

Helpdesk Research Report: RCTs for empowerment and accountability programmes

Date: 01.04.2011

Query: To what extent have randomised control trials been used to successfully measure the results of empowerment and accountability processes or programmes?

Enquirer: DFID

Contents

1. Overview
2. General Literature
3. Accountability Studies
4. Empowerment Studies
5. Additional Information

1. Overview

The use of randomised controlled trials (RCTs) to measure the impact of development programmes has traditionally focused on short-term interventions in the fields of health, education and agriculture. In recent years, however, a growing number of studies have used experimental methods to assess the impact of a range of governance programmes and processes.

This report examines the extent to which RCTs have been used successfully to measure empowerment and accountability processes and programmes. It identifies the main studies in this field and draws together current thinking about the utility and appropriateness of these studies, highlighting a range of issues associated with using RCTs to examine empowerment and accountability processes.

The studies reviewed here focused on a number of areas:

- **Education** (Banerjee et al 2010, Banerjee & Duflo 2008, Nguyen & Lassibille 2008, Pandey et al 2009)
- **Corruption** (Olken 2007, Reinikka & Svensson 2006, Ferraz & Finnan 2009)
- **Elections** (Banerjee et al 2010a, Humphreys & Weinstein 2010, Chong et al 2010)
- **Economic Empowerment** (Ashraf et al 2009, Dupas & Robinson 2010, Banerjee & Duflo 2008)

- **Community Empowerment** (Gugerty & Kremer 2008, Casey et al 2010, Rassekh & Segaren 2009)
- **Health** (Bjorkman & Svensson 2007)
- **Road Safety** (Habyarimana & Jack 2009)

Most of the studies where accountability was the primary focus were concerned with government accountability (particularly its accountability as a provider of services), or with the accountability of specific local officials or MPs (measured either electorally or in terms of resource allocation). Empowerment studies covered a range of empowerment mechanisms including the economic empowerment of individuals or households, the empowerment of community organisations and the empowerment of public transport users.

The majority of these studies are based on randomised sampling, but some exploit the fact that interventions by government agencies or NGOs were already delivered on a randomised basis (these include Reinikka & Svensson 2006, Gugerty & Kremer 2008, Ferraz & Finnan 2009). Most papers were focused on the community level: they were conducted across a number of villages (between 50 and 775) in several districts. One study sought to measure empowerment processes beyond the community level, by measuring the behaviour of MPs (Humphreys & Weinstein 2010). Another two studies explicitly considered the behaviour of local officials or politicians alongside a broader investigation into community-based corruption reduction strategies (Olken 2007, Ferraz & Finnan 2009).

Suitability of RCTs for measuring empowerment and accountability programmes

Scholars are divided about whether or not RCTs are suitable for measuring the impact of empowerment and accountability programmes. Moehler (2010, p.42), whilst highlighting a number of potential issues with the use of randomized methods in governance and democracy programmes, argues that field experiments present 'immense opportunities' 'to understand something universal and generalizable about a particular phenomenon'. A number of authors have made a more general case *against* the use of RCTs to measure development outcomes. Bardhan (2011), for example, argues that '[a] successful experiment tells us what works, but not why or how; this, in many cases, makes the experimental approach less useful for policy design or broader application' (no page number). Another common critique is that RCTs are more suited to measuring short-term results with short causal chains and less suitable for complex interventions where many factors produce change (Jones 2009). Empowerment and accountability programmes are likely to fit in the latter category. Blattman (2011) is sceptical about the use of RCTs in community-driven development programmes, corruption and electoral reform: 'Even if done well (most experiments are not) I expect inconsistent and erratic results' (no page number).

This hunch appears to be at least partly borne out by the studies assessed in this paper: while some studies demonstrate that accountability leads to improvements in, for example, health services in Uganda (Bjorkman & Svensson 2007) or education services in Madagascar (Nguyen & Lassibille 2008), similar interventions are shown to have little effect on education services in India (Banerjee et al 2010, Pandey et al 2009). Similarly, while a study in India shows that voter information led to greater voter turnout (Banerjee et al 2010a), the opposite conclusion is drawn from a study in Mexico (Chung et al 2010). J-PAL (no date) have argued that one reason for these divergent results may be differences in programme design: in the Bjorkman & Svensson (2007) study, for example, action plans were agreed upon whereas in the Banerjee et al (2010) study, communities were encouraged to develop their own solution to the problem.

Issues and challenges

The literature highlights a number of issues associated with the use of RCTs to assess the effectiveness of accountability and empowerment programmes. Some of these issues apply more broadly to the use of RCTs in measuring governance or other kinds of development interventions:

- **Scale:** Most evaluations focus at the village or neighbourhood level, while most empowerment and accountability processes involve higher-order institutions and actions (Moehler 2010). RCTs need large sample sizes, which may not be available for some more specific political economy questions (for example, questions focused on the behaviour of one ministry) (Humphreys & Weinstein 2009). RCTs will therefore be unsuitable for measuring accountability within a single organisation, which will require qualitative methods (HW comments).
- **Context:** The unique characteristics of the village or neighbourhood can interfere with the ability to generalize from micro-level results. This issue of 'external validity' is arguably the central weakness of RCTs in development (Posner 2011). Rules for decision-making that work well among acquaintances may 'not yield optimal results beyond such intimate settings' (Moehler 2010, 39). Furthermore, while many RCTs focus on the individual or household level, research has shown that most individual behaviour is strongly shaped by social pressure from the community (Posner 2011). As a result, accountability mechanisms may not work in the same way once they are scaled up. The external validity issue implies a need to 'assess the biases that may result from the selection of sites into the pool of studies' (Humphreys & Weinstein 2009, p.375).
- **Complexity:** Governance programmes tend to combine a range of activities targeted at a number of levels. As a result it is often difficult to decide on the appropriate frame for random assignment. If the intervention targets a bundle of activities, it is often difficult to determine the impact of individual activities (Moehler 2010).
- **Time Frame:** Experiments take place in real time, which can make it difficult to use RCTs to measure long-term processes (Humphreys & Weinstein 2009). Many accountability and empowerment processes occur over long time frames.
- **Coordination:** Donors often try to create synergies between governance interventions, making it difficult to identify cause and effect of specific programmes (Moehler 2010).
- **Selection:** Implementing agencies are more likely to volunteer their programmes for evaluation if they suspect they are effective. Researchers also prefer to evaluate programmes where they expect to generate conclusive results (Moehler 2010).
- **Theoretical Gap:** Most studies do not sufficiently explain why an intervention has or has not worked, which limits their relevance to other contexts (HW comments).
- **Assignment of treatment:** How a treatment is assigned may be a relevant feature of the treatment (Humphreys & Weinstein 2009, HJ comments). 'For example, a message provided to a voter may have a different effect if the voter knows a politician strategically targeted him or her to receive that message than if it was randomly assigned' (Humphreys & Weinstein 2009, p.374).

Steps towards improved methods

It has been widely argued that researchers using RCTs should make more effort to understand contextual issues, consider how experiments can be scaled up to measure higher-order processes, and focus more on learning (Blattman 2008, 2011; Humphreys & Weinstein 2009, Moehler 2010). Some of these lessons appear to already have been learnt and these concerns are reflected in many of the studies reviewed in this report. Most studies included a rigorous theoretical analysis of the potential causal mechanisms at play and controlled for other variables. As mentioned above, however, most studies focus on the village level, with only a few seeking to examine higher-order processes.

A number of additional strategies for overcoming these limitations have been proposed in the literature. Many authors have argued that experimental approaches should be combined with qualitative methods (Blattman 2008, Jones 2009). Others suggest that existing studies should be replicated in a number of contexts, although this approach would be more complex and more expensive (Posner 2011). A few

studies reviewed here have aggregated the findings of several RCTs in order to draw more generalizable conclusions about, for example, which empowerment strategies are most effective in improving health delivery (see Rassekh & Segaren 2009, Banerjee & Duflo 2008, J-PAL no date). While this approach provides a way around the generalizability problem, it also tends to report a variety of outcomes, implying that contextual factors are more significant than the type of intervention (Rassekh & Segaren 2009, Blattman 2008). Another potentially promising modification of existing approaches comes from Casey et al (2010), who used randomised methods in conjunction with a 'set of novel field activities designed to measure how real world collective action and participation changed over the four year programme' (p.1).

Outline of this report

Section two presents a range of general papers that, taken together, provide an overview of current thinking on the use of RCTs in development (and more specifically for accountability and empowerment programmes). *Sections three and four* provide an overview of studies using randomised methods that seek to analyse the impact and effectiveness of empowerment and accountability initiatives. These studies have either been selected after searching RCT databases of the main impact and evaluation organisations (J-PAL and 3ie), recommended by experts, or identified in a recent bibliography of the literature on measuring impact of transparency and accountability initiatives (Barrett et al 2010). It should be noted that the division between accountability and empowerment is slightly artificial and a number of studies evaluate programmes that seek to enhance both accountability and empowerment.

2. General Literature

Moehler, D., 2010, 'Democracy, Governance and Randomised Development Assistance', *Annals of the American Academy of Political and Social Science*, 628: 30.

<http://ann.sagepub.com/content/628/1/30.full.pdf>

This article reviews the recent field experiments on democracy and governance (DG) interventions. Twelve out of 41 DG field experiments were focused on electoral interventions; more than any other category. The next largest category was community-driven development interventions (ten out of 41). Nine studies were in the area of 'public service delivery and government performance'. Several of these studies tested the hypothesis that greater information about officials increases accountability and as a result improves government performance. The article notes that 'to date the evidence on the hypothesis is mixed' (p.37). The article notes that USAID has a new pilot impact evaluation programme in the Office of Democracy and Governance and states that this is the only donor programme that is exclusively targeted at impact evaluations of democracy assistance programmes.

The article highlights a number of challenges or issues that are particularly associated with conducting field experiments in the DG area.

- **Scale:** Field experiments are typically conducted at the village level. Many DG interventions are targeted at higher levels of government. Furthermore, most 'theories of political economy of development are typically based on higher-order institutions and interactions' (p.39).
- **Context:** 'The unique contextual features of a village or neighbourhood could interfere with our ability to generalize from micro-level results. In villages or neighbourhoods, citizens often know each other, share strong informal norms, can observe outcomes, and have reason to expect repeated interactions. Voter rules or mechanisms for decision-making that work well among acquaintances may not yield optimal results beyond such intimate settings. Lessons about citizen participation and monitoring at the point of service delivery may not translate into best practices of building accountability at higher levels of government. The context of interventions needs to be considered when scholars or practitioners draw inferences from the evidence' (p.39).
- **Complexity of DG interventions:** DG programmes tend to 'combine many different activities targeted at a diverse array of beneficiaries' (p.40). 'Complex interventions tend to occur at

multiple levels or with a variety of different target populations, making it difficult to decide on an appropriate frame for random assignment. At higher levels, the numbers of units are often too small to yield sufficiently precise estimates, and at lower levels, random assignment cannot capture the effects of higher-level interventions due to the diversity of targets. Third, if the field experiment is designed to measure the impact of the bundle of activities, then it can be hard to determine whether or not specific activities had an impact' (p.40). Measuring outcomes of complex interventions is also more difficult because of the diversity of targets. One alternative would be to test the effect of individual activities. However, the resources committed to one particular activity in DG programmes are often very small and the expected impact is therefore hard to measure.

- **Coordination issues:** Donors often try to create synergies in their programmes. This makes it difficult to identify cause and effect.
- **Selection bias:** Implementing agencies are more likely to volunteer their programmes for evaluation if they suspect they are effective. Researchers also prefer to evaluate programmes where they expect to generate conclusive results. The costs of evaluation are easier to justify for larger and more expensive programmes.
- **Research planning under ambiguity:** RCTs require careful testing of pre-agreed hypotheses. In reality, programme objectives and hypotheses will change over time and are often only specified after implementation has begun. Flexibility of this kind can lead to better programming but creates problems for the evaluator. Many programmes fail not because of faulty hypotheses but because of poor implementation. A failed project does not therefore imply that the underlying hypothesis is wrong.

The article concludes that despite these constraints field experiments present 'immense opportunities' 'to understand something universal and generalizable about a particular phenomenon' (p.42). It suggests that experimentalists should 'push above the village, neighbourhood or point of service delivery and focus on interventions with higher levels of local government and administrative units' (p.42).

Humphreys, M. and Weinstein, J., 2009, 'Field Experiments and the Political Economy of Development', *Annu. Rev. Polit. Sci.* 2009. 12:367–78 (available from http://www.columbia.edu/~mh2245/papers1/HW_ARPS09.pdf)

This paper reviews field experiments designed to address fundamental questions about the political economy of development. The paper identifies a number of studies that address issues of accountability such as corruption, monitoring of teachers' attendance, the use of community report cards to improve local health care delivery, and the performance of MPs. It notes that 'as such studies multiply, research can shift from asking whether different interventions are effective to asking under what conditions one strategy may be more effective than another' (p.372).

The paper highlights a number of limitations associated with field experiments:

- **Experiments take place in real time:** This can be problematic if the processes being studied take place over a long period of time.
- **Constraints on sample size:** Most experimental treatments require large numbers of units to generate precise estimates. In some political economy questions, however, such as the impact of increased salaries on corruption in a public works ministry, the number of units is constrained.
- **Assignment to treatment:** How a treatment is assigned may be a relevant feature of the treatment. 'For example, a message provided to a voter may have a different effect if the voter knows a politician strategically targeted him or her to receive that message than if it was randomly assigned' (p.374).

- **Spillovers:** There is often a risk that a treatment intended for one unit might affect another.
- **External Validity:** Field experiments are often criticised for suffering from weak external validity (it is difficult to know whether the results found in one setting apply in another). 'There is... a need to assess the biases that may result from the selection of sites into the pool of studies' (p.375).
- **Ethical concerns:** At present there are no clear guidelines for social scientists to follow. General principles such as 'do no harm' are not clearly codified.

The article concludes by warning that 'field experiments must not "transform the field from a search for the underlying forces of development into a form of policy analysis"' (p.376). It argues that studies should 'move beyond the very micro level at which many now operate' since 'much of the political economy of development happens at the elite level, among ministers, members of parliament, leaders of business, and international donors' (p.376).

Blattman, C., 2008, 'Impact Evaluation 2.0', Presentation to the Department for International Development', 14th February 2008.

http://www.chrisblattman.com/documents/policy/2008.ImpactEvaluation2.DFID_talk.pdf

This transcript of a presentation made to DFID staff describes the shift away from a traditional approach to impact evaluation (Impact Evaluation 1.0), which stressed cost-benefit analysis and assessing the return on investment of a particular intervention, towards a new approach (Impact Evaluation 2.0), which is more focused on why a programme works. One of the main limitations of the old approach is the fact that the variability of returns of various interventions is sometimes greater than average return itself. As a result, the more important question may be not 'what works', but 'how and where it works'. Another problem with Impact Evaluation 1.0 is that it 'is not a system of performance management or process learning, and thus of too-limited use to implementers' (no page number).

Version 2 evaluations often vary programme components experimentally. For example, 'many micro-finance programs are a mixture of skills training, access to capital, and social network creation. These components can be mixed and matched, or provided in different combinations, to understand the marginal contribution and productivity of each factor'. This unpacking of causal mechanisms can also be done non-experimentally or qualitatively.

Blattman concludes by stating that while economic, educational and health interventions have been heavily evaluated, some of the least evaluated are governance programmes (including accountability programmes). He suggests this may be because governance is 'outside the focus of most economists'. He also states that 'qualitative and quantitative methods can combine for much more powerful evidence than either one alone'. Moving from evaluation 1.0 to 2.0 increases costs and time negligibly.

Posner, D., 2011, 'Individual and Community', Boston Review

http://www.bostonreview.net/BR36.2/daniel_n_posner_behavioral_economics_global_development.php

This short article responds to an article by Rachel Glennerster and Michael Kremer advocating the use of RCTs for evaluating development interventions. It argues that the central weakness of RCTs 'lies in their external validity, that is, in the ability to generalize a study's findings to other contexts' (no page number). Emphasising the importance of context is not simply an argument for making more modest claims about general patterns, but 'also a call for a deeper understanding of the mechanisms that generate the outcomes we seek to affect, along with a recognition that these mechanisms may be situationally dependent'. The article notes that while most RCTs focus on the individual and household, most development issues relate to the community 'either because we are interested in community-level outcomes per se or because we cannot really understand the individual or household-level response

apart from the community'. Some studies have shown how individual behaviour is strongly shaped by the social pressure exerted by community members. The article concludes that 'what we need, therefore, are not just more randomised studies of all sorts, but replications of existing studies, carried out in settings selected for the variation they offer in community-level characteristics. This way we can target research to contextual factors that we hypothesize will shape the peer effects that, in turn, affect the individual behaviours we hope to alter'. This will be more complex and more expensive.

Bardhan, P., 2011, 'Experimental Fad', Boston Review, March/April 2011.

http://www.bostonreview.net/BR36.2/pranab_bardhan_behavioral_economics_global_development.php

This short article critiques the current trend towards more widespread use of RCTs in development policy. Bardhan argues that '[a] successful experiment tells us what works, but not why or how; this, in many cases, makes the experimental approach less useful for policy design or broader application' (no page number). He notes that it is important to bear in mind that 'economists have a professional inclination toward false universality and that our departures from narrow rationality can be highly culture-specific'.

The main argument presented here is that while experimental methods can make causal identification 'much cleaner', they can distract researchers from the bigger questions such as 'various historical and structural explanations for why poverty persists in some contexts more than in others'. '[M]any of these explanations can't be found through random evaluations or general experimental methods'. Governments often make decisions in spite of available evidence about what works. RCTs are usually confined to short-run impacts of relatively small projects: 'they cannot shed much light on the enduring effects of projects that are supposed to last, nor can they teach us a great deal about scaled-up projects where political entities necessarily get involved and distributional issues...may be at least as important as finding out the average impact of a given treatment'.

Blattman, C., 2011, 'Behavioral economics and randomized trials: trumpeted, attacked and parried', 24th March 2011, Chris Blattman Blog

<http://chrisblattman.com/2011/03/24/behavioral-economics-and-randomized-trials-trumpeted-attacked-and-parried/>

This blog posting summarises some of the recent debate over the use of RCTs in measuring development interventions. It is sceptical about the usefulness of RCTs in community development and other, more complex forms of development intervention:

'I suspect a half dozen de-worming experiments or a half dozen vocational training program evaluations will yield somewhat consistent and generalizable results, at least on the same continent (assuming anyone ever gets around to serious and consistent replication). But experiments with community driven development programs? Corruption control? Electoral reform? Even if done well (most experiments are not) I expect inconsistent and erratic results' (no page number).

Blattman adds that 'every trial ought to be registered, and economics journals ought to enforce the practice'.

Bruhn, M. and McKenzie, D., 2009, "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." American Economic Journal: Applied Economics, 1(4).

http://www.iza.org/conference_files/ELMPDC2009/mckenzie_d2795.pdf

This paper provides an overview of how RCTs have been used in development economics to date. It finds that 'many papers do not describe the randomization in detail, implying that better reporting is needed. Our simulations suggest that in samples of 300 plus, the different methods perform similarly' (p.1). Typical sample sizes in the literature are 100 to 300 units. The paper provides useful guidance on which methods are most likely to generate robust results, based on a survey of existing studies.

Jones, H., 2009, 'The 'gold standard' is not a silver bullet for evaluation', Opinion, Overseas Development Institute.

<http://www.odi.org.uk/resources/download/2811.pdf>

This short opinion piece summarises some of the broader dangers associated with over-reliance on RCTs to measure impacts. It argues that privileging RCTs and other experimental or quasi-experimental evaluation methods above others (such as quantitative methods) is problematic and may undermine learning and accountability in the long term. It argues that RCTs are more suited to measuring the impact of programmes with relatively short-term aims and short causal chains such as cash transfers or the provision of vaccines. They are less suitable to measuring programmes where 'many complex, interacting factors produce change' (p.2). Accountability and empowerment programmes are likely to fit into this latter category.

Additional Resources:

Barrett, Gregory, Richard Calland, Ruth Carlitz, Anuradha Joshi, Rosemary McGee, Andrés Mejía Acosta and the Society for Participatory Research in Asia (PRIA), 2010, 'Review of Impact and Effectiveness of Transparency and Accountability Initiatives: Abstracts of the Relevant Literature', Institute of Development Studies, University of Sussex <http://bit.ly/fdKKxC>

This bibliography provides a summary of studies that seek to measure impact of transparency and accountability initiatives by IDS. It includes details of a number of RCTs, all of which have been included in this report.

E-Gap Website

<http://e-gap.org/>

'Experiments in Governance and Politics' was founded in April of 2009 as a network for scholars and practitioners actively engaged in field experiments on the topics of governance, politics, and institutions. Its website provides a list of scholars and practitioners engaged in field experiments on politics and governance topics. A 'Research' page listing ongoing research projects and publications is forthcoming.

3. Accountability Studies

J-PAL, no date, 'Power to the People?', J-PAL website <http://www.povertyactionlab.org/policy-lessons/governance/community-participation>

This short article provides an overview of the findings of four J-PAL studies designed to assess the extent to which greater community participation has improved the quality of public projects and services. It finds that evidence from these RCTs has produced very mixed results about their effectiveness. It states that 'while it is clear that the details and context matter for this type of program, some common themes about what works are beginning to emerge. Programs where the community had more direct control over service providers tended to work better' (no page number).

Education

Banerjee, A., Duflo, E., Glennerster, R. and Khemani, S., 2010, 'Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation of Education in India.' American Economic Journal: Economic Policy 2(1): 1-30. <http://poverty-action.org/sites/default/files/pitfalls.pdf>

This randomised study assesses the impact of public participation in improving public services. It takes the case of school committees in India, which are made up of locally-elected leaders and parents of children enrolled in public schools. These committees have power over resource allocation and monitoring the school's performance. The study evaluates three different interventions to encourage

beneficiaries' participation. It finds that these interventions had no impact on community involvement in the schools and no impact on learning outcomes or teacher effort. It found that one intervention (training volunteers to teach children to read) had a large impact on activity outside of school. The study concludes that citizens face substantial constraints in participating to improve the public education system, even when they are willing to do something to improve it.

The study was done at the village level in the Jaunpur district of Uttar Pradesh. Each of the three interventions were conducted in 65 villages, with a fourth group of 85 villages used as a control group. The study concludes that much of the variation in outcomes can be explained by the role of individual actions in generating collective benefits.

Nguyen, T. and Lassibille, G., 2008, 'Improving Management in Education: Evidence from a Randomized Experiment in Madagascar', unpublished paper.

http://www.povertyactionlab.org/sites/default/files/publications/93_Nguyen_Improving_Management.pdf

This paper evaluates several interventions in Madagascar that sought to use top-down and local monitoring mechanisms to improve education quality. 15 school districts and 207 sub-districts were randomly selected and some were provided with teacher tools and parent-teacher meetings centred on a school report card. The study found little impact of targeting district and sub-district administrators on their behaviour, although the interventions led to an improvement in performance at a school level.

Reinikka, Ritva and Jakob Svensson. 2006, 'The Power of Information: Evidence from a Newspaper Campaign to Reduce Capture of Public Funds'. Unpublished Manuscript, Institute for International Economic Studies, Stockholm.

<http://people.su.se/~jsven/information2006a.pdf>

This study documents the impacts of an information campaign designed to minimize the diversion of educational funding meant for local schools in Uganda. It exploits the fact that the Ugandan government has been implementing a randomised policy experiment to encourage parents and schools to monitor local officials' handling of a large school grant programme. It relies on the fact that there was good baseline data on spending levels before the campaign began in 1996. The information campaign involved publishing data on monthly grants in national newspapers. The study finds that schools in geographic areas with newspaper access received significantly more of the intended transfers.

Pandey, P., Goyal, S. and Sundararaman, S., 2009, 'Community Participation in Public Schools: Impact of Information Campaigns in Three Indian States.' Education Economics 17(3): 355-75.

<http://siteresources.worldbank.org/SOUTHASIAEXT/Resources/223546-1192413140459/4281804-1215548823865/informationcampaigns2.pdf>

This summary document outlines an RCT designed to assess the impact of providing information to the community about public schools and services to children on learning outcomes. The study was conducted in 340 treatment villages and 270 control villages in three Indian states (Madhya Pradesh, Uttar Pradesh and Karnataka) . A baseline survey was conducted in 2006, followed by the intervention in 2006 and a blinded follow-up survey in 2007. The intervention resulted in improved educational outcomes in all states.

Health

Björkman, M. and Svensson, J., 2007, 'Power to the People: Evidence from a Randomized Field Experiment of a Community-based Monitoring Project in Uganda.'

http://www.cid.harvard.edu/neudc07/docs/neudc07_s2_p11_bjorkman.pdf

This paper analyses the extent to which efforts to increase community-based monitoring can improve health service provision. They find evidence that community-based monitoring does lead to an improvement in health outcomes. The study examines the use of citizen report cards in 50 rural communities across Uganda. Citizen report cards were designed to strengthen providers' accountability to citizen-clients by enhancing communities' ability to monitor providers. The study explores a number of alternative ways in which the interventions could have affected health outcomes, and finds no support for any of these alternatives in the data. It notes that there is a potential for spill-over effects from one catchment area to another. It also states that these issues were minimal since catchment areas for control and intervention groups were at least thirty kilometres apart. The paper concludes by stressing that 'before scaling up, it is also important to subject the project to a cost-benefit analysis and relate the cost-benefit outcomes to other possible interventions' (p.26). It also notes that this was a controlled experiment and that scaling up may require delegating tasks to local actors in the various communities, which may result in capture.

Elections and Corruption

Banerjee, A., Kumar, S., Pande, R., Su F., 2010a, 'Do Informed Voters Make Better Choices? Experimental Evidence from Urban India', J-PAL

<http://www.povertyactionlab.org/sites/default/files/publications/142%20Informed%20Voters%20Nov%2010.pdf>

This paper presents a field experiment conducted in the run up to the 2008 elections in Delhi. Slum dwellers were provided with newspapers that included report cards about candidates. The study found that access to report cards increased voter turnout and voter sophistication – 'voters make comparisons across spending categories and candidates to overcome political agency problems and reward better performing incumbents' (p.1), thus increasing electoral accountability. There was a reduction in cash-based vote-buying by around 20 per cent.

Humphreys, M. and Weinstein, J., 2010, 'Policing Politicians: Citizen Empowerment and Political Accountability in Africa', unpublished paper.

<http://www.columbia.edu/~mh2245/papers1/scorecard2010.pdf>

This study uses an RCT to assess whether greater transparency improves governmental performance and increase political accountability. It uses a randomly assigned scorecard in 147 constituencies in Uganda to assess whether providing information about what happens in Parliament helps to increase political accountability. It found strong evidence that voters are responsive to information on politician behaviour, but only weak evidence that politicians are responsive to their concerns. The scorecard includes measures of MPs' effort (which measures MPs' participation, attendance and initiative), position (which records the areas in which MPs are most active) and community activities. The article includes an extensive discussion of the causal channels through which the scorecard may have influenced accountability and draws on evidence from the literature to build theoretical understanding of the problem.

Chong, A., De La O, A., Karlan, D., Wantchekron, L., 2010, 'Information Dissemination and Local Governments' Electoral Returns: Evidence from a Field Experiment in Mexico', unpublished paper.

<http://www.povertyactionlab.org/sites/default/files/publications/266%20Information%20Dissemination%20and%20Local%20Governments%20Electoral%20Returns%20in%20Mexico.pdf>

This article examines the effects of an information campaign on electoral participation and the vote share of incumbent parties in municipal elections of 2009 in Mexico. 12 municipalities were randomly assigned and received information about municipalities. The results show that 'voters respond to information about excessive corruption by staying away from the voting booth. This drop in participation translates into

losses for the incumbent party. Conversely, information about overall expenditure increases turnout and incumbent parties' vote share'.

Ferraz, C. and Finan, F., 2009, 'Electoral Accountability and Corruption: Evidence from the Audits of Local Governments', Working Paper 14937, National Bureau of Economic Research.
http://www.3ieimpact.org/admin/impact_evaluations/Electoral%20Accountability%20and%20Corruption.pdf

This paper uses audit reports from an anti-corruption programme in Brazil to test whether electoral accountability affects the corruption practices of politicians. Municipalities in Brazil were randomly audited by the Brazilian government as part of an anti-corruption strategy launched in 2003. 60 out of around 5500 municipalities are audited every year. The study finds less corruption in municipalities where mayors can get re-elected and thus concludes that electoral rules that enhance political accountability play a crucial role in constraining politician's corrupt behaviour. It examines a number of alternative mechanisms which might explain why first term mayors might be less corrupt.

Olken, B., 2007, 'Monitoring Corruption: Evidence from a Field Experiment in Indonesia.'
Journal of Political Economy 115(2): 200-49. <http://www.nber.org/~bolken/corruptionexperiments.pdf>

This paper presents an RCT which examines interventions to reduce corruption in over 600 Indonesian village road projects and was conducted in East and Central Java. It finds that 'increasing government audits from 4 percent of projects to 100 percent of projects reduced missing expenditures, as measured by discrepancies between official project costs and an independent engineers' estimate of costs, by eight percentage points' (p.200). It finds that increasing grassroots participation through 'accountability meetings' has little impact, 'reducing missing expenditures only in situations with limited free-rider problems and limited elite capture' (p.200). The study was able to assess differences in impact between different types of grassroots participation. One of the points made in the conclusion is that this kind of experiment has demonstrated only the efficacy of a short-term intervention. An adapted form of this experiment would need to be conducted to measure the long-term impact of these interventions on corruption.

4. Empowerment Studies

Economic Empowerment

Dupas, P., and Robinson, J., 2011, 'Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya', Economics Department, University of California.
<http://www.econ.ucla.edu/pdupas/SavingsConstraints.pdf>

This paper presents the findings of a study examining whether a lack of access to formal financial services impedes business growth in poor countries. Part of a randomly selected sample of 279 self employed men and women in Kenya were provided with access to an interest-free bank account while others were not. The research found that bank accounts had large positive impacts on the total amount saved by market women in our sample. Four to six months after having gained access to the account, the daily private expenditures of market women in the treatment group were 27% to 40% higher than those of market women in the comparison group. Overall, the research suggests that informal savings mechanisms in Kenya are ineffective due to present-biased preferences (i.e. they spend the money they hold) and due to demands for money from others. The formal savings account allowed the market women in the treatment group to save as individuals, leading to increased investment and spending. The study was careful to exclude those who might have anticipated receiving a loan before the intervention. It also examines a range of alternative mechanisms for the perceived impact of the treatment.

Ashraf, N. et al, 2009, 'Female Empowerment: Impact of a Commitment Savings Product in the Philippines', Harvard Business School, Working Paper 09-100.

http://www.3ieimpact.org/admin/impact_evaluations/Impact%20of%20a%20Commitment%20Savings%20Product%20in%20the%20Philippines.pdf

This research paper presents the findings of an RCT that aimed to examine whether access to an individually-held savings product called SEED (Save, Earn and Enjoy Deposit) led to an increase in female decision-making within the household. The strongest results related to married women assigned to the treatment group, who displayed a significant increase in the research project's decision-making index. This change was not observed within the married men in the treatment group. A further finding to emerge was that access to a savings account positively impacts women's self-perception of savings behaviour, with time-inconsistent females reporting increased discipline with regards their savings. The paper concludes that possessing a savings account could encourage a woman to save. In the conclusion, the study explores a range of mechanisms via which the commitment savings product could change household dynamics.

Banerjee, A. and Duflo, E., 2008, 'Mandated Empowerment Handing Antipoverty Policy Back to the Poor?', Annals of the New York Academy of Sciences 1136, pp.333-341.

http://www.3ieimpact.org/admin/impact_evaluations/mandated%20empowerment.pdf

This paper reviews a number of studies (including several RCTs) to question the assumption that poor people want to take responsibility for service delivery and run their own businesses. They find that there is little evidence that most poor entrepreneurs want to expand their businesses and that most parents opt not to get involved in the school system despite having the opportunity to do so. This study does not reflect on the use of RCTs to measure these kinds of empowerment interventions but illustrates the way in which RCTs can be aggregated to challenge the assumptions that underpin policy decisions.

Community empowerment

Gugerty, M. and Kremer, M., 2008, 'Outside Funding and the Dynamics of Participation in Community Associations', American Journal Of Political Science, Vol. 52, No. 3, pp. 585-602

http://www.wcfia.harvard.edu/sites/default/files/Kremer_Outside.pdf

This study takes advantage of random assignment in a programme supporting women's community associations in Kenya. The project was designed and conducted by a Dutch NGO - International Child Support. The NGO identified 100 operational women's groups in two poor rural districts. 80 groups were selected and of these 40 were supported and 40 were told that they would be supported in the future. It finds little evidence that outside funding expanded organisational strength, but found 'substantial evidence that funding changed group membership and leadership, weakening the role of the disadvantaged' (p.585). The study uses a stylised model of group dynamics to interpret the data, which was based on the empirical findings. The study generates some findings about the conditions under which funded groups will attract higher-status members and leaders.

Casey, K., Glennerster, R., Miguel, E., 2010, 'Experimental Evidence on External Aid and Community Governance in Sierra Leone', Draft Paper.

http://mitsloan.mit.edu/neudc/papers/paper_257.pdf

This paper evaluates the impact of a 'community driven development' (CDD) programme in Sierra Leone (the World Bank-funded 'GoBifo' Project) using a randomised methodology. It uses data combining survey instruments with a 'set of novel field activities designed to measure how real world collective action and participation changed over the four year programme' (p.0). The study finds that there is evidence for positive impacts 'on the establishment of local development committees, the stock and quality of local public goods, strengthened links between communities and local government officials, household

economic welfare, and village-level market activity. We do not find any programme impacts on community social norms, the role of women and youths, more egalitarian decision making or the capacity for collective action beyond the immediate project sphere' (p.0). The study concludes that CDD projects 'may leave communities materially better off but may be less effective in fundamentally transforming local institutions or power dynamics' (p.0).

The study is based on an unusually large sample size of 236 villages and 2,832 households. It also includes follow-up data gathered in 2009, four years after the project was launched, which 'is especially important given the potentially slowly evolving nature of the institutional and social outcomes we study' (p.3). The study uses three novel field activities 'designed to directly observe and measure communities making real-world decisions in three dimensions: (i) a community choice between two relatively comparable alternatives... (ii) the community's ability to manage an asset they were given for free ... and (iii) the community's ability to raise funds in response to a matching grant opportunity' (p.3). 'We feel these field activities capture actual local collective action capacity more accurately than lab experiments, hypothetical vignettes, or survey reports alone, and we are unaware of other studies that have used them before' (p.3).

Rassekh, B. and Segaren, N., 2009, 'Chapter 4: Review of Community Empowerment Strategies for Health', in Peters, D., El-Saharty, S., Siadat, B., Janovsky, K., Vujicic, M. (eds.) 'Improving Health Service Delivery in Developing Countries: From Evidence to Action', World Bank, Washington DC
http://www.enrecahealth.dk/calendar/events/news/e-learning/Improving_Health_Service_Delivery.pdf/

This chapter evaluates the published literature on community empowerment strategies with respect to health outcomes in low and middle income countries based on a systematic literature search. Nearly all (89%) of the good quality studies examined involved an educational component, with educating community health workers the most common strategy. More than 60% of the studies involved training community health workers. Less than one fifth of the studies involved a mechanism to improve accountability, with the most common approach involving joint monitoring of services (9% of studies). The chapter aggregates the findings of a large number of RCTs to draw broad conclusions about which empowerment strategies are most effective. It also examines a number of contextual factors (such as strong local leadership, involvement of non-state providers) and assesses which of these are most closely correlated with success.

The chapter acknowledges that there are some limitations to this kind of analysis. First, a systematic review tends to 'show what has worked for a common set of interventions rather than how or why more complex strategies have worked' (p.139). The studies showed a variety of outcomes across a number of different contexts, which were not accounted for in the analysis presented here. It notes that the review's premise 'that certain factors may be generalizable to empowerment strategies as a whole is overly simplistic (p.139). The chapter concludes by stating that there is a need to understand better how different strategies work.

Passenger empowerment

Habyarimana, J. and Jack, W., 2009, 'Heckle and Chide: Results of a Randomized Road Safety Intervention in Kenya', Center for Global Development, Working Paper No. 169.
http://www.cgdev.org/files/1421541_file_Habyarimana_Jack_Heckle_FINAL.pdf

This paper presents a field experiment that seeks to assess the impact of a passenger empowerment intervention on road safety. The experiment involved placing evocative messages in minibuses encouraging passengers to speak out against bad driving. The study found that insurance claims in the treatment minibuses fell by between a half and two-thirds, with claims involving injury or death falling by at least 50 percent. A driver survey helped to suggest that this reduction was due in part to passenger heckling. Unlike most of the current randomised studies on empowerment or accountability, which focus

on the provision of public services such as health or education, this study is unusual because it focuses on the impact of consumer empowerment on the delivery of a privately provided service.

4. Additional information

Author

This query response was prepared by Oliver Walton, oliver@gsdrc.org

Selected websites visited

3ie, Boston Review, Centre for Global Development, Chris Blattman's Blog, Columbia University, E-gap, Google, Google Books, Google Scholar, Harvard University, IDS, Informaworld, J-PAL, MIT, NBER, Sage Publications, University of California, World Bank, Yale University

Experts consulted

Benjamin Olken, MIT
Hugh Waddington, 3ie
Harry Jones, ODI
Michael Clemens, Centre for Global Development
Devra Moehler, University of Pennsylvania
Chris Blattman, Yale University
Clara Delavallade, J-PAL
Kelly Bidwell, J-PAL
Anjini Mishra, 3ie
Birte Snilstveit, 3ie
Jessica Gottlieb, Stanford University

About Helpdesk research reports: Helpdesk reports are based on two days of desk-based research. They are designed to provide a brief overview of the key issues; and a summary of some of the best literature available. Experts are contacted during the course of the research, and those able to provide input within the short time-frame are acknowledged.

Need help finding consultants? If you need to commission more in-depth research, or need help finding and contracting consultants for additional work, please contact consultants@gsdrc.org (further details at <http://www.gsdrc.org/go/consultancy-services>).