Impact Evaluation Report 3

# The GoBifo project evaluation report

## Assessing the impacts of community-driven development in Sierra Leone

Katherine Casey, Rachel Glennerster and Edward Miguel

October 2013

International Initiative for Impact Evaluation

3ie

**About 3ie**

The International Initiative for Impact Evaluation (3ie) was set up in 2007 to meet growing demand for more and better evidence of what development interventions in low- and middle-income countries work and why. By funding rigorous impact evaluations and systematic reviews and by making evidence accessible and useful to policymakers and practitioners, 3ie is helping to improve the lives of people living in poverty.

**3ie Impact Evaluations**

3ie-supported impact evaluations assess the difference a development intervention has made to social and economic outcomes. 3ie is committed to funding rigorous evaluations that include a theory-based design and that use the most appropriate mix of methods to capture outcomes and impact usefully in complex development contexts.

**About this report**

3ie accepted the final version of this report, *The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone,* as partial fulfilment of requirements under grant OW1.65 issued under Open Window 1. The content has been copyedited and formatted for publication by 3ie. Due to unavoidable constraints at the time of publication, a few of the tables or figures may be less than optimal. All of the content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its Board of Commissioners. Any errors and omissions are also the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Any comments or queries should be directed to the corresponding author, Katherine Casey at kecasey@stanford.edu

Funding for this impact evaluation was provided by 3ie's donors, which include UKaid, the Bill and Melinda Gates Foundation, Hewlett Foundation and 12 other 3ie members that provide institutional support. A complete listing is provided 3ie website at http://www.3ieimpact.org/en/about/3ie-affiliates/3ie-members/

Suggested citation: Casey, K., Glennerster, R. and Miguel, E., 2012. *The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone. 3ie Impact Evaluation Report 3.* New Delhi: International Initiative for Impact Evaluation (3ie)

3ie Impact Evaluation Report Series executive editors: Jyotsna Puri and Beryl Leach
Managing editors: Radhika Menon and Howard White
Assistant managing editor: Kanika Jha
Production manager: Lorna Fray
Assistant production manager: Rajesh Sharma
Copy editor: Lucy Southwood
Proofreader: Proof Reading Service
Designer: John F McGill
Cover design: Via Interactive
Cover photo: UK Department for International Development

**The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone**

October 2013
3ie Impact Evaluation Report 3

**Katherine Casey**
Stanford University Graduate School of Business

**Rachel Glennerster**
Abdul Latif Jameel Poverty Action Lab (J-PAL)

**Edward Miguel**
University of California, Berkeley

International Initiative for Impact Evaluation

## Acknowledgements

This evaluation was carried out by the Evaluation Unit of the Institutional Reform and Capacity Building Project (IRCBP) with technical advice from the Jameel Poverty Action Lab/Innovations in Poverty Action (JPAL/IPA) and the University of California (Berkeley)'s Center for Evaluation of Global Action (CEGA). Katherine Casey served as the Evaluation Specialist at the IRCBP for the first two years of this research project. The research would not have been possible without generous financial support from the GoBifo Project, the IRCBP, the World Bank Development Impact Evaluation initiative, the Horace W. Goldsmith Foundation, the International Growth Centre, the International Initiative for Impact Evaluation and the National Bureau of Economic Research African Successes Project (funded by the Gates Foundation).

We wish to thank the GoBifo Project staff – Minkahil Bangura, Kury Cobham, John Lebbie, Dan Owen and Sullay Sesay – and IRCBP staff – Liz Foster, Emmanuel Gaima, Alhassan Kanu, S.A.T. Rogers and Yongmei Zhou – without whose cooperation this research would not have been possible. We are grateful for excellent research assistance from John Bellows, Mame Fatou Diagne, Mark Fiorello, Philip Kargbo, Angela Kilby, Gianmarco León, Tom Polley, Tristan Reed, Arman Rezaee, Alex Rothenberg and David Zimmer. Jim Fearon, Brian Knight, Kaivan Munshi, Gerard Roland, Ann Swidler and seminar audiences at the Center for Global Development, Brown, MIT, Northeastern Universities Development Consortium Conference, Working Group on African Political Economy and U.C. Berkeley have provided helpful comments.

## Executive summary

The World Bank and other donors dedicate sizeable portions of their portfolios to community-driven development (CDD) projects, yet until recently there has been little rigorous evidence regarding the efficacy of this approach. By emphasising local participation in and control over project implementation, CDD has come to be seen as an efficient and accountable mechanism for delivering local public goods. However, CDD aims to achieve much more than this. Through intensive, long-term facilitation, CDD aims to strengthen local institutions, enable them to become more democratic and inclusive of marginalised groups and enhance the capacity of communities to engage in collective action.

This evaluation tests the extent to which CDD has achieved these goals in Sierra Leone and has several key features. First, by randomly assigning project participation across a large pool of eligible communities, the experimental design provides rigorous evidence regarding the causal effects of the programme. Second, the research team followed communities over four years, allowing it to capture changes in behaviour that are likely to evolve only slowly over time. Third, by using a rich set of survey techniques, as well as creating a series of real-world decisions and opportunities to act collectively through 'structured community activities' (SCAs), we approach these important but elusive concepts of social dynamics from a variety of angles. Fourth, to avoid data mining, the research and project teams jointly agreed to a set of hypothesised areas of impact in 2005 before the project began, and then in 2009 the research team defined exactly which outcome measures would be used to evaluate success before analysing any of the post-project data. Finally, our relatively large and diverse sample enables us to make precise statements about even subtle changes. We use this framework to estimate (i) the direct impacts of the project during its implementation, as well as (ii) potential spillovers onto other non-project realms of local affairs and (iii) the lingering effects that persisted after the project itself had ended.

More specifically, one might expect a range of impacts from any development intervention, ordered by the reach of their influence and the difficulty of their attainment. The first stage is actually to implement projects in communities, which GoBifo accomplished quite successfully (see outcome Family A in Table 1 below). The programme achieved what it intended: it established village-level structures and tools to plan and manage development projects; it provided communities with financing and guidance to implement small-scale projects; and it created links between these processes and local government institutions. Moreover, the contributions to and benefits from the sponsored projects were distributed broadly and equitably, and the leakage of project resources appears to be minimal. The extreme poverty, recent recovery from civil war and endemic struggles against corruption in Sierra Leone make these achievements impressive.

**Table 1 Summary of evaluation results**

| Hypotheses (H) by family | | GoBifo mean effect (standard error) |
|---|---|---|
| **Family A: Project implementation** | | **0.687** (0.062)** |
| H1 | GoBifo creates functional development committees (7 outcomes) | 0.687** (0.062) |
| **Family B: Development infrastructure or 'hardware' effects** | | **0.273** (0.032)** |
| H2 | GoBifo increases the quality and quantity of the local public services infrastructure (16 outcomes) | 0.164** (0.040) |
| H3 | GoBifo improves general economic welfare (14 outcomes) | 0.399** (0.047) |
| **Family C: Institutional and social change or 'software' effects** | | **0.029 (0.019)** |
| H4 | GoBifo increases collective action and contributions to local public goods (15 outcomes) | 0.041 (0.042) |
| H5 | GoBifo enhances inclusion and participation in community decisions, especially for vulnerable groups (43 outcomes) | 0.001 (0.031) |
| H6 | GoBifo changes local systems of authority (25 outcomes) | 0.048 (0.036) |
| H7 | GoBifo enhances trust (11 outcomes) | 0.042 (0.064) |
| H8 | GoBifo builds groups and networks (12 outcomes) | 0.033 (0.044) |
| H9 | GoBifo increases access to information about local governance (19 outcomes) | 0.003 (0.039) |
| H10 | GoBifo increases participation in local governance (15 outcomes) | 0.114** (0.047) |
| H11 | GoBifo reduces crime and conflict (8 outcomes) | 0.028 (0.054) |
| H12 | GoBifo fosters more liberal political and social attitudes (9 outcomes) | 0.034 (0.041) |

Note:
The GoBifo mean effect indices calculate the average treatment effect across all outcomes under a given hypothesis and are expressed in standard deviation units.
Significance levels indicated by + $p < 0.10$; * $p < 0.05$; ** $p < 0.01$
Specifications include robust standard errors clustered by village for panel data, fixed effects for the local council wards (unit of stratification) and control variables (total households per community, distance to nearest motorable road, index of war exposure and index of history of domestic slavery.
These mean effects are limited to full sample outcomes and thus exclude conditional outcomes (i.e. those that depend on the state of another variable – for example, the quality of the infrastructure depends on the existence of the infrastructure).

Moving a step beyond this, successfully injecting financial and human resources into communities should have immediate impacts on the 'hardware' aspects of development, leaving them better off in terms of direct economic measures (see Family B in Table 1). We find that the treatment communities have a larger stock of local public goods that are of higher quality than the control areas. We find further that beneficiaries are better off in terms of household assets and that there is more market activity – for example, more petty traders and goods on sale – in their villages. This suggests that the programme resources were invested in projects and activities that improved the general welfare of the recipient communities. GoBifo-funded projects further attracted the attention and involvement of local leaders. We find that ward development committee (WDC) members and chiefdom representatives play a more active role in the planning, construction and oversight of local public goods in the treatment areas.

Turning from the proximate impacts tied directly to programme activities to the indirect influences on local norms and institutional practices, we find no evidence that GoBifo led to fundamental changes in the way in which community members interact or in their capacity to act collectively outside the immediate sphere of the project (see Family C in Table 1). Using a wealth of measures, we find no impacts on any of the five proxies for social capital – trust, collective action, groups and networks, access to information, and inclusion and participation – emphasised in the project objectives. As a specific example of broader social change, CDD emphasises a greater role for women in local decision-making as part of its focus on inclusion and participation. GoBifo facilitators thus encouraged women's participation in meetings, required women to serve on the village development committee (VDC) and as co-signatories to community bank accounts and provided resources for them to manage their own projects. Nevertheless, despite these experiences, women in GoBifo areas are no more likely to speak up in a general community meeting. Overall, we find no evidence that the GoBifo process – effective as it was in establishing participation and collective action within the project – led to greater participation or empowerment of marginalised groups in local affairs and decisions outside the project.

The second major component of CDD's more far-reaching objectives involves teaching communities how to help themselves in the long run by giving them the experiences and skills necessary to undertake initiatives independently after the programme has ended. The idea is that once communities have the institutions in place – a VDC, a development plan, a bank account and experience in budgeting and management – they should be better able to take advantage of the new opportunities that arise. However, we find that the treatment communities were no more likely to organise themselves and take up a real-world building materials subsidy programme than the controls. While GoBifo was successful in generating collective action within the project, there is no evidence that GoBifo served as a catalyst for collective action beyond the activities stipulated by the project itself.

The results here reflect the position in communities four years after the project was launched. While four years may be a short time in comparison with the lifetimes over which the current institutions developed, it is not a short time in comparison with the time scales of these types of projects. Other interventions (quotas for female leaders in India, anti-caste-based voting campaigns) have had measurable, and dramatic, impacts on institutions (including the participation of women) in similar time frames, albeit in different contexts. While the long-term impact of CDD has yet to be rigorously evaluated, there is a real question regarding how realistic it is to expect that changes in social norms will be stronger years after the project has ended than they are shortly after the period of intense facilitation. Moreover, the project chose two districts balanced along key dimensions – including ethnic composition, region and political affiliation – and we find very similar results in the diverse operational settings.

Overall, the experience in Sierra Leone suggests that CDD is a reasonable approach to deliver small-scale local public goods in a way that is equitable, accountable and low cost. In this sense, an expanded programme could serve as an effective community-level complement to existing development efforts, including those by the National Commission for Social Action (NaCSA) and the local councils. Our evidence suggests that CDD was not, in this context, an effective agent of institutional and social change. As such, it did not advance the presidential 'agenda for attitudinal change' nor did it specifically empower women and youths (adults aged 18 to 35 years) in local development processes outside the specific GoBifo programme. It is important to note that we cannot rule out the focus on participation and inclusion *within* GoBifo activities being in part responsible for its achievements in building local public goods and enhancing economic welfare. What we can say is that influencing the way in which communities organise themselves *outside* the project activities – including how members make decisions and act collectively in other realms of local affairs and in times after the project has ended – is incredibly difficult, especially from the outside. These results raise serious questions about whether we yet know how to do this effectively.

## Contents

## List of figures and tables

**Figures**



**Tables**



## Abbreviations and acronyms

| | |
|---|---|
| CDD | community-driven development |
| CEGA | Center for Evaluation of Global Action (University of California, Berkley) |
| H | hypothesis |
| IDPs | internally displaced persons |
| IPA | Innovations for Poverty Action |
| IRCBP | Institutional Reform and Capacity Building Project |
| J-PAL | Abdul Latif Jameel Poverty Action Lab |
| KDP | Kecamatan Development Program |
| LC | local councillor |
| LPG | local public goods |
| NaCSA | National Commission for Social Action |
| NGO | non-governmental organisation |
| NSAP | National Social Action Project |
| PHU | public health unit |
| SCAs | structured community activities |
| SLL | Sierra Leonean Leone |
| SSL | Statistics Sierra Leone |
| SUR | seemingly unrelated regressions |
| TBA | traditional birth attendant |
| VDC | village development committees |
| VDP | village development plan |
| WDC | ward development committee |

# 1. Background

## 1.1 Community-driven development and Sierra Leone

International donors, governments and non-governmental organisations (NGOs) invest substantial amounts of resources in the community-driven development (CDD) approach, yet until recently there has been little rigorous evidence regarding its efficacy. For the World Bank alone, over 9 per cent of its total lending supports CDD projects, placing the magnitude of investment in billions of dollars (World Bank 2007). By emphasising local participation in and control over project implementation, CDD aims to provide public goods through a process that empowers the poor. While advocates promise a long and varied list of benefits ranging from more efficient and cost effective infrastructure construction to the transformation of authoritarian institutions, critics hold concomitant concerns that participation requirements function as a regressive tax and that project benefits are easily captured by local elites. This study contributes to a concerted effort to evaluate the claims on both sides of this debate and gather rigorous evidence regarding the impacts of CDD.

While the emphasis on participation stems from the intrinsic value of civic engagement and empowerment, it further relates to a broader theme in development that institutions matter for economic growth. Academics and development practitioners agree that strengthening the transparency, accountability and inclusiveness of institutions could be an important precondition for development – as well as being important objectives in their own right. However, they also acknowledge that the types of interventions that could successfully make progress towards these goals remain unclear. CDD has become one very popular mechanism in the attempt to build such institutions, aiming to influence community dynamics and authority structures beyond the immediate sphere of the project activities. Its success in enhancing local governance in particular also remains largely unsubstantiated. The GoBifo Project (which means 'move forward' in the national *lingua franca*, Krio) and impact evaluation thus aim to assess the overall effectiveness of CDD as well as its specific impacts on institutions in Sierra Leone.

The recent history of poor governance and civil war, coupled with widespread and chronic poverty, makes Sierra Leone a challenging and appropriate place for the CDD approach. In particular, three typical CDD components – the provision of basic infrastructure, a focus on empowering marginalised groups in local decisions and explicit ties to local governance structures – align well with the recent history and current priorities of the country. First, the contemporary United Nations Human Development Index – which compiles welfare measures like life expectancy, education and standards of living – ranks the country 180th out of 182 (United Nations 2004). In response to the underlying gaps in service provision and low household incomes, GoBifo provided block grants to communities to contribute to building local public goods like schools, latrines and grain drying floors and/or sponsoring skills training and income-generating activities. These grants were deposited in community accounts that GoBifo facilitators helped communities to open and that were then available for other community projects. Second, as many observers suggest that the disenfranchisement of young men created frustration and anger, which helped fuel Sierra Leone's brutal civil war (1991–2002), GoBifo facilitators placed special emphasis on enhancing the voice of youth (defined as consisting of adults aged 18 to 35 years) in local decision-making. Along similar lines, since women have historically held less power in local governance than men, GoBifo encouraged women to participate actively in all aspects of project planning and implementation, including managing their own sub-projects. Third and finally, the Government of Sierra Leone reconstituted its system of local councils in 2004, over 30 years after they were abolished under the one-party state of President Siaka Stevens. These democratically elected politicians represent wards, which are sub-district

administrative units each comprising roughly 10,000 citizens. Housed within the Decentralisation Secretariat, GoBifo provided 'bottom-up' support to the decentralisation process and local government by giving funding and technical assistance to both village development committees (VDCs) and ward development committees (WDCs). By coordinating development planning and activities between the two levels, GoBifo created links between the citizens and the local politicians who represent them.

### 1.2 Precursors to the GoBifo project

The GoBifo project designers aimed to develop innovative solutions to the challenges that two existing World Bank-supported programmes were encountering and provide a more localised complement to their ongoing operations. Specifically, the original grant proposal to the Japanese Social Development Fund emphasised the need to pilot innovative approaches aimed at strengthening village-level institutions and empowering women and youths. An assessment of the National Social Action Project (NSAP) – a major vehicle for reconstructing basic infrastructure destroyed in the war 'identified the lack of community social capital as the main challenge for implementation of CDD in Sierra Leone … Ex-combatants, internally displaced persons (IDPs), and in particular women and youth, are marginalised and unable to participate in development activities and decision-making processes at the local level' (World Bank 2004). In response, GoBifo aimed to 'strengthen social capital by enhancing the capacity of villages and local governments to design and implement strategic development plans at village and ward levels' (World Bank 2004).

In addition, the Institutional Reform and Capacity Building Project (IRCBP) was created to support the newly elected local councils, providing them with technical assistance and funding as well as guidance in establishing their respective WDCs. GoBifo aimed to extend this work down a step further, to see how effectively village-level structures could manage development resources and how best to link them up with their ward-level representatives. The idea was to foster a village-level development planning process that fed directly upwards into ward- and council-level activities. As such, it aimed to 'provide focused information, facilitation, training and technical support to ensure that vulnerable groups are able to participate fully in local development planning and to improve governance, transparency and accountability mechanisms'. Taking these two strands together, the project 'complements both 'supply' and 'demand' initiatives already ongoing in Sierra Leone' (World Bank 2004).

Led by the Evaluation Unit of the IRCBP, with technical assistance from JPAL, IPA and CEGA, this impact evaluation assesses the extent to which GoBifo achieved these goals. Specifically, the research investigates the project's impacts on the creation of local institutions, the connections between villages and ward-level government, the stock and quality of local public goods, general economic welfare, social capital and participation in local governance. For a presentation of results tailored to an academic audience, see Casey, Glennerster and Miguel (2011).

### 1.3 CDD evaluation literature

In their critical review, Mansuri and Rao (2004) note that the active involvement of community members in the design and implementation of development projects is the defining feature of CDD. Such participation aims to leverage local knowledge and information to improve programme targeting, ensure equitable distribution of benefits, align investment with local preferences, enhance cost-effectiveness, reduce corruption and empower beneficiaries through learning by doing. At the same time, participation is costly: it effectively shifts some of the burden of service provision onto beneficiaries, and its effort-intensive social transformation aspect may be readily abandoned under programmatic pressures to deliver infrastructure quickly and efficiently. The authors' main point is that the

accumulated evidence regarding the efficacy of CDD 'lags well behind the rate at which projects are being implemented and scaled up'. A number of studies have taken up this challenge, with mixed results.

Considering the Philippines, Labonne and Chase (2008) use propensity score matching and household panel data to estimate the impacts of a World Bank-funded CDD project on social capital. Their analysis suggests that the project preparation cycle increased the participation in village assemblies, the frequency of interaction between village leaders and residents, and trust in strangers. At the same time, they find evidence that CDD may serve as a substitute for other associational activities and contributions as they see negative impacts on group membership, collective action and village-level investments in other projects.

Investigating what many consider a flagship CDD project, Voss (2008) analyses the impacts of the Kecamatan Development Program (KDP) in Indonesia on household welfare and access to services using propensity score matching on panel data. While the KDP had no overall impact on consumption per capita, he finds positive and significant gains for the poorest quintile matched by significant losses for the richest quintile. Notably, despite strong project emphasis on incorporating women into the development process, this analysis uncovers negative consumption impacts for female-headed households. The author further finds evidence of improvements in access to outpatient health care for the sick, reductions in unemployment and no impact on school enrolment. Focusing on roads constructed under the same KDP project, Olken (2007) finds that enhanced top-down project monitoring – through guaranteed government audits – was more effective in reducing corruption than increased grassroots participation in village-level accountability meetings between residents and project officials. A related set of papers exploring the impacts of community mobilisation on public service providers similarly finds mixed results with strong positive effects seen for health care in Uganda (Bjorkman and Svensson 2009) but no effect on education in India (Banerjee *et al.* 2010).

Most similar to this study in terms of context and empirical methods, Fearon, Humphreys and Weinstein (2009) present preliminary results from a randomised field experiment of a community-driven reconstruction project in northern Liberia implemented by the International Rescue Committee. While their full academic paper is not yet available, their basic result of positive impacts on collective action and social cohesion – as measured by greater contributions to an experimental public goods game in the mixed-gender treatment arm (although there were no impacts in the women-only treatment arm) and reduced self-reports of inter-group tensions – accompanied by little effect on hardware appears quite the opposite to our findings. However, closer inspection reveals commonalities. Fearon *et al*. do find positive impacts on female employment and positive though insignificant effects on total household assets. In addition, their public goods game results are driven mainly by high contributions from IDPs, while there are few remaining IDPs in our research sites. More speculatively, the Liberia programme operated in what was the 'epicentre' of the latter years of that country's civil war, and thus may have faced more disruption to local institutions than the Sierra Leone programme did. Attempts to create new institutions and norms where formal structures have broken down may encounter less resistance than efforts to persuade the existing authorities to adopt new practices. Crucially, though, neither our study nor that of Fearon *et al*. finds compelling evidence of programme spillovers on real-world, non-project collective activities including contributing to existing public goods (such as road maintenance, schools and wells) and attending or speaking up in community meetings.

This research project makes five main contributions to the literature. First, by randomly assigning communities to treatment (who received the project) and control groups (who participated in the research only), it generates scientifically rigorous evidence regarding the causal effects of CDD. Second, the research uses both rich standard surveys and real-world

tests of collective action and community decision-making to tackle difficult-to-measure concepts relating to community dynamics and social change. In so doing, we develop innovative new tools that we call structured community activities (SCAs), which we hope will be useful in other settings. Third, implementing the final surveys and SCAs four years after the project began allows us to explore evidence concerning longer-run effects. This is especially important given the slowly evolving nature of the outcomes of interest, for which no one expects fundamental local structures and the ways in which community members interact to change overnight. The flip side to this is that the clear lack of effects found concerning social capital cannot be dismissed as hazards of a short-term study. Fourth, the research and project teams agreed to a set of hypotheses regarding the likely areas of programme impacts in 2005 before the intervention began. As the project came to a close in 2009, they fleshed out this document with the exact outcome measures and econometric specifications that we would use to evaluate success and archived this *ex ante* analysis plan before analysing the follow-up data. Our decision to adhere rigorously to this plan eliminates the risk of data mining or other selective presentation of empirical results and generates correctly sized statistical tests, bolstering the scientific credibility of the findings. Registering *ex ante* analysis plans is standard in medical trials, but, to our knowledge, this is among the first economics studies to adopt this approach. Finally, the experiment covers a relatively large sample – a panel of 236 villages and 2,832 households. The large sample size lends statistical precision to the treatment effect estimates, lending confidence to both the positive impacts found on local public goods and household welfare and the precisely measured zeros seen in other areas. In other words, it is highly unlikely that our positive hardware effects are due to random chance and we are able to rule out even relatively small changes in software.

In terms of contributing to larger debates concerning the effectiveness of CDD, this study suffers from two key limitations. First, it cannot test directly whether or not CDD is more accountable or efficient than standard top-down models. The research framework includes no direct comparison with other approaches: it explores only the impact of CDD compared with a control group that received nothing. While the project under study successfully delivered local public goods and the communities appear to have made modest grants spread a long way, the results do not prove whether this mechanism of delivery was any better or worse than more centralised provision would have been. Related to this, the following analysis cannot distinguish the effect of facilitation from the effect of the cash grants. While we would have liked to test the effect of money plus facilitation against money alone (or against facilitation alone), this particular treatment was a uniform package.

## 2. GoBifo project description

GoBifo project managers worked in cooperation with international consultants from the Royal Tropical Institute to develop an intervention that combined intensive community facilitation in development planning with untied block grants to implement the projects identified during the planning process. Their resulting design aimed to deliver local public goods while at the same time strengthening local institutions and making community decisions more inclusive and democratic.

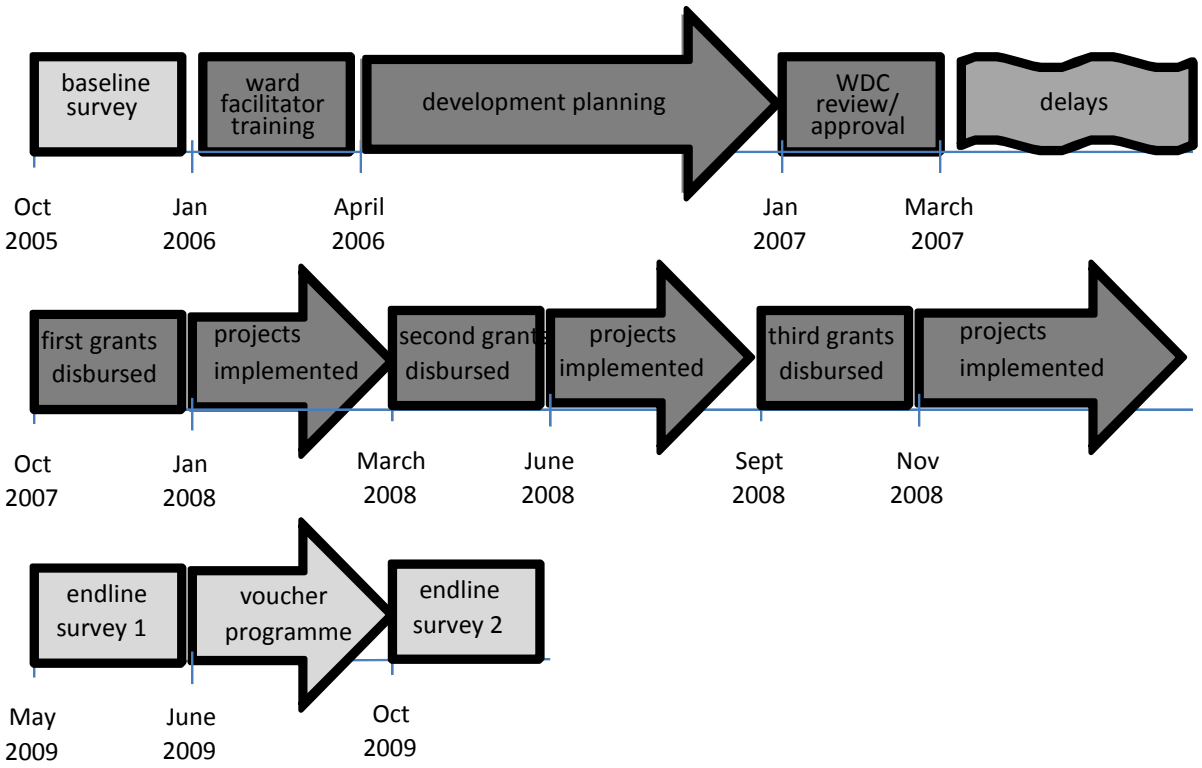
In terms of the overall budget shares, just under half of the total budget was dedicated to village- and ward-level block grants (USD896,000 or 47 per cent) with the balance covering capacity development in village- and ward-level planning (USD589,732 or 30 per cent), project management and contingencies (USD255,320 or 14 per cent), and monitoring and evaluation (USD177,300 or 9 per cent). Thus, for every dollar spent directly on community

projects, roughly USD1 was spent ensuring the money was used well through facilitation, administration and oversight.

Each GoBifo front-line staff member, or ward facilitator, covered six village-level and one ward-level intervention. The facilitators were required to reside in one of the six villages assigned to them and spend approximately one day per week in each of the remaining villages. They began work in January 2006 and completed all the village-level projects by July 2009. This implies that each village received roughly six months of direct facilitation over a period of three and a half years.

To summarise the main objectives of each village and ward, the GoBifo Ward Facilitators helped community members: i) to establish a VDC or WDC; ii) to articulate a development plan; and iii) to implement projects identified in the plan using a series of block grants (see the timeline in Figure 1). Regarding the first goal, the project viewed the VDCs and WDCs as the central institutions to guide and manage local development, thus strengthening these institutions became a primary objective. As background, Richards *et al*. (2004) suggest that many VDCs were introduced by humanitarian assistance groups in the 1990s, and at the time the project began roughly half of the villages under study had one. The WDCs were set up in conjunction with the re-establishment of the local councils in 2004. The Local Councillor representing the ward serves as the WDC chair. The GoBifo project operated this three-stage approach – institution building, development planning and project implementation – at the village and ward levels simultaneously. Note, however, that the experimental evidence presented in this report concerns only the village-level programming, as this was the only feasible level of randomisation. Thus, the discussion below also focuses exclusively on the village-level interventions.

**Figure 1 Project and evaluation timeline**

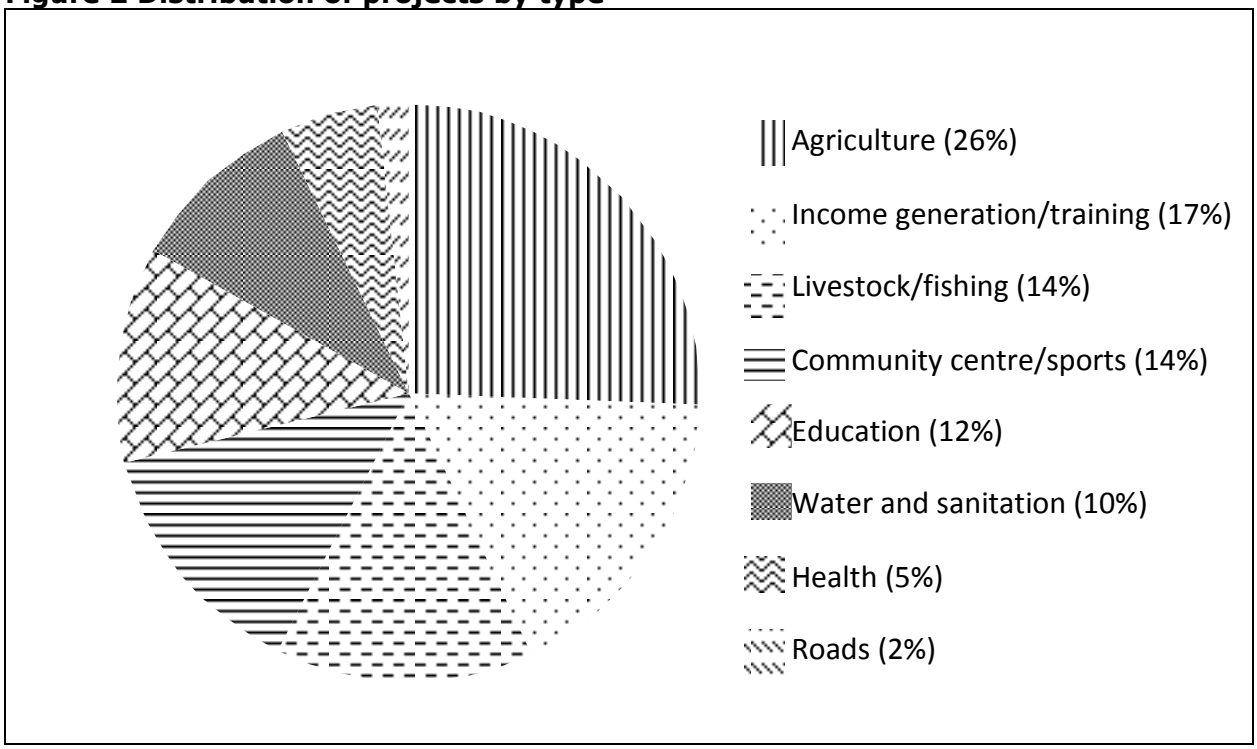


Where the VDC was absent or defunct, the first step taken by the GoBifo facilitators was to help community members choose who would represent them on the VDC and formalise the structure. Where it was present but less functional, the GoBifo Ward Facilitators refreshed the institution by training the VDC members in their roles and responsibilities. The GoBifo facilitators further trained the VDC members in bookkeeping, budgeting, financial management, procurement, project planning and oversight. They also helped each GoBifo community to open a bank account and select three members (including at least one woman) to serve as account co-signatories.

Drafting a village development plan (VDP) was the second major initiative undertaken in each village. Over a period of several months, the GoBifo facilitators used a variety of participatory tools to assist community members in identifying their own resources, priorities and constraints. Beginning with a 10-year vision, the process gradually narrowed in focus to a three-year strategic plan for local development and then a one-year action plan that included specific uses for the GoBifo grants. The participation and inclusion of marginalised groups served as cornerstones of this process. As an example, the facilitators divided the community into different social groups – women, youths, adult men – and allowed each group to come up with its own development plan. They then helped the community as a whole to combine these into a single unified vision. This approach allowed women and youths to voice their concerns and ideas during the break-out sessions, which they may have been uncomfortable articulating in the general meeting. Throughout the process, the facilitators guided the community members through methods to identify and leverage their social capital, as measured by five dimensions of trust, collective action, inclusion, networks and information. Once completed, the VDC submitted the development plan to the appropriate WDC for review, endorsement and onward transmission to the district councils for approval. The project thereby explicitly linked the village-level planning process to the lowest tier of the elected government.

The third and final stage of the GoBifo intervention involved the use of untied block grants to implement the projects identified in the Village and Ward Development Plans. Each community received a total of Sierra Leonean Leone (SLL) 14 million (USD4,667) divided into three separate disbursements. Typical projects selected include: i) local public goods construction, like schools, road rehabilitation, latrines, community stores and community centres; ii) agricultural production and livestock management, for example communal farms, grain drying floors and goat herding; iii) skills training and income-generating activities, like soap making and carpentry; and iv) social projects, for example youth football clubs and equipment. Figure 2 below provides the distribution of projects across categories. For the second and third disbursements, the GoBifo Ward Facilitators monitored the community's progress in implementing projects and the transparency of financial management for funds already received before releasing the next tranche. Note that there was a substantial delay between the WDC approval of projects in early 2007 and the first release of funds late that same year due to difficulties encountered in recruiting financial management staff and completing the operations manual. On the downside, this delay created frustration and impatience for the participants and staff; on the upside, it lengthened communities' exposure to facilitation well beyond that envisioned in the original project plan.

**Figure 2 Distribution of projects by type**




A key question regarding the external validity of these results is how representative the GoBifo project is of CDD initiatives in general. The project implementation stages – establishing a local development committee, providing technical support and allocating block grants – are quite standard, as is the pervasive emphasis on inclusive, transparent and participatory processes. Compared with projects studied in other countries (Olken 2007; Labonne and Chase 2008), the most notable programmatic difference is that the village-level component of GoBifo did not involve any inter-community competition for funding. Regarding the scale of the funding, GoBifo disbursed grants worth a little under USD5,000 to communities with 50 households, or 300 residents, on average (so roughly USD100 per household, or USD16 per capita over three and a half years). The Fearon *et al*. (2009) Liberia project provided roughly USD20,000 to 'communities' that comprised around four villages with 2–3,000 residents, so USD8 per capita over two years, and villages received USD8,800 in Indonesia (Olken 2007). While the difference in the total grant size may affect the maximum feasible project scale, the per capita funding differences are not substantial.

## 3. Research design

### 3.1 Hypotheses

The research and project management teams together took great care to state explicitly and up front what they expected GoBifo to accomplish and how they would measure success. Given the long and wide-ranging list of potential impacts, this step is important to avoid *ex post* rationalisation that highlights only positive impacts and hides any negative or zero effects discovered during the analysis. Thus, in 2005 (before the project began), the GoBifo managers and the evaluation team agreed to a set of hypotheses regarding the specific areas on which they expected GoBifo to impact. These hypotheses can in turn be

grouped into three major families: i) project implementation, which considers how successful GoBifo was in delivering its stated activities in communities; ii) proximate impacts or 'hardware' effects, which are more immediate consequences of building local public goods and injecting financial resources into communities; and iii) social change or 'software' effects, which ask whether or not the GoBifo experience triggered spillovers or changes in other realms of community life, including the way in which community members interact and make decisions. Table 2 below presents the specific hypotheses. Note that the research team submitted a document that lists for each hypothesis all the indicators from the surveys and participatory experiments to be considered *before* beginning the data analysis (see the details in Appendix C).

**Table 2 Research hypotheses by family**

| | |
|---|---|
| **Family A**: Project implementation | H1. This is an implicit hypothesis that by delivering its interventions successfully, GoBifo creates local institutions like VDCs and equips them with development plans, bank accounts and connections to their respective WDCs. |
| **Family B**: Development infrastructure or 'hardware' effects | H2. Participation in GoBifo improves the quality and quantity of the local public services infrastructure.<br><br>H3. Participation in GoBifo improves general economic welfare. |
| **Family C**: Institutional and social change or 'software' effects | H4. Participation in GoBifo increases collective action and contribution to local public goods.<br><br>H5. GoBifo increases inclusion and participation in community planning and implementation, especially for poor and vulnerable groups; GoBifo norms spill over into other types of community decisions, making them more inclusive, transparent and accountable.<br><br>H6*. GoBifo changes the local systems of authority, including the roles and public perception of traditional leaders (chiefs) versus the elected local government.<br><br>H7. Participation in GoBifo increases trust.<br><br>H8. Participation in GoBifo builds and strengthens community groups and networks.<br><br>H9. Participation in GoBifo increases access to information about local governance.<br><br>H10. GoBifo increases public participation in local governance.<br><br>H11. By increasing trust, GoBifo reduces crime and conflict in community.<br><br>H12**. GoBifo changes political and social attitudes, making individuals more liberal towards women, more accepting of other ethnic groups and 'strangers' and less tolerant of corruption and violence. |

Note:
*Not an explicit objective of the GoBifo project leadership itself, but is a plausible research hypothesis. **Not part of the original programme hypotheses document, but relates closely to the GoBifo project objectives.

### 3.2 Value of randomisation

In relation to most of the existing research on CDD impacts and considering the aim of accumulating rigorous evidence regarding the specific effects of GoBifo in Sierra Leone, the most important design feature of this evaluation is random assignment. After the programme managers determined the basic eligibility requirements, a random lottery selected communities from the several hundred eligible to receive the project and to participate in the research as controls. Just as in a medical trial to test a new drug, this lottery process is critical for substantiating claims about the causal effects of GoBifo.

More specifically, random selection ensures that those participating in GoBifo are representative of all the communities in the eligible pool – that is, on average, they have the same observable and unobservable characteristics as those left out. This further implies that what we discover about the effectiveness of GoBifo in these particular communities also reflects its effectiveness in any of the others. By also randomly selecting communities to participate in the research but not the programme, we further ensure that the members of the treatment group are on average the same as those in the control group. Thus, at the start, the treatment and control communities are statistically indistinguishable from one another (as a check on this, note the lack of significant differences between the two groups in Table 3 on page 13). Over time, we would therefore expect these communities to evolve in similar ways. However, by operating GoBifo only in the treatment communities, we introduce a single systematic difference between the two groups: the project itself. Therefore, any differences we detect between the treatment and the control communities after the programme has ended can be directly attributed to the GoBifo intervention. A couple of examples illustrate why this framework is important.

Consider first the strategy of only following the GoBifo communities over time. The Ministry of Finance estimates that Sierra Leone was growing by roughly 5 per cent per year over the duration of the project. If we did not also follow the control communities, we would wonder whether the positive impacts we find on household welfare are due to participation in GoBifo or whether they simply reflect the fact that everyone in the country was gradually becoming better off. With this framework, however, we are able to isolate the positive economic impacts of the project above and beyond the general improvements that also occurred in non-GoBifo communities.

Next consider alternative ways of selecting communities. Suppose that project managers chose the communities they thought would benefit most from GoBifo activities or allowed communities to apply for entry. Following these communities and their neighbours over time reveals a positive impact on the stock of public goods. Without the lottery, however, we would not be able to tell whether this difference was due to GoBifo or instead caused by differences in the underlying aptitude for collective action that made communities attractive to the managers or enabled them to put together a successful application. Critics would thus be justified in suggesting that GoBifo actually made no difference because the treatment communities would have performed better anyway, even without the programme.

### 3.3 Sampling details

**Local councils**: Since GoBifo was a small-scale pilot with scarce resources, the project team carefully selected the local council areas, wards and communities where it would work. The selection of two districts for operation struck a balance along a number of dimensions. Starting at the regional level, the GoBifo project team wanted one area from the north and a second area from the south to maximise the political and ethnic variation. Within these regions, they targeted poorer districts using estimates of the incidence and severity of rural poverty from the Sierra Leone Integrated Household Survey. Amongst the poorest, they further prioritised relatively service-deprived areas, those struggling with higher levels of

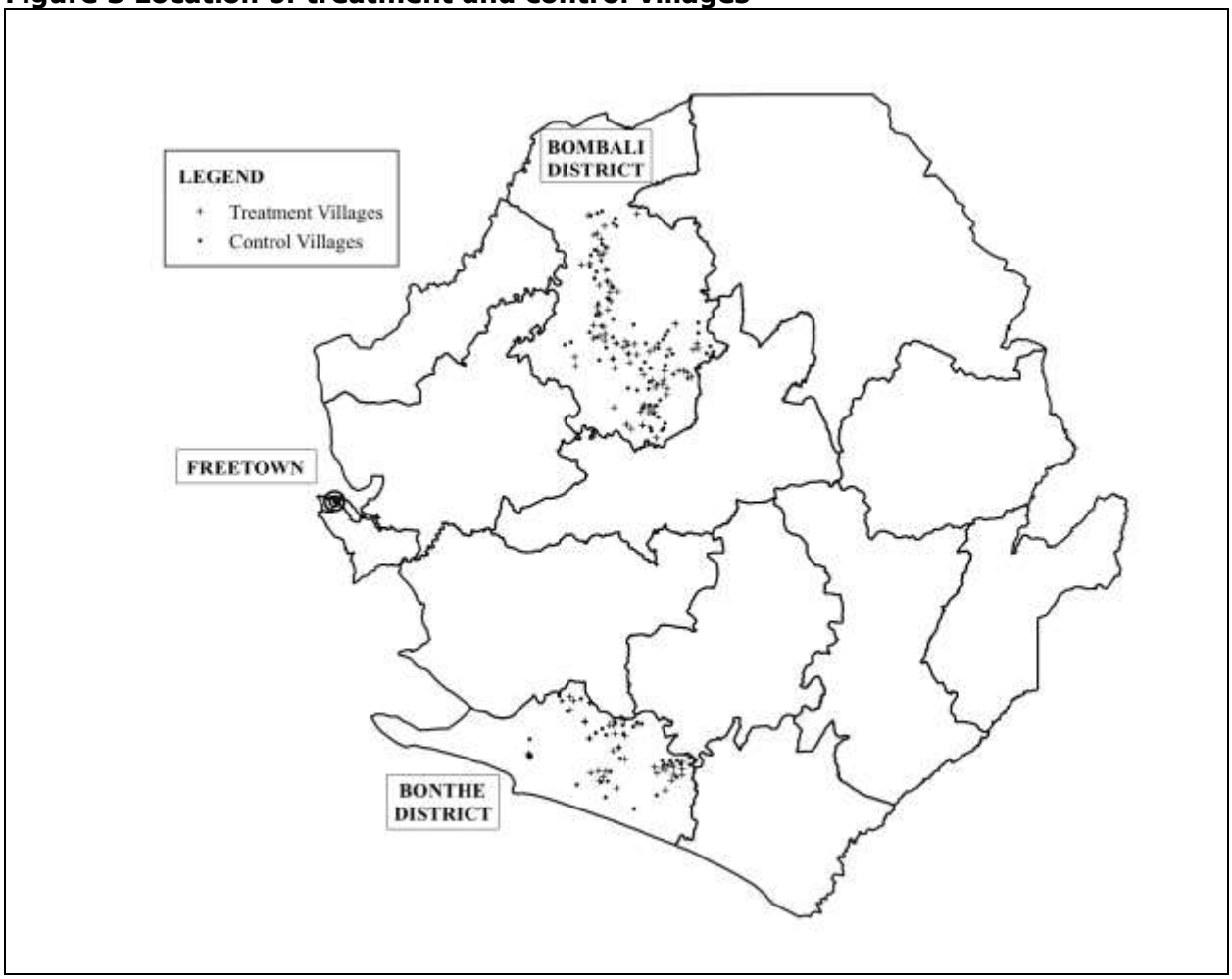
internal conflict and those with the fewest other donor/NGO activities to avoid duplication of effort. As a result, the larger district chosen, Bombali, is located in the north, the region dominated by the Temne, Limba and Loko ethnic groups and traditionally allied with the All People's Congress (APC) political party. The smaller district, Bonthe, is in the south, where the Mende and Sherbro ethnic groups dominate and are historically aligned with the Sierra Leone People's Party (SLPP). As GoBifo operates within and aims to support decentralisation, the project works in wards under each of the four relevant local councils (Bombali District, Bonthe District, Makeni City and Bonthe City).

**Wards**: The GoBifo project management selected operational wards in collaboration with the NSAP. The GoBifo managers agreed to work in all the wards under the four councils where the NSAP was not scheduled to work in 2006. This means that every ward in these four local council areas received assistance from either GoBifo or the NSAP. Since the NSAP selected the most vulnerable wards as classified by its services and opportunities mapping exercise, the GoBifo wards are likely to represent better-off communities in these councils, although by any measure the project areas are very poor. In total, 78 villages within 13 wards in Bombali District, six community development committees in four wards in the Makeni City Council, as well as 40 villages within eight wards in Bonthe District and Island benefited from the project. However, as a densely populated urban area, Makeni Town posed a distinct set of challenges for the design of GoBifo interventions, choice of operational areas and evaluation. As a result, the community development committees were not randomly selected and are therefore not part of the evaluation.

**Communities**: Within each ward, the GoBifo project and research teams worked together to select 12 communities using a two-stage process. First, the project team narrowed the list of all the communities in the 2004 Population and Housing Census to those located in their target wards and of an appropriate size for a GoBifo project: between 20 and 200 households in Bombali and between 10 and 100 households in Bonthe. Then the research team combined this restricted list with data on the distance to the nearest paved road and used a computer programme to assign six communities randomly to the treatment group and six to the control group within each ward. Particular care was taken to ensure that this process was random; within each ward, every community had an equal chance of being selected. Note that these communities are not representative of the entire country, nor even of Bombali and Bonthe. They are instead representative of all similarly sized communities in the sample of wards.[1] Figure 3 maps the location of all 236 communities participating in the research.

[1] There were two minor data issues in measuring community size and ward location that led to a partial resampling of villages; however, this did not affect the integrity of the randomisation process (see Appendix A for details).

**Figure 3 Location of treatment and control villages**




Note:
Some of the marks on this map overlap because they take up more space than the actual communities.
There is no overlap between the treatment and control communities on the ground. Indeed, rural communities in Sierra Leone tend to be relatively geographically distinct compared with those in more densely populated countries.

**Respondents**: Before the fieldwork began, Statistics Sierra Leone staff members randomly selected 12 households per community using the census household listings. Given the research interest in dynamics of exclusion and empowerment, the identification of respondents within the targeted households aimed to capture a broad range of perspectives. Rotating by household, enumerators thus interviewed four different types of respondents: non-youth male, youth male, non-youth female and youth female. All the respondents are at least 18 years old, and the Government of Sierra Leone's definition of youth includes people up to 35 years of age. This selection strategy means that for each community, and for the overall sample, the responses are roughly balanced across the four distinct groups.

**Table 3 Baseline comparison of treatment and control communities**

| Outcome measured at baseline (2005) | Mean for controls | T-C Difference | Standard error | Observations |
|---|---|---|---|---|
| **Community characteristics** | | | | |
| Total households per community | 46.76 | 0.30 | (3.67) | 236 |
| Distance to nearest motorable road in miles | 2.99 | –0.32 | (0.36) | 236 |
| Index of war exposure (range 0–1) | 0.68 | –0.01 | (0.02) | 236 |
| Historical legacy of domestic slavery (range 0–1) | 0.36 | 0.03 | (0.06) | 236 |
| Average respondent years of education | 1.65 | 0.11 | (0.13) | 235 |
| **Selected outcomes from project implementation Family A** | | | | |
| Proportion of communities with a VDC | 0.55 | 0.06 | (0.06) | 232 |
| Proportion visited by a WDC member in the past year | 0.15 | –0.01 | (0.05) | 228 |
| **Selected outcomes from 'hardware' Family B** | | | | |
| Proportion of communities with a functional grain drying floor | 0.23 | 0.05 | (0.05) | 231 |
| Proportion of communities with a functional primary school | 0.41 | 0.08 | (0.06) | 230 |
| Average household asset score | –0.06 | 0.11 | (0.08) | 235 |
| Supervisor assessment that the community is 'better off' than others nearby | 0.31 | 0.04 | (0.06) | 201 |
| Proportion of communities with any petty traders | 0.54 | –0.01 | (0.06) | 226 |
| **Selected outcomes from 'software' Family C** | | | | |
| Respondent agrees that chiefdom officials can be trusted | 0.66 | –0.01 | (0.02) | 235 |
| Respondent agrees that local councillors can be trusted | 0.61 | 0.00 | (0.02) | 235 |
| Respondent is a member of a credit/savings group | 0.25 | –0.03 | (0.02) | 235 |
| Respondent is a member of a labour–sharing gang | 0.50 | –0.01 | (0.03) | 235 |
| Among males who attended a community meeting, the respondent spoke publicly | 0.59 | –0.02 | (0.04) | 235 |
| Among females who attended a community meeting, the respondent spoke publicly | 0.29 | 0.03 | (0.04) | 229 |
| The respondent claimed to have voted in the last local elections | 0.85 | –0.01 | (0.02) | 235 |

Note:
Significance levels indicated by + $p < 0.10$; * $p < 0.05$; ** $p < 0.01$
Robust standard errors.
The T-C difference is the pre-programme 'treatment effect' run on the baseline data aggregated to the village-level mean, using a minimal specification that includes only fixed effects for the district council wards (the unit of stratification) and the two balancing variables from the randomisation (total households and distance to road).
The regressions for the two balancing variables in rows 1 and 2 exclude the outcome from the set of controls.

**Two balanced groups**: Table 3 shows the average value of the community characteristics and selected outcomes of interest under each family for the treatment and control villages before the programme began. The lack of statistically significant differences for all nineteen indicators suggests that the randomisation process successfully generated two extremely similar groups and increases our confidence that any post-programme differences we find are directly attributable to GoBifo. Specifically, in late 2005 both sets of communities on average: contained 47 households; were located roughly 3 miles from the nearest road; had similar exposure to violence during the civil war and legacies of domestic slavery; and demonstrated similar levels of respondent education (1.65 years of schooling on average). Regarding implementation family A, slightly over half of the villages had a VDC and about 15 per cent had been visited by a WDC member in the past year. In terms of economic welfare family B, treatment and control communities had similar levels of public

infrastructure, household assets and petty trading. Regarding software family C, the respondents in both groups of villages reported similar levels of trust in local officials, participation in groups (25 per cent were members of credit or savings clubs and half were members of collective labour gangs) and voting (85 per cent claimed to have voted in the last local government elections). The baseline level of participation in local affairs was higher for men in both treatment and control villages: on average, twice as many men and women reported that they spoke up in the last community meeting they attended. Note that the analysis below includes the baseline values of the outcomes considered whenever possible, which controls for any spurious pre-programme differences between treatment and control areas.

### 3.4 Data collection

This analysis draws on three major sources of data: household surveys from 2005 and 2009; village-level focus group discussions held in 2005 and 2009; and three SCAs conducted in 2009. For both the baseline and the endline exercises, extensive fieldwork using qualitative methods – including open-ended interviews, focus group discussion and participant observation – was carried out to develop and refine a set of quantitative survey measures for data collection.

**Household surveys**: Before the project began, research teams conducted in-depth household-level interviews that collected extensive data on the baseline levels of the outcomes of interest – including trust, participation in community activities and groups, knowledge about government, market activities and conflict resolution – as well as detailed information on demographics, education, household assets and experiences during the war to serve as control variables. Many of the outcomes that interest us are difficult to measure: the underlying concepts are amorphous; the terms used to describe them are subject to wide differences in interpretation; and people often sense that there is a 'right' answer that the enumerator wants to hear. In response to such challenges, the research team devoted substantial effort to developing concrete, specific examples grounded in the reality of rural life in Sierra Leone. For instance, asking someone how much they 'trust' other members of their community is less likely to elicit meaningful responses than asking whether they have ever given money to a neighbour to buy something on their behalf in the local market. Several months of qualitative fieldwork informed the development of these contextually specific examples as well as the most appropriate phrasing of questions.

The field teams sought out the same respondents during the 2009 endline survey as they had interviewed in 2005. Examining how the responses of the same individual change over time lends power to the analysis by controlling for unobservable differences between people that affect the way they answer questions and view the world. Given the four-year gap between the baseline and the endline survey, the attrition rate was reasonably low: overall, 96 per cent of the same households were located and 76 per cent of the same individual respondents.

**Village focus groups and physical assessment of infrastructure**: During data collection visits, the field team supervisor assembled key opinion leaders to answer questions about issues that affect the community as a whole. These items include the stock of public goods, management of community teachers and communal farming activities. At a minimum, the group assembled included the town chief, women's leader, youth leader, teacher, religious leader, VDC member, master farmer and town speaker or deputy town chief. The supervisors also completed a tour of the village and made their own assessments of the materials used in and quality of construction of common public assets, like the primary school, water well, grain drying floor and latrine. As with the household survey, the research team translated qualitative fieldwork into quantitative survey instruments, focusing on concrete examples and behaviours, and tailored the questions to the local context.

**Structured community activities (SCAs)**: Given the centrality of inclusion, participation and collective action to GoBifo's objectives, and the difficulty of gauging these dynamics through survey responses alone, the third main element of data collection involved conducting three SCAs. During these activities, the research teams introduced a standardised, real-world decision, asset and opportunity, and observed how the communities responded (see Appendix B).

The first SCA examined whether GoBifo had lingering effects on the capacity for collective action and whether it served as a catalyst for group activity beyond the life of the project itself. Each community received six vouchers they could redeem at a nearby building materials store if they raised matching funds. Specifically, each card was worth SLL50,000 only if accompanied by SLL100,000 from the community. Topping up all six cards generated SLL900,000 or approximately USD300 for use in the store. During the final endline survey five months later, the research team explored the relative take-up of the programme, how inclusive and transparent the management of the resulting project was and the quality of the final construction. This exercise captures the degree to which the experience of project management under GoBifo enhanced the capacity of villagers to act collectively and take up a development opportunity outside the direct sphere of the project. If CDD has long-term impacts on the communal ability to come together and 'help themselves', we would expect greater take-up and better management of the building materials programme in the treatment communities.

The second SCA measured the extent to which a specific community decision was inclusive, participatory and democratic. The day before the survey work, the research teams met the Village Head and asked him/her to assemble the entire community for a meeting the next morning. At the subsequent meeting, the field supervisor presented the community with a choice of one of two small gifts – a carton of batteries useful for radios and flashlights or many small bags of iodised salt – as a token of appreciation for participating in the research programme. The supervisor emphasised that the members of the community should choose the gift and decide how to share it in any way they saw fit and then withdrew to observe and take notes on the decision-making process. Throughout the meeting, the enumerators remained behind the circle of the community meeting and observed how the deliberation evolved and exactly who participated. Of the four enumerators, one focused on the participation of youths, one on women and one on all adults, and the fourth kept careful track of each person who spoke publicly. Among other things, the researchers recorded who participated in any side-meetings ('hang heads'), the degree to which the Village Head and elders dominated the process, the extent of the debate and the relative influence of different sub-groups on the final outcome. As one example, this exercise provided concrete data on the relative frequency of female versus male speakers and youths versus non-youths in an actual community meeting. It is important to note that this is exactly the same metric as the GoBifo Ward Facilitators were required to track during project meetings as part of their own performance assessment (GoBifo Project 2008).

Finally, for the third SCA, the research team left each village with a large plastic tarpaulin, which is frequently used as a makeshift shelter or roof or in agriculture to dry grain or protect it from rain. This activity examines elite capture, a common concern and risk inherent in CDD's emphasis on devolving control to local power structures. During the follow-up visit, the enumerators explored the distribution of access to the tarp across households, as well as who received any salt/batteries and who contributed funds to and received benefits from the building materials. This exercise also has a collective action component, as the teams gauged whether or not the village had come up with a use for the tarp and whether they put it towards public or private ends.

A few cross-cutting points about these activities are worth noting. First, they are standardised across all the villages, including treatment and control, enabling us to observe the way in which different communities handle exactly the same decision or opportunity. As an alternative, observing only communal activities that the village itself creates suffers from the problem that particular decisions or projects exist in some communities but not in others, and it can be difficult to compare the dynamics of participation across different settings. For example, women are likely to participate differently in a communal farm meeting in one village compared with a school oversight meeting in another for reasons unrelated to the underlying inclusiveness of the two local societies. Second, the exercises each involve direct participant observation of community members engaged in a concrete, real-world activity. We feel that these SCAs capture the actual local collective action capacity and uncover the decision-making processes that underlie it more accurately than lab experiments, hypothetical vignettes or survey reports alone. Third and finally, the fact that these activities were carried out *after* the GoBifo programme (and its financial resources) had ended allows us to isolate any persistent impacts on collective action and institutional performance generated by the programme.

### 3.5 Analysis

For each outcome under each hypothesis, we estimate the difference between the average value of the outcome in the treatment group and the average value in the control group. We further calculate the standard error of this difference to ascertain whether or not the difference is statistically distinguishable from zero. To enhance the precision of our estimates and account for any spurious pre-programme differences between the two groups, we calculate these differences while controlling for the baseline value of the outcome in question as well as key characteristics of the communities and individuals under study. Specifically, we control for the distance from a road, the total number of households, an index of conflict experienced during the war, a measure concerning the historical presence of domestic slavery and a set of variables for the local council ward (the administrative level on which the randomisation was stratified). This is a straightforward linear regression model (for complete details, see Appendix C).

For each hypothesis, this report explores the effect of GoBifo on several different outcome measures. As an example, Hypothesis 2 (H2) considers the effect of GoBifo on the presence of primary schools, traditional birth attendant (TBA) huts, drying floors, grain stores, community centres and latrines one by one. However, a single global measure would provide a useful summary of how GoBifo affects a particular hypothesis *overall*. Continuing our example, we would like to know whether GoBifo increases or decreases the total stock of local public goods in the community. A natural choice for such a summary index is the *average* of all the specific relationships between GoBifo and each outcome under a given hypothesis – i.e. on average, what is the effect of GoBifo across this large bundle of distinct public goods? Thus, for each hypothesis we estimate: i) the impact of GoBifo on each individual outcome; as well as ii) the mean effects index, which provides a single summary measure for the average effect of GoBifo on the entire family of outcomes. (For a full exposition of the mean effects approach, see Appendix D.) This is especially important given the large number of outcomes under consideration: the treatment effects for approximately 16 of the 318 unique outcomes would be statistically significant due purely to random chance. In response, the mean effect approach reduces the number of effective statistical tests to only 12.

We collected data at the individual, household and community levels. For simplicity and comparability inside the mean effects index, we analyse all the outcomes at the community level. To do so, we first calculate the community-level mean of any outcome measured at the household or individual level before estimating the regression models. As an example, we asked individuals whether or not they had participated in road brushing during the past

one month. Thus, we first calculate the average number of people who participated out of the 12 respondents in the community, before estimating the regression models discussed above.

Finally, we explored whether or not the impact of GoBifo was smaller or larger depending on particular community or individual characteristics. We thus considered whether or not GoBifo had different impacts in Bombali as compared with Bonthe; in small versus large communities; and in remote versus easily accessible areas. As an example of why this might be important, GoBifo staff suggested that smaller communities were better able to adapt the CDD model than larger ones. In addition, for individual-level outcomes, we asked whether women benefited more or less than men and whether youths benefited more or less than their elder counterparts. Given the programme's emphasis on empowering marginalised groups, we might expect women and youth to benefit more. At the same time, since men and elders are relatively powerful compared with women and youth in many communities, we might expect them to capture disproportionate benefits for themselves. While we looked for these differences for all the outcomes and hypotheses below, we did not find evidence for any such differential treatment effects.

# 4. Results

## 4.1 Summary of results

Figure 4 below provides a succinct summary of our results and serves as a graphical analogue to Table 1 (in the Executive summary). The figure displays the estimated mean effect index and accompanying confidence interval for the impact of GoBifo on each outcome family. As explained in the previous section, the mean effects indices measure the average treatment effect across all the indicators listed under a particular family or hypothesis. Regarding the interpretation, the index co-efficients are in standard deviation terms. As an example, Figure 4 suggests that GoBifo caused an increase of nearly 0.7 standard deviation units on average for the project implementation outcomes in family A and an increase of nearly 0.3 standard deviation units for the hardware measures in family B. As a point of reference, if the data are normally distributed, an increase of one standard deviation unit would move an observation from the $50^{th}$ to $68^{th}$ percentile, while an increase of two standard deviation units would move an observation from the $50^{th}$ to $95^{th}$ percentile. These would be very large changes. Note that while the point estimate of the mean effect index for software family C is positive, its confidence interval straddles zero, implying that we cannot reject the null hypothesis that the true effect is in fact zero.

**Figure 4 GoBifo impacts by outcome family**

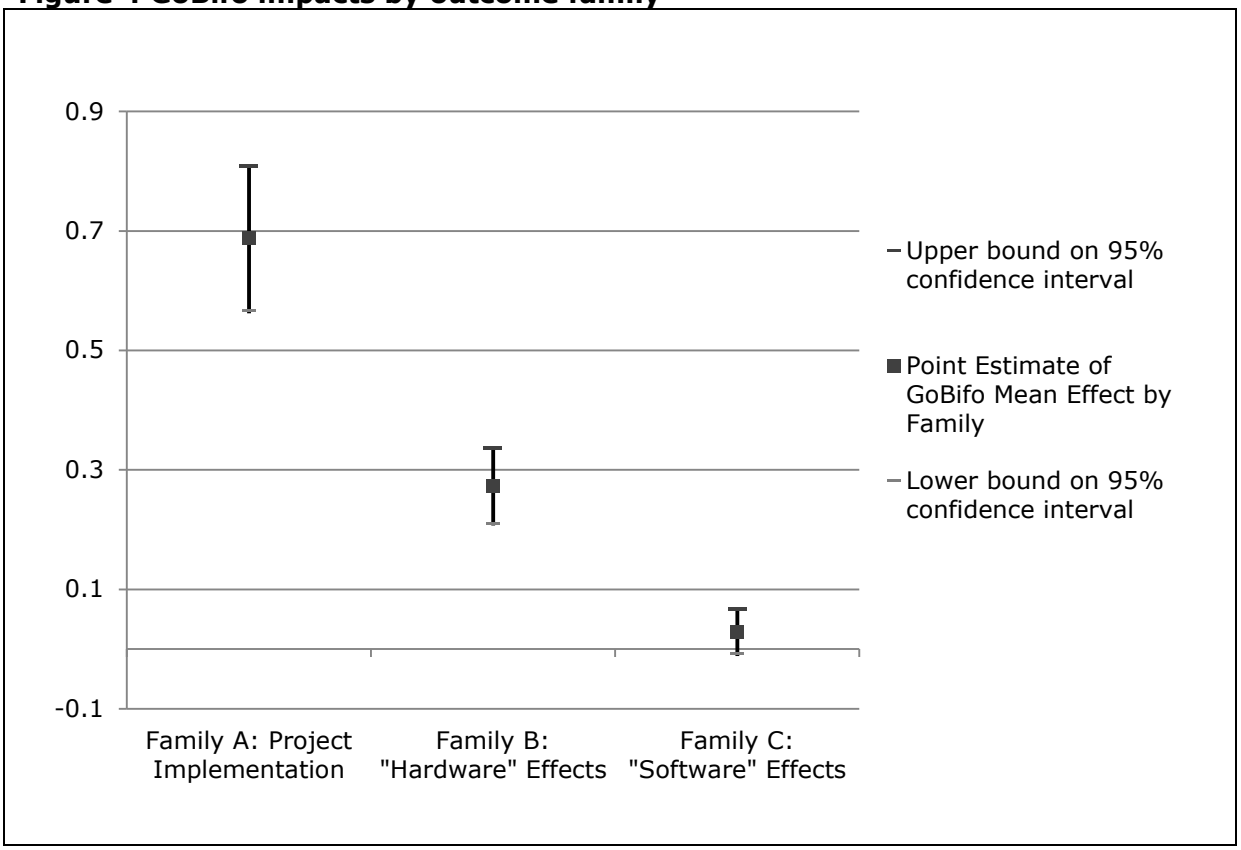



Table 4 fleshes out these family-level summary statistics with the mean effect index for each hypothesis. The table lists every single outcome variable under each hypothesis (recall that these outcomes were specified in the *ex ante* analysis plan before analysing any endline data) and then presents two different mean effect indices. The first index covers only the outcomes that apply to all the villages (unconditional ones, like whether or not the community has a primary school) – this is the 'full sample' outcomes index – and the second index combines these full-sample outcomes with conditional ones that depend on the existence of a particular item (i.e. given that the community has a primary school, what is the quality of construction?) – this is the 'all' outcomes index. We separate out the second set because the conditionality at times means that the measure applies only to a select subset of treatment and control villages, which may not be representative of the whole and may also suffer from small-sample estimation problems. The overall sample for these estimates is 236 villages at two points in time (yielding 471 observations in total, as one village is missing baseline data).

The sections that follow Table 4 discuss the individual outcomes that compose these summary statistics in greater detail and provide a better sense of the real-world significance and magnitude of these effects.

**Table 4 GoBifo impacts by hypothesis**

| **List of full sample outcomes** | **Full sample mean effect index** | **List of conditional outcomes** (conditional outcomes are added to the full sample set of outcomes when estimating the All mean effect index) | **All mean effect index** |
|---|---|---|---|
| **Family A: Project implementation** | | | |
| H1: GoBifo creates functional local institutions | | | |
| VDC; community has a VDP; community has a bank account; community visited by a WDC member; community visited by a local councillor; average household has met the local councillor; average resident has attended a WDC meeting. | 0.687** (0.062)<br><br>7 outcomes | Given the presence of local public goods in the community, WDC/local councillor (LC) involved in the planning, construction, maintenance or oversight of the local public goods (LPG) project (primary school, health clinic, TBA hut, water well, dry floor, grain store, community centre, latrine, sports field). | 0.507** (0.054)<br><br>16 outcomes |
| **Family B: Development infrastructure or 'hardware' effects** | | | |
| H2: GoBifo increases the quality and quantity of the local public services infrastructure | | | |
| Community has functional local public goods (primary school, public health unit (PHU), TBA hut, water well, dry floor, grain store, community centre, market, latrine, seed bank, football field and uniforms); community submitted a proposal to NGO or donor for support; community used the tarp for a public purpose; nearest footpath is clear; footpath was recently brushed. | 0.164** (0.040)<br><br>16 outcomes | Given presence of four common local public goods in the community (primary school, latrine, dry floor, water well), quality of materials used in the construction, and overall quality of infrastructure; given presence (of primary school, PHU, TBA hut, water well, dry floor, grain store, community centre, latrine, football field), community financial contributions to local public good; given that cards were redeemed, used them for a public project. | 0.118** (0.037)<br><br>34 outcomes |
| H3: GoBifo improves general economic welfare | | | |
| Community has petty traders; supervisor's overall assessment is that the community is better off than those surrounding; community has a bank account; total number of petty traders in the community; total (out of ten) common goods on sale in the community; total number of new businesses in the community; average household principal components analysis asset score; average household rank in the principal components analysis asset quintiles; average number of income sources per household; average total income of households; average respondent has attended a skills training course; average respondent has started a new business; average household has sold agricultural goods in the last month; average household has sold non-agricultural goods in the last month. | 0.399** (0.047)<br><br>14 outcomes | Given farmers, the average household has marketed any agricultural produce from the last harvest (check); given farmers, average agricultural income; given children aged 5 to 18 years in the household, average days in school; given some sales, average household has sold agricultural goods outside the community; given some sales, average household has sold non-agricultural goods outside the community. | 0.285** (0.038)<br><br>19 outcomes |

| **List of full sample outcomes** | **Full sample mean effect index** | **List of conditional outcomes** (conditional outcomes are added to the full sample set of outcomes when estimating the All mean effect index) | **All mean effect index** |
|---|---|---|---|
| **Family C: Institutional and social change or 'software' effects** | | | |
| H4: GoBifo increases collective action and contributions to local public goods | | | |
| Community redeemed any building materials vouchers; number of vouchers redeemed; household expectation of the total amount the community will collect for the vouchers; average amount households expect to contribute to the building materials vouchers; community has VDC; community submitted a proposal to NGO or donor for support; community has a communal farm; respondents work on a communal farm; community has any community teachers; community used the tarp; community used the tarp frequently; community has a communal agricultural marketing group; nearest footpath is clear; footpath was recently brushed; respondents participated in road brushing within the last month. | 0.041<br>(0.042)<br><br>15 outcomes | Given the presence of communal teachers, teachers have been trained; any support given to community teachers; total Leones paid to community teachers; given presence, community financial, labour and local materials contributions to local public goods (primary school, PHU, TBA hut, water well, dry floor, grain store, community centre, latrine, football field); given membership, respondent financial and labour contributions to group (credit/savings group, communal labour gang, school committee, social club, savings for special events); given work on communal farm, days worked; given redeemed vouchers, community used the materials and brought them back to the village; given something was built with the vouchers, quality of construction. | –0.008<br>(0.034)<br><br>59 outcomes |

| **List of full sample outcomes** | **Full sample mean effect index** | **List of conditional outcomes** (conditional outcomes are added to the full sample set of outcomes when estimating the All mean effect index) | **All mean effect index** |
|---|---|---|---|
| H5: GoBifo enhances inclusion and participation in community decisions, especially for vulnerable groups | | | |
| Total number of adult, women and youth attendees at the gift meeting; total number of adult, women and youth public speakers during the gift meeting; community held a vote during the gift meeting; there was debate during the gift meeting; opinion leaders did not decide the salt/batteries in private; there was no separation off of a smaller group to decide about the salt/batteries ('hanging heads'); the salt/battery choice was democratic; overall women played an equal role in the salt/battery choice to men; overall youths played an equal role in the salt/battery choice to non-youths; duration of the gift meeting; the gift chosen reflected the views of the majority; proportion of respondents who attended the initial gift meeting; respondents feel the choice between salt/batteries was made democratically; respondents feel that everyone had an equal say in the salt/batteries choice; respondents feel the decision on what to do with the tarp will be made democratically and that everyone will have an equal say; hypothetically, if the big ones in the village had wanted a different gift from the others, they would have gone with the majority opinion; after the research team left, the community held a meeting about the salt/batteries, tarp and vouchers; respondents attended a meeting to discuss what to do with the salt/batteries, tarp and vouchers; respondents feel the decision about what to do with the salt/batteries, tarp and vouchers was made democratically; respondents feel that everyone had an equal say in the decision about what to do with the salt/batteries, tarp and vouchers; community opinion leaders report that everyone had an equal say in the final decision about how to use the salt/batteries, tarp and vouchers; respondents agree hypothetically that projects should be decided democratically; proportion of respondents who attended a community meeting in the past year; minutes were taken during the last community meeting; disabled people attended the last community meeting; disabled people hold leadership positions in the community; community was able to produce the tarp during the follow-up visit; the tarp was being stored in a public place; community did not have any recent episodes of financial mismanagement. | 0.001<br>(0.031)<br><br>43 outcomes | Given that community members broke off into a separate group during the salt/batteries decision, the group included women and youths and what proportion of the sub-group was women and/or youth; given redeemed vouchers, the community can produce the receipt from the building materials supply store, kept written records about the materials project, made a public presentation of the materials upon returning from the store, is able to produce the materials for the survey team and stored the materials in a public place; given attendance, the respondent spoke publicly during the meeting (community, teacher pay, communal farm, salt/battery, tarp, vouchers); given membership, the respondent attended a group meeting (credit/savings group, communal labour gang, school committee, social club, savings for special events); given community teachers, the respondent attended a meeting to decide how much to pay them; given a communal farm, the respondent attended a meeting about what to do with the harvest; given presence, the respondent attended a meeting about a local public good (primary school, PHU, TBA hut, water well, dry floor, grain store, community centre, latrine, football field). | 0.015<br>(0.029)<br><br>72 outcomes |

| **List of full sample outcomes** | **Full sample mean effect index** | **List of conditional outcomes** (conditional outcomes are added to the full sample set of outcomes when estimating the All mean effect index) | **All mean effect index** |
|---|---|---|---|
| H6: GoBifo changes local systems of authority (not a project hypothesis, but a relevant research question only) | | | |
| The community voted during the gift meeting; the observed behaviour indicates that the chief or elders did not decide about the salt/battery gift without consulting people more broadly; respondents feel that everyone had an equal say in the salt/batteries choice; respondents expect that the person who will have the most say in how to use the tarp will not be the chief or an elder; respondents think the best place to store the tarp is not the chief's house; respondents want to store the tarp in a public place; during the follow-up visit, respondents report that the chief alone did not decide about the salt/batteries choice, how to share the salt/batteries, how to use the tarp and vouchers; respondents felt that the person with the most say in the salt/batteries choice, how to share the salt/batteries, how to use the tarp and how to use the vouchers was not a village-level chiefdom figure or elder; local opinion leaders say that the person with the most say in how to share the salt/batteries, how to use the tarp and how to use the vouchers was not a village-level chiefdom figure or elder; tarp was not stored in the chief's house; when presented with a choice of two opposing statements, respondents agree that citizens should question authorities, women can be good leaders and youths can be good leaders; respondents agree hypothetically that projects should be decided democratically; respondents have not taken a dispute to traditional authorities for resolution; respondent's relative trust, spending effectiveness and belief that they listen to people in the area for the local councillors as compared with chiefdom authorities. | 0.048<br>(0.036)<br><br>25 outcomes | Given community teachers, it is not the chief alone who decides how they should be remunerated; given that some parents default on fees for the community school, respondents do not report the defaulter to the chief; given vouchers redeemed, building materials are stored in a public place; given vouchers redeemed, someone other than a traditional authority was sent to the store on behalf of the community. | 0.038<br>(0.034)<br><br>29 outcomes |

| **List of full sample outcomes** | **Full sample mean effect index** | **List of conditional outcomes** (conditional outcomes are added to the full sample set of outcomes when estimating the All mean effect index) | **All mean effect index** |
|---|---|---|---|
| H7: GoBifo enhances trust | | | |
| Average respondent reports that he/she trusts people from his/her own community, people from outside the community, chiefdom officials, local councillors, central government officials, NGOs/donor projects; average respondent is a member of a credit/savings group; average respondent would give money to a neighbour to buy something on their behalf in the market (and has actually done so); average respondent feels that if he/she left money at a community meeting it would still be there an hour later; community has a communal agricultural marketing group. | 0.042<br>(0.064)<br><br>11 outcomes | Given that they ever left money at a community meeting, respondents report that the money was still there when they came back for it. | 0.027<br>(0.060)<br><br>12 outcomes |
| H8: GoBifo builds groups and networks | | | |
| Respondent is a member of each particular group (credit/savings group, communal labour gang, school committee, social club, savings for special events, fishing cooperative, seed multiplication group, women's group and youth group); respondent helped to thatch another person's house; community has a fishing cooperative and a communal agricultural marketing group. | 0.033<br>(0.044)<br><br>12 outcomes | Given membership, respondent attended a meeting, contributed financially and contributed labour to the group (credit/savings group, communal labour gang, school committee, social club, savings for special events, fishing cooperative, seed multiplication group, women's group and youth group); given the need to rethatch one's own house, respondent received help from a neighbour. | –0.000<br>(0.038)<br><br>40 outcomes |
| H9: GoBifo increases access to information about local governance | | | |
| Respondents attended the gift meeting; respondents can name the two choices offered (salt or batteries) and which one was chosen; respondents can correctly name local leaders and facts (local councillor, council chair, section chief, paramount chief, date of next election, tax rate, a council project); respondents listen to the radio for news; number of gifts respondents can recall (salt/batteries, tarp, vouchers); respondents know whether the community has used the tarp, and if not, what the plan is; respondents know whether any cards were redeemed; respondents know the specific (or planned) use of the tarp; community visited by a WDC member, the local councillor and the paramount chief; information publicly displayed in village (index of 6: awareness campaigns, financial info, VDP, meeting minutes, government policies, election info). | 0.003<br>(0.039)<br><br>19 outcomes | Given that they sell goods in the market, respondents know who has the right to spend their market dues; given vouchers redeemed, respondents know details about how the cards were used (index of 6: number of cards redeemed, total Leones raised, who went to store, saw materials, saw receipt, can name items purchased) | 0.000<br>(0.039)<br><br>21 outcomes |

| **List of full sample outcomes** | **Full sample mean effect index** | **List of conditional outcomes** (conditional outcomes are added to the full sample set of outcomes when estimating the All mean effect index) | **All mean effect index** |
|---|---|---|---|
| H10: GoBifo increases participation in local governance | | | |
| Respondents voted in 2007 first-round presidential elections, 2007 second-round presidential elections and 2008 local elections; a community member stood for paramount chief, section chief, local council and WDC; respondents have met a local councillor; respondents have attended a WDC meeting; respondents feel like they could change an unjust chiefdom law and an unjust LC policy; respondents believe the LC listens to what their community says; community has a VDC; community has a VDP; respondents discuss politics. | 0.114** (0.047)<br><br>15 outcomes | Respondents can produce a voter identity card with an appropriate hole punch documenting their vote in first-round presidential, second-round presidential and local elections; given a VDC, how many women and youths are members, and what is the proportion of women and youths in the total committee; respondents would like to be a member of the VDC; given a VDP, the plan is written down, the tarp and building materials were used toward projects in the plan; given the presence of local public goods in the community, WDC/LC involved in the planning, construction, maintenance or oversight of the LPG project (primary school, health clinic, TBA hut, water well, dry floor, grain store, community centre, latrine, sports field); given the presence of local public goods in the community, the section chief or paramount chief was involved in the planning, construction, maintenance or oversight of the LPG project (primary school, health clinic, TBA hut, water well, dry floor, grain store, community centre, latrine, sports field). | 0.139** (0.039)<br><br>43 outcomes |
| H11: GoBifo reduces crime and conflict | | | |
| When presented with a choice of two opposing statements, respondents agree that violence is never justified in politics, it is not okay to beat one's wife and it is not okay to punish children physically; household has had no conflict over financial matters in the past year; household has not suffered theft of goods; household members have not been engaged in a conflict with physical fighting in the past year; household has not suffered any witchcraft (*juju*); in the past year, household has not had any conflict that needed an external agent to help resolve it. | 0.028 (0.054)<br><br>8 outcomes | Given a conflict over money, there was no interpersonal violence involved; given a conflict in which an external mediator was needed, the conflict was resolved by agents within the community. | 0.035 (0.052)<br><br>10 outcomes |
| H12: GoBifo fosters more liberal political and social attitudes (not a project hypothesis, but a relevant research question only) | | | |
| When presented with a choice of two opposing statements, respondents agree that it is not okay to abuse one's wife, women can be good leaders, youths can be good leaders, people outside the community can be good leaders, it is not okay for local leaders to coerce labour, it is wrong to pay government officials bribes and youths are treated well in their community; communities have ever had a female village head; communities have ever had a youth village head. | 0.034 (0.041)<br><br>9 outcomes | Given membership, respondents report that the group contains more diverse members (index of mixed gender, mixed age groups and multiple tribes for credit/savings group, labour gang, social club, special event savings); given the presence of a communal farm, the group that works on the farm includes both genders, both age groups, multiple tribes and no children; given that the community has selected a new village head since 2006, the village head is female and the village head is a youth. | −0.033 (0.052)<br><br>19 outcomes |

Note: Significance levels are indicated by + $p < 0.10$, * $p < 0.05$ and ** $p < 0.01$

Robust standard errors clustered at the village level. The full sample contains 236 villages measured at 2 points in time for a total of 471 observations (one village is missing in the baseline data). Three variables from the ex-ante analysis plan have been omitted due to an insufficient sample size: community contributions to market and peripheral health unit (H2 and H4) and presence of football equipment (H2).

**Family A: Project implementation**

The first family of GoBifo impacts measures the extent to which the project successfully implemented its stated activities. The positive and highly significant treatment effects document the fact that GoBifo achieved what it intended: it established VDCs; helped communities to draw up development plans and open bank accounts; and created meaningful links between the participating villages and their local government representatives. These results suggest that GoBifo achieved its primary objective of creating the institutions, tools and linkages to facilitate local development. Specifically, the co-efficient on the full sample mean effects index for H1 in Table 4 implies that overall, GoBifo caused a 0.687 standard deviation unit increase in the presence of these institutions and links in the treatment communities. This effect is large in magnitude and significant at the 99 per cent confidence level.

To provide a better sense of the specific accomplishments and corresponding size of these effects, Table 5 below presents the results from each individual outcome regression. First, note that 15 of the 16 included treatment effects are greater than zero, and 12 of these are significant at conventional levels. This suggests that the positive mean effect shown earlier is not driven by one or two select outcomes, but instead represents a broad set of programmatic achievements. In contrast to Table 4 above, these co-efficients refer to the original units that the outcome was measured in – not the standardised unit of the mean effects index – so are more immediately interpretable. As some examples, the first row reveals that GoBifo caused the proportion of villages with a VDC to increase by 34.1 percentage points. By the post-programme period, 86.3 per cent of GoBifo communities had a VDC compared with 45.8 per cent of controls – a large effect. The corresponding co-efficient in the second row indicates that GoBifo increased the likelihood that a community was visited by a member of its WDC in the past year by 15.6 percentage points. Row 6 shows a positive treatment effect on the existence of VDPs by 29.6 percentage points, nearly a 50 per cent increase on the base of 61.7 per cent in the controls. Row 7 reveals an increase in the presence of a village bank account of 70.6 percentage points, capturing a tenfold increase.

Turning to the set of 'conditional' outcomes, these apply only to villages where a particular local public good exists. For each of nine common resources listed in rows 8 to 16, the household survey asked the respondents whether a member of the WDC or local council was 'directly involved in the planning, construction, maintenance or oversight' of the particular good. Note that the treatment effect is positive for all but one of these goods, and statistically significant for seven. This is powerful evidence that GoBifo successfully linked the village-level projects with the work of local politicians, a key objective given the position within the Decentralisation Secretariat. Overall, the aggregated impact of GoBifo on the project implementation family is positive and significant regardless of whether only the full sample or a broader set of all outcomes is considered.

**Table 5 Hypothesis 1 – project implementation individual outcomes**

| Row | Outcome variable | | 2009 mean in controls | Treatment effect | Standard error | N | Specification |
|---|---|---|---|---|---|---|---|
| **H1 – Full sample outcomes** | | | | | | | |
| 1 | VDC | | 0.458 | 0.341** | (0.077) | 467 | Panel |
| 2 | Visit by WDC member | | 0.212 | 0.156* | (0.070) | 462 | Panel |
| 3 | Visit by local councillor | | 0.322 | 0.042 | (0.080) | 464 | Panel |
| 4 | Met local councillor or attended LC meeting | | 0.184 | 0.032 | (0.030) | 471 | Panel |
| 5 | Met WDC member or attended WDC meeting | | 0.090 | 0.045+ | (0.023) | 471 | Panel |
| 6 | VDP | | 0.617 | 0.296** | (0.048) | 221 | Cross section |
| 7 | Community bank account | | 0.081 | 0.706** | (0.045) | 226 | Cross section |
| **H1 – Conditional outcomes** | | | | | | | |
| 8 | [Given functional infrastructure in the village] A WDC or local council member was involved in the planning, construction, maintenance or oversight of the resource | Primary school | 0.415 | 0.181** | (0.055) | 138 | Cross section |
| 9 | | PHU | 0.615 | –0.222 | (0.189) | 26 | Cross section |
| 10 | | TBA hut | 0.399 | 0.002 | (0.106) | 70 | Cross section |
| 11 | | Water well | 0.354 | 0.199** | (0.044) | 150 | Cross section |
| 12 | | Grain drying floor | 0.243 | 0.140* | (0.061) | 115 | Cross section |
| 13 | | Grain store | 0.144 | 0.295** | (0.076) | 71 | Cross section |
| 14 | | Community centre | 0.251 | 0.244** | (0.053) | 95 | Cross section |
| 15 | | Latrine | 0.219 | 0.155** | (0.040) | 169 | Cross section |
| 16 | | Sports field | 0.163 | 0.081* | (0.035) | 181 | Cross section |

Note:
Significance levels indicated by + p < 0.10, * p < 0.05, ** p < 0.01
The relevant treatment effect for the panel specifications is the co-efficient on GoBifo*Post, while the relevant treatment effect for the cross-section (follow-up data only) specifications is on GoBifo (see Appendix C for details).
Robust standard errors in parentheses, clustered by village for panel specifications.
Includes fixed effects for the district council wards (the unit of stratification) and control variables (total households per community, distance to nearest motorable road, index of war exposure and index of history of domestic slavery).
'Conditional' outcomes are conditioned on the existence of a public good.

### 4.2 GoBifo process analysis

Before turning to the next family of project impacts, this section examines GoBifo's processes in greater depth. Note that these are the views and opinions of randomly selected households and respondents. This means that they are representative of the typical resident of GoBifo villages, and are therefore no more or less likely to have participated directly in project activities or held leadership roles than anyone else.

The benefits of GoBifo village projects were widely enjoyed: 90 per cent of the respondents reported that their household had directly benefited from at least one GoBifo project. Similarly, the contributions were shared broadly: 83 per cent of households contributed something – money, labour, materials, food or record keeping – to at least one project. By far the most common form of contribution was unskilled labour, followed by cooking food for workers, building materials and money. Note further that there were no significant differences between women and men, or between youths and non-youths, in terms of who benefited and who contributed.

Choices about GoBifo projects – including the selection of the village facilitator, VDC membership and types of projects – were made in remarkably similar ways. The respondents reported that roughly 38 per cent of these decisions were made by the chief and other 'big ones' alone (including traditional opinion leaders); for 43 per cent everyone helped decide; in 1 per cent an outsider chose; and 18 per cent did not know. Direct participation in the early set-up activities varied: while 64 per cent of the respondents said they had attended the meeting at which the village facilitator was chosen, only 41 per cent of the respondents said they had attended a meeting to draft the VDP and 38 per cent reported that they had helped draw a dream map, an early activity in the planning process.

While nearly everyone (95 per cent) could name at least one GoBifo project, the amount of detailed knowledge they held about the projects varied considerably. At one end of the spectrum, 76 per cent of the respondents said they knew which people within the community were responsible for maintaining GoBifo projects (where applicable), while at the other end, only 18 per cent of the respondents said they knew how much money the community had received from GoBifo. Of those who thought they knew, only 10 per cent (23 people) correctly reported SLL14 million as the specific amount.

Table 6 below summarises the information collected during focus group discussions with community leaders about which GoBifo projects tended to be more or less successful than others. Water wells, skills training and school construction projects appear to have been highly successful: the majority of communities that implemented these projects had only positive things to say about them. Livestock rearing and seed multiplication tended to involve more problems. For the former, community leaders reported issues with the animals becoming sick and the lack of veterinary services in the area to tend to them; for the latter, funding delays interfered with seed multiplication projects.

**Table 6 Focus group reports of problems and successes by project type**

| Project type | Total reports | Problems reported | Positive reports (N) | Positive reports (%) |
|---|---|---|---|---|
| Water well construction and rehabilitation | 6 | 1 | 5 | 83.3 |
| Skills training | 43 | 8 | 35 | 81.4 |
| Latrine construction | 36 | 9 | 27 | 75.0 |
| School construction and rehabilitation | 23 | 6 | 17 | 73.9 |
| Construct multipurpose centre, court barrie, store | 46 | 15 | 31 | 67.4 |
| TBA house construction | 23 | 10 | 15 | 65.2 |
| Livestock (goat rearing, animal husbandry) | 62 | 25 | 37 | 59.7 |
| Seed (multiplication, rice, groundnut) | 23 | 10 | 13 | 56.5 |

To help estimate the potential leakage of project funds, the GoBifo field office coordinators provided the research team with detailed accounts of all the projects implemented by and financial transfers to each village. During the final research visit to communities, the research field team supervisors asked the village leaders to verify the receipt of funds listed in these records. Of 273 financial transactions cross checked, in 13.5 per cent of cases the community members were unable to confirm receipt. These discrepancies fall into two categories: i) the amount in the community records was markedly less than that in the project accounts; ii) community members reported receiving building materials in kind and could not estimate their total worth. Given the low levels of education and potential recall problems over the three-and-a-half-year project cycle, some of these discrepancies are likely to be reporting errors.

In response to the first issue, the GoBifo project accountant, field office coordinators and ward facilitators produced the hard copy payment vouchers for each of the disputed transactions. These receipts are signed by both a village representative (either the VDC chair or the village finance officer) and a project field staff member. Regarding the second issue, the project managers are investigating the cases in which village members claimed to have received materials without financial documentation. To prevent any potential collusion between the village representative and the ward facilitator in future, the project has devised a more intensive system of safeguards. These include i) public display within the village of all procured project materials with accompanying receipts; ii) establishment of safe storage facilities within the villages with written stock record cards; iii) more intensive training of VDC members in record keeping and financial management; and iv) greater sensitisation of all community members to the total amount of block grants and specific sub-totals by tranche. Overall, the potential leakage of funds between project headquarters, field offices and recipient villages seems to be minimal.

In general, the community members have a positive opinion of GoBifo. A large majority (77 per cent) of the respondents said they were satisfied or very satisfied with the work GoBifo has carried out in their community. As seen above, women and men, as well as youths and non-youths, all expressed the same levels of satisfaction with the project.

**Family B: Development infrastructure or 'hardware' effects**

The second family of impacts makes the logical leap from the first set of results: given that GoBifo successfully implemented its interventions, what are the most immediate impacts we would expect to see? For the first hypothesis in this family, recall that the communities used the majority of GoBifo grants for small-scale infrastructure projects, which should thus lead to an improvement in the stock of local public goods. The mean effect index for H2 in Table 4 shows that this is indeed the case: on average, GoBifo caused a 0.164 standard deviation increase in the stock and quality of goods.

Many of the outcomes under this hypothesis cluster into three natural sub-groups: the stock of local public goods, the quality of such goods and community financial contributions to the construction and upkeep of the stock. Table 7 below breaks the full sample aggregate measure into its individual components, in which the first 12 rows explore the impacts in the first sub-group. Note that 10 out of 12 of these treatment effects are positive, and half are statistically significant. Because the communities chose their own projects, one would expect differences in local needs to translate into more muted effects dispersed across the stock of goods as opposed to large increases in one particular good (as one would expect for, say, an education project). Along these lines, there are marked increases in the proportion of villages with a functional TBA house by 17.5 percentage points, a community centre by 24.1 percentage points, a latrine by 21.0 and a seed bank by 17.2. It is not surprising that there is no impact on peripheral health units, as they are large projects managed by the local government, well beyond the scope of the GoBifo grants. Calculating a mean index for this sub-group reveals a highly significant increase of 0.258 standard deviations (with a standard error of 0.049, not shown).

Turning to the next sub-group, rows 17 to 24 in Table 7 reveal that GoBifo had a strong positive impact on the quality of the materials used and overall construction of four common goods – primary school, grain drying floor, water well and latrine. All of these individual effects – save those for the well – are positive and significant, as is the sub-group index overall (an increase of 0.296 standard deviation units with a standard error of 0.077, not shown). Note further that there is some evidence that households responded positively to these improvements in the stock and quality of key public goods with higher utilisation. Specifically, there are significantly higher utilisation levels of latrines and concrete drying floors in treatment communities. However, there are no average changes in the utilisation of water wells, frequency of community meetings or number of days that young children spend at school.

Lastly, note that when we estimate the nine indicators concerning the community financial contributions to the existing infrastructure (listed in rows 25 to 33 of Table 7) together as a group, we see a negative but not statistically significant treatment effect (–0.113 standard deviations with a standard error of 0.104, not shown). Combined with the negative and marginally significant treatment effect on whether the community approached another NGO or donor for development support in row 13, this provides suggestive evidence that the GoBifo funds may have served as a substitute, rather than a complement, for the community's own resources. At the minimum, it clearly suggests that the GoBifo money did not serve as a catalyst for additional fund-raising, nor did the project experience encourage participants to seek development assistance beyond the project itself.

**Table 7 Hypothesis 2 stock of local public goods individual outcomes**

| | Outcome variable | | Mean in controls | Treatment effect | Standard error | N | Specification |
|---|---|---|---|---|---|---|---|
| **H2 – Full sample outcomes** | | | | | | | |
| 1 | Existence of functional local public good in the community | Primary school | 0.462 | –0.007 | (0.050) | 464 | Panel |
| 2 | | Peripheral health unit | 0.060 | 0.009 | (0.029) | 462 | Panel |
| 3 | | TBA house | 0.079 | 0.175** | (0.035) | 235 | Cross section |
| 4 | | Water well | 0.459 | –0.068 | (0.081) | 451 | Panel |
| 5 | | Grain drying floor | 0.237 | 0.104 | (0.066) | 459 | Panel |
| 6 | | Grain store | 0.136 | 0.001 | –0.051 | 468 | Panel |
| 7 | | Community centre | 0.212 | 0.241** | (0.063) | 469 | Panel |
| 8 | | Market | 0.017 | 0.000 | (0.021) | 467 | Panel |
| 9 | | Latrine | 0.462 | 0.210** | (0.059) | 234 | Cross section |
| 10 | | Seed bank | 0.170 | 0.172** | (0.048) | 226 | Cross section |
| 11 | | Sports field | 0.444 | 0.071+ | (0.040) | 236 | Cross section |
| 12 | | Sports uniforms | 0.100 | 0.103* | (0.048) | 225 | Cross section |
| 13 | Community took proposal to NGO | | 0.292 | –0.156+ | (0.081) | 460 | Panel |
| 14 | Tarp put toward public purpose | | 0.857 | 0.023 | (0.050) | 161 | Cross section |
| 15 | Nearest bush path is clear (index 0–1) | | 0.482 | –0.001 | (0.034) | 228 | Cross section |
| 16 | Fewer days since bush path last cleared | | –41.263 | –8.199 | (6.168) | 192 | Cross section |
| **H2 – Conditional outcomes** | | | | | | | |
| 17 | Supervisor physical assessment of construction quality (index 0–1): | Primary school | 0.583 | 0.116* | (0.055) | 123 | Cross section |
| 18 | | Grain drying floor | 0.375 | 0.142+ | (0.076) | 101 | Cross section |
| 19 | | Water well | 0.464 | 0.007 | (0.041) | 224 | Cross section |
| 20 | | Latrine | 0.270 | 0.177** | (0.055) | 154 | Cross section |
| 21 | Supervisor physical assessment of overall appearance (index 0–1): | Primary school | 0.482 | 0.122** | (0.045) | 123 | Cross section |
| 22 | | Grain drying floor | 0.426 | 0.093+ | (0.056) | 99 | Cross section |
| 23 | | Water well | 0.426 | –0.025 | –0.033 | 221 | Cross section |
| 24 | | Latrine | 0.417 | 0.060+ | (0.031) | 153 | Cross section |
| 25 | [Given infrastructure in the community] Money and supplies were provided at least in part by the community | Primary school | 0.554 | –0.007 | (0.112) | 242 | Panel |
| 26 | | PHU | 0.142 | 0.092 | (0.318) | 30 | Panel |
| 27 | | TBA house | 0.449 | –0.013 | (0.095) | 81 | Cross section |
| 28 | | Water well | 0.375 | –0.136 | (0.131) | 217 | Panel |
| 29 | | Grain drying floor | 0.105 | 0.086 | (0.124) | 184 | Panel |
| 30 | | Grain store | 0.000 | –0.028 | (0.150) | 88 | Panel |
| 31 | | Community centre | 0.808 | –0.234 | (0.143) | 127 | Panel |
| 32 | | Latrine | 0.761 | –0.197* | (0.093) | 126 | Cross section |
| 33 | | Sports field | 0.335 | 0.028 | –0.039 | 182 | Cross section |

Note:
Significance levels indicated by + p < 0.10; * p < 0.05; ** p < 0.01
The relevant treatment effect for the panel specifications is the co-efficient on GoBifo*Post, while the relevant treatment effect for the cross-section (follow-up data only) specifications is on GoBifo.
Robust standard errors in parentheses, clustered by village for panel specifications.
Includes fixed effects for the district council wards (the unit of stratification) and control variables (total households per community, distance to nearest motorable road, index of war exposure and index of history of domestic slavery); and vi) 'conditional' outcomes are conditioned on the existence of a public good.

The second hypothesis in the 'hardware' category turns to general measures of economic welfare for households and the community overall. There are two main reasons to expect GoBifo to improve economic welfare. First, note in Figure 2 that 17 per cent of sub-projects were dedicated to skills training and income-generating initiatives, like soap-making, weaving, blacksmithing and carpentry training. If well implemented, these should translate into new businesses, and if these are in turn profitable, greater take-home income for the proprietors. Similarly, a quarter of the communities implemented agricultural projects, which should generate greater post-harvest profits if successful. Second, GoBifo injected close to USD5,000 into each community (roughly equivalent to USD100 per household on average), and the money is fungible. As an example, the discussion above suggests that GoBifo grants may serve as a substitute for the existing communal contributions to public goods, which would translate into higher household savings. In either or both cases, the analysis reveals strong positive impacts on economic welfare for H3 in Table 4, using either the narrow full sample or the wider set of all outcomes. Using the former, on average GoBifo caused a highly significant 0.399 standard deviation unit increase in the outcomes concerning household assets, entrepreneurship and market activities.

Table 8 unpacks this figure to examine the individual components of the core index. Nearly all of the treatment effects are positive in sign and six are statistically significant at the conventional levels. The first six outcomes refer to village-level dynamics, for which we see a 30 per cent increase in the number of petty traders in row 4 (i.e. 0.7 more traders on a base of 2.4 traders in the control group) and a 13 per cent increase in goods locally available for sale in row 5 (i.e. 0.56 additional goods on a base of 4.45 common items). The final eight outcomes are aggregated up from the household data. We observe improvements in an asset ownership score (derived using principal components analysis) in row 7, in which the underlying assets include common household durables (for example radios, mobile phones), amenities like drinking water source and sanitation, and the materials used in the roof, walls and floor of the dwelling. The project tripled the proportion of respondents who had recently participated in skills training: an 11.9 percentage point increase on a base of 6.1 per cent in the control communities (row 11). We find no impact on the total household income in 2009; however, this is difficult to measure among households engaged in subsistence agriculture and the treatment effect estimate is relatively imprecise (row 10).

Assessing the two hypotheses in this family, we conclude with confidence that communities are better off in very tangible ways due to their participation in GoBifo.

**Table 8 Hypothesis 3 – economic welfare individual outcomes**

| | Outcome variable | Mean (2009) in controls | Treatment effect | Standard error | N | Specification |
|---|---|---|---|---|---|---|
| **H3 – full sample outcomes** | | | | | | |
| 1 | Any petty traders in village | 0.441 | 0.106 | (0.079) | 455 | Panel |
| 2 | Better off than other villages nearby | 0.263 | 0.090 | (0.087) | 432 | Panel |
| 3 | Community has a bank account | 0.081 | 0.706** | (0.045) | 226 | Cross section |
| 4 | Total number of petty traders in village | 2.432 | 0.719* | (0.344) | 225 | Cross section |
| 5 | Total goods on sale of 10 | 4.449 | 0.560* | (0.240) | 236 | Cross section |
| 6 | Total number of new businesses | 1.745 | 0.091 | (0.310) | 207 | Cross section |
| 7 | Household asset score | –0.170 | 0.212* | (0.090) | 471 | Panel |
| 8 | Household asset quintile | 2.835 | 0.158+ | (0.094) | 471 | Panel |
| 9 | Total number of household income sources | 1.543 | 0.102 | (0.105) | 472 | Panel |
| 10 | Income from top three cash-earning sources (in SLL1,000) | 746.94 | –21.773 | (73.069) | 236 | Cross section |
| 11 | Attended skills training | 0.061 | 0.119** | (0.018) | 235 | Cross section |
| 12 | Started a new business | 0.072 | 0.014 | (0.012) | 236 | Cross section |
| 13 | Sold agricultural goods in last month | 0.507 | –0.002 | (0.026) | 236 | Cross section |
| 14 | Sold other non-agricultural goods in last month | 0.186 | 0.018 | (0.018) | 236 | Cross section |

Note:
Significance levels indicated by + p < 0.10, * p < 0.05, ** p < 0.01
The relevant treatment effect for the panel specifications is the co-efficient on GoBifo*Post, while the relevant treatment effect for the cross-section (follow-up data only) specifications is on GoBifo.
Robust standard errors in parentheses, clustered by village for panel specifications.
Includes fixed effects for the district council wards (the unit of stratification) and control variables (total households per community, distance to nearest motorable road, index of war exposure and index of history of domestic slavery).
'Conditional' outcomes are conditioned on the existence of a public good.

### Family C: Institutional and social change or 'software' effects

This section explores whether the experience of working together on successful community development initiatives strengthens the bonds between individuals and changes the way in which they interact in other, non-project settings. As we turn from the proximate impacts tied directly to programme activities to these more indirect influences on local norms and institutional practices, we find no evidence that GoBifo led to fundamental changes in the way that communities make decisions, the voice of women and youths or the capacity to act collectively outside the immediate sphere of the project. Table 9 presents some illustrative treatment effect estimates for individual outcomes in this family that are drawn from the SCAs.

Altogether there are 59 outcome variables relating to collective action and contributions to local public goods (H4 in Table 4). Of this set, only seven treatment effects meet the conventional levels of statistical significance, with five positive in sign and two negative, leading quite conclusively to a zero effect when considered as a group. These measures fall into a few different categories. The first set concerns the real-world exercises of the building materials vouchers and the tarp. Rows 1 and 2 of Table 9 show that there was no differential take-up of the voucher programme – exactly 62 treatment villages (53 per cent) and 64 control communities (54 per cent) redeemed cards at the local supply stores – nor is there any difference in the number of cards redeemed, as most of the villages that cashed in cards used all six. Similarly, if anything, there is a slight negative effect on the likelihood that the community had put the tarp to use by the time of the endline follow-up visit (row 14). The next set of outcomes considers household contributions to the existing public goods discussed above under H2. The only difference here is that we expand the consideration from just financial contributions to those of labour and local materials/food for workers. However, these also yield no treatment effects. Lastly, there are no differences in contributions to membership groups, participation in communal farms or support of community teachers.

This finding concerning collective action is significant and troubling for the long-term impacts of the project. Clearly, community members gained experience in working together to implement local development projects successfully and engaged in collective action as part of the GoBifo project. However, their GoBifo-specific experiences did not lead them to take advantage of the new collective action opportunity presented by the building material voucher programme. Thus, while GoBifo created the structures to manage local development – the VDC, a plan, a bank account – and communities gained experience in implementation, the project left them no better able to take advantage of a simple real-world subsidy programme. This suggests that while GoBifo created immediate positive benefits for the community, it did not send the communities onto a stronger development trajectory that would continue to accrue benefits beyond the life of the project.

Because the inclusion of women and youths held such prominence in the project objectives and was directly incorporated into the Ward Facilitators' operating procedures, it received special attention in the research programme. Despite an exhaustive battery of measures, there is no evidence of impacts of GoBifo on the role of women and youths in local decisions or on the transparency and accountability of decision-making more generally. The findings are very consistent across the indicators – of the 72 measures considered, only six were statistically significant, divided equally between positive and negative treatment effects. Among the large set of outcomes listed in Table 4, three related groups are noteworthy. The first group of measures recorded concrete behaviours during the meeting to decide between the salt and batteries gift. Panel B of Table 9 reveals no treatment effects on the total number of adults, women and youths who attended the meeting or spoke publicly during the deliberation. To illustrate: on average, 25 women attended these meetings but just two of them made a public statement during the discussion about which item to choose. The difference between the number of women who spoke in treatment versus control communities is only –0.19 (standard error 0.22), and the proportion of males who spoke during the meeting remained twice as high as the proportion of females in the treatment villages, the same as at the baseline. We similarly find no impact regarding whether any smaller 'elite' groups broke off from the general meeting to make the gift choice without broader consultation; the duration of the deliberation; or how democratic the decision-making process appeared to the enumerators (for example by holding a direct vote).

**Table 9 Illustrative treatment effect estimates for structured community activities**

| | SCA outcome | Mean for controls | Treatment effect | Standard error | N |
|---|---|---|---|---|---|
| **Panel A. Collective action and building materials vouchers** | | | | | |
| **GoBifo mean effect for SCA #1 (13 outcomes in total)** | | **0.00** | **–0.057** | **(0.053)** | **236** |
| 1 | Proportion of communities that redeemed any vouchers at the building material supply store | 0.54 | –0.01 | (0.06) | 236 |
| 2 | Average number of vouchers redeemed at the store (out of six) | 2.95 | 0.11 | (0.35) | 236 |
| 3 | Proportion of communities that held a meeting after the research team left to discuss what to do with the vouchers | 0.98 | –0.05* | (0.02) | 231 |
| **Panel B. Participation in gift choice deliberation** | | | | | |
| **GoBifo mean effect for SCA #2 (32 outcomes in total)** | | **0.00** | **0.005** | **(0.036)** | **236** |
| 4 | Duration of gift choice deliberation (in minutes) | 9.36 | 1.60 | (1.13) | 225 |
| 5 | Total adults in attendance at the gift choice meeting | 54.51 | 3.50 | (3.20) | 236 |
| 6 | Total women in attendance at the gift choice meeting | 24.99 | 1.99 | (1.68) | 236 |
| 7 | Total youths (approximately 18 to 35 years old) in attendance at the gift choice meeting | 23.57 | 2.10 | (1.38) | 236 |
| 8 | Total number of public speakers during the deliberation | 6.04 | 0.24 | (0.40) | 236 |
| 9 | Total number of women who spoke publicly during the deliberation | 1.88 | –0.19 | (0.22) | 236 |
| 10 | Total number of youths (approximately 18 to 35 years old) who spoke publicly | 2.14 | 0.23 | (0.24) | 236 |
| 11 | Proportion of communities that held a vote during the deliberation | 0.10 | 0.07 | (0.04) | 236 |
| **Panel C. Community use of tarpaulin** | | | | | |
| **GoBifo mean effect for SCA #3 (18 outcomes in total)** | | **0.00** | **–0.032** | **(0.045)** | **236** |
| 12 | Proportion of communities that held a meeting after the research team left to discuss what to do with the tarp | 0.98 | –0.03 | (0.02) | 233 |
| 13 | Proportion of communities that stored the tarp in a public place | 0.06 | 0.06 | (0.04) | 225 |
| 14 | Proportion of communities that had used the tarp by the follow-up visit (5 months after receipt) | 0.90 | –0.08+ | (0.04) | 233 |
| 15 | Given use of the tarp, proportion of communities that used the tarp in a public way | 0.86 | 0.02 | (0.05) | 161 |

Note:
Significance levels denoted by + p < 0.10; * p < 0.05; ** – p – < 0.01
Robust standard errors
Treatment effects estimated on follow-up data
Includes fixed effects for the district council wards (the unit of stratification).
Each specification includes the following control variables: total households per community, distance to nearest motorable road, index of war exposure and index of history of domestic slavery.

The second group of outcomes under H5 concerns the respondents' experience and opinions regarding how the choice was made, which were recorded immediately after the meeting, including the proportion of these randomly selected individuals who attended the meeting, whether they spoke publicly and their opinions about who had the final say and how much

the decision was dominated by local elites. The third set was measured during the follow-up visit and concerns respondents' and local leaders' views on how the decisions were made concerning how to share the salt/batteries, what to do with the tarp, whether to use the building materials vouchers and if so how to mobilise funds, choose items and manage the resulting construction. For both of these sets, the analysis shows no systematic differences between how decisions are made in treatment as compared with control communities.

These findings about participation and inclusion deserve careful consideration. As mentioned earlier, GoBifo ward facilitators actively encouraged women and youths to participate in all aspects of project planning, meetings and implementation, and measured the frequency and intensity of their participation in internal review processes. Women and youths were further required to serve in leadership positions for the project (VDC and account signatories) and had the opportunity to manage their own sub-projects. However, despite all of this and the perception among project staff that women were participating more, we do not see any change in their role in community decision-making outside the project itself. Even for a simple measure like speaking up in meetings, the ongoing encouragement over more than three years of project facilitation did not translate into a greater voice in deciding whether to take the salt or the batteries, how best to use the tarp or how to take advantage of the voucher opportunity.

Referring to a research question outside the explicit project objectives, the mean effect index for H6 in Table 4 shows no systematic effect of GoBifo on the local systems of authority. Most outcomes under this hypothesis explore the extent to which the village head and elders dominated the decisions about whether to choose the salt or the batteries, how to share the resulting item amongst the community members, how to use the tarp and what to do with the building materials vouchers. While we find large variation in how these decisions are made – at one extreme, in two villages the chief decided between the salt and the batteries in less than one minute without anyone else's input, while at the other an open discussion lasted nearly an hour and was followed by a formal vote – there are no systematic differences in whether people feel that everyone had an equal say in these matters, or if someone had a disproportionate say, whether that person was a traditional authority or an elder. However, considering the pool of outcomes in which the respondents answered the same questions about local council/WDC members as well as chiefdom officials reveals some interesting trends about the relative perceptions of the two spheres of local government. This set includes questions about how much respondents feel that officials in the council [chiefdom] listen to what people in their community say, how effectively the council [chiefdom] spends public funds, how much they trust council [chiefdom] representatives and the involvement of the council [chiefdom] officials in the implementation of nine local public goods. In absolute terms, the respondents generally answer these questions more favourably with respect to chiefdom officials: they trust chiefdom officials more, think that they listen more closely to their needs and report that they are more involved in overseeing the community's stock of public goods. Interestingly, however, in GoBifo areas the gap in perceptions of chiefdom versus council/WDC officials is narrowing: while the traditional authorities still hold a popular margin in these areas, the council and WDC members are catching up. This suggests that the links and opportunities for community members to work together with their local elected officials created by GoBifo leaves respondents with a higher opinion of local council and WDC members than they had before.

A general risk inherent in the CDD emphasis on devolving control over project finances and choices is that local elites will use their authority and influence to capture benefits for themselves. The research programme thus explored this issue through the third structured community activity involving the tarp. While the analysis finds no treatment effects on the extent of elite capture, it also reveals that the level of elite capture is relatively low in the communities under study (at least according to this measure). As an example, for the 90

per cent of communities that had used the tarp by the time of the second visit (row 14 in Table 9), 86 per cent had put the tarp towards a public purpose (row 15). The most obvious example of elite capture is to use the tarp to patch the roof of a single individual's house, which happened in less than 3 per cent of these communities. That said, several communities had not yet used the tarp and were storing it at a private residence. The risk of capture for these communities is higher: by storing the tarp where it is not publicly available, local elites may be preventing the community from benefiting from the tarp, even if they have not yet put it to use for themselves directly. Alternatively, the storage may simply reflect the failure to make a final decision on the tarp's use or the ability to raise additional funds to put it to a use supported by the community.

Despite exploring a wealth of measures, the analysis reveals no GoBifo impacts on any of the remaining three proxies for social capital – trust, groups and networks, and access to information – emphasised in the project objectives. Beginning with trust, of the 12 measures collected, only one treatment effect is statistically distinguishable from zero (H7 in Table 4). Encouragingly, this significant positive effect is on respondents' reports about how much they trust NGOs and donor projects, which suggests that GoBifo field staff gained the trust of the communities in which they operated and enhanced their confidence in this class of external actors. There are no other effects on the remaining indicators, which combine respondents' self-reports regarding how much they trust various groups with concrete examples of trusting behaviour, like giving money to a neighbour to purchase goods on their behalf.

For groups and networks (H8 in Table 4), the research team collected multiple measures of household-level engagement for a set of nine different groups. Specifically, for each kind of group, the enumerators asked the respondents whether they were a member, and if so, whether in the past one month they had attended a meeting, contributed financially or contributed any labour. The list of groups included a credit/savings group, communal labour gang, school committee, social club, savings for special events like funerals or weddings, fishing cooperative, seed multiplication group, women's group and youth group. The analysis combines these indicators with a couple of more general network measures, like helping a neighbour rethatch their house. Of the 40 treatment effects estimated under this hypothesis, only three are significant at traditional levels and two of these work to cancel each other out: positive impacts on membership of seed multiplication and women's groups, plus a negative impact on contributing labour to women's group activities. Taken as a whole, this is robust evidence that GoBifo did not exert any spillover influence on group membership or the intensity of group participation.

Turning to the final social capital proxy (H9 in Table 4), the analysis finds no evidence of treatment effects on participants' access to information about local governance. Of the 21 treatment effects considered, only one – an increase in the proportion of villages visited by a WDC member – is statistically significant. The collection of zero effects includes measures of how much respondents know about what the community is doing with the vouchers and tarp; whether they can name their leaders in the local council and chiefdom system; and how much people know about the taxes collected and how they are used.

In contrast to the rest of the hypotheses in this family, the mean effect index for H10 in Table 4 reveals a positive and significant impact on participation in local government. This positive finding is driven by project impacts in two major areas: establishing local institutions and creating links between communities and their local leaders (as seen in Family A: Project Implementation). As expected from project activities, the analysis reveals strong impacts on the existence of VDCs and VDPs. Notably, however, there are no effects on the overall proportion of women and youths with VDC membership, nor were the treatment communities any more likely to use the tarp and building materials for the activities specified in their development plan. As shown for participation and inclusion, this

latter disconnection between the establishment and the use of the VDP for activities outside GoBifo suggests that the communities are not using the project structures and tools for initiatives beyond the programme. Expanding on earlier findings regarding linkages, GoBifo attracted the attention of not only WDC members but also chiefdom officials. For the nine local public goods considered, the treatment effect concerning the involvement of the section or paramount chief in its planning, construction and oversight is positive in sign for seven and at least marginally significant for three goods. For the third significant grouping in this category – voting and running for local office – there are no impacts. More specifically, GoBifo had no impact on the likelihood that respondents voted in the national and local elections or stood for offices within the chiefdom or local government system.

The survey evidence suggests no impact on the level of crime and conflict in communities or the mechanisms through which they are resolved (H11). Of the 10 indicators considered, only one – a reduction in household reports of physical fighting over the past one year – is statistically significant. While the nine null results imply that the project's efforts to enhance the conflict management capacity may not have created large impacts, on the positive side it provides some reassurance that the infusion of cash grants into the community did not exacerbate conflicts over limited resources. That said, there is a marginally significant negative effect on the proportion of respondents who reported that they had ever had a conflict over a loan or other 'money business'. Considering dispute resolution, we find no change in the probability that a community resolves a conflict itself without the assistance of external actors.

The twelfth and final hypothesis in the software family concerns political and social attitudes. While not a direct project objective, the programme's emphasis on the empowerment of women and youths and the transparency and accountability of local institutions may have engendered a more liberal outlook toward politics and social dynamics more generally. The enumerators gauged attitudes using pairs of largely opposing statements, for example 'As citizens, we should be more active in questioning the actions of leaders' versus 'In our country these days, we should have more respect for authority', and by asking the respondents which one they agreed with more. There are no systematic treatment effects on such opinions.

### 4.3 Robustness

The previous section highlights strong positive impacts on the creation of local development institutions, the stock and quality of local public goods, economic welfare and linkages between communities and local leaders. It further presents evidence that the programme did not have an impact on social capital or the dynamics of decision-making. What are the possible threats to the validity of these results?

To start, let us examine the typical threats to randomised experiments that would imply that the positive results are biased upwards. Fortunately, there were no problems with treatment non-compliance: all the communities assigned to the treatment group received the programme and none of those in the control group participated. Also, the results in Table 3 suggest that the randomisation process successfully created two quite similar groups. If there is concern that the treatment group looked slightly better off at the start, remember that the analysis used the baseline values of the outcome of interest as a control variable wherever panel data were available. Thus, in order for spurious differences between the two groups to explain the positive impacts, the treatment group would have to have been better off enough to be on a steeper development trend than the controls. None of the differences in Table 3 appear large enough to substantiate this argument.

Next, we consider the reasons why the abundance of zero results might be underestimates. First, if there were significant spillovers from the treatment to the control communities, the results above would underestimate the impact of the programme, since the control

communities were also receiving benefits. However, for this to be true, we would expect the co-efficient on the post time period to be positive and significant for the outcomes of interest. This is not the case: across all the outcomes in Table 4 where panel data are available, there are exactly as many (21) positive as negative co-efficients on the time trend that are significant. Thus, it seems unlikely that these results are downward biased due to spillovers.

A second concern is that the projects GoBifo implemented at the ward level systematically benefited the control group at the expense of the treatment group. This would be perfectly understandable if, for example, local council members took into account the placement of GoBifo village-level projects in deciding where to locate ward projects and targeted those areas that did not already benefit. First, even if this is true, note that there are several hundred villages in these wards that did not receive a GoBifo project yet were also not part of the control group. Thus, these ward-level projects would need to have very broad coverage and large impacts to swamp the village-level impacts in the treatment group. Second, if anything, it appears that the treatment villages knew more about and received more benefits from the ward-level projects than the controls. Specifically, 27.7 per cent of the respondents in the treatment villages were able to name a GoBifo ward project compared with only 16.6 per cent of the respondents in the control areas. Similarly, while 15.2 per cent of the respondents in the treatment areas reported that a member of their household directly benefited from a ward-level project, only 6.1 per cent of the respondents in the control areas reported benefits. Thus, it also seems unlikely that GoBifo ward-level projects led to an underestimation of the village-level project impacts.

The third and final major critique would be that the measures we used for the various dimensions of social capital and community dynamics were simply not good enough or too blunt to detect subtle differences between the treatment and the control communities. While some measures are certainly better than others, our main strength lies in the diversity and multiplicity of the instruments used and the fact that they all produce similar results. First, we combined survey and non-survey data collection methods. As an example, the tables above report both the percentage of male and female respondents who said they spoke up during the salt/batteries deliberation from the survey data and enumerator accounts of how many men and women they saw speaking from direct participant observation. Second, the research teams gathered information from a variety of sources. In each community, the field teams interviewed men and women in the privacy of their own homes, held a focus group discussion with key opinion leaders, observed a community decision as it unfolded and made their own assessment of the materials used in and the quality of construction of local public goods. Lastly, for each hypothesis, we examine a large number of outcomes. Taking these aspects together, the zero treatment effects we estimate for the hypotheses in Family C are quite precisely estimated. This means that any true effect that we incorrectly ruled out would have been very small in magnitude. To illustrate with a specific outcome, note that on average approximately six people made a public statement during the salt/batteries deliberation and that 29 per cent of these speakers were women. The *largest* true positive treatment effect that we may have incorrectly rejected at the 95 per cent confidence level is a 0.038 increase in the proportion of public speakers who were women, or one additional female speaker per every 4.3 villages visited, which is quite small.

## 5. Policy recommendations

The World Bank and government of Sierra Leone requested this project and research agenda to test innovative approaches to building social capital and linking citizens and community-level development efforts with local government. They indicated that if the results were sufficiently favourable, they would consider scaling up the programme.

Our findings suggest that GoBifo is a reasonable mechanism to deliver local public goods efficiently and at a low cost. Indeed, the magnitude of the improvements in public infrastructure and economic outcomes, the even and widely dispersed distribution of benefits across the community and the low leakage of project funds are noteworthy in the challenging environment of rural Sierra Leone. These results suggest that GoBifo could serve as an effective complementary mechanism to deliver smaller-scale goods at the community level under the NaCSA umbrella currently operating at the level of chiefdom section/local council ward. It has also been highly successful in connecting citizens and community development projects with local leaders in the WDCs and chieftaincy system. These achievements align well with its placement under the Decentralisation Secretariat and suggest that it could be an effective bottom-up approach to nurturing local government.

However, despite a wealth of measures, the evaluation finds no evidence that GoBifo is an effective method to strengthen social capital or fundamentally alter local hierarchies and decision-making processes in this context. In this sense, it does not advance the 'attitudinal change' agenda of President Ernest Bai Koroma. It should further not be seen as a vehicle specifically to empower women or youths in local development. While we find no treatment effects on participation, inclusion and collective action *beyond* the direct project sphere, it is important to note that we cannot rule out whether GoBifo's emphasis on participatory methods was in part responsible for the positive impacts on public goods and economic welfare accomplished *within* the project. Building on this point, it would be extremely valuable to assess the relative benefit of facilitation as opposed to block grants both for GoBifo and for CDD more generally. The key question that this evaluation cannot address is whether the programme would be just as effective if the budget balance was shifted towards less facilitation and more grants to communities. As discussed early in this report, the current rate of spending one dollar on facilitation, administration and oversight for every dollar given to communities is quite intensive. The question is whether that first dollar is necessary, or would the communities perform just as well with only 50 cents' worth of facilitation? Finding the right balance could be easily achieved by varying the intensity of facilitation across areas as the programme expands.

Beyond Sierra Leone, this evaluation speaks to the broader debate regarding the effectiveness of CDD. Recognising that this study concerns one project, in one country, it is worth pausing to compare our results with those from other evaluations of CDD and with other approaches to institutional change pursued elsewhere. We find that CDD boosted public goods and economic welfare, yet did not enhance collective action, the voice of marginalised groups or participation in decision-making in Sierra Leone. In sharp contrast, the preliminary results from a randomised study of CDD in Liberia find little evidence for hardware benefits accompanied by gains in software as measured by greater contributions to a public goods game. Some of the differences in results may be due to differences in context (the Liberia project operated in areas with more internally displaced people and greater disruption to local institutions), and it is important to note that neither study found evidence of CDD spillovers onto real-world, non-project collective activities including contributions to existing public goods and attending or speaking up in community meetings. Looking around the world at examples of institutional change that have been found to be successful by randomised impact evaluations, perhaps the most promising are those generating change through elections – mandating quotas for female leaders and reducing caste (ethnicity)-based voting in elections. Both of these studies were conducted in India

and it would be worth testing whether similar approaches would be effective in an African context.

The results from our study suggest that the comparative advantage of the World Bank and similar external donors may lie in implementing development hardware, not instigating social change, at least not as currently delivered through CDD. Changing social norms and community dynamics is important but arguably incredibly difficult. This evaluation of what has become one of the standard approaches for trying to influence social norms and community dynamics suggests that there is much that we, as outsiders, do not yet know about how to undertake this effectively.

## Appendix A Sampling details

This section discusses the sample selection process in more detail. It explains: i) the on-site randomisation process for the eight research communities in Bonthe Town; and ii) the partial resampling of the research areas. Importantly, none of these issues systematically affect either treatment or control communities, and thus should not bias the treatment effect estimates.

While nearly all of the randomisation was conducted electronically, the randomisation process for the island communities of Bonthe Town (which comprise 3.4 per cent of our sample) was conducted manually by a public lottery. Since there was no community list from the Statistics Sierra Leone (SSL) 2004 Population and Housing Census available for the town to use in the computerised process, the estimation of community size and selection into treatment and control groups was completed on-site. As the town sections are roughly comparable in size to villages in Bonthe District, the project team treated each section as a separate community. Also, since the wards in Bonthe Town are substantially smaller than those in the district, GoBifo decided to intervene in only two (instead of six) communities per ward. The research team thus wrote the names of all the sections in the target wards on individual pieces of paper and drew the four project and four control sites from a box in the presence of the Town Section Heads, district councillors and two independent observers.

As explained in the paper, community-level eligibility for the GoBifo programme was determined by: i) the total number of households (20 to 200 households in Bombali District and 10–100 in Bonthe); and ii) location within a targeted ward in one of the two districts. At the time of sample selection, the most up-to-date information on community size was from the 2004 census. As the census data entry process was still ongoing, the only electronically available measures were the pre-census cartographic team estimates of the total households per locality. We thus used this measure to eliminate communities that were too small or large, and on 17 October 2005 conducted an initial randomisation on the resulting eligibility pool to select 228 villages, composed of 114 treatment and 114 control communities (or six treatment and six control communities in each of 19 wards). The first three days of fieldwork surveyed 32 communities from this initial list. However, the field team reported non-trivial differences in the community size estimates from the cartographic team and what they encountered in the villages, frequently off by 50 or more households. The research team thus manually generated a new list of total households for all communities in the target wards using the hard copies of the 2004 census enumeration area summary books. Using this more accurate measure of the total households per village to define a new pool of eligible villages, while retaining the 32 villages already surveyed due to budget reasons, we conducted a second randomisation on 18 November 2005 for the remaining 196 villages.

There were 11 communities in this second sample for which SSL was unable to locate the full census books, and was therefore unable to compile household listings. One further village was found to be empty as it was not a permanent settlement. Replacements for these 12 communities were randomly sampled from the respective wards and randomly divided into treatment and control.

Nine selected communities in one particular chiefdom were assigned to the wrong local council ward. As background, since the 2004 census had not yet been completed, the National Electoral Commission had to rely on old census data in drawing ward boundaries for the 2004 district council elections. This process created some confusion on the ground concerning which chiefdom sections individual councillors represented. Direct reports from the relevant district councillors revealed that nine communities from our sample needed to be replaced in order to retain the balance of six treatment and six control communities in

each of the three targeted wards in one chiefdom (Gbendembu Ngowahun). The research team made the necessary replacements by randomly selecting communities from the respective wards and randomly dividing them into treatment and control.

It is important to note that all of these steps reflect adjustments to the sample and none of them compromise the integrity of the random assignment.

## Appendix B SCA protocols

**Community Development Survey 2009: Gift experiment module – supervisor instructions**

**STEP 1: Start the meeting when the community is ready.** Allow the community to tell you when it is ready to start the meeting. If appropriate, suggest that the meeting starts with prayers.

**STEP 2: Introduce yourselves.** After the prayers, introduce yourselves briefly.

**STEP 3: Ask the community to introduce themselves.** Allow the community to introduce as many of its members as it chooses. This is when your enumerators should count how many women/adults/youths are present at the beginning of the meeting.

**STEP 4: Present the gift choice.** Read the following script:

> *We are from Statistics Sierra Leone, in Freetown. You may remember that in 2005, our team came to your community and asked people some questions as part of a research project. After this gathering, we would like to ask some questions again, to see how things have changed in the past 4 years.*
>
> *But first, we would like to give you a gift. Apart from helping us with our work today, we do not expect anything from you in exchange for this gift. The gift is just a 'thank you' for helping us with our research. Please note that we will NOT be coming back to your community with any development projects; this is just a thank you for your help today.*
>
> *We have two different gifts, but only have enough supplies to give you one. We would like the community to choose between the two gifts. You should make this decision however you want. How the gift is divided among community members is also your choice – we do not want to tell you what to do with the gift or how to share it.*
>
> *While you are deciding which gift you would like, we will stand to the side. Take as much time as you need, and then tell us when you have reached a decision.*
>
> *The first gift is salt. There are 4 large bags of salt, each of which contains 20 smaller bags. Each small bag contains 2 pounds of salt. (Show one small bag of salt). This salt is different from locally made salt in an important way. This salt has been treated with very small amounts of a chemical called iodine. Iodine is not there in salt that is boiled down from seawater, which is how local salt is made. Iodine is very important for health. If pregnant women or small children do not eat enough iodine, it is bad for their brain and for their intelligence. Also, not enough iodine can cause goiter (Krio: gege) for older*

*people (explain goiter if necessary). Therefore, iodized salt has important health benefits for everyone.*[2]

*The second gift is a carton of batteries. There are 144 large vinnic batteries (show one battery).*

**STEP 5: Step back dramatically from the meeting and observe.** After you have presented the gift choice, move to the edge of the meeting area. Remain there and observe the decision-making process until the community reaches its decision. As you are observing the decision-making process, pay special attention to any public discussions.

**STEP 6: Keep track of all public speakers.** Use the tally sheet here on your Supervisor Section survey (G-S) form to keep track of each person who speaks during the meeting. When a person speaks for the first time, write down the colour of his or her clothing to help you remember who it is, and use a check to indicate the age group (18–35 or 35+) and gender of the person.

**STEP 7: Give the community the tarpaulin.** After the decision is announced, read the following script:

*Okay, great. On top of this gift, we would also like to give the community a tarpaulin. This is yours to use in any way you choose. Do you have any questions about the tarpaulin?*

**STEP 8: Present cards.** After any discussion of the tarpaulin concludes, read the following script:

*We have another gift for the community. Here are 6 cards that the community can use to buy building materials (show six cards).*

*This is how the cards work. Each card is worth SLL50,000, but they are like phone top-up cards. You must activate them by adding SLL100,000 of your own money. You go to the store with the card, and you bring SLL100,000 cash, and you can buy SLL150,000 of building materials. Without SLL100,000 of your money, the card is not activated and is worth nothing.*

- *If you go to the store with 1 card and SLL100,000 cash, you can buy SLL150,000 of building materials.*
- *If you go to the store with 2 cards and SLL200,000 cash, you can buy SLL300,000 of building materials.*
- *If you go to the store with 3 cards and SLL300,000 cash, you can buy SLL450,000 of building materials.*
- *If you go to the store with 4 cards and SLL400,000 cash, you can buy SLL600,000 of building materials.*
- *If you go to the store with 5 cards and SLL500,000 cash, you can buy SLL750,000 of building materials.*
- *If you go to the store with all 6 cards and SLL600,000 cash, you can buy SLL900,000 of building materials.*

*But the cards are not good forever. You must use them on or before July 1, otherwise they will expire. After July 1, you can throw away any cards you have not used.*

[2] Note that early on in the fieldwork the supervisors were instructed to equalise the amount of things they said about the salt and the batteries. They thus shortened the description of salt in this paragraph and added a discussion about how the batteries were useful for powering flashlights and radios.

*It is for you to decide how to raise the money to activate the cards, which building materials to buy and what to use the materials for.*

*You can take the card to [shop names] in [Makeni/Bo]. The names and addresses of the shops are also written on the cards. When you go to the shop and use the cards, they will take the card, so it can only be used once. Also, after you use the cards to buy building materials, the store will give you a receipt which will list all the things you have bought.*

*Just so we know you understand everything, we would like one community member to explain back to us how the cards work.*

**STEP 10: Quiz the community about the cards.** Ask one member of the community to explain the card system back to you. Make sure the explanation includes the following points:

1. That each card is worth SLL50,000 and that the community has to contribute SLL100,000 to activate each card.
2. That without matching funds, the cards are not valid.
3. That the cards will expire after July 1.

If there are any points you think they may not fully understand, re-explain these points. Once you are satisfied that the community understands the card system, read the following script.

*Great. In case you have any questions or problems, we will give you a phone number which you can call for more information [this is on the community receipt]. Now, who wants to take the cards?*

Give the cards to whomever the community designates. If you do not know already, make sure you ask what position the person holds within the community. This is when your enumerators should count how many women/adults/youths are present at the end of the meeting.

**STEP 11: Have the community sign for the goods.** Fill in the receipt for the community, and ask someone from the community to sign for receipt of the goods. Make sure you record the position this person holds within the community. Sign and give the community their portion of the gift receipt (note the contact information listed).

**STEP 12: Explain the remainder of the interview process.** Tell the community about both the village and the household questionnaires.

**STEP 13: Record your Observations: Fill in the rest of the survey here**

After the meeting, fill in the rest of the Supervisor Section (G-S) based on your own memory and observations. Confer with the other enumerators if necessary.

## Appendix C *Ex ante* analysis plan and econometric specifications

In this report, we present the results for the specific hypotheses described in our *ex ante* analysis plan, a document that was finalised before we analysed any follow-up data. The genesis of the plan was a pre-programme 2005 agreement between the research and project teams that set out the areas GoBifo was likely to have an impact on and how success in these areas would be assessed. Building on this early document, we drafted a formal analysis plan that specified the exact outcomes under each of 11 hypothesised areas of impact and the econometric specifications to be used, which we archived with the Abdul Latif Jameel Poverty Action Lab randomised evaluation archive in August 2009. This approach limits data mining or *ex post* rationalisation that selectively highlights only positive impacts (or negative) effects discovered during analysis.

Towards this end, the plan has several components. First, it defines both the sets of explanatory and dependent variables (Leamer 1983) and the econometric models (Leamer 1974) before the start of the data analysis. While the randomised framework naturally imposes much of this narrowing (i.e. the treatment indicator is the leading explanatory variable), the plan also details the set of interaction terms we would use to explore heterogeneous treatment effects. Second, the large number of outcome variables we consider means that several individual treatment effects will be statistically significant due simply to random chance. To account for this, the plan commits us to a mean effects approach that reduces the effective number of tests we conduct by identifying in advance which outcome variables would be grouped together to identify jointly the different hypotheses laid out in the 2005 document (see the detailed description of the mean effects approach in the next section). While the mean effect index is the primary metric by which we evaluate a given hypothesis, we also provide results for the outcome measures individually to provide a better sense of the magnitude and economic significance of our results. Third, for further transparency, we disclose the complete results for all 318 outcome variables considered, including the exact wording of the survey question, in a supplementary web appendix.

There are two minor deviations from the original *ex ante* analysis plan in the presentation in this report. We added a twelfth hypothesis (called H1 above) by pulling together project implementation outcomes that had already been explicitly included as outcomes within the original 11 hypotheses. Thus, no new outcome measures were added to or excluded from what we present below. Those who wish to consider only the results as exactly laid out *ex ante* can ignore H1. However, we feel it was an oversight to exclude a project implementation hypothesis beforehand and thus still find the results of H1 useful to consider. Perhaps more important is that we group the 12 hypotheses into three 'families' for ease of comprehension and to facilitate links to the theory. While we did not specify these families beforehand, we believe that the groupings – project implementation (family A), development 'hardware' (family B) and the 'software' of local collective action (family C) – are intuitive.

Under each hypothesis, we evaluate specific treatment effects using the following model:

$$Y_c = \beta_0 + \beta_1 T_c + X'_c \Gamma + W'_c \Pi + \varepsilon_c \qquad (11)$$

where $Y_c$ is a given outcome (i.e. local road maintenance) in community $c$; $T_c$ is the GoBifo treatment indicator; $X_c$ is a vector of the community-level controls, including those used to assess treatment versus control group balance in the original computer randomisations; $W_c$ is a fixed effect for geographic ward, the administrative level on which the randomisation was stratified; and $\varepsilon_c$ is the usual idiosyncratic error term. Elements of $X_c$ include the distance from a road, the total number of households, an index of violence experienced during the recent civil war and a measure capturing the historical extent of domestic

slavery. The parameter of interest is $\beta_1$, the average treatment effect. Note that while some outcomes are measured at the household (for example radio ownership) or individual level (for example political attitudes), for simplicity we measure all the variables at the village level, taking village averages as necessary; analysis at the household level yields nearly identical results (not shown).

For the subset of outcome variables that were collected in both the baseline 2005 survey and the 2009 follow-up surveys, the analysis exploits the panel data structure:

$$Y_{ct} = \beta_0 + \beta_1(T_c * POST_t) + \beta_2 T_c + \beta_3 POST_t + X'_c \Gamma + W'_c \Pi + \varepsilon_{ct} \quad (12)$$

Where $Y_{ct}$ is a particular outcome for community *c* at time *t,* where *t = 0* in the 2005 baseline survey and *t = 1* in the 2009 follow-up. The additional indicator variable *POST* denotes the follow-up period. The parameter of interest is again $\beta_1$, the average treatment effect, and here the disturbance terms are clustered at the village level. The results are robust to the exclusion of the vector of community controls and to limiting our analysis to only the post-programme data. We further assess the degree of heterogeneous treatment effects by including interaction terms of treatment with gender, age, village remoteness, community size, war exposure, the local history of domestic slavery and location in each of the two study districts. As we do not find any evidence of heterogeneous effects along any of these dimensions, we have excluded this discussion from the main text.

## Appendix D Explanation of the mean effects indices

For each area of interest, this report investigates several distinct outcome measures and uses a mean effects index to estimate the average treatment effect of GoBifo on the hypothesis overall. The usefulness of a single summary measure in the presence of multiple related outcome variables has roots in the medical trial literature, in which for example O'Brien (1984) estimates the overall efficacy of a new diabetes treatment on 34 different measures of nerve function. The exposition below follows Kling and Liebman (2004), who apply this strategy to an evaluation of the 'Moving to Opportunity' demonstration programme implemented in several U.S. cities in the 1990s. The programme provided subsidised housing vouchers to families living in high-poverty public housing projects, thereby creating an opportunity to move into relatively more affluent neighbourhoods. Focused on youth welfare, the authors examine the impact of the programme on multiple indicators grouped into 'families' of wellbeing. For example, the category of risky behaviour includes reductions in drug use, smoking, alcohol consumption and pregnancy. Each variable is evaluated on its own and as a component of the mean effects index for its respective family.

Correctly estimating an average effect requires a method that accomplishes two main things. First, it must standardise the binary (i.e. 'yes' or 'no' outcomes) and continuous outcomes (i.e. total Leones contributed to an *osusu* group or the number of days since the last community meeting) into comparable units. This is straightforwardly achieved by translating each outcome into standard deviation units (i.e. by subtracting its mean and dividing by its standard error as measured for the control group) before regressing each outcome on the vector of independent variables. One can thus calculate the *magnitude* of the mean effects index without further manipulation by simply taking the average of the $K$ co-efficients on GoBifo across the standardised outcome equations.

Secondly, the estimation method must calculate the standard error of the average index itself, which involves co-efficients from multiple equations. Note that the ordinary least squares equation-by-equation approach does not provide any information about the covariances between estimators from the $K$ distinct equations. The variance of the mean index, however, depends on both the variances of each individual $\beta_k$ as well as any covariances between $\beta_k$ and $\beta_{\neg k}$. To thus obtain the complete variance–covariance matrix requires a seemingly unrelated regressions (SUR) system approach, which stacks each of the $K$ outcome equations on top of one another and estimates the entire system simultaneously. With the complete covariance matrix in hand, we can then test the cross-equation hypothesis that the average index of $K$ co-efficients does not equal zero.

More specifically, let $Y_k$ refer to the $k$th outcome within a single hypothesis. As usual, the effects on each individual outcome are estimated by regressing the outcome on the independent variable of interest, $Z$, as well as a vector of control variables, $X$.

$$Y_k = X\beta_k + Z\pi_k + \nu_k = W\theta_k + \nu_k$$

The summary index for the hypothesis overall, $\tau$, is constructed by averaging the standardised treatment effect sizes from each of the $K$ outcome equations, where

$$\sigma_k^2 = Var(Y_k)$$

$$\tau_k = \frac{\pi_k}{\sigma_k}$$

$$\tau = \frac{1}{K}\sum_{k=1}^{K} \tau_k$$

Then the complete variance–covariance matrix across the $K$ equations is computed using SUR estimation on the following system:

$$Y = (I_K \otimes W)\theta + \nu$$

where $I_K$ is a $K$ by $K$ identity matrix and $Y = (Y_1', Y_2', \ldots, Y_K')'$. This technique stacks each of the individual outcome equations on top of one another and estimates the entire system simultaneously.

$$\begin{matrix} Y_1 \\ \vdots \\ Y_K \end{matrix} = \begin{pmatrix} W_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & W_K \end{pmatrix} \begin{matrix} \theta_1 \\ \vdots \\ \theta_K \end{matrix} + \begin{matrix} \nu_1 \\ \vdots \\ \nu_K \end{matrix}$$

While this stacking approach does not change the estimated magnitude or variance of any individual treatment effect, it now incorporates potential covariance between the error terms (and estimated co-efficients $\theta_k$) across equations. Since these multiple outcomes are all measured for the *same* households or communities, this approach allows us to leverage cross-equation correlation in the error terms due to unobservable differences between units.

Finally, using the $\tau_k$ elements within $\theta_k$, the variance of the mean effects index is calculated straightforwardly as:

$$Var(\tau) = Var\left(\frac{1}{K}\sum_{k=1}^{K} \tau_k\right) = \frac{1}{K^2}\sum_{i=1}^{K}\sum_{j=1}^{K} Cov(\tau_i, \tau_j)$$

Note that if there is no covariance across equations, the variance of the mean index would simply be $1/K$ times the average variance of the individual treatment effects.

## References


Banerjee, A.V., Banerji, R., Duflo, E., Glennerster, R. and Khemani, S., 2010. Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India. *American Economic Journal: Economic Policy*, 2 (1), pp.1–30.

Bjorkman, M. and Svensson, J., 2009. Power to the people: Evidence from a randomized field experiment on community-based monitoring in Uganda. *Quarterly Journal of Economics,* 124 (2), 735–769.pp.

Casey, K., Glennerster, R. and Miguel, E., 2011. *Reshaping institutions: Evidence on external aid and local collective action*. NBER Working Paper No. 17012.

Fearon, J., Humphreys, M. and Weinstein, J.M., 2009. *Development assistance, institution building, and social cohesion after civil war: Evidence from a field experiment in Liberia*. Center for Global Development Working Paper 194.

GoBifo project, 2008. *Results towards goals and objectives*. Project mimeograph.

Kling, J.R. and Liebman, J.B., 2004. *Experimental analysis of neighborhood effects on youth*. Princeton University manuscript.

Labonne, J. and Chase, R., 2008. *Do community-driven development projects enhance social capital? Evidence from the Philippines*. The World Bank Policy Research Working Paper 4678.

Leamer, E.E., 1974. False models and post-data model construction. *Journal of the American Statistical Association*, 69 (345), pp.122–131.

Leamer, E.E., 1983. Let's take the con out of econometrics. *American Economic Review,* 73 (1), pp.31–43.

Mansuri, G. and Rao, V., 2004. Community-based and -driven development: A critical review. *The World Bank Research Observer,* 19 (1), pp.1–39.

O'Brien, P.C., 1984. Procedures for comparing samples with multiple endpoints. *Biometrics,* 40 (4), pp.1,079–1,087.

Olken, B.A., 2007. Monitoring corruption: Evidence from a field experiment in Indonesia. *Journal of Political Economy,* 115 (2), pp.200–249.

Richards, P., Bah, K. and Vincent, J., 2004. *Social capital and survival: Prospects for community-driven development in post-conflict Sierra Leone.* Community-Driven Development Conflict Prevention and Reconstruction Paper No. 12, pp.1–65.

United Nations, 2004. *Human development report 2004*. New York, United Nations Development Program, Oxford University Press.

Voss, J., 2008. *Impact evaluation of the second phase of the Kecamatan development program in Indonesia*. The World Bank, Jakarta Research Paper.

World Bank, 2004. *Japan Social Development Fund grant proposal: Capacity development to strengthen social capital in Sierra Leone*. Project mimeograph.

World Bank, 2007. *World development report 2008: Agriculture for development.* Washington D.C., The World Bank.

## Publications in the 3ie Impact Evaluation Report series

The following reports and other 3ie publications are all available from http://www.3ieimpact.org/publications/3ie-impact-evaluations/

*The promise of preschool in Africa: A randomised impact evaluation of early childhood development in rural Mozambique*, *3ie Impact Evaluation Report 1*. Martinez, S., Naudeau, S. and Pereira, V. (2012)

*A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana*, *3ie Impact Evaluation Report 2*. Burwen, J. and Levine, D.I. (2012)

*The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone*, *3ie Impact Evaluation Report 3*. Casey, K., Glennerster, R. and Miguel, E. (2013)

*Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India, 3ie Impact Evaluation Report 4*. Meenakshi, J.V., Banerji, A., Mukherji, A. and Gupta, A. (2013)

*Impact evaluation of the non-contributory social pension programme 70 y más in Mexico*, *3ie Impact Evaluation Report 5.* Rodríguez, A., Espinoza, B., Tamayo, K., Pereda, P., Góngora, V., Tagliaferro, G., Solís, M. (2013)

*The impact of daycare on maternal labour supply and child development in Mexico*, *3ie Impact Evaluation Report 6*. Angeles, G., Gadsden, P., Galiani, S., Gertler, P., Herrera, A., Kariger, P. and Seira, E. (2013)

*Social and economic impacts of Tuungane: final report on the effects of a community-driven reconstruction programme in eastern Democratic Republic of Congo*, *3ie Impact Evaluation Report 7*. Humphreys, M., Sanchez de la Sierra, R., van der Windt, P. (2013)

*Paying for performance in China's battle against anaemia*, *3ie Impact Evaluation Report 8*. Zhang, L., Rozelle, S. and Shi, Y. (2013)

*No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia*, *3ie Impact Evaluation Report 9*. Ashraf, N., Bandiera, O. and Jack, K. (2013)

*Truth-telling by third-party audits and the response of polluting firms: Experimental evidence from India, 3ie Impact Evaluation Report 10*. Duflo, E., Greenstone, M., Pande, R. and Ryan, N. (2013)

**Impact Evaluation Report Series**

International Initiative for Impact Evaluation
c/o Global Development Network
2nd Floor, West Wing, ISID Complex
Plot No. 4, Vasant Kunj Institutional Area
New Delhi – 110070
India

3ie@3ieimpact.org
Tel: +91 11 4323 9494

**www.3ieimpact.org**