As a measure of prior effectiveness, the perception by the supporting NGO can step in. RDRS built for Federations that it considered effective, and with modest equity requirements, brick-and-concrete centers and warehouses. These “pukka” buildings created a demonstrative powerful presence in the local community, but for the purposes of this model, they proxy for prior effectiveness. As indicated before, since the RDRS approved and largely funded these buildings, we treat their presence in 2003 as external institutional factor even though hundreds of executive committee and other meetings held in these premises give them the mark of a precious input.

Internal conflicts: Conflicts within associations of the poor depress their effectiveness – hence the importance of unity in the eyes of the members -, and will likely harm also the efficiency (inputs will be husbanded less carefully). Although conflicts are not directly observed for all federations, I take advantage of a particular aspect of the NGO-supported association history. The federations had a checkered career as RDRS borrowers and loan intermediaries that defies brief summary. One facet, though, is relevant here: In federations that traded accusations over loan corruption or where RDRS field staff pressurized executive committees to collect old loans, paralysis ensued. During the period in point, RDRS wrote off vast amounts of unserviceable loans (US\$ 117,000 in 2004-5 alone), but did not notify executive committees until later. I use the magnitude of loans extended, between the start of the RDRS microfinance program and October 2004, as a proxy for the potential of internal conflict. Because RDRS approved these loans, and their positive worth to produce the outputs was either absent or captured in the “liquid asset” variable, they are not rated among the inputs.

Environmental factors

Local relevance: The 260 federations fulfill similar functions. Variation in their relevance occurs in member density. Since the enumerator (members or active members) enter the input-output model, only the denominator, the size of the local Union population is considered. I use the number of households recorded in the 2001 population census, expecting a negative relationship with efficiency.

Poverty level: Poverty impacts the opportunities for association activities. Very poor women may not be able to give time and make the expected group savings. A lack of literate women narrows the executive committee candidate pool. In the particular context, the poorest Unions comprise fragile river sandbar island communities, with few stable roads for the kinds of roadside tree plantations in which the federations have a

comparative advantage over private producers. A negative relationship with efficiency is expected. On second thought, however, such an assumption appears poorly justified. Since poor women know that the federations offer a measure of protection from violence and destitution, incentives for members in poorer communities to make the best out of limited inputs could be particularly strong. On balance, the effect appears unpredictable.

Quality of the NGO support

Higher quality of the support that allied organizations extend to local associations should have, tautologically, a positive effect on their outcomes, and, one would naively think, on their input-output rapport (efficiency) as well. But this reasoning is tricky, particularly when it comes to subsidized inputs. Greater subsidies may boost effectiveness, but discourage efficiency. In our model, we circumvent the issue by exploiting the dynamic that the Federation Capacity Building Project introduced in this period. Support to individual federations was mediated essentially by field staff each assigned to all the federations of one sub-district. Other things being equal, the quality of their work was reflected in the growth of membership in their respective areas. We thus measure the relative support quality benefiting each federation as the mean, over all other federations in the same sub-district, of the magnitude of membership change from mid-2003 to mid-2005¹³.

Descriptive statistics of input, output and second-stage variables are given in the appendix. Similarly, the technical aspects of model estimation are summarized there.

Results

The distribution of efficiency

The efficiency estimates from my DEA model yield a noteworthy result conspicuous in the first graph on the next page. There are numerous federations that, in this selection of inputs and outputs, come out as very inefficient. There is another, smaller group on the strong side of efficiency. The mid-field is atrophied.

One immediately looks for statistical artifacts. Should the distribution be one-peaked, if perhaps skewed to one or the other side? One possibility is that business income is subject to irregular fluctuations, particularly from tree sales; thus, the gap in the 0.45 - 0.65 range should be filled from the left, by some “incorrectly” scored as less efficient. The +0.25 correlation coefficient between income and tree holdings leaves this open. From known executive committee rationales, it is just as likely that some of the federations with a strong social development record deliberately kept a wary distance from business ventures. If for that they were “penalized” with a lower score, then the model did not incorporate their objectives correctly.

¹³ Statisticians may object that this measure violates the iid assumption if its arguments have measurement errors. Similarly, my prior effectiveness indicator is not iid; the decision by RDRS to build a *pukka* center for federation X in year t depended on budgets and through them on the effectiveness judgments reserved for other federations.

Figure 7: Distribution of efficiency scores

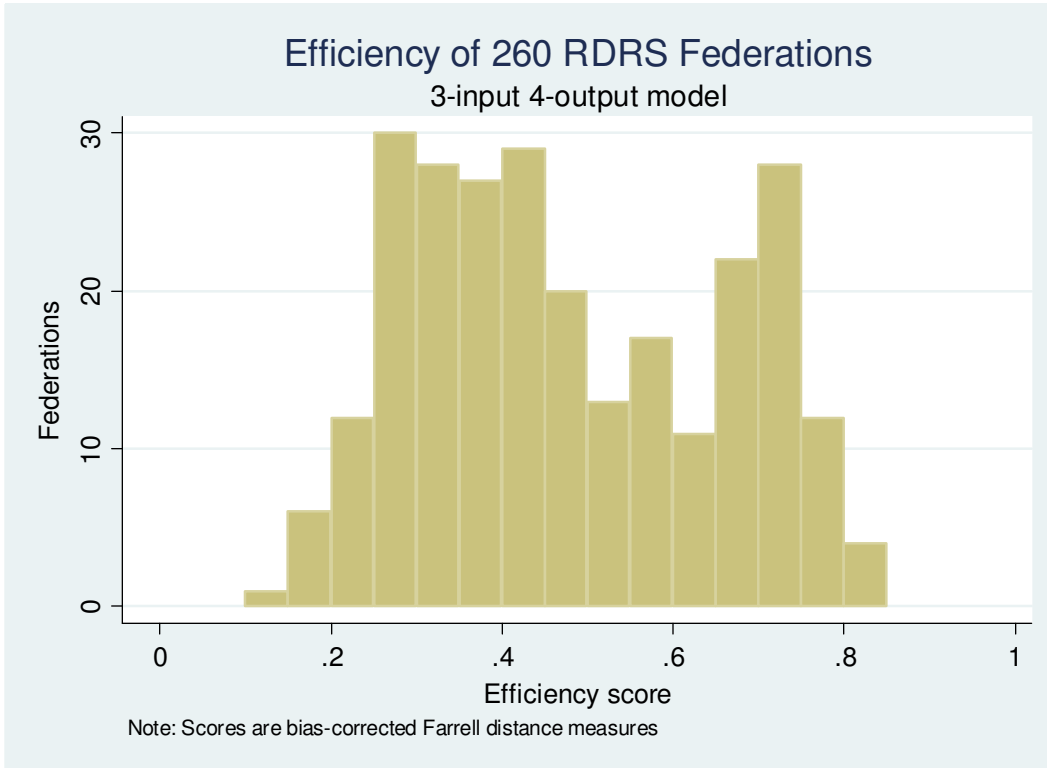
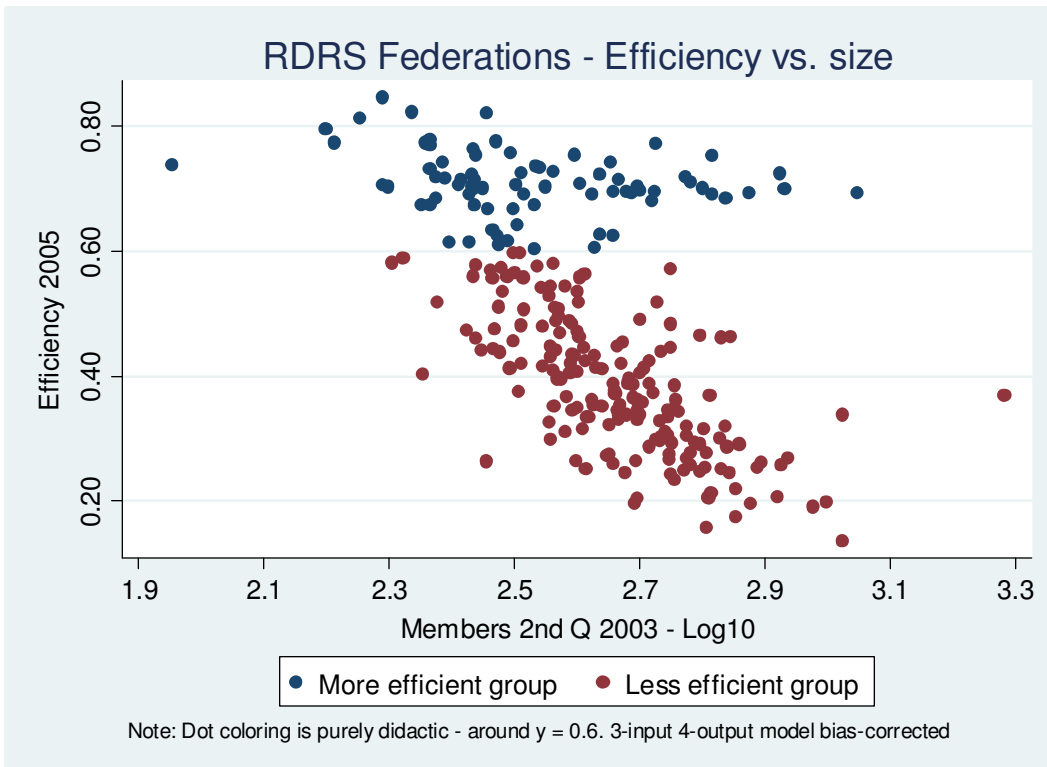


Figure 8: Efficiency in response to membership



However, a different explanation is just as plausible: prior to receiving more focused support under the Capacity Building Project, many federations kept admitting new members from groups graduated by RDRS. These members may have existed on paper only, or they may simply have been so many that the federations were not been able to leverage their swelling ranks into better results.

The second graph on the previous page suggests a bifurcation around $10^{2.5} \approx 300$ members in 2003. A smaller group (blue) experiences very little efficiency loss across the member size spectrum. For the larger group (brown), greater size is associated with a straight slide into inefficiency. In DEA theory, such federations appear “congested” (Cooper, Seiford et al. 2007: 349) with less productive members. However, when we compare technical efficiency and scale efficiency with size, the loss in efficiency is on the technical, not on the scale side¹⁴. There were large, efficient federations, but there were many others that, while struggling with deadwood or freeloading members (e.g. over proceeds from tree sales) enjoyed less efficient cooperation internally.

The bimodal (= two-peak) efficiency distribution demands much more research before a satisfactory explanation appears. Some significant causes may be hidden in the history of the RDRS – federation partnership. Federations often had a facilitating or implementing role for specific projects, some of which were limited to parts only of the large RDRS working area. Another suspect is the interaction with RDRS credit workers, particularly over access to savings and over liability for debts contracted by former executive committee members, driving disgruntled groups into passivity. By and large, though, it is unknown how these interactions worked themselves into the inputs and outputs selected for our model. Similarly, unobserved local federation successes and failures may have impacted their efficiency, such as when spectacular protests mobilized members into greater activism, or criminal gangs cut down trees with impunity, robbing federations of income.

It will be worthwhile comparing this distribution with that for other organizational populations. Pending that, we need to remind ourselves that a peak of highly *efficient* federations does not imply that a minority was fostered and show-cased. If so, this would show up in the *effectiveness* measure. To this difference, we now turn.

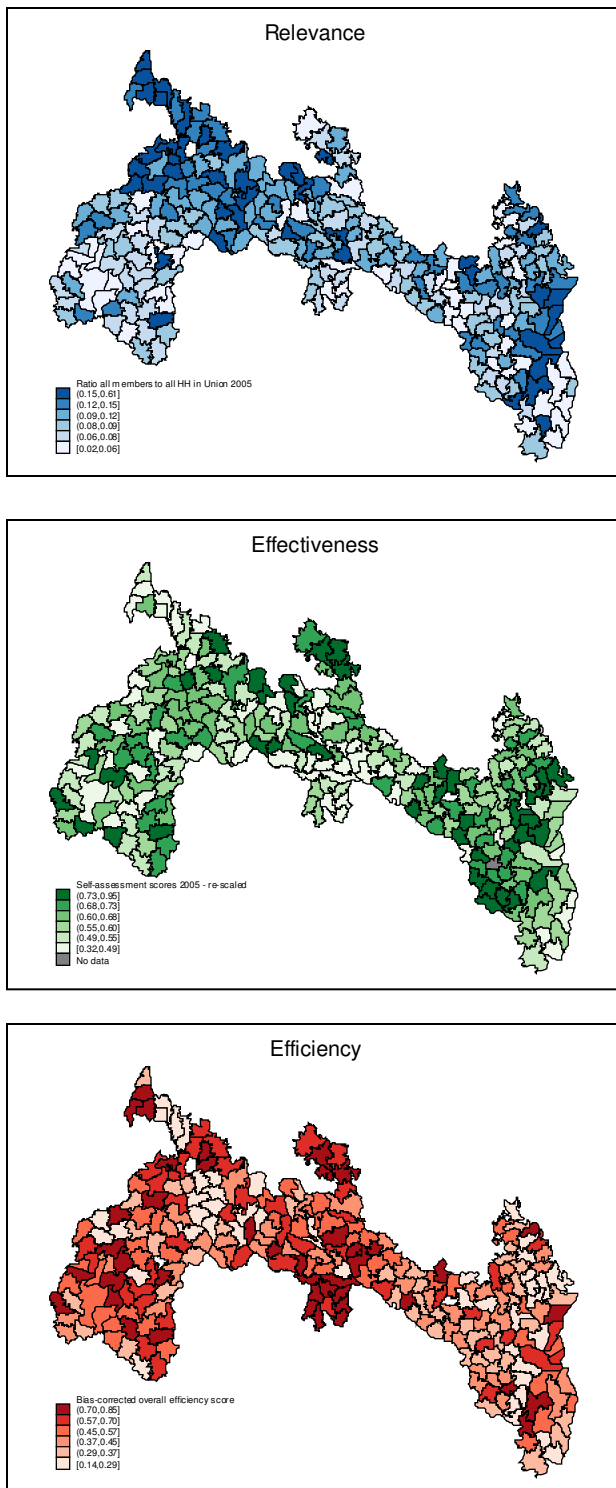
Efficiency and effectiveness

The effectiveness scores that RDRS calculated for the 260 federations and my estimates of their efficiency are slightly *negatively* correlated. If RDRS’ own federation effectiveness score (last calculated in 2004) is used, the correlation is statistically more significant; it is negative, but not very significantly so when we rely on the federations’ self-assessment exercise in 2005¹⁵. A non-statistical formulation is to say that “effective” and “efficient” are indifferent to each other, and perhaps in mild opposition.

¹⁴ This assertion is possible only when we use the efficiency estimates prior to bias correction. The correlation coefficient for the 2003 membership (logged) with technical efficiency is -0.27 , with scale efficiency +0.23.

¹⁵ The correlation coefficients are -0.19 (p<0.002), resp. -0.10 (p<0.094).

Figure 9: Geographical distribution of relevance, effectiveness and efficiency scores



Without detailed knowledge of the local geography, the reader may nevertheless be able to see distinct clusters of Unions – local government areas - that score particularly high or low on select measures of the co-extensive RDRS Federations.

These measures capture, in the manner explained in the text, the relevance, effectiveness and efficiency of the 260 federations. Darker hues signify higher values.

In terms of the restricted definition of relevance (member density), two clusters stand out. One covers a large part of the northwestern part of RDRS' then working area (upper left on the map); the other drives a wedge into the eastern zone (right hand side). The northwestern cluster almost completely fills Panchagarh District, the region that first experimented with federations, and one that historically had a stronger social movement tradition among the poor. The eastern cluster is formed chiefly of Brahmaputra river sandbar island communities, in which RDRS has had a more consistent presence than other NGOs, and where the lack of alternative institutions is particularly grave.

Clusters of federations that rated themselves as highly effective are concentrated in the eastern zone. While the one in the Southeast (lower right) has been well-known to RDRS insiders, there are several in outlying sub-districts that are surprising. Field staff in such areas have been harder to support and supervise, but this has not apparently prevented their federations from remaining effective associations.

Efficiency clusters are of different nature. The high cluster shown as appendix in the bottom center is artificially so, by a policy that created women-only federations in this sub-district [Kishoreganj] (and nowhere else). On the opposite side (center top), an

efficient cluster overlaps with an equally effective one. The factors responsible for this convergence in this remote sub-district [Patgram] are not known; and traditionally Patgram federations have not enjoyed a corresponding reputation.

However, the discussion of these differences can be fruitfully advanced once the effects of the institutional and environmental variables are seen side by side for both effectiveness and efficiency, and in the same breath also for relevance.

Institutional and environmental determinants

The table below summarizes the effects estimated in models with identical institutional and environmental indicators for all three of the key measures. We discuss them for each measure separately, and then look more closely at the opposing patterns for effectiveness and efficiency. The symbolism for the direction and strength of the effects (e.g., +++ for a strong positive effect) is patterned, in simple, perhaps dull manner, on the sign and significance of the statistical output, but should be adequate for the shorthand review here.

Table 1: Effects of institutional and environmental variables

Concept	Indicator	Relevance	Effectiveness	Efficiency
Prior effectiveness	RDRS paid for brick building		+++	
Potential for internal conflict	Cumulative RDRS loans		++	-
Potential member pool	Households in Union	---		-
Poverty	WFP poverty estimate			--
Quality of NGO support	Member growth in sub-district	+++	-	+

Concerning the relevance of the federations, our measure captures only their representativeness, not the importance of the kinds of services they provide. With this proviso in mind, it is almost tautological that larger Unions should have lower membership ratios. “Almost” is important here; historically, RDRS had more extension workers in larger Unions, although probably not fully in step with population. Thus one assumes that more primary group members were graduated over time, and more should have joined the federations. In the empowerment context, the negative effect of community size should be noted; there may be inherent restraining factors at work, either through the dilution of benefits, or in management and communications.

Support quality and relevance

The strong effect that the quality of RDRS support had on member density, however, is not tautological. Between 2003 and 2005, the success of the Capacity Building Project clearly varied with the quality of field staff assigned to sub-districts. The abilities of individual workers who regularly met with the executive committees to contribute something of value to their federations, rather than merely asserting RDRS presence and moving papers, was a continuous concern and topic in coordination meetings, and they transferred into differing rates of member growth.

What is just as noteworthy, though, is the absence of significant effects of prior effectiveness, conflict potential and poverty levels. Whether a Federation boasted a *pukka*

building or not in 2003, is not related to member density in 2005. Similarly, on balance, the loan history with RDRS, and the numerous conflicts over repayment and responsibility, do not impact density. And, finally, the attractiveness, if this is an appropriate term here, of the federations did not significantly differ between poorer and better-off local communities. These findings, however, are difficult to generalize beyond empowerment programs that created similar member recruitment mechanisms for their supported associations as RDRS did.

Loan histories and effectiveness

The effectiveness measure in this table is the one based on the federations' self-assessment exercise. Federations that had been highly effective prior to 2003 (because RDRS perceived them thus and rewarded them with a brick building) strongly tended to rate themselves more effective in 2005. Moreover, larger cumulative loans by RDRS were associated with significantly higher self-rated effectiveness. These ratings did not involve credit-related indicators. The positive effects from loans thus overrode the negative ones from the internal paralysis and repayment disputes¹⁶.

Flying in the face of all anecdotal evidence, better quality of the federation support that RDRS extended between 2003 and 2005 led to lower effectiveness. The effect is statistically significant, if mildly so. However, this may be an artifact of the self-assessment process, in which a delegate from another federation in the sub-district would participate. The raters may thus have considered local outcomes looking also at neighboring federations. If they saw strong member growth in these, it may have muted their appreciation for their own achievements. This explanation is purely ad-hoc, but not unrealistic. Using the bureaucratic measure, the effect almost disappears.

This last effect is reversed when we study efficiency. Federations enjoying more competent support by RDRS staff were more efficient. The down-to-earth interpretation may be simple: during the Capacity Building Project, staff aggressively identified and encouraged committee members for specific trainings; accounting was reinforced and used for better transparency with the rank-and-file. Quality was uneven, particularly in technical advice for income-earning projects, but a positive relationship between support qualities, measured through the growth of *neighboring* federations, and efficiency is significant.

Reverse effect of loans on efficiency

Large cumulative loans too had a reverse effect compared to that on effectiveness. Conflicts are probably only part of this mechanism. Loans may have been extended for projects that the federations did not want (this is known about cattle-fattening projects, more or less imposed by RDRS staff), were not competent to manage, or were distracted by from other, more consensual or profitable ones. In a distantly similar line, the presence of a brick building had a small negative effect on efficiency, not significant but for the conclusion that a prior effective track record does not bid higher efficiency later.

¹⁶ The effect size in an alternative model that uses a bureaucratic measure of effectiveness is about half the size of the one under self-assessment.

Poverty does depress efficiency

Both environmental variables – Union population and poverty – are negatively associated with efficiency. The mechanisms of translation can only be guessed. In larger Unions, small federations may have enjoyed lesser influence with the public and the elected local government body; large federations may be prone to weaker discipline and more conflict. Federations in poorer communities are significantly less efficient but no less effective or representative. River erosion in the poorer eastern part of the RDRS working area, for example, may wash away trees or cause able leaders to migrate, but these disasters would not necessarily lower the federations' self-assessed effectiveness or the incentives for stayers to remain active.

These numerous findings, and particularly the disconnect between effectiveness and efficiency, cry out for an overarching interpretation. For a meaningful one, it is worth considering a specific output that has a plausibly strong and direct effect on the welfare of members and even some non-members: A high proportion of the disputes that are brought before the federations concern violence against women. Therefore it is tempting to test how relevance, effectiveness and efficiency influence the number of cases brought in over a given period. We expect positive effects from all three. But it turns out that greater member density has no effect in either direction, greater effectiveness strongly attracts cases, and the relationship with efficiency is statistically significantly *negative*. I use this finding, among others, for the overall interpretation¹⁷.

Conclusions

Empowerment through second-best institutions?

All the local associations with whose efficiency we are concerned here were founded and fostered by the same NGO. By 2005, the year to which our efficiency estimates apply, their cooperation with RDRS had lasted, for all but a handful, for more than a decade. During this period RDRS handed out rewards that varied widely among the 260 federations. Over half of them held committee meetings and community events in solid brick buildings that document their established community status; others met in simpler bamboo structures or under trees. Almost all had, once or numerous times, received loans, totaling sums that ranged from a paltry Tk. 1,800 (less than USD 30) to a million. Much of this generosity seeped into the pockets of executive committee members; some may have been used productively. At any rate, losses from federation loans added to RDRS' financial hemorrhaging for many years. There was certainly no market-driven discipline built in for productive investment and consistently good credit.

Subsidies and continuity

However, even when loans were used inefficiently, they contributed to more intense interaction between RDRS and the debtor federations, and presumably between committees and member groups to whom some funds were lent onward, or to whom

¹⁷ I did not include the caseload in the DEA outputs because the decision to approach the federation mostly comes from afflicted parties and is not, or only slightly, controlled by the committees who subsequently work the cases. The statistics referred to are detailed on page 46.

physical projects were parceled out. This presumption holds also for other kinds of subsidies, again almost regardless of efficient use, because the incentive to extend them calls for more networking, or more hierarchy (e.g., NGO field staff de-facto presiding over meetings of association committees that they were supposed to help empower), but probably fewer projects that are profitable in the market.

Similarly on the side of the NGO. Donor support (in our case the EU grant for federation capacity building) boosts employment that is not controlled by its marginal utility to the federations, but which nevertheless is necessary to the stated capacity goals. These workers become more productive slowly and in functions that in some measure could be filled by the more experienced ones among the federation leaders, and less expensively so. However, the continuity is necessary to keep the aid chain predictable, through suitable monitoring arrangements. It helps steer the use of grants and subsidies towards issues that promise greater continuity in aid flows, although some of them may not immediately tally with federation or member group preferences. Others may be hailed enthusiastically.

Common to most is a very careful maneuvering in a multi-level network in which efficiency may be important in selective areas, but may remain foreign to others critical for continued cooperation. For an illustration, recall the last of the specific findings above: it takes an effective federation to attract substantial numbers of disputes for mediation. Thus, it takes many such federations to produce the kind of case volume that lends credence to the rights-based approach that NGOs opt, or are told by their donors, to follow. Efficiency concerns are subordinate, and not least because the alternatives, such as official courts of law, are less effective.

Second-best institutions

This tentative diagnosis has a parallel in some provocative arguments advanced by Harvard political economist Dani Rodrik (2008). Under the “constraints and challenges” of poor nations - his argument runs - practices that, when viewed from elsewhere seem efficient, may not be the most effective. Relational contracting, i.e. long-term relationships between the same parties, may give better results than shopping around for less expensive providers and then finding out, at great cost, that contracts with newcomers are not enforceable.

Similarly, obstacles to new entrants that in theory make markets less efficient can have a necessary protective function. Entrepreneurs, in Rodrik’s world of “second-best institutions”, invest in searches for what may work under the difficult circumstances of their poor economies. If anyone could immediately copy their solutions, no one would be willing to invest in searches. Entry obstacles, such as licensing, create rents for the incumbents, rents that are nevertheless necessary for certain investments to be made.

Applied to the federations, RDRS controlled the flow of recruits, through a credentialing process. This prevented many poor people from joining, but it also set a strong bar to local potentates taking over. The arrangement has rewarded federation leaders with social relations capital that motivates them and others to contribute voluntary services. A significant portion of effective federations may be relatively inefficient, but they get

useful work done. As Rodrik also warns, over time the major constraints may shift, “requiring a change in focus as well”. Whether an NGO – local association network, largely run on aid, is capable of continued discovery and reform is an open question.

A moral dilemma

In the empowerment context, the disconnect between effectiveness and efficiency nevertheless leaves a moral dilemma. What is to be done to local associations that manage their resources relatively efficiently in input-output terms, but remain ineffective vis-à-vis the larger objectives? Their low effectiveness may be due to the accident of cooperation (e.g., RDRS ran out of budget before they received their brick building) or to local circumstances (the village priest is so apt at dispute mediation that the federation is not needed for this) or may otherwise be path-dependent.

How much attention can the supporting NGO reserve for a bunch of disadvantaged, remote communities who may be husbanding their modest local resources very conscientiously, but may not in any foreseeable future rise to notable success? In organizational arrangements that claim to transport “empowerment”, the relationship between the efficient pedestrian and the effective automobile is not an easy one.

DEA for multi-objective NGOs

This exercise has attempted to differentiate, in the world of community empowerment, the concepts of effectiveness and efficiency. For the former, existing NGO monitoring systems provide measurements, however embryonic or untested. Virtually nothing has been done for the measurement of efficiency, the relationship between program outputs and resources expended. The RDRS data on its 260 federations is rich enough to experiment with a technique already familiar in a “neighboring” realm, microfinance. The results may be stimulating, but the reader will feel that many questions of model validity and interpretation remain unclear, let alone settled.

Not only apples and pears, but a philosophical change?

However, DEA, the technique applied here to come to terms with the “apples-and-pears” situation of diverse, unweighted, unpriced inputs and outputs, should be further investigated. It is barely known among people working in NGO program monitoring and evaluation, but it has potential to penetrate some of the less tractable problems that they face. In empowerment programs, and also in others in which inputs and outputs vary in kind and proportion, program reporting is at a serious loss for a meaningful balance sheet of what was spent and sown vs. what was gained and harvested.

DEA can contribute to a greater appreciation of what multi-objective, multi-activity programs can do, and at what price. There are serious technical obstacles, stemming from unreliable and missing data and from low data literacy, but the major change required will be philosophical. Should development NGOs be exempt from efficiency considerations, living with a split personality (micro-finance run on efficiency criteria, other parts much less so)? Or, rather, should they try to shed light on the efficiency of all programs, the better to weight its value against other values?

Appendix

DEA considerations

Basic model

From the large variety of DEA models and their different formalizations, I present only the very first step of formalization of one of the most basic models, the output-oriented CCR model. CCR stands for the three authors Charnes, Cooper and Rhodes, who proposed it in 1978.

More important is the fundamental idea underlying the formal approach. DEA assumes decision making units (DMUs) that each have discretion over the use of a variety of inputs in the production of outputs. In our case, the federations are the DMUs. Inputs and outputs are observed. But their prices, values or importance weights are not known prior to the estimation of efficiencies¹⁸. This defect, before the advent of DEA, frustrated efficiency measures and their comparisons.

The genius of DEA lies in the combination of three assumptions:

- Virtual weights, on both the input and output sides, can be used to calculate a combined “virtual output” as well as a combined “virtual input”.
- The ratio of “virtual output” to “virtual input” validly expresses DMU efficiency.
- Maximizing this ratio for each DMU while respecting constraints for all leads to the identification of weights that are fair to all, and hence to fair efficiency scores.

In the notation of Cooper et al. (2007: 23), a maximization program is run for n DMUs, $j=1, 2, \dots, n$, with the individual DMU whose ratio is being maximized designated as DMU_o . m different kinds of inputs $x_1 \dots x_m$ are observed and weighted by input weights v_i , $i = 1, 2, \dots, m$. Similarly, s different kinds of outputs have weights u_r , $r = 1, 2, \dots, s$.

By convention, the ratio being maximized is denoted by the small letter theta (θ) in the Greek alphabet. The maximization runs, for each DMU, over the admissible space of v and u :

$$\max_{v, u} \theta = \frac{u_1 y_{1o} + u_2 y_{2o} + \dots + u_s y_{so}}{v_1 x_{1o} + v_2 x_{2o} + \dots + v_m x_{mo}}$$

The restrictions are:

$$\begin{aligned} \frac{u_1 y_{1j} + \dots + u_s y_{sj}}{v_1 x_{1j} + \dots + v_m x_{mj}} &\leq 1 \quad (j = 1, \dots, n) \\ v_1, v_2, \dots, v_m &\geq 0 \\ u_1, u_2, \dots, u_s &\geq 0. \end{aligned}$$

¹⁸ This does not hold for models of allocative efficiency for which prices are important.

The first means that the maximum efficiency value cannot exceed 1. The other restrictions impose non-negative virtual weights; in other words, all inputs and outputs have positive worth or are at least not harmful¹⁹.

For estimation purposes, this fractional model is transformed to a linear one. The technicalities are presented in most basic texts on DEA (op.cit., ibd., and, e.g. Coelli 1996: 10-12). For a broader theoretical framework, see e.g. Fried, Lovell, Schmidt (2008).

Input- vs. output-oriented models

One of the main purposes of DEA is to reveal relative efficiencies among DMUs. Inefficient units are projected on the frontier; the projected distances are fundamental to the score calculation. The projection can be done in different directions. For example, in Figure 3 in the intuitive introduction to DEA (page 19), the unit F is projected *horizontally* to the frontier segment AB and thence to point P on the y-axis. This perspective *minimizes input* for a given output – F would be efficient if it were at point R in the AB segment. Thus Figure 3 illustrates input orientation. Had I chosen an output-oriented perspective, F would have been projected vertically, to a point between D and E. This point would *maximize output*, given F's input.

Besides input-oriented and output-oriented models, there are others that optimize both jointly. Cooper et al. (op.cit.: 115) give an overview of basic models and of some of their most important formal properties.

This study uses an *input*-oriented model. This choice was perhaps naive; it was motivated by the origin of the federations as NGO-induced associations and by their continuous dependency on various kinds of support from RDRS.

However, it would be easy to supply a devil's advocate's rationale for an output-oriented model. Anecdotal evidence (and some theoretical reasons) suggests that the constraints on the output side of the 260 federations were more severe than those limiting inputs. RDRS supplied credentialed members in a more or less steady stream over many years. Savings and investments held in 2003 were in large part the result of RDRS subsidies, not of market enterprise profits. By contrast, activism among members is dampened by (unobserved) multiple social and motivational factors, such as limited mobility of women. On the business results side, federations do not enjoy comparative advantage over private producers, with the exception of roadside tree plantations, where their numbers and power deter theft, at least by neighbors. Comparisons with an output-oriented model await future work.

It should also be stated that the simple CCR model is not fully appropriate for the set of output variables used in this study. The number of executive committee seats that women can fill in a federation is limited to 9. This would argue for a bounded-variable model (Cooper et al., op.cit., 224). However, since most federations were far from reaching that number, the bias from using an unbounded-variable model is likely to be minor.

¹⁹ Some classes of DEA models can incorporate “bads”, outputs to be minimized (e.g. air pollution).

Issues

Apart from specific model selection issues that may affect this study, there are a number of known weaknesses common to basic DEA models.

Statistical artifacts

An important issue is that the number of 100% efficient units increases with the number of inputs and outputs in the model, and adding irrelevant variables may change the results obtained (Pedraja-Chaparro, Salinas-Jimenez et al. 1999).

Violating independence assumptions

The DEA model violates one of the basic statistical assumptions, that the measurements on the members of a sample are independent the ones from the others and are drawn from an identical distribution. This is so because the projected distances to the frontier depend, by definition, on the positions of the efficient DMUs. If the set of efficient units changes, so do the scores of some of the inefficient ones.

Measurement error

The main drawback of DEA (and other nonparametric techniques) is that it assumes that there is no random noise in the data. In other words, that the inputs and outputs were measured without error. This is unlikely to be the case, and particularly unwarranted in the case of local associations in poor countries, with limited accounting facilities, or with incentives to over- or understate some of the values.

Outlier identification and bootstrapping

Following Nghiem (2007: Chapter 5), those issues can be mitigated by outlier identification and re-estimation of the model thereafter, as well as by bootstrapping. The latter technique means the draw of repeated samples from the set of DMUs to compare and the averaging of efficiency scores obtained for each DMU over repeated estimates. Theoretical background is given by Simar and Wilson (2007); practically, Wilson made available a suite of procedures that identify outliers and calculate bootstrapped scores (see below). The scores so produced are, however, not immediate efficiency scores; they rely on a distance concept for multi-output productions pioneered by Shephard (1953). The reciprocal of this measure (without bootstrapping) is identical to the technical efficiency score in the basic CCR model²⁰.

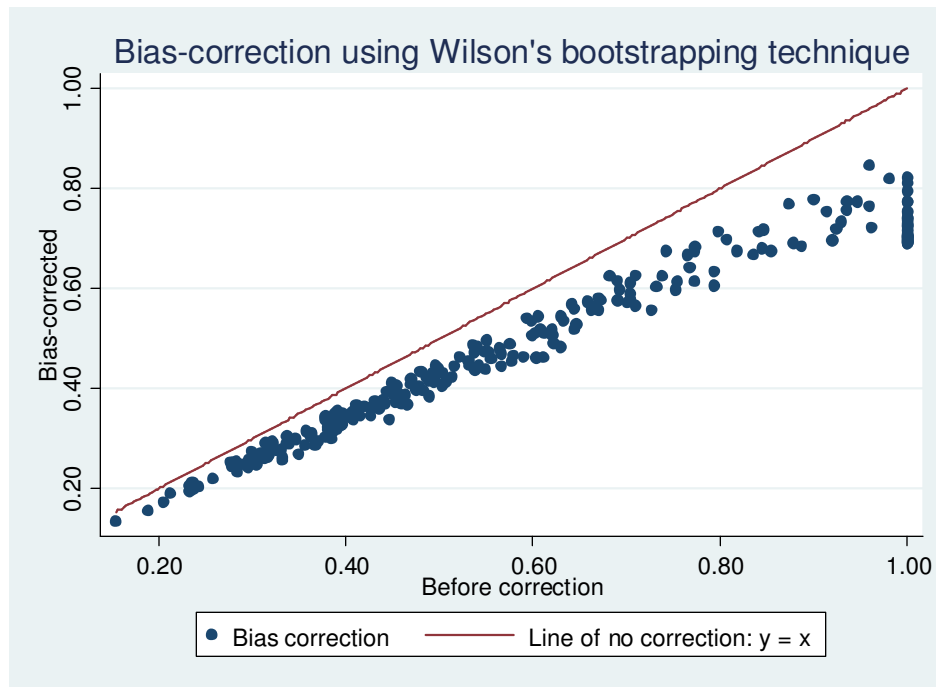
The price paid for the suppression of spuriously efficient units is both conceptual and didactic: The Wilson procedures lose the difference between technical and scale efficiency, supplying only a technical efficiency measure. Second, bootstrapping penalizes most or all formerly efficient members with some bias correction; as a result, the naive user may be bewildered to find that the bias-corrected distribution no longer holds any efficient units.

²⁰ According to Nghiem (personal communication), these reciprocals are technically known as “Farrell distances”, according to Wilson as “Farrell-Debreu measures” (Wilson 2008a: 11). However, this terminology may differ from that used by Cooper et al. (op.cit., 46 and 73), who present the Farrell efficiency model as one that treats slacks differently from the classic CCR model.

Bias-corrected results

For the convenience of the reader, I show some of the mitigation results here, before I present other statistics further below. The graph below makes it clear that the biases of the naive estimates were substantial. On average, the bootstrapping corrected the efficiency score downward by no less than 19 percent. The biases were stronger in the units estimated to be highly efficient. None of the 39 federations initially considered efficient retained a score of 1; they were downgraded, on average, by 28 percent, more than the cramped graph area on the far right ($x = 1$) suggests. Obviously, all had taken their initial places on the hyperplane in ways that were not robust to frontier perturbations.

Figure 10: Extent of bias correction using bootstrapping



Also, I followed advice to recalculate the scores after eliminating the worst outliers. Wilson's outlier detection procedure identified 8 candidates. However, the efficiency scores, after bootstrapping, of the 252 remaining units are highly correlated (+0.99). Therefore, I did not further pursue analyses using reduced sets.

Second-stage regressions

As is often done in DEA studies, the first-stage results (the efficiency scores) are investigated for their dependency on environmental variables that were not part of the set of input and output variables. Cooper et al. (op.cit.) devote an entire chapter to the "Multi-stage use of parametric and non-parametric models". Typically, however, researchers are content to regress the first-stage score on environmental variables; these models then are known as second-stage regressions.

The discussion of appropriate models is wide-ranging. Following Simar and Wilson (op.cit: 57), I used truncated regressions throughout (Cong 1999), and not Tobit models, which they resolutely condemn in the DEA context.

Applications used

I estimated the initial input-oriented, variable returns-to-scale model using Coelli's DOS-based DEAP software (Coelli 1996). To identify outliers and obtain bias-corrected estimates, I used four procedures in Wilson's R-based FEAR suite (Wilson 2008a; ; Wilson 2008b):

- ap (Outlier Detection for Non-parametric Frontier Models)
- ap.plot (Produce log-ratio plot for outlier analysis)
- boot.sw98 (Homogeneous Bootstrap for Shephard (1970) Distance Functions)
- dea (Compute DEA Efficiency Estimates (general version))

Descriptive statistics, truncated regression estimates and graphs were obtained using the statistical package STATA Version 10.

Data sources and descriptive statistics

Sources

Most of the data originated in RDRS quarterly federation monitoring reports. Profits from federation businesses were obtained from external auditors that RDRS had commissioned. Union-wise poverty estimates were obtained from the United Nations World Food Program (WFP). Household figures are from the Government of Bangladesh 2001 population census reports.

All data were integrated in a master table already used for, and are described in, Benini (2006). The federation monitoring system was analyzed in Benini (2007).

Descriptive statistics

Inputs and outputs

Table 2: Descriptive statistics [several sub-tables]

Variable	Obs	Mean	Std.Dev.	Min	Max
<i>Inputs (second quarter 2003)</i>					
Male members	260	198.9	103.6	0	519
Female members	260	254.7	151.1	41	1,720
Capital (RDRS measure, B'desh Taka ²¹)	260	72,660.2	261,274.6	20	2,450,437
<i>Outputs (second quarter 2005)</i>					
Members active female groups	260	121.0	96.5	0	585
Female executive committee members	260	4.1	1.8	0	9
Gross business income (July 2004-June 2005, Taka)	260	29,521.9	93,885.3	0	815,321
Trees owned by federation	260	3,973.3	5,570.2	0	43,000

²¹ The exchange rate used in my 2006 study is US\$ 1 = Taka 70.

Efficiency measures

Variable	Obs	Mean	Std. Dev.	Min	Max
DEAP estimates (before bootstrapping)					
Overall efficiency	260	0.500	0.256	0.071	1.000
Technical efficiency	260	0.589	0.247	0.154	1.000
Scale efficiency	260	0.831	0.166	0.091	1.000
FEAR bootstrapping					
Shephard input distance	260	2.047	0.935	1.000	6.487
Bias-corrected Shepard input distance	260	2.446	1.049	1.183	7.395
Bias of bootstrapping SID - 500 replic.	260	-0.399	0.161	-1.107	-0.135
Estimated variance of bias	260	0.032	0.032	0.002	0.249
Reciprocals of Shephard distance					
Farrell efficiency	260	0.589	0.247	0.154	1.000
Farrell efficiency - bias-corrected	260	0.478	0.177	0.135	0.846

Note the identity of technical efficiency (DEAP) with the Farrell efficiency obtained as the reciprocal of Shepard input distances (FEAR).

Second-stage regressions

Regressands

Separate models were estimated for relevance, effectiveness and efficiency of the federations. All used the same set of covariates (see next para).

Variable	Obs	Mean	Std.Dev.	Min	Max
Relevance					
Member density	260	0.110	0.065	0.018	0.613
Effectiveness					
RDRS-assessed (2004)	260	0.618	0.109	0.349	0.901
Self-assessed (2005)	259	0.609	0.125	0.320	0.947
Efficiency					
Farrell measure	260	0.589	0.247	0.154	1.000
Farrell measure - bias-corrected	260	0.478	0.177	0.135	0.846

The effectiveness measures were rescaled from their original theoretical ranges to [0, 1]. The five measures are correlated among themselves as follows:

	M.density	Eff.-RDRS	Eff.-self	Farrell	F. bias-c.
M.density	1				
Effect.-RDRS	-0.1192	1			
Effect.-self	-0.1640	0.6591	1		
Farrell effic.	0.0283	-0.1902	-0.0999	1	
F. bias-c.	-0.0077	-0.1918	-0.1043	0.9804	1

Regressors

The substantive regressors are discussed in the section “A model of federation efficiency” (page 27 sqq.). I added two auxiliary regressors for error correction.

Kishoreganj sub-district is unique in the RDRS Federation world for the fact that some of its eight federations began as women-only associations. The number of male members remained low, and in 2005 two federations had nine female executive committee members. “Kishoreganj” is a dummy for federations situated in this sub-district.

Second, in the second quarter of 2005, RDRS Micro-Finance field staff did not collect savings from member groups in twelve Unions of two sub-districts, Boda and Rajarhat. This caused the federation monitoring system to report federations in such Unions to have zero active members in this period, an incorrect measurement affecting one of the output variables. “Active member error” is a dummy with value 1 if a federation was in one of those sub-districts and the number of active members in that quarter was reported as zero, else 0.

Variable	Obs	Mean	Std.Dev.	Min	Max
Institutional					
Had pukka building 2 Q 2003	260	0.596	0.492	0.000	1.000
Magnitude all loans from RDRS up to Oct 2004 - log10(x+1)	260	4.390	1.771	0.000	6.000
Environmental					
Households in Union in 2005 (extrapol. from 2001 census)	260	5159.917	1831.058	326.165	15562.780
Fraction households poor (WFP est)	260	0.389	0.083	0.000	0.616
Quality of NGO support					
Index of RDRS support in Upazela 2003-05	260	0.041	0.080	-0.135	0.266
Error correction					
Federation in Kishoreganj sub-dist.	260	0.031	0.173	0.000	1.000
Active member error	260	0.046	0.210	0.000	1.000

Results of the second-stage regressions

Table 3: Second-stage regression results

Variable	Models :				
	relevance [member dens.]	effectivela [RDRS index]	effective2a [self-assessed]	efficient1 [deterministic]	efficient2 [bias-corr]
Model variables:					
Prior effect.	.01051683	.08248455***	.08035904***	-.06937954*	-.04236268
Loans (log)	.00165909	.00598143	.0101299**	-.00731949	-.01536664*
Pot.memb.pool	-.00003404***	6.015e-06	3.830e-06	-4.944e-06	-.00001268*
Poverty	-.02330456	.01568126	.00780997	-.1259776	-.39542001**
Qual.NGO supp	.17287687***	-.02245221	-.24765233*	.34971768	.32470381*
Error corrections:					
Kishor. subd.	-.01267794	-.12243563***	.03297962	.63422839***	.17510445***
Act.memb.err.	.01343203	-.01051846	.04096396	-.14148533*	-.14369586**
_cons	.2613265***	.5108168***	.50183551***	.6628338***	.77710067***

```

[cont. from previous page:]
sigma |      relevance      effectivela      effective2a      efficient1      efficient2
   _cons | .05640281***      .09262379***      .11329387***      .18613712***      .16089348***
-----+-----
Statistics |
   N |          260          260          259          221          260
   ll | 430.37015      249.83824      197.74278      71.043303      113.47001
  chi2 | 28.151059      157.29883      97.10428      104.88982      208.87911
   p | .00020645      1.187e-30      4.271e-18      1.052e-19      1.508e-41
   aic | -842.74029      -481.67648      -377.48555      -124.08661      -208.94001
   bic | -810.69416      -449.63035      -345.4741      -93.503141      -176.89388
-----+-----
                                          legend: * p<0.05; ** p<0.01; *** p<0.001
Note: effectivela = 2004 RDRS rating rescaled to [0, 1], 2a = 2005 self-assess to [0, 1]

```

Auxiliary model of dispute caseloads

in response to relevance, effectiveness and efficiency. Findings from this model are briefly used on page 36. For background, see Benini (2006: op.cit.).

Table 4: Dispute caseload model - descriptive statistics and regression results

Variable	Obs	Mean	Std. Dev.	Min	Max
-----+-----					
Disputes heard in the two years ending June 2005					
	260	12.41538	9.605166	0	56
Member dens.	260	.109876	.0651638	.0183015	.6131865
Effectiveness (self-assessed 2005, rescaled to [0, 1])					
	259	.6093694	.1253092	.32	.9466667
Efficiency measure (bias corrected)					
	260	.4779659	.1774698	.1352245	.8455203
-----+-----					

A negative-binomial model was estimated, with robust standard errors:

```

Negative binomial regression      Number of obs =      259
Dispersion = mean                Wald chi2(3) =      27.41
Log pseudolikelihood = -889.22095 Prob > chi2 <      0.0001

```

Disputes	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
Member dens.	.2600936	.5873027	0.44	0.658	-.8909985	1.411186
Effectiveness	1.425693	.3819389	3.73	0.000	.6771064	2.174279
Efficiency	-.6400364	.245999	-2.60	0.009	-1.122186	-.1578874
_cons	1.903691	.311734	6.11	0.000	1.292704	2.514678
-----+-----						
/lnalpha	-.7737159	.1053393			-.9801771	-.5672546
-----+-----						
alpha	.4612958	.0485926			.3752446	.5670802
-----+-----						

Likelihood-ratio test of alpha=0: chibar2(01) = 879.43 Prob>=chibar2 = 0.000

Bibliography

- Alsop, R. and Heinsohn, N. (2005). *Measuring Empowerment in Practice: Structuring Analysis and Framing Indicators* [World Bank Policy Research Working Paper 3510]. Washington DC, The World Bank.
- Ashcraft, K. L. (2001). "Organized dissonance: Feminist bureaucracy as hybrid form." *Academy of Management Journal* 44(6): 1301-1322.
- AusAid (2000). *Assisting Local Communities: Evaluation of Government Funded NGO Projects in Vietnam* [Quality Assurance Series No. 18]. Canberra, The Australian Government's Overseas Aid Program.
- Bartunek, J. M. and Spreitzer, G. M. (2006). "The Interdisciplinary Career of a Popular Construct Used in Management: Empowerment in the Late 20th Century." *Journal of Management Inquiry* 15(3): 255-273.
- Basso, A. and Funari, S. (2003). "Measuring the performance of ethical mutual funds: a DEA approach." *Journal of the Operational Research Society* 54(5): 521-531.
- Bebbington, A. (2005). "Donor-NGO Relations and Representations of Livelihood in Nongovernmental Aid Chains." *World Development* 33(6): 937-950.
- Benini, A. (2006). *A Shelter for the Poor. The long-term viability of NGO-supported local associations* [Field research support by Bhabatosh Nath and Ayesha Siddika, Responsive to Integrated Development (RIDS), Dhaka]. Rangpur and Washington DC, RDRS Bangladesh and North Bengal Institute.
- Benini, A. (2007). "A Monitoring System for Numerous Local Organizations of Poor People: A Case from Bangladesh." *Electronic Journal on Information Systems in Developing Countries (EJISDC)* 31(8): 1-41.
- Benini, A. (2008). *Does Empowerment Work? Underlying concepts and the experience of two community empowerment programs in Cambodia and Tanzania.* Washington DC.
- Benini, A. and Benini, J. (1997). *Federations - A Shelter for the Poor? The Experience of RDRS Bangladesh.* San Luis Obispo - Dhaka, RDRS Bangladesh.
- Bröckling, U. (2007). *Das unternehmerische Selbst. Soziologie einer Subjektivierungsform.* Frankfurt a.M., Suhrkamp.
- Carnap, R. (1936). "Testability and meaning." *Philosophy of Science* 3: 439-53.
- Coelli, C. M. (1996). *A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program* [CEPA Working Papers No.8/96]. Armindale, Australia, University of New England, Department of Econometrics, Centre for Efficiency and Productivity Analysis (CEPA).
- Cong, R. (1999). "Truncated regression." *Stata Technical Bulletin* 52: 47-52.
- Cooper, W. W., Seiford, L. M. and Tone, K. (2007). *Data envelopment analysis : a comprehensive text with models, applications, references and DEA-solver software.* New York, Springer.
- Epstein, G. S. and Gang, I. N. (2005). *Contests, NGOs and Decentralizing Aid* [Discussion Paper No. 1711]. Bonn, IZA.
- Fried, H. O., Lovell, C. A. K. and Schmidt, S. S. (2008). *The measurement of productive efficiency and productivity growth.* New York, Oxford University Press.

- Gsänger, H. (1994). The Future of Rural Development: Between the Adjustment of the Project Approach and Sectoral Programme Design, Frank Cass & Co.
- Gutiérrez-Nieto, B., Serrano-Cinca, C. and Mar Molinero, C. (2007). "Microfinance institutions and efficiency [page no. cited after 2004 discussion paper version]." Omega 35(2): 131-142.
- Hassan, K. M. and Tufte, D. R. (2001). "The X-Efficiency of a Group-Based Lending Institution: The Case of the Grameen Bank." World Development 29(6): 1071-1082.
- Karim, A. (2008). Progress at a glance (Social Organization) [Fact sheet and personal communication with A. Benini, 24 September 2008]. Rangpur and Johannesburg, RDRS Bangladesh.
- Khwaja, A. I. (2001). Can Good Projects Succeed in Bad Communities? Collective Action in the Himalayas [John F. Kennedy School of Government Harvard University Faculty Research Working Papers Series RWP01-043], SSRN.
- Khwaja, A. I. (2006). Measuring Empowerment at the Community Level: An Economist's Perspective. Measuring empowerment : cross-disciplinary perspectives. D. Narayan-Parker. New Delhi ; New York and Washington, D.C., Oxford University Press ; World Bank: 267-283.
- Kumbhakar, S. and Lovell, C. A. K. (2000). Stochastic frontier analysis. Cambridge [England] ; New York, Cambridge University Press.
- Lensink, R., Meesters, A. and Hermes, N. (2008). Outreach and Efficiency of Microfinance Institutions: is there a trade-off? [draft]. Groningen, Finance Department , University of Groningen
- Ménard, C. (2004). "The Economics of Hybrid Organizations." Journal of Institutional and Theoretical Economics 160(3): 345-376.
- Narayan-Parker, D. (2006a). Conceptual Framework and Methodological Challenges. Measuring empowerment : cross-disciplinary perspectives. D. Narayan-Parker. New Delhi ; New York, Washington, D.C., Oxford University Press ; World Bank: 3-38.
- Narayan-Parker, D. (2006b). Measuring empowerment : Cross-disciplinary perspectives. New Delhi ; New York and Washington, D.C., Oxford University Press ; World Bank.
- Nghiem, H. S. (2003). Efficiency and Effectiveness of Microfinance in Vietnam: Evidence from NGO Schemes in the North and the Central Regions, Centre for Efficiency and Productivity Analysis (CEPA), School of Economics, University of Queensland. 2.
- Nghiem, H. S. (2007). The efficiency and effectiveness of microfinance : Evidence from programs of non-governmental organizations in Vietnam. Brisbane, Australia, University of Queensland. PhD Thesis.
- Nghiem, H. S., Coelli, T. and Rao, P. (2006). "The Efficiency of Microfinance in Vietnam: An Empirical Analysis." International Journal of Environmental, Cultural, Economic and Social Sustainability 2(5): 71-78.
- Niskanen, W. A. (1968). "Nonmarket Decision Making. The Peculiar Economics of Bureaucracy." American Economic Review 58: 293-305.

- OECD DAC (2002). Evaluation and Aid Effectiveness: Glossary of Key Terms in Evaluation and Results Based Management. Paris, OECD Development Assistance Committee (DAC).
- Pedraja-Chaparro, F., Salinas-Jimenez, J. and Smith, P. (1999). "On the quality of the data envelopment analysis model." Journal of the Operational Research Society 50(6): 636-644.
- Pestieau, P. and Tulkens, H. (1993). "Assessing and Explaining the Performance of Public Enterprises: Some Recent Evidence from the Productive Efficiency Viewpoint." Finanzarchiv Neue Folge 50: 293-323.
- Rodrik, D. (2008). Second-Best Institutions [Discussion paper #6764]. London, Centre for Economic Policy Research (CEPS).
- Ruttan, V. W. (1984). "Integrated Rural Development Programmes: A Historical Perspective." World Development 12(4): 393-401.
- Sen, A. K. (2001). Development as freedom. Oxford ; New York, Oxford University Press.
- Shephard, R. W. (1953). Cost and production functions. Princeton,, Princeton University Press.
- Simar, L. and Wilson, P. W. (2007). "Estimation and inference in two-stage, semi-parametric models of production processes." Journal of Econometrics 136(1): 31-64.
- Stern, N. H., Dethier, J.-J. and Rogers, F. H. (2005). Growth and empowerment : making development happen. Cambridge, Mass., MIT Press.
- Wilson, P. W. (2008a). FEAR 1.11 User's Guide. Clemson, South Carolina, Clemson University, Department of Economics.
- Wilson, P. W. (2008b). FEAR (Frontier Efficiency Analysis with R) 1.11.
- World Bank (2001). World Development Report 2000-2001. Washington DC, The World Bank.
- Yaron, J. (1994). "Successful rural financial institutions." The World Bank Research Observer 9(1): 49-70.
- Zaman, H. and al., e. (2006). Economics and Governance of Nongovernmental Organizations in Bangladesh [Report No. 35861-BD]. Washington DC, The World Bank.

About the author

Aldo Benini has a dual career in rural development, with a focus on Bangladesh and another on organizations of the poor, and in humanitarian action. In the latter capacity, he has worked for the International Committee of the Red Cross and for the Global Landmine Survey. He has a Ph.D. in sociology from the University of Bielefeld, Germany, based on field research in community development in West Africa.

Between 1983 and 1986, Benini was program coordinator of RDRS Bangladesh, the NGO that founded and supported the 260 grassroots organizations of this study. Since 1996, he has assisted RDRS in various advisory capacities.

Benini is a citizen of Switzerland and an independent researcher based in Washington DC.

He can be contacted at abenini@starpower.net. Publications are available at <http://aldo-benini.org>.



Published by

